

FUNCIÓN BIOLÓGICA Y CONTENIDO MENTAL

TESIS DOCTORAL

ANDRÉS LUIS JAUME RODRÍGUEZ

DIRECTORES:

DR. MIGUEL ÁNGEL QUINTANILLA FISAC
Catedrático de Lógica y Filosofía de la Ciencia
Universidad de Salamanca

DR. ANTONIO MANUEL LIZ GUTIÉRREZ
Profesor Titular de Lógica y Filosofía de la Ciencia
Universidad de La Laguna

Universidad de Salamanca. Facultad de Filosofía
Departamento de Filosofía y Lógica y Filosofía de la Ciencia

Programa de doctorado de Lógica y Filosofía de la Ciencia

MMIX

Contenidos

Introducción	4
---------------------	----------

Capítulo I. Nueva y vieja teleología	11
---	-----------

1. Sobre el concepto de teleología en general
2. Esbozo histórico de las doctrinas teleológicas
 - 2.1. La teleología clásica: Platón y Aristóteles
 - 2.1.1. La teleología platónica
 - 2.1.2. La teleología aristotélica
 - 2.2. La recepción escolástica de la teleología
 - 2.3. Teleología y Modernidad
 - 2.4. Pensamiento darwiniano y teleología
3. La pervivencia de la teleología en el pensamiento filosófico y científico contemporáneo
 - 3.1. El problema de las explicaciones funcionales en Filosofía de la ciencia
 - 3.1.1. Teleología y explicación funcional
 - 3.1.2. El modelo nomológico deductivo de Hempel y Oppenheim y sus problemas con las explicaciones teleológicas y funcionales
 - 3.1.3. La solución de Nagel: la inexistencia de explicaciones genuinamente teleológicas o funcionales y la traducción al modelo de cobertura legal
 4. De explicaciones funcionales a enunciados adscriptores de función en Biología
 - 4.1. El enfoque histórico-etiológico de L. Wright
 - 4.2. El enfoque sistémico de R. Cummins
 - 4.3. La herencia de Wright y Cummins
 5. Comentarios finales sobre teleología y funcionalidad

Capítulo II. Sobre el concepto de función biológica. El estado actual de la cuestión	73
---	-----------

1. La polémica entre las dos teorías y el debate actual
2. La teoría disposicional de Bigelow y Pargetter
3. Los continuadores de la teoría histórico-etiológica
 - 3.1. La crítica de Boorse
 - 3.2. La teoría de las funciones propias de Millikan
 - 3.3. La estrategia analítica de K. Neander
 - 3.4. Godfrey-Smith. Explicaciones funcionales en términos de historia reciente
 - 3.5. Bedau o la dimensión valorativa de la teoría de las funciones
4. Los intentos de instanciación
 - 4.1. Griffiths
 - 4.2. La teoría relacional de Walsh
 - 4.3. La teoría débil de Buller
5. Los desarrollos recientes de la teoría sistémica: P.S. Davies
6. ¿Pluralismo o unificación?
7. El análisis de Mayr

Capítulo III. Adaptaciones y normatividad

127

1. Adaptación y función biológica
 - 1.1. Adaptación y Selección Natural
 - 1.2. Exadaptaciones y Adaptacionismo
 - 1.3. La relación entre las adaptaciones y las funciones
 - 1.4. Conclusiones sobre las relaciones entre los conceptos de función y adaptación
2. Funciones y normatividad
 - 2.1. El rechazo de la normatividad dentro de la teoría sistémica
 - 2.1.1. La crítica de Davies a la neoteleología: la imposibilidad de las ejecuciones erróneas de las funciones
 - 2.1.2. Crítica de Cummins a la neoteleología: ni adscripciones funcionales ni normatividad
 - 2.1.3. Consecuencias del rechazo de las nociones normativas
 - 2.2. Nomologitud y normatividad
 - 2.2.1. Sobre el concepto de ley científica en el neopositivismo
 - 2.2.2. La crítica de Dretske a la concepción neopositivista de las leyes de la naturaleza
 - 2.3. ¿Por qué es necesaria una teoría normativa de las funciones?
 - 2.4. Funciones sistémicas normativizadas

Capítulo IV. Una caracterización del contenido mental

180

- Introducción
- I. Esbozo histórico del problema: La caracterización tradicional del contenido mental
 - II. Una presentación sistemática del contenido y sus dimensiones
 1. Contenido mental, contenido lingüístico y lenguaje del pensamiento
 - 1.1. La caracterización de nuestras actitudes
 - 1.2. La teoría representacional de la mente y el lenguaje del pensamiento
 - 1.3. Pensamiento y lenguaje
 2. La individuación del contenido mental
 - 2.1. La teoría ideacional del significado y la caracterización internista tradicional del contenido
 - 2.2. Los argumentos externistas
 - 2.3. Internismo y externismo a debate
 3. Normatividad
 4. La caracterización informacional del contenido y la dicotomía entre lo conceptual y lo no conceptual
 5. Sistemática, composicionalidad y holismo
 6. Recapitulación

Capítulo V. Teorías teleosemánticas del contenido mental

244

- I. Introducción: la caracterización teleológica del contenido
- II. Estudio de cuatro teorías teleosemánticas

1. La semántica causal o del indicador: Dretske
 - 1.1. La primera teoría causal de Dretske: cuando Dretske no era teleosemántico
 - 1.1.1. Información y conocimiento
 - 1.1.2. Información y contenido. La definición del contenido informacional de una señal
 - 1.1.3. Contenido informacional y actitudes proposicionales
 - 1.2. La caracterización teleológica del contenido en Dretske
 - 1.2.1. Tipos de sistemas representacionales
 - 1.2.2. La explicación del error
 - 1.2.3. El papel de los procesos de aprendizaje
 - 1.2.4. Cómo se fija el contenido
2. Semánticas basadas en el consumidor *down up*
 - 2.1. La teleosemántica de Millikan
 - 2.1.2. Signos, intencionalidad y representaciones
 - 2.1.3. ¿Cómo fijar el contenido? Propiofuncionalidad y contenido mental
 - 2.2. Neander: *Low Church* frente a *High Church*
 - 2.2.1. La teleología *Low Church*
 - 2.2.2. La determinación del contenido
 - 2.2.3. La explicación del error
 - 2.2.4. Comentarios finales
3. Papineau: el enfoque *top down*
 - 3.1. La alternativa *top down*
 - 3.2. La determinación del contenido
4. Comentarios finales: Una comparación de las teorías expuestas en este capítulo

Capítulo VI. Las limitaciones del proyecto teleosemántico

298

- Introducción: Los problemas de las teorías teleológicas del contenido
1. La inadecuación del concepto de función biológica
 2. El problema de la indeterminación
 3. La independencia de la polémica entre internismo y externismo y el caso del hombre del pantano
 4. La cuestión del contenido abstracto
 - 4.1. Naturaleza del contenido abstracto
 - 4.2. Contenido abstracto y aprendizaje
 5. Utilidad biológica y contenido mental
 6. La cuestión del contenido no conceptual
 7. La teleosemántica no explica la vida mental al completo

Capítulo VII. Conclusiones

334

Apéndice

345

Bibliografía

355

Introducción

1. Motivaciones y tesis principal

Las publicaciones en lengua española sobre el concepto de función biológica y la caracterización teleológica del contenido mental son escasas, tanto tomando ambos tópicos conjuntamente como por separado¹. De hecho, ambos pueden estudiarse de manera autónoma; así el análisis del concepto de función biológica pertenece a la Filosofía de la Biología y el estudio de las teorías teleológicas del contenido a la Filosofía de la mente, rama que en nuestro país ha sido cultivada con mayor profusión. Sin embargo, esta tesis estudia ambos temas. La razón hay que buscarla en el uso que hacen las teorías teleológicas del contenido mental del concepto de función biológica. Para estas teorías el referido concepto es fundamental, pues permite tanto una naturalización de la intencionalidad como mantener las propiedades normativas de conceptos y actitudes. Una crítica de las teorías teleológicas del contenido, o teorías teleosemánticas, debe proceder a través de un análisis exhaustivo del concepto de función biológica del que dependen. Así, la tesis principal que defiende es la siguiente: el concepto de función biológica es insuficiente para explicar la naturaleza y fijación del contenido mental. Esto me permite, por una parte y tras haber clarificado el estado actual de la cuestión en Filosofía de la Biología, desarrollar un análisis substantivo del concepto de función biológica y de la normatividad que le adscribimos y, por la otra, una serie de críticas paralelas a las diversas teorías teleosemánticas que presento y clarifico para su análisis conforme a un criterio constituido por seis desiderata a propósito de lo que cabe esperar de una teoría del contenido mental.

¹ Sirvan de ejemplo las siguientes referencias: Gomila (1995), Quesada (1995) o Jaume (2007).

2. Estructura del trabajo

Este trabajo tiene dos partes. La primera de ellas, que comprende los tres primeros capítulos, versa sobre el concepto de función biológica y constituye el punto de partida para el estudio que se desarrolla en la segunda parte. En esta primera parte cada capítulo presenta un estadio de la reflexión en torno a la problemática del concepto de función biológica que sirve como preliminar para el siguiente. En este sentido, creo que la estructura en la que están dispuestos responde a un determinado orden de las razones. El primero de ellos, *Nueva y vieja teleología* presenta tanto una tematización muy general como unos lineamientos históricos en torno a la teleología. Así, los problemas más generales con los que nos encontramos en el análisis del concepto de función biológica no son sino los problemas suscitados por la teleología que ya tienen una larga historia.

El segundo capítulo, *Sobre el concepto de función biológica. El estado actual de la cuestión* tiene un carácter eminentemente expositivo. Mi objetivo al redactarlo ha sido triple. En primer lugar presentar la génesis contemporánea de las diferentes teorías científico-filosóficas que sobre el concepto de función biológica se han formulado. En este respecto se puede afirmar con toda seguridad que dichas teorías proceden, para bien o para mal, del análisis que en su día Hempel presentara de la explicación teleológica y funcional. Así, desde Hempel los filósofos no hemos hecho otra cosa que reformular un argumento explicativo, perdiendo cada vez más de vista la situación metateórica desde la que Hempel se movía. En cierto sentido hemos dado un paso ilegítimo al querer solventar los problemas intrínsecos al modelo nomológico-deductivo y nos hemos metido de lleno en un terreno más propio de la Biología teórica. Aunque quizás este libertinaje científico-filosófico no nos haya ido tan mal... En segundo lugar, he creído oportuno presentar un estado de la cuestión lo más completo posible que no sólo respondiera al criterio histórico, sino que también lograra dar cuenta de las principales teorías actuales que se han vertido sobre el concepto de función biológica. En este respecto mi exposición parte de los trabajos seminales de Wright y Cummins en la

década de 1970 para finalizar con algunas contribuciones recientes ya de nuestro siglo. Por último, me ha parecido necesario introducir una breve referencia hacia otros ámbitos en los que el concepto de función acuñado desde la Biología ha merecido cierta notabilidad; el ámbito de lo artefactual y el problema del contenido mental. El primero de ellos pertenece por completo a la Filosofía de la tecnología y en este trabajo no hago referencia alguna al mismo, pues además de que se aparta de la línea de investigación general, ya lo abordé en otro trabajo académico presentado en 2005. En cuanto al segundo, constituye la segunda parte de la presente tesis doctoral.

En el tercer capítulo, *Adaptaciones y normatividad*, se abordan algunos tópicos que he creído importantes y que no podían ser pasados por alto en una reflexión filosófica en torno al concepto de función biológica. Pero, por otra parte, también constituye mi particular contribución al debate. Así, en este capítulo desarrollo dos tesis substantivas como son la equiparación del concepto de función al de adaptación biológica y la explicación naturalista de la normatividad que adscribimos a las primeras.

La segunda parte está por completo dedicada al estudio de la teleosemántica una vez ya aclarado el concepto de función biológica que es central para estas teorías. El estudio de este conjunto de teorías lo desarrollo de la siguiente manera. En el capítulo IV, *Una caracterización del contenido mental* presento este tópico ateniéndome a una doble estrategia. Mi caracterización tiene dos partes: una histórica y otra que denomino sistemática. En la parte histórica he querido resaltar la dependencia que las actuales teorizaciones acerca del contenido tienen respecto de la tradición moderna. Cuando oí por vez primera la palabra "contenido mental" me pareció una absoluta novedad y a día de hoy todavía me sorprende que cuando empleo dicho término ante alumnos o colegas procedentes de otras tradiciones filosóficas, me observen como si hablara de algo que no pertenece a su mundo filosófico. La presentación histórica tiene por objeto destruir esa impresión. El tema del contenido mental muy bien puede encajar en lo que algunos manuales del s. XIX consignaban como *ideología* y es obvio que la importancia del racionalismo y el empirismo no se agota en considerar a ambos movimientos como simples antepasados de nuestras reflexiones actuales. La parte sistemática tiene por objeto

presentar de una manera organizada aquellos aspectos que se consideran relevantes para una teoría del contenido. En los textos de Filosofía de la mente al uso no es frecuente encontrar un capítulo que desarrolle sistemáticamente la agenda de problemas de la teoría del contenido, así que he creído necesario hacerlo yo mismo.

En el capítulo V, *Teorías teleosemánticas del contenido mental* mi proceder es paralelo al que ya empleé en el capítulo II. El objetivo de este capítulo es eminentemente descriptivo, ya que trato de presentar una taxonomía suficientemente exacta acerca de las diferentes teorías teleosemánticas del contenido, aunque también se procura poner de relieve algunas dificultades. El objeto de estudio de este capítulo se circunscribe a las aportaciones de cuatro filósofos muy representativos en este campo como son Dretske, Millikan, Neander y Papineau. Aunque también existen otros enfoques teleosemánticos como puedan ser los de Sterelny o Price, pero la discusión actual se ha centrado de manera específica y sistemática sobre la obra de estos cuatro.

El capítulo VI, *Las limitaciones del proyecto teleosemántico* presenta una crítica sistemática de las teorías teleológicas del contenido mental. Esta crítica depende de los seis *desiderata* que se proponían en el capítulo IV y que una teoría del contenido mental debería satisfacer. Así abordo diversas cuestiones acerca de las carencias del proyecto teleosemántico como puedan ser la inadecuación del concepto de función biológica para el problema de la determinación del contenido, la independencia respecto de la polémica entre internismo y externismo, o las diversas dificultades con las que estas teorías se encuentran para explicar otros aspectos relevantes de nuestra vida mental. Finalmente, en el capítulo VII presento las conclusiones de esta investigación.

3. Metodología

A pesar de que este trabajo analiza un determinado enfoque naturalista del contenido mental, es este un trabajo eminentemente filosófico. Significa esto que no se va a encontrar en él referencia a ningún dispositivo experimental nuevo. No hago *Filosofía experimental* ni creo que ésta exista. Mi proceder es el propio de la tradición filosófica; el análisis de los argumentos que otros filósofos y científicos

presentan. Si bien aquí el punto de partida es la ciencia natural, hay que señalar que es una ciencia natural muy alejada de lo que hacen la gran mayoría de los científicos y que se circunscribe a una gran teoría como es la teoría sintética de la evolución, teoría que se adopta como marco de la explicación dentro de las Ciencias de la vida pero que ha sido apenas abordada desde un punto de vista estrictamente teórico por la gran mayoría de los científicos. Así, mi trabajo se ha centrado en el análisis conceptual de las interpretaciones mayoritariamente acometidas por filósofos. Aunque también en esta discusión las teorizaciones llevadas a cabo por los teóricos de la Biología han sido tenidas muy en cuenta.

4. Agradecimientos

Miguel Ángel Quintanilla ha sido la condición de posibilidad de esta tesis y a él debo el poderme haber dedicado todos estos años a mi investigación. Ha tenido la desgracia de tener que aguantar como doctorando a un aprendiz híbrido entre metafísico y filósofo analítico que se presentaba en su despacho para pedirle que le dirigiera la tesis a la vez le decía que no quería saber nada de CTS. Durante todos estos años ha contribuido con buenos comentarios oportunos y una disposición de ánimo excelente hacia mis inquietudes y diversas empresas filosóficas.

Conocí a Manuel Liz por mediación de Miguel Ángel Quintanilla, que desde el principio supuso que teníamos intereses comunes y que muy bien podría orientar mi trabajo académico. Manuel Liz ha dirigido impecablemente mis pasos en la investigación y con muy buenas ideas. Margarita Vázquez, anfitriona excelente en mis visitas a La Laguna, ha visto crecer esta tesis desde la nada y siempre me ha alentado. A ambos les estoy agradecido en grado sumo, me han tratado siempre muy generosamente y con enorme comprensión y cariño. Igualmente a los miembros del grupo LEMA (María Ponte, David Pérez, Juanjo Colomina, Gemma Robles, Luis Artiles y Rafa Herrera) de la Universidad de La Laguna, con quienes disfruto enormemente de hacer filosofía, y a Ulises, que tiene ya una solución a la paradoja del mentiroso.

A José Luis Luján le debo que me iniciara en la Filosofía de la Biología nada más empezar primero de carrera en 1998 y que en 2007 facilitara mi adscripción al departamento de Filosofía de la Universidad de las Islas Baleares. Jamás sospeché que buena parte de mi actividad académica la dedicaría a este campo y sin su introducción al mismo muy posiblemente no hubiera tomado la decisión de abordar los problemas que aquí presento. Desde que estoy en la Universidad de las Islas Baleares me he beneficiado enormemente de sus consejos y buen olfato filosófico. A Oliver Todt, que ha soportado pacientemente los nervios de un doctorando desde los inicios de mi etapa en Mallorca y que siempre me ha animado durante mi investigación. Mis compañeros de área, Andreu Berga, uno de los primeros en presentar en este país una tesis sobre Filosofía de la Biología, y Tomeu Adrover, con quienes paso muy buenos ratos.

Sebastián Álvarez Toledo no sólo me recibió como decano en el año 2000 cuando llegué a Salamanca, sino que, además, guió mis primeros pasos en Filosofía de la Ciencia y en todos estos años de doctorado ha constituido una valiosa ayuda y un estímulo constante.

A Ignacio Vicario, que amablemente me permitió dar clases de la que ahora es mi asignatura y que comparte el mismo interés que yo por la filosofía analítica.

Debo agradecer que algunos buenos profesores como Lluís Pujadas, Jesús Vega y Paco Salto me iniciaran en su día en la reflexión filosófica propia de la tradición analítica, tradición y metodología que ha guiado mi trabajo en los últimos años. Tampoco me gustaría dejarme en el tintero a otros profesores y colegas, tanto de mi período en Baleares como en Salamanca que, aunque no lo crean, con sus clases y comentarios han aportado su granito de arena como Miquel Beltrán, Joan Lluís Llinàs, Antoni Gomila, Paco Torres, Diego Sabiote, Roberto Albares, Mariano Álvarez, Enrique Bonete, Cirilo Flórez, José Luis Fuertes, Antonio Heredia, María Teresa López, María Manzano, José Luis Molinuevo, María del Carmen Paredes, Ricardo Piñero, J.L. Rodríguez Molinero (†), Laureano Robles o David Teira. Así como a mis amigos y compañeros de Filosofía de la Universidad de Salamanca, Julio Ostalé, Adrián Pradier, César Tejedor y Giuseppe Tuffano. A todos ellos les estoy muy agradecido.

A Ernest Sosa, que posibilitó mi estancia en Brown y supuso un gran avance en este trabajo.

En mis queridos amigos, Juan Zurita, Pedro Ribas, Juan Antonio Horrach, Toni Deyá, José Antonio Fuster y Mercedes Prieto que siempre me han animado con palabras de aliento y confianza.

Paco Funcia, Antonio Hermelando y Nana Villalonga, que contribuyeron a que mi estancia en Salamanca fuera maravillosa. A Manoli Escribano y Baldo Oreja y su estupenda familia, mi *familia* de Salamanca.

A mis padres, Pura y Joaquín, mi querida familia y mi tío Luis, que ha facilitado muchos de los viajes necesarios para la realización de este trabajo. A Vega, que pacientemente ha aguantado todo el proceso de elaboración y que sin su cariño y comprensión no habría sido posible. Todos ellos se han ocupado de mí en las diferentes facetas de mi vida y siempre han creído en lo que yo hacía. A Leo y a Atina, constante compañía en las horas de trabajo en solitario y fuente de inspiración en mis reflexiones sobre el pensamiento sin palabras y la fijación del contenido.

Me licencié en Filosofía el mes de junio de 2003 en la Universidad de Salamanca. Desde hacía unos cuantos años, y a medida que avanzaba en mis estudios, me fui interesando por cuestiones metafísicas, si bien nunca acabó de gustarme el modo en que los problemas que yo creía que correspondían a esta disciplina se trataban. Por otra parte, la Filosofía de la Ciencia me fue interesando cada vez más y pude ver cómo muchos de los problemas que hasta entonces me habían atraído iban encontrando un espacio más afín a mis intereses en el seno de las discusiones que estudiaba en esta disciplina. Finalmente decidí escribir la presente tesis doctoral tras presentar un trabajo de grado en septiembre de 2005 sobre el concepto de función en ciencia y tecnología dirigido por Miguel Ángel Quintanilla. Desde entonces he tratado de moverme entre dos aguas: la Filosofía de la ciencia, más concretamente de la Biología y la Filosofía de la mente. Fruto de este doble interés es la presente tesis doctoral que dedico a la memoria de mis dos abuelos difuntos, Andrés Jaime Rovira y Luis Rodríguez Sosa, de quienes llevo sus respectivos nombres y a quienes con toda seguridad les habría gustado verlo, colmando muchas de sus inquietudes vitales como padres de familia.

I. Nueva y vieja teleología

Contenidos

1. Sobre el concepto de teleología en general
2. Esbozo histórico de las doctrinas teleológicas
 - 2.1. La teleología clásica: Platón y Aristóteles
 - 2.1.1. La teleología platónica
 - 2.1.2. La teleología aristotélica
 - 2.2. La recepción escolástica de la teleología
 - 2.3. Teleología y Modernidad
 - 2.4. Pensamiento darwiniano y teleología
3. La pervivencia de la teleología en el pensamiento filosófico y científico contemporáneo
 - 3.1. El problema de las explicaciones funcionales en filosofía de la ciencia
 - 3.1.1. Teleología y explicación funcional
 - 3.1.2. El modelo nomológico deductivo de Hempel y Oppenheim y sus problemas con las explicaciones teleológicas y funcionales
 - 3.1.3. La solución de Nagel: la inexistencia de explicaciones genuinamente teleológicas o funcionales y la traducción al modelo de cobertura legal
4. De explicaciones funcionales a enunciados adscriptores de función en Biología
 - 4.1. El enfoque histórico-etiológico de L. Wright
 - 4.2. El enfoque sistémico de R. Cummins
 - 4.3. La herencia de Wright y Cummins
5. Comentarios finales sobre teleología y funcionalidad

Nuestra imagen científica del mundo a menudo se presenta como incongruente con nuestros planteamientos de sentido común. El sentido común nos da buenas razones para confiar en cómo es y de qué manera funciona nuestro mundo más inmediato; el conocimiento científico, el hermano rico del primero, desconfía de la apariencia externa de los fenómenos y nos propone una imagen de la realidad más refinada. En muchos casos esta imagen es difícil de creer si no es porque, con frecuencia, aparece como garante del éxito en nuestra actividad práctica. Si tenemos que hacer caso a Bacon, éste fue uno de los argumentos principales que contribuyeron a modificar una cierta concepción del mundo mayoritaria entre muchos de los académicos de su época. En su *Novum Organum* (1620) vemos reflejada la idea de que el conocimiento no tiene un valor meramente contemplativo tal como lo presentaba Aristóteles, sino también práctico: conocer es poder. La idea de Bacon acerca de la naturaleza del conocimiento y, de manera más particular, de su poder, llama a una aceptación de los métodos que permiten obtenerlo. El recurso al poder como argumento es tremendamente seductor y pocos pueden escapar a sus encantos.

Hoy, queremos pensar, sabemos más que ayer y, por supuesto, menos que mañana. Las ideas germinales del humanismo renacentista cuajan en la Ilustración en torno a un ideal de progreso del conocimiento y de la vida humana. Curiosamente el finalismo de la historia reclamado por la Ilustración y que podemos rastrear hasta los inicios del cristianismo y que quizás adquiere plena entidad filosófico-teológica con el *De civitate Dei* de Agustín de Hipona, sólo afecta a las obras de los hombres en su desarrollo histórico y no a la naturaleza tal como se tematizaba en la filosofía aristotélica. La Modernidad extirpa la teleología inherente a la naturaleza y la injerta en el desarrollo de la historia de los hombres. Así, entre las teorías que el progreso científico ha ido dejando en la cuneta está la teleología, la idea de que muchas entidades y procesos están dirigidos hacia un fin. La teleología ha sido una doctrina defenestrada en numerosas ocasiones que, sin embargo, vuelve a resurgir de un modo más o menos explícito, según la ocasión y el autor, a la hora de justificar nuestras asunciones acerca del mundo vivo.

En el siglo XX y en el contexto de la Filosofía analítica de la ciencia y de la mente, la teleología ha protagonizado un debate muy

interesante sobre el contenido mental y su explicación en términos de función biológica. Ambas disciplinas, de manera más o menos explícita, reproducen una tensión que muy fácilmente puede explicarse como la discusión entre los dos hermanos que al principio hacía referencia; el sentido común y nuestra imagen científica del mundo. Así, lo que está en juego en esta disputa familiar tiene que ver con cómo conciliar cierta visión de sentido común sobre la realidad biológica con la perspectiva científica. El debate es interesante no porque pueda suponer una caza de brujas o un proceso inquisitorial sobre el criptoteleologismo de algunos de los conceptos principales de la Biología actual, sino porque afecta a nuestra comprensión sobre lo mental.

En este capítulo presento un análisis de las principales doctrinas teleológicas, su desarrollo histórico y los problemas actuales que éste nos lega. Mi objetivo, así, es preparar el terreno para una discusión filosófica sobre una serie de tópicos más concretos.

1. Sobre el concepto de teleología en general

Ojalá la tarea filosófica de aclarar términos se resolviera acudiendo a un buen diccionario. Si alguien busca en un diccionario de uso del español probablemente no logre tener una idea muy clara de lo que este término significa y, cuando vea la referencia a la palabra "Filosofía" junto al término, salga espantado y cierre de golpe el libro. La definición normativa del término circunscribe éste a la doctrina de las causas finales². De manera más explícita, María Moliner³ sitúa el uso del término en el contexto de la Metafísica y señala que esta subdisciplina se ocupa de las causas finales o de la finalidad en general y la contrapone a las concepciones mecanicistas. Ninguna de las dos definiciones son suficientes para el estudio filosófico, aunque ambas señalan lo que ya sabíamos; la teleología tiene que ver con los fines, de manera más, explícita, con cualquier fin.

Por teleología habitualmente entendemos un proceso o cosa que está dirigido a un fin. Sin embargo, conviene aclarar qué procesos o cosas son susceptibles de estar dirigidos a fines y en qué sentido

² Cf. la vigésimo segunda edición del DRAE (2001).

³ Cf. la segunda edición del DUE de María Moliner (1999).

decimos que lo están. Las ideas teleológicas o finalistas pueden interpretarse desde puntos de vista diversos, estableciéndose dos grandes grupos polarizados. Por una parte tenemos una teleología o finalismo antropomórfico que como doctrina metafísica sostiene que el mundo está intencionalmente organizado o es la creación de una inteligencia divina. El calificativo de “antropomórfico” quizás no es del todo afortunado y se emplea para referirse más bien a lo que podría denominarse “teomórfico”. En cualquier caso lo que se quiere subrayar es la acción intencional por parte de un ser inteligente y con capacidad volitiva de disponer un orden preestablecido en el mundo. Históricamente buenos ejemplos lo constituyen el *Timeo* de Platón o el *De ordine* de Agustín de Hipona. Este tipo de finalismo antropomórfico encuentra defensores entre creacionistas y partidarios del argumento del diseño inteligente. Tesis más cercanas a la creencia religiosa que a la especulación científico-filosófica y que nacen únicamente de la voluntad de creer. No pienso que deban ser tenidas en cuenta en la presente reflexión ni posiblemente constituyan un problema filosófico relevante. Por la otra parte se puede señalar un amplio grupo de doctrinas teleológicas no antropomórficas, al menos no en el sentido señalado anteriormente, aunque no siempre exentas de intencionalidad. Este segundo grupo es muy numeroso y comprende un dominio de objetos muy amplio que es estudiado por las diferentes disciplinas empíricas. Sobre este segundo grupo creo que sí vale la pena filosofar y va a constituir la reflexión principal de este capítulo.

Usamos términos finalistas en una diversidad de contextos bastante amplia. Una investigación sobre el uso del término puede indicarnos qué parcelas de la realidad todavía hoy emplean nociones teleológicas. Considérense los siguientes ejemplos:

- (1) Voy a la nevera para coger una fruta
- (2) Salió disparado en cuanto oyó la noticia.
- (3) Su intención es llegar a ser ministro.
- (4) Las ratas huyen de los gatos por pura supervivencia.
- (5) La finalidad de los riñones consiste en filtrar la sangre.
- (6) La lordosis lumbar de la rata indica su disposición para la cópula.

(7) Los registros del órgano tienen diferentes funciones expresivas, para accionarlos hay que tirar del accesorio correspondiente al mismo situado a uno u otro lado del teclado.

(8) Un destornillador sirve para ajustar o desajustar tornillos.

Acudir a fines es una estrategia explicativa habitual que cubre gran parte de nuestra experiencia cotidiana. Los ejemplos citados anteriormente muestran la diversidad de ámbitos en los que aplicamos esta estrategia. Sin embargo, no es nueva: ha sido seguida por Platón, Aristóteles, Leibniz, Kant... y así hasta nuestros días.

Admitimos que los órganos, las conductas, los artefactos tienen funciones y las entendemos en sentido finalista. Seres vivos y artefactos son, pues, el tipo de entidades susceptibles de adscripciones funcionales o finalistas. No cambia nada decir que la función de X es Y o afirmar que el fin de X es Y. Los ejemplos aducidos al principio muestran diferentes registros de nuestro uso de conceptos teleológicos en los distintos ámbitos del saber.

Así (1), (2) y (3) son explicaciones de la conducta en términos de fines. Es el uso preteórico o popular de la teleología como explicación de la conducta intencional. Los enunciados (4), (5) y (6) muestran el uso de estas nociones en un ámbito científico como es el de las ciencias de la vida. Por último, (7) y (8) se aplican al ámbito de la tecnología o a sus creaciones.

En el ámbito intencional esto no parece representar ningún problema. La conducta humana propositiva es una conducta dirigida a fines. La explicación intencional, en la que se involucra la causación mental, es un tipo de explicación teleológica particular. Las razones, desde el sentido común, son causas. Por analogía extrapolamos este tipo de finalismo al ámbito de la ciencia natural, los ejemplos (4), (5) y (6) son suficientemente ilustrativos. Sin embargo, es aquí donde empiezan los problemas filosóficamente relevantes. ¿Qué garantías tiene esta presunta finalidad de ser intrínseca a las entidades de las que se predica y no una mera analogía respecto de la experiencia cotidiana? El riesgo de este tipo de argumentos radica en un sesgo antropocéntrico. Si admitimos que las razones son causas de la conducta parece que también vamos a conceder que otros seres vivos tienen algún tipo de "razones para actuar". Si se trata de animales

superiores no parece tan problemático. Pero, a medida que descendemos a través de la *Scala Naturae* van aflorando nuevos problemas. Invirtamos ahora el orden. No vamos a mirar desde *arriba a abajo* sino desde *abajo a arriba*. Voy a adelantar una de las cuestiones que motivan la presente investigación. Supongamos pues que en la naturaleza hay fines, es obvio que no pueden ser intencionales ¿cómo explicamos entonces la propia intencionalidad? De otra manera y ya en el ámbito de la Filosofía de la ciencia, ¿qué hay de bueno en recurrir a argumentos teleológicos y cómo pueden cohabitar con una visión científica del mundo.

2. Esbozo histórico de las doctrinas teleológicas

Parece ser que el término “teleología” fue empleado por primera vez por Ch. Wolff⁴. Sin embargo, hay que señalar que la idea se remonta hasta mucho atrás. Lo que presento en las siguientes páginas no es una historia de la teleología, sino una serie de reflexiones sobre algunos autores relevantes con el fin de ilustrar el problema que creo ver en el tratamiento actual de la misma. A saber, ¿existe algo así como una teleología natural o simplemente es un procedimiento heurístico al que no hemos podido renunciar ni parece que lo vayamos a hacer?

2.1. La teleología clásica: Platón y Aristóteles

La idea de Bien pasó a la tradición como uno de los denominados trascendentales. De lo que es, según admitían los escolásticos, podía predicarse el bien, la verdad y la belleza; *bonum, verum et pulchrum*. La idea de bien es de raigambre platónica. Sin embargo, en tanto ésta no es sino una concreción del *eidos* podemos pensar que Aristóteles es deudor de la misma y, aunque de una manera más inmanente, también la emplea. Tanto uno como otro admitirán respectivamente que bien los diferentes eventos que acaecen, en la medida en que son, tienden a un bien propio o que nada ocurre siguiendo una causalidad ciega; todo ocurre con fines al cumplimiento de un fin natural.

⁴ Wolff *Philosophia rationalis sive logica* III, §83. Cf. Ferrater Mora (1981: 3205).

Con la idea de bien o de finalidad podemos ver una injerencia de los asuntos humanos en la esfera natural. No es la primera. El mito constituye un intento previo de dar razón de los fenómenos acudiendo a diversas antropomorfizaciones. De algún sitio tenemos que sacar los conceptos para una explicación del mundo y es obvio que la única posibilidad es el reciclaje de los conceptos cotidianos. El hecho de que se recurra a materiales del ámbito cotidiano más inmediato para la explicación de lo que acaece debería darnos qué pensar a propósito de las relaciones entre sentido común y ciencia. La idea de bien no tiene la exclusiva en estas relaciones. Un ejemplo tomado del ámbito jurídico y extrapolado a la ciencia y la filosofía es la causalidad. La causalidad no deja de ser un concepto antropocéntrico de origen jurídico. *Aitía* es causa, pero también culpa, *aitiós* es el culpable. El responsable cuando se comete un delito. Como se ve no es ninguna solución neutra.

2.1.1. La teleología platónica

Platón aborda el tema en el *Fedón*, el pasaje conocido como segunda navegación, y en el *Timeo*. Si bien no puede decirse que haya una doctrina teleológica unitaria en Platón, puesto que tanto el *Fedón* como el *Timeo* pertenecen a períodos distintos dentro de la filosofía platónica. Así, mientras que el *Fedón* es una obra del Platón medio, el *Timeo* pertenece a los diálogos de madurez.

Aunque la temática del *Fedón* es principalmente el tema de la inmortalidad del alma a raíz de los momentos finales de Sócrates, también encontramos en este diálogo una defensa de la doctrina de las ideas y, de modo más concreto, una defensa de la teleología. Así, en este diálogo encontramos un importante pasaje (95e y ss.) en la que se argumenta favor de la consideración de la finalidad. Esta argumentación, de carácter autobiográfico, es subsidiaria respecto de la tesis principal del diálogo, si bien puede ser considerada de modo autónomo. En este pasaje Sócrates relata a Cebes -discípulo del pitagórico Filolao- sus experiencias con las doctrinas materialistas y su posterior desencanto; Sócrates fue partidario de Anaxágoras y su teoría del *nous* que todo lo ordena, pero pronto se desencantó:

"Pero de mi estupenda esperanza, amigo mío, salí defraudado, cuando al avanzar y leer veo que el hombre no recurre para nada a la inteligencia ni le atribuye ninguna causalidad en la ordenación de las cosas, sino que aduce como causas aires, éteres, aguas y otras muchas cosas absurdas." Platón, Fedón, 98b-c.

El problema de las doctrinas materialistas es que sólo explican cómo suceden las cosas y no por qué suceden. Sócrates/Platón se aparta así del ámbito de los fenómenos para estudiar las razones a la luz de la teoría de las ideas. Es lo que se conoce como segunda navegación (*deuteros plouos*).

"En cambio, la facultad para que estas mismas cosas se hallen dispuestas del mejor modo y así estén ahora, ésa ni la investigan ni creen que tenga una fuerza divina, sino que piensan que van a hallar alguna vez un Atlante más poderoso y más inmortal que éste y que lo abarque todo mejor, y no creen para nada que es de verdad el bien y lo debido lo que cohesiona y mantiene todo. Pues yo de tal género de causa, de cómo se realiza, habría sido muy a gusto discípulo de cualquiera. Pero después de que me quedé privado de ella y de que no fui capaz yo mismo de encontrarla ni de aprenderla de otro (...) Opiné, pues, que era preciso refugiarme en los conceptos para examinar en ellos la verdad real. Ahora bien, quizás eso a lo que lo comparo no es apropiado en cierto sentido. Porque no estoy muy de acuerdo en que el que examina la realidad en los conceptos la contemple más en imágenes, que el que la examina en los hechos." Platón. Fedón 99e- 100a.

El anterior pasaje es muy rico en ideas. En él critica las doctrinas materialistas y se decanta por el finalismo. El argumento a favor del finalismo lo encontramos en las primeras líneas al afirmar que el mundo exhibe una organización que no puede justificarse por medio del azar, sino que claramente apunta a un bien determinado. Este rasgo de organización óptima de la realidad no puede ser explicado adecuadamente basándose en el recurso a las explicaciones materialistas, sino que requiere de un principio organizador que en este caso es la teoría de las ideas. En la afirmación "*En cambio, la facultad para que estas mismas cosas se hallen dispuestas del mejor modo y así estén ahora (...)*." encontramos los primeros atisbos de la teleología platónica que será desarrollada de un modo más amplio en el Timeo. El finalismo platónico no es intrínseco a la naturaleza sensible, sino que depende de la existencia de las ideas o formas como causas. El argumento que emplea en este caso Platón es muy sutil y establece

una analogía entre actuar por causas y actuar por razones⁵ y sus principales tesis epistemológicas⁶. En el pasaje del Fedón 99b encontramos como Sócrates considera insuficiente la apelación a la materia como justificación de las acciones, para ello es necesario también una causa formal o final. Así, del mismo modo que Sócrates condenado a beber cicuta no huye sino que permanece por emulación a la idea de justicia, la realidad también imita a esas formas que son necesarias para la justificación del conocimiento y de la acción. De este modo las formas o ideas son tanto los principios del conocimiento como de la realidad, así se justifica también la segunda navegación como estrategia antimaterialista. Pese a las dudas manifestadas por Sócrates/Platón respecto de la estrategia de la segunda navegación al final del fragmento citado cabe pensar que el recurso a la teoría de las ideas parece estar mejor justificado que otras estrategias basadas en el conocimiento sensible⁷.

Los argumentos del Fedón están claramente tintados por una determinada concepción casi religiosa de origen místico. Aunque el tema principal sea demostrar la inmortalidad del alma no hay que obviar las cuestiones estrictamente filosóficas que se plantean. Entre ellas no solamente está la defensa de la teleología, sino que podemos encontrar también importantes argumentos epistemológicos como la diferencia entre genuino conocimiento y opinión. El genuino conocimiento no puede encontrar su instancia justificatoria en los sentidos, sino en las razones. Así Sócrates, con su rechazo a las doctrinas materialistas estaría presentando una versión de lo que hoy conocemos como el mito de lo dado.

En el *Timeo*, Platón es más crítico con la teoría de las ideas. Las ideas o formas ya no son causa formal, eficiente y final como sucedía en el Fedón, sino sólo causa formal, habiéndose convertido la causalidad eficiente y final en los designios de un Demiurgo ordenador

⁵ Cf. Platón, *Fedón* 99a-b.

⁶ El pasaje de la segunda navegación también justifica una nueva estrategia epistemológica; no se trata tanto de fiarse del conocimiento sensible como del conocimiento inteligible. Los sentidos son engañosos, en ellos no hallamos además el carácter de necesidad que pretendemos para la realidad, así, la garantía epistémica de la que nos proveen es muy débil.

⁷ Cf. Guthrie (1996) Vol. IV, pág. 339.[1962].

del mundo a partir de un material originario⁸. En este diálogo el elemento dialéctico que presenciábamos en el anterior es todavía más débil, relegando la explicación acerca del origen del mundo a un mito de carácter ciertamente esotérico. Platón presenta en este diálogo al mundo creado por el Demiurgo como un gran organismo. Así el mundo incorpora belleza (29 a) y orden (30 a); el mundo es así producto de la inteligencia del Demiurgo que copia un modelo perfecto (28 a). De este modo hay fines conforme a las intenciones del arquitecto del universo, el mundo no es producto del azar.

La teleología platónica es principalmente de corte antropomórfico. El recurso a los fines en el parte de argumentos epistemológicos, las formas son garantía de inteligibilidad de lo real, porque hay formas hay conocimiento y ese conocimiento es posible porque la realidad sensible participa de algún modo de las formas inteligibles. Las formas, además, aúnan a diferencia de los argumentos materialistas, tanto la causa formal como la final, es decir, explican tanto el qué es una cosa en tanto que cosa como el para qué de esa cosa, conceptos que para Platón van íntimamente ligados. En el caso del *Timeo* la teleología se hace depender del Demiurgo que ordena el mundo conforme a sus propios designios.

La teleología platónica la heredamos a través de Agustín y el pensamiento cristiano tamizada por el neoplatonismo tanto en la Tardoantigüedad y el Renacimiento y sobrevive en la idea que por ejemplo está presente en autores como Paley y su *Teología Natural* (1802).

2.1.2. La teleología aristotélica

Aristóteles, a diferencia de Platón, estudia la realidad a través de las entidades o substancias; no hay una duplicidad de realidades, sino sólo una única realidad. Así, todo lo que es, es substancia o es predicado de una substancia. Aristóteles aborda el problema de la teleología en varias de sus obras. Quizás de todas ellas la más importante en lo que a teleología concierne sea la *Física* (ca 324-323 A de N.E.) La *Física* es el texto en el que elabora más generalmente lo

⁸ Recuérdese que la idea de *creatio ex nihilo* es ajena al pensamiento griego.

que ha venido en llamarse la Filosofía Natural de Aristóteles y, a diferencia de la *Metafísica*, en la que también se habla de teleología⁹, tiene una unidad estructural mayor. También hay que mencionar el tratado *Acerca de las partes de los animales* (*De partibus animalium*) en el que, a diferencia de otras obras eminentemente descriptivas como la *Historia de los Animales* (en la que los argumentos teleológicos están ausentes), desarrolla una serie de argumentos a favor de la teleología de carácter previo a las investigaciones fisiológicas que componen el tratado¹⁰.

Todos los argumentos teleológicos que presenta Aristóteles van dirigidos en contra de las explicaciones materialistas de los fenómenos naturales. En este sentido Aristóteles también se une a la crítica de Platón al materialismo que aparece reflejada en la “segunda navegación” del Fedón en el que Sócrates/Platón argumenta duramente contra las tesis materialistas.

El gran hallazgo de Aristóteles frente a sus predecesores fue la causalidad final. Así lo expone en el libro A de su *Metafísica* en la que presupone y hace referencia a lo tratado en el libro B de la *Física*. Con ello pretendía haber completado el esquema explicativo de la realidad. La causalidad es un principio explicativo en tanto que da respuesta a un porqué. La explicación consiste en la respuesta a un determinado porqué. Explicar un fenómeno determinado implica establecer un trasfondo causal que permita dar razón del mismo¹¹. Al igual que el ser, la causalidad se dice de muchas maneras, Aristóteles identifica cuatro¹²: la causa material, formal, eficiente y final. Sin embargo,

⁹ La causalidad final, fundamento de la teleología aristotélica, se expone en varios pasajes de la *Metafísica*, si bien, atendiendo a la hipótesis genética de Jaeger, la *Metafísica*, es posterior a la *Física*.

¹⁰ El título de éste es engañoso, ya que antes de ser una descripción anatómica, que es lo que se parece sugerir, explica las causas de las partes de los animales.

¹¹ “No creemos conocer algo si antes no hemos establecido en cada caso el por qué (lo cual significa captar la causa primera) (...)” Arist., Phys. 194b 20. El traductor, en la nota al pie correspondiente, señala la conveniencia de interpretar *causa primera* como causa próxima.

¹² Las famosas cuatro causas tratan de dar razón de lo real, Aristóteles recoge los cuatro sentidos que cree suficientemente explicativos en conjunto. Cualquier respuesta a un determinado porqué remitirá, al menos, a una de las cuatro causas que a continuación se señalan: “En este sentido se dice que es causa (1) aquel

considera que su gran descubrimiento frente a las caracterizaciones de los presocráticos y los materialistas contemporáneos. Los argumentos en contra de las tesis materialistas tienen por finalidad evidenciar la insuficiencia explicativa de aquellas frente a los argumentos basados en la causalidad final. La postulación de la causalidad final es, en consecuencia, un recurso contra el mecanicismo atomista¹³. El procedimiento que seguirá Aristóteles consistirá en señalar la insuficiencia de las teorías materialistas.

Aristóteles quiere demostrar en todo momento el carácter necesario de la causa final en la explicación de lo orgánico. A lo largo del Libro II de la *Física* presenta varias argumentaciones dirigidas a mostrar la pertinencia de su esquema explicativo. Estos argumentos

constitutivo interno de lo que algo está hecho, como por ejemplo, el bronce respecto de la estatua o la plata respecto de la copa, y los géneros del bronce o de la plata. En otro sentido (2) es la forma o el modelo, esto es, la definición de la esencia y sus géneros (como la causa de una octava es la relación del dos al uno, y en general el número), y las partes de la definición. En otro sentido (3) es el principio primero de donde proviene el cambio o el reposo, como el que quiere algo es causa, como es también causa el padre respecto de su hijo, y en general el que hace algo respecto de lo hecho, y lo que hace cambiar algo respecto de lo cambiado. Y en otro sentido (4) causa es el fin, esto es, aquello para lo cual es algo, por ejemplo, el pasear respecto de la salud.” 194 b 25-35.

El texto anterior presenta las cuatro causas que la tradición ha denominado respectivamente material, formal, eficiente y final. Obsérvese el *se dice* con el que comienza el texto, está recogiendo el uso ordinario del concepto, está tratando de responder a la pregunta ¿de cuántas maneras decimos que algo relevantemente cuenta como causa? Aristóteles insiste en la completud de su esquema de cuádruple causalidad, no hay ni más ni menos que cuatro causas, todos los demás ejemplos que podamos aducir recaerán sobre uno de los cuatro tipos señalados anteriormente. La argumentación de Aristóteles se basa en un supuesto previo muy importante que desarrolla en el primer capítulo del mismo libro II, la distinción entre lo artificial y lo natural. Por naturaleza Aristóteles entiende los animales, las plantas y los elementos en virtud de que todos ellos tienen un principio intrínseco de movimiento, es decir, pueden pasar de la potencia al acto, lo que, además, los diferencia de las entidades artificiales cuyo ser es derivativo. Esto es lo que Aristóteles entiende por naturaleza, probablemente no resulte extraña la inclusión de animales y plantas, pero sí la de los elementos. El análisis se complica más cuando debemos admitir una determinada finalidad en los propios elementos.

¹³ Cf. Aristóteles. *Phys.* 198b 10-15.

son cuatro: (1) Aquellos argumentos basados en el desarrollo adaptativo de los seres vivos, (2) El argumento por analogía con el arte, (3) el denominado argumento contra la causalidad y (4) el argumento de la necesidad hipotética.

El argumento basado en el desarrollo adaptativo de los seres vivos trata de demostrar la necesidad de la causalidad final por medio de la referencia explícita al desarrollo adaptativo de los seres vivos. Con este recurso se quiere hacer frente otras argumentaciones que, como las de Empédocles o Anaxágoras, prescinden de ella al explicar la realidad orgánica. De hecho, lo que propone es que el recurso a la causalidad final explica mejor lo que de suyo tiene lo orgánico que cualquier otro tipo de argumento, por lo tanto, cualquier otra explicación resultaría insuficiente. Con esto se lograría insertar en el ámbito no estrictamente propositivo la causalidad final. El objetivo es, pues, situar la causalidad final en la naturaleza¹⁴. Pero, ¿cómo se llega a postular la necesidad de la causalidad final en la naturaleza? Aristóteles recurre a la observación de la adaptación entre diseño y función que está presente en la disposición anatómica de los seres vivos. La disposición anatómica de los animales y plantas a través de su ontogénesis¹⁵ exhibe propositividad. Los dientes afilados sirven para desgarrar, los ojos para ver y los pulmones para respirar. La consideración de cualquier órgano incluye su descripción funcional o propositiva, es decir, su explicación en términos de causas finales.

Lo vivo constituye un dominio ontológico cualitativamente distinto respecto de lo material. La postulación de causas materiales y eficientes en la explicación de determinados eventos físicos como la caída de un grave o la ocurrencia de una tormenta puede explicarse suficientemente recurriendo a los dos tipos de causalidad mencionados. Pero tan pronto como tratamos de explicar un fenómeno orgánico como la disposición actual de las piezas dentales de un animal (es el ejemplo que él mismo aduce) o el diseño y adaptación al medio de las plantas, las dos causas citadas resultan insuficientes.

¹⁴ Ésa es la pregunta con la que inicia el capítulo: “*Tenemos que decir, primero, por qué razón incluimos a la naturaleza entre las causas que son para algo.*” 198b 10.

¹⁵ No podemos hablar de filogénesis ya que Aristóteles era fijista.

Aristóteles trae a colación diversos ejemplos para tratar de argumentar su tesis. El pasaje de 198b 10 – 199a 30 es especialmente ilustrativo. En él Aristóteles presenta dos tipos de ejemplos para concluir la necesidad de la presencia de fines en la naturaleza. El primer ejemplo atañe a entidades inanimadas y presenta la explicación de un fenómeno como la lluvia en los términos admitidos por los atomistas, es decir, argumentos basados en el azar y la necesidad que operan sobre un substrato material¹⁶. Aristóteles estaría de acuerdo con la explicación que sobre el fenómeno de la lluvia pudiera dar un materialista, aquí es obvio que no se necesita postular causa final alguna para explicar el fenómeno. El hecho de que se arruinara la cosecha a causa de la lluvia no conllevaría ninguna propositividad, sería sólo un accidente. Por otra parte, no es menos obvio que no llueve *para que crezca el trigo*, sino que la lluvia es un fenómeno independiente. El crecimiento del trigo depende de la lluvia, aunque la conexión que pueda existir entre la lluvia y el desarrollo del trigo sea accidental. Pero que una causalidad final no sea necesaria explicativamente no implica que no esté presente.

Si analizamos un caso procedente del ámbito de lo vivo, nos percataremos de que el azar y la necesidad no pueden dar cuenta de la disposición de los seres vivos como los animales y las plantas. Para Aristóteles el azar y la necesidad, los principios de la física materialista, son insuficientemente explicativos¹⁷.

¹⁶ “¿Qué impide que la naturaleza actúe sin ningún fin, ni para lo mejor, que sea como la lluvia de Zeus, que no cae para que crezca el trigo sino por necesidad? Porque lo que se evapora tiene que enfriarse y cuando se ha enfriado tiene que transformarse en agua y descender, y el hecho de que crezca el trigo es algo accidental. Y, de la misma manera, cuando el trigo se pudre sobre la era, no ha llovido para que se pudra, sino que eso ha ocurrido por accidente.” Aristóteles, *Phys.* 198 b 15-25.

¹⁷ “¿Y qué impide que las partes de la naturaleza lleguen a ser también por necesidad, por ejemplo, que los dientes incisivos lleguen a ser por necesidad afilados y aptos para cortar, y los molares planos y útiles para masticar el alimento, puesto que no surgieron así por un fin, sino que fue una coincidencia? La misma pregunta se puede hacer también sobre las otras partes en las que parece haber un fin. Así, cuando tales partes resultaron como si hubiesen llegado a ser por un fin, sólo sobrevivieron las que “por casualidad” estaban convenientemente constituidas, mientras que las que no lo estaban perecieron y continúan pereciendo, como los terneros de rostro humano de que hablaba Empédocles.” 198b 25-35.

Hablar del diseño y funciones actuales de una entidad o parte de una entidad natural como un órgano parece implicar el establecimiento de un fin que regimenta la actuación de los mismos, fin que en un esquema mecanicista como el de Demócrito está completamente ausente y que con toda probabilidad se negaría a aceptar. La presencia de dientes, así como de otros rasgos anatómicos en el organismo animal parece estar justificada por la utilidad que le confiere. La coincidencia no explica esta disposición anatómica concreta. De otra manera, la disposición anatómica de los organismos es útil y así se perpetúa en la naturaleza, siendo en la mayoría de los casos de este modo y en función del beneficio de los organismos. Una disposición alternativa, a los ojos de Aristóteles, haría perecer a los distintos organismos¹⁸.

La estrategia argumentativa que ha seguido Aristóteles es una inferencia a la mejor explicación (Ariew, 2002). La teleología se infiere abductivamente a partir del hecho de que la disposición anatómica de los seres vivos contribuye a su desarrollo, la posesión de estos rasgos no es accidental, sino en virtud de un fin determinado. La estrategia explicativa de los materialistas es insuficiente. Así, sentada la existencia de la causalidad final, Aristóteles descarta los argumentos materialistas por medio del siguiente silogismo disyuntivo¹⁹:

“Pues no parece un resultado de la suerte ni de una mera coincidencia el hecho de que llueva a menudo durante el invierno, pero sí en verano; ni que haga calor en verano, pero sí en invierno. Así pues, ya que se piensa que las cosas suceden o por coincidencia o por un fin, y puesto que no es posible que sucedan por coincidencia ni que se deban a la casualidad, sucederán entonces por un fin. Ahora bien, todas estas

¹⁸ Salvo que no fuera un equivalente funcional. Aunque Aristóteles no parece prestarles excesiva atención.

¹⁹ Silogismo disyuntivo o *modus tollendo ponens* (quitando se afirma). La premisa menor niega el primer miembro de la disyunción de la premisa mayor. Reformulada la argumentación aristotélica, quedaría del modo siguiente: Premisa Mayor: Todas las entidades naturales suceden o por coincidencia o por un fin. Premisa Menor: No es posible que ninguna cosa suceda ni por coincidencia ni que se deba a la casualidad. Conclusión: En todas las cosas que llegan a ser y son por naturaleza está presente una causa final.

cosas y otras similares son por naturaleza, como lo admitirían los que mantienen la anterior argumentación. Luego en las cosas que llegan a ser y son por naturaleza hay una causa final.” Aristóteles, *Phys.* 199a

La conclusión es que la causalidad final se predica de toda entidad natural.

También el capítulo 8 (199a 10 y ss.) encontramos un argumento por analogía con el arte. La analogía refrenda la argumentación principal anteriormente expuesta. Al principio del libro II de la Física se encuentra una distinción importante en la ontología aristotélica cuya tesis principal es la defectividad ontológica de las entidades artificiales. El argumento abordado en esta sección se va a sostener en la referida distinción. La tesis principal es que la naturaleza imita al arte, siendo éste posterior a aquella. Por lo tanto, sostiene Aristóteles, si pudiéramos reproducir las entidades orgánicas, éstas no diferirían de cómo son por naturaleza. Implícita se encuentra la tesis de que la naturaleza hace las cosas del mejor modo. Sin embargo, con lo anteriormente dicho, no se defiende ningún tipo de finalismo. Es aquí donde juega su papel determinante la analogía, que es la siguiente: en la medida en que las entidades artificiales poseen intrínsecamente fines, es decir, están diseñadas conforme a un plan previo y para cumplir una función determinada, las entidades naturales, cuya existencia es ontológicamente anterior y no defectiva, también están dirigidas hacia un fin. Ahora hemos inferido la finalidad de la naturaleza a partir de la finalidad de lo artificial.

El principal recurso que le queda al atomista contra la finalidad intrínseca de la naturaleza defendida por Aristóteles es la explicación de la disposición actual de animales y plantas a través de la coincidencia causal. El partidario del materialismo podría contrargumentar que por coincidencia plantas y animales tienen la disposición que tienen y que, por lo tanto, no opera ningún principio teleológico en las entidades naturales. Contra estas razones Aristóteles procede a aplicar su análisis de la casualidad que anteriormente había desarrollado en el capítulo 6²⁰ del libro II de la Física. Por casualidad

²⁰ En el capítulo 6 del libro II de la Física presenta un análisis en el que distingue entre suerte y casualidad. Para Aristóteles la casualidad es un concepto más amplio que la suerte, estando esta última circunscrita únicamente al ámbito de lo humano. La

no ocurre lo que ocurre siempre o en la mayoría de los casos²¹. De la casualidad no se infieren las regularidades que observamos en la naturaleza, de una bellota no nace un olivo, del mismo modo que de un hombre no nace un gato. Las entidades naturales siguen una pauta de desarrollo regular que queda insuficientemente explicada por medio del recurso a la casualidad.

Obsérvese la relevancia del argumento de Aristóteles, habitualmente en la contemporánea filosofía de la ciencia se insiste en el carácter contrafáctico de las leyes. Que la explicación tenga un carácter legaliforme implica que pueda apoyarse en contrafácticos, es decir, que no dé lugar al azar o la simple ocurrencia casual de fenómenos. Aristóteles también examina el carácter normativo de su cuádruple esquema y señala la importancia de descartar el azar en el marco de la explicación de los fenómenos naturales. Si bien el azar sí está presente e incluso concurre con los propósitos de las distintas entidades. Hay que señalar que para Aristóteles las tres últimas causas son la misma, a saber, la final, en muchos casos²². De hecho la causa formal o esencia es lo que hace que la entidad en cuestión llegue a ser lo que es y alcance, en consecuencia su fin, o realización plena (entelequia)²³. De ahí se justifica la importancia de la finalidad y la defensa enconada que hace.

casualidad implica la ausencia de todo propósito que sobreviene a la entidad y que no es resultado del movimiento que imprime a la entidad su principio interno o *eidos* en el que está contenido el fin, es, por tanto, algo externo a la misma que ejerce un poder causal sobre ella.

²¹ La expresión “*en la mayoría de los casos*” aparece repetidas veces en los textos aristotélicos y no debe ser pasada por alto. Aristóteles admite la posibilidad de errores en la naturaleza, de entidades que por algún motivo extraño no llegan a desarrollarse o a ejecutar sus funciones como es debido. Es importante reparar en el importante contenido normativo que conlleva lo anteriormente mencionado.

²² “Y puesto que las causas son cuatro, es tarea propia del físico conocerlas todas, pues para explicar físicamente el “por qué” tendrá que remitirse a todas ellas, esto es, a la materia, a la forma, a lo que hace mover y al fin. **Las tres últimas se reducen en muchos casos a una, pues la esencia y el fin son una misma cosa**, y aquello de lo que primeramente proviene el movimiento es específicamente lo mismo que éstas, (...).” Aristóteles. *Phys.* 198a 25-30. (El subrayado es mío)

²³ Esta concepción aristotélica, a menudo denostada por lo que tiene de teleológico, no creo que en absoluto esté obsoleta, ya que fácilmente puede reinterpretarse a la

Finalmente encontramos el argumento de la necesidad hipotética o condicional expuesto en el capítulo 9 del libro II de la *Física* que también va en contra del materialismo, pero, a diferencia de los anteriores, no centra su argumentación sobre la necesidad o la casualidad, sino sobre la materia. La materia, junto con los dos principios citados anteriormente, constituía un recurso explicativo de la realidad no sólo para los atomistas, sino también para los primeros filósofos jonios como Tales, Anaximandro y Anaximenes. Aristóteles, en el conocido pasaje del libro A de su *Metafísica*, concede a sus antecesores el mérito de haber identificado la causa material, aunque sólo con la causalidad final quede completo el esquema explicativo de la realidad.

La necesidad hipotética o condicional debe ser diferenciada de la necesidad incondicionada²⁴. La primera es la condición necesaria para la realización de un fin. Aristóteles considera que, aunque las entidades naturales están compuestas de materia, la materia por sí misma no da razón de que sean como son o de que sean para algo. En cualquier caso la materia es necesaria para que las cosas sean, en la medida en que es necesario que haya un soporte para instanciar un determinado diseño y que aquel responda a las exigencias funcionales de éste. Para Aristóteles la materia es, pues, necesaria, pero sólo condicionalmente, es decir, en aras a un fin determinado que viene dado por la definición o forma de la entidad.

¿Son legítimas las conclusiones aristotélicas anteriormente estudiadas? Es obvio que Aristóteles no presenta ningún argumento

luz de la idea de programa genético. Otros autores como Mayr (2004) refrendan la misma opinión. A propósito de la teleología aristotélica señala lo siguiente: "*Algunas dificultades de los filósofos de deben a que interpretan erróneamente los escritos de los grandes filósofos del pasado. Por ejemplo, a menudo se ha calificado a Aristóteles como finalista, y a la teleología cósmica se le ha denominado perspectiva aristotélica. Grene tiene toda la razón cuando señala que el telos de Aristóteles no tiene que ver con un propósito*" p.65.

A continuación señala la relación de la teleología aristotélica con los problemas de ontogenia y adaptación. "*Los especialistas modernos en Aristóteles (Balme, Gotthelf, Lennox y Nusbaum) están todos de acuerdo en que su aparente teleología tiene que ver con problemas de ontogenia y de adaptación en organismos vivientes, en lo que las opiniones de Aristóteles resultan notablemente modernas.*" Ibid.

²⁴ Cf. Aristóteles, *Met.* V, cap. 5.

deductivo del que podamos colegir la pretendida obviedad de sus planteamientos. Los cuatro argumentos presentados anteriormente proceden por medio de inferencias abductivas. En todos los casos la conclusión será que la mejor explicación que podemos ofrecer es aquella que postula fines en la naturaleza.

Sin embargo, una inferencia a la mejor explicación no compromete ontológicamente. Durante la Edad Media hubo quien interpretó la finalidad aristotélica bajo un determinado sesgo subjetivista o metafórico como Escoto, y Ockham desterró la causalidad final de los seres inanimados. Pero las inferencias a la mejor explicación de Aristóteles sí parecen estar comprometidas con su realismo metafísico. La existencia de la causalidad final es algo objetivo que se corresponde con la realidad²⁵. Así como la existencia de las entidades materiales es para Aristóteles un hecho indubitable, tampoco se puede dudar de la finalidad.

En cuanto a la teleología aristotélica hay que insistir en que no sería correcto interpretarla como una teleología universal, simplificación excesiva que a menudo se comete con las doctrinas aristotélicas. En ningún momento parece sostenerse que todo aquello que sucede por naturaleza persiga intrínsecamente un fin, sino sólo algunas entidades o procesos naturales están diseñados para un fin. La ocurrencia de una marea o una tormenta no tienen finalidad alguna, al igual que la existencia de la piedra o el hecho de que mi gata tenga un pelaje blanco y negro. Sin embargo, los órganos de los seres vivos o sus conductas sí tienen claramente finalidad, tal como he presentado los argumentos de Aristóteles.

Volvamos a la cuestión de las entidades naturales. Que animales y plantas posean algún tipo de finalidad intrínseca no parece problemático en exceso. Pero sí que a los elementos, en la medida en que participan de lo natural, se les aplique la conclusión ya señalada de que en todas las cosas que llegan a ser y son por naturaleza hay una finalidad. Si esto es así estaría en clara contradicción con el ejemplo 1 que señalaba que en una tormenta o cualquier otro fenómeno natural inanimado no había propositividad alguna. La

²⁵ La interpretación de Wieland discrepa del compromiso ontológico que aquí presento. Para Wieland la teleología aristotélica es un “como si...”.

respuesta cabe encontrarla en un análisis más sutil. En Met. 1044b 8-12 podemos encontrar una distinción bastante acertada; hay cosas que son naturales pero no son entidades. Un eclipse o una tormenta responden claramente a la descripción anterior. ¿Y la finalidad de los elementos, aún así, como entenderla?

Si las entidades naturales que, por antonomasia son los vivientes, necesitan explicativamente de causalidad final, también podrá establecerse su presencia en las entidades inanimadas

2.2. La recepción escolástica de la teleología

Algunos historiadores como Gilson (1952) introducen el estudio de la Filosofía Medieval a partir de la Patrística del siglo II. El primer autor cristiano que muestra un claro interés por la Filosofía es Agustín (s. IV), más tarde obispo de Hipona y padre de la Iglesia de Occidente. Con Agustín de Hipona nos encontramos al final de la Antigüedad -lo que historiográficamente se ha denominado Tardoantigüedad- y a las puertas de la Alta Edad Media. La idea de la naturaleza de Agustín está dominada por el influjo neoplatónico y cabe ver en su producción filosófica el legado del *Timeo* de Platón bajo la perspectiva soteriológica cristiana. Un buen ejemplo de ello lo constituye una de sus primeras obras, el diálogo *De ordine* en la que defiende que toda la creación-concepto ajeno al pensamiento griego- es fruto del designio divino:

"¿Quién negará, ¡oh Dios grande!, que todo lo administras con orden? ¡Cómo se relacionan entre sí en el universo todas las cosas y con qué ordenada sucesión van dirigidas a sus desenlaces!" Agustín de Hipona, *De ordine* cap. V.

La herencia platónico-agustiniana en el pensamiento europeo será clave hasta el siglo XIII con el advenimiento del estudio de Aristóteles. Las obras aristotélicas de las que hasta el momento se había dispuesto eran las traducciones de Boecio de los tratados de lógica y poco más. El aristotelismo medieval cobra fuerza con la escolástica y particularmente con la vigorosa y monumental obra de Tomás de Aquino. El tomismo refleja muy bien cuál es la concepción de la naturaleza que defiende una facción importante del mundo académico

medieval. En este debate la teleología desempeña un papel muy importante, tanto por su conceptualización, que a mi juicio difiere en algunos matices de la aristotélica, como por los problemas que evidencia y que serán tematizados por otros autores posteriores como Occam o Suárez dentro de la misma tradición.

Las investigaciones de Tomás de Aquino sobre la teleología natural aparecen ya en su obra más temprana, de modo particular el opúsculo intitulado *De Principiis Naturae* (ca.1255), si bien el tema lo retoma en su *Summa contra Gentes*, aunque de manera menos exhaustiva. El opúsculo *De principiis Naturae* procede a una clasificación de los principios de la naturaleza, es decir, de las entidades que componen el mundo o la creación. Se trata ante todo de una reflexión centrada en el concepto aristotélico de entidad o substancia. En esta reflexión ocupa un papel preeminente el concepto de finalidad como propio de la noción de substancia conforme a la doctrina aristotélica. Así, encontramos formulado lo que la tradición ha denominado el principio de finalidad; el principio que afirma que todo agente obra por un fin²⁶:

"Y debe saberse que todo agente, tanto natural como voluntario, pretende un fin; sin embargo de esto no se sigue que todo agente conozca el fin o que delibere sobre el fin. Pues conocer el fin es necesario cuando las acciones no están determinadas, sino que pueden ser cosas opuestas, como es el caso de los agentes voluntarios; y por esto es necesario que conozcan el fin por el que determinan sus acciones. Pero en los agentes naturales las acciones están determinadas, por lo que no es necesario elegir lo que conduce al fin." Tomás de Aquino, *De Principiis Naturae*, 3.

La finalidad, que en las entidades naturales está determinada, es considerada por Tomás de Aquino un principio extrínseco a la substancia. Tomás de Aquino distingue claramente entre principios intrínsecos y principios extrínsecos, es decir, principios que actúan desde el interior de la substancia y principios que actúan desde el exterior de la misma. Los principios intrínsecos son la materia y la forma, mientras que los principios extrínsecos son la causalidad eficiente y la causalidad final. Si bien los principios mencionados se corresponden con las cuatro causas aristotélicas, la clasificación anteriormente señalada parece ajena al pensamiento de Aristóteles. No

²⁶ Cf. Aristóteles, *Met.* B.

quiere esto decir que con ésta se violenta de algún modo la teoría aristotélica, pero sí es cierto que implica ya una interpretación, bien sea para aclarar la concepción aristotélica, bien para contaminarla.

Para Tomás de Aquino el fin es *causa causarum*, el fin es la causa de la causalidad de todas las causas:

"Por tanto, el fin es causa de la causalidad eficiente por cuanto hace a la eficiente ser eficiente; igualmente hace a la materia ser materia y a la forma ser forma, ya que la materia sólo recibe la forma por el fin y la forma sólo perfecciona a la materia por el fin. Por eso se dice que el fin es la causa de las causas, porque es la causa de la causalidad de todas las causas." Tomás de Aquino, *De Principiis Naturae*, 4.

Obsérvese que, además, Tomás se centra en la finalidad de la substancia; son las entidades las que incorporan una finalidad intrínseca en virtud de su esencia y no el mundo como copia de algún misterioso arquetipo.

Por su parte, Guillermo de Ockham en su *Summula philosophiae naturalis* discrepa respecto de la concepción finalista. Ockham es de los primeros en criticar el concepto de causalidad final y lo restringe a los seres capaces de actos volitivos. A los agentes que no actúan consciente y voluntariamente difícilmente pueden atribuírseles una causa final.

Finalmente, ya en el Renacimiento, con Francisco Suárez, autor escolástico con mentalidad de hombre moderno, encontramos una curiosa crítica a la teleología o a las causas finales. Suárez en la *Disputatio XXIII* de sus *Disputationes Metaphysicae* hace notar la dificultad acerca de cómo es posible que un resultado, un fin, sea causa de algo, pues el efecto es posterior a la causa. En el texto de Suárez tenemos ya una insinuación del problema de la causación retroactiva que tanto va a preocupar a los filósofos de la ciencia de la concepción heredada a propósito de las explicaciones funcionales y teleológicas.

2.3. Teleología y Modernidad

En los inicios de la Modernidad e incluso bien entrado el siglo XVII conviven diversas tendencias filosóficas en lo que concierne al estudio

de la denominada Filosofía Natural. La aceptación del esquema explicativo galileano no es unánime y en el ámbito académico como muy bien ha estudiado Kristeller (1979) predomina antes el aristotelismo escolástico que la nueva ciencia o la filosofía de los modernos. Tanto una como otra rechazan las causas finales. Tal rechazo tiene orígenes cartesianos. El mecanicismo cartesiano eliminaba de raíz cualquier pretensión organicista o vitalista a través de la que pudiera vislumbrarse cualquier atisbo de teleología. Mucho se ha escrito sobre el peso de la escolástica y, más concretamente, del aristotelismo en Descartes, pero es obvio que su ciencia natural rechaza de plano cualquier dependencia de estas dos fuentes.

De este modo tenemos una ciencia galileana, cartesiana y baconiana que trata de explicar la realidad sensible a través del mecanicismo. Mientras que de algún modo pervive cierto aristotelismo residual en las ciencias de la vida: medicina y método experimental. Kant constituye un hito en el estudio de la teleología. Así, siguiendo a Rossi (1986), hay varias historias posibles de la ciencia. No creo que se trate, pues, de hacer una apologética de la racionalidad científica a través de sus logros históricos, sino de presentar diferentes estadios de la racionalidad científica que no responden en modo alguno a una imagen triunfalista.

Si hay un filósofo que se ha opuesto radicalmente en la historia de la filosofía a la autoridad de Aristóteles éste ha sido Descartes tanto por su concepción dualista del alma que bien le valió la condena eclesiástica²⁷ y que propone una antropología radicalmente opuesta al aristotelismo y a la doctrina católica, como por su rechazo a la causalidad final, que suplantó por su mecanicismo. Para Descartes, además de Dios, a quien conviene especialmente el concepto de substancia, la realidad creada se divide en dos substancias; la *res cogitans* y la *res extensa*. No hay algo así como un hilemorfismo aristotélico o la concepción igualmente aristotélica de los seres vivos como entidades dotadas de un principio anímico intrínseco que informa la materia. La justificación de la ontología cartesiana es claramente deudora de sus tesis epistemológicas. Si el atributo esencial de la *res cogitans* es el pensamiento o autoconciencia, el de la *res extensa* es la

²⁷ Cf. Müller (1998: 113-114).

extensión²⁸. La extensión tiene, además, dos modos o afecciones, la figura y el movimiento. Si las cosas materiales tienen por atributo esencial la extensión un estudio de estas substancias debe hacerse sobre su principal atributo. La extensión es susceptible de un análisis geométrico. De este modo para estudiar la física es necesario únicamente considerar su extensión sin requerir nada más. El mecanicismo es suficiente como explicación de la naturaleza. El movimiento lo crea Dios conjuntamente con la materia; las leyes de la naturaleza son leyes mecánicas pues no es necesaria más causalidad que la causalidad eficiente. De hecho, Descartes critica la causalidad final en la IV de sus *Meditaciones Metafísicas*:

"Al considerar esto con atención, me ocurre ante todo pensar que no debo extrañarme de no poder comprender por qué hace Dios lo que hace, y que no por eso hay que dudar de su existencia; porque acaso veo por experiencia muchas otras cosas que existen, sin poder comprender por qué ni cómo las ha hecho Dios; pues sabiendo, como sé, que mi naturaleza es en extremo endeble y limitada y que la de Dios, por el contrario, es inmensa, incomprensible e infinita, ya no me es difícil desconocer que hay infinitas cosas en su poder cuyas causas exceden al alcance de mi espíritu; y basta esta razón para persuadirme de que todas esas causas, que suelen sacarse de los fines, no tienen aplicación alguna en las cosas físicas o naturales; pues no me parece que se pueda, sin temeridad, indagar y tratar de descubrir los fines impenetrables de Dios."

Descartes, *Meditaciones Metafísicas* IV.

Obsérvese que la crítica cartesiana a las causas finales parte de un posicionamiento teológico y se acomete tras un proceso de garantía epistémica acerca de la existencia divina. Las causas finales no tienen lugar porque los designios divinos son incognoscibles, apelar a las causas finales supone afirmar más de lo que es legítimo afirmar para un entendimiento finito. La crítica cartesiana a la causalidad final es, pues, una crítica desde el presupuesto de la trascendencia divina y la asunción de los recursos cognitivos humanos. No se discute si hay o

²⁸ Descartes ejemplifica la res extensa con un trozo de cera. El atributo que persiste a todos los cambios imaginables del trozo de cera es siempre la extensión, de este modo afirma en la segunda de sus *Meditaciones Metafísicas*: *"Pero ¿qué es, hablando con precisión, lo que yo imagino cuando lo concibo de esta suerte? Considerémosle atentamente y, separando todas las cosas que no pertenecen a la cera, veamos lo que queda. No queda ciertamente más que algo extenso, flexible y mudable."*

no hay causas finales o, lo que es lo mismo, designios divinos en la creación, sino que se pone en duda que puedan ser conocidos. La negación de la finalidad es desde el ámbito epistemológico, no ontológico. En cualquier caso poco debía importarle a Descartes cuando concibe una ciencia física que no las tiene presentes y que, además, no las necesita para explicar la realidad sensible. La física cartesiana tuvo poco éxito en la ciencia moderna, siendo sustituida por la de Newton. Sin embargo, sí perduró de manera sistemática su mecanicismo y con él la completa erradicación del finalismo de la ciencia física. Con Descartes la Modernidad hereda la primacía de la epistemología respecto de la ontología. Desde este momento saber cómo es el mundo estará subordinado a saber cómo me es dado conocerlo.

En Leibniz el peso de la tradición escolástica es más explícito que en Descartes y no duda en rehabilitar el finalismo que Descartes había criticado retomando la advertencia de la segunda navegación socrática en el Fedón de Platón tal como se indica en el §20 de su *Discurso de Metafísica*. En el §22 de esta misma obra intenta encontrar un equilibrio entre mecanicismo y finalismo y señala que ambas son vías adecuadas y compatibles en contra de Descartes que anulaba el recurso a las causas finales en contra de las tesis cartesianas. Las causas finales, señala Leibniz, son especialmente idóneas para la explicación en ciencias naturales especialmente cuando consideramos la realidad como el producto de la creación voluntaria de Dios. Los argumentos de Leibniz a favor de la reconsideración de las causas finales oscilan entre una inferencia a la mejor explicación y una analogía²⁹ en la que se compara la acción creadora de Dios con la acción intencional de un artesano. Si bien la inferencia a la mejor explicación es aquí un principio epistemológico subordinado a una tesis más fuerte de carácter metafísico acerca de la existencia de fines de manera intrínseca en la

²⁹ "Sin embargo, encuentro que la vía de las causas eficientes, que es más profunda, en efecto, y en algún sentido más inmediata y a priori, es en cambio bastante difícil cuando se llega al detalle, y creo que nuestros filósofos están, la mayoría de las veces, todavía muy lejos de ella. Pero la vía de las finales es más fácil y no deja de servir con frecuencia para adivinar verdades importantes y útiles, que habrá que buscar mucho tiempo por otro método más físico; de lo cual la anatomía puede proporcionar ejemplos considerables." Leibniz, *Discurso de Metafísica* §22.

naturaleza conforme a los lineamientos de una metafísica teológica que incide en la propositividad divina³⁰. En este aspecto el lineamiento de Leibniz con las tesis de la metafísica escolástica es claro. La consecuencia inmediata de este planteamiento dentro de las corrientes modernas³¹ del siglo XVII es que el materialismo cartesiano se ve complementado con un argumento a favor de la necesidad de incluir en las sustancias materiales algún principio metafísico intrínseco que explique los fenómenos en los que éstas se ven involucradas.³²

Finalmente cabe decir algo respecto de la perspectiva kantiana. Con Kant se cierra la Modernidad y se inaugura la filosofía contemporánea. En cierto sentido acrisola los dos grandes posicionamientos en torno a la naturaleza existentes en la Modernidad; la posición teleológica que sigue la tradición platónico-aristotélica y el mecanicismo que se inaugura con la nueva ciencia. De otra manera, el divorcio existente entre un posicionamiento centrado en la consideración de la incipiente nueva física tras la Revolución científica y la tradición más centrada en el estudio de los organismos que no verá su propia revolución hasta bien entrado el siglo XIX.

La perspectiva kantiana a propósito del tema de la teleología en la ciencia natural plantea dos cuestiones muy actuales en Filosofía de la Biología. Por una parte la autonomía de las ciencias de la vida. Por la otra, la justificación de un lenguaje teleológico en las mismas.

³⁰ Cf. Leibniz, *Discurso de Metafísica* § 1 titulado "De la perfección divina, y que Dios lo hace todo del modo más deseable" Cf. También §19 "Dios se propone siempre lo mejor y lo más perfecto."

³¹ No hay que olvidar que durante toda la Modernidad pervive una importante corriente escolástica que además es enseñada de manera sistemática en las universidades. La filosofía de la Modernidad no es sólo *Filosofía Moderna*.

³² "Pero esta fuerza es algo diferente de la magnitud, de la figura y del movimiento, y por ello se puede juzgar que no todo lo que se concibe en los cuerpos consiste únicamente en la extensión y sus modificaciones, como opinan nuestros modernos. Así nos vemos obligados de nuevo a restablecer algunos entes o formas que han desterrado. Y parece, cada vez más, aunque todos los fenómenos particulares de la naturaleza se puedan explicar matemática o mecánicamente por los que los entienden, que, no obstante, los principios generales de la naturaleza corpórea y de la mecánica misma son más bien metafísicos que geométricos, y corresponden más bien a algunas formas o naturalezas indivisibles como causas de las apariencias, que a la masa corpórea o extensión." Leibniz, *Discurso de Metafísica* §18.

Si bien en Aristóteles señalé que no creía posible una interpretación al margen de cualquier compromiso ontológico, es obvio que en Kant la presencia de dicho compromiso es, cuando menos, problemática, ya que antes que un problema acerca de cómo sea la realidad, Kant plantea un problema sobre cómo la conocemos.

El análisis de la teleología aparece en la tercera de sus críticas o *Crítica del Juicio* [K.U.]³³ (1790). Para Kant el juicio es una facultad mediadora entre el mundo sensible o la naturaleza y la razón o el reino de la libertad. Así el juicio debería posibilitar pensar la naturaleza de modo compatible con la libertad, que es otra manera de introducir causalidad en el ámbito natural³⁴. Los fines humanos determinados libremente por la voluntad deben poder realizarse en el mundo sensible. De este modo el juicio es la unión entre la filosofía teorética y la filosofía práctica. En caso contrario nos encontraríamos con una imagen totalmente escindida del hombre. Nótese que con ello se está dando respuesta de alguna manera al problema del dualismo cartesiano. Para Kant existen tres facultades: la facultad cognoscitiva, la facultad de sentir placer o displacer y la facultad de desear, la segunda de ellas tiene que ver con el sentimiento, de modo que es

³³ Habitualmente en español a partir de la traducción de Morente (1914) se ha traducido *Kritik der Urteilskraft* como *Crítica del Juicio*, sin embargo la traducción puede ser confundente, ya que propiamente lo que se aborda es un análisis sobre la facultad de juzgar.

³⁴ "Pero si bien se ha abierto un abismo infranqueable entre la esfera del concepto de la naturaleza como lo sensible y la esfera del concepto de libertad como lo suprasensible, de tal modo que del primero al segundo (por medio del uso teórico de la razón) ningún tránsito es posible, exactamente como si fueran otros tantos mundos diferentes, sin poder el primero tener influjo alguno sobre el segundo, sin embargo, debe éste tener un influjo sobre aquel, a saber: el concepto de libertad debe realizar en el mundo sensible el fin propuesto por sus leyes, y la naturaleza, por tanto, debe poder pensarse de tal modo que al menos la legalidad de su forma concuerde con la posibilidad de los fines, según leyes de libertad, que se han de realizar en ella. Tiene, pues, que haber un fundamento para la unidad de lo suprasensible, que yace en la base de la naturaleza, con aquello que el concepto de libertad encierra de práctico; el concepto de ese fundamento, aunque no pueda conseguir de él un conocimiento ni teórico ni práctico, y por tanto, no tenga esfera característica alguna, sin embargo hace posible el tránsito del modo de pensar según los principios de uno al modo de pensar según los principios del otro." Kant, K.U. Introducción, II.

mediadora³⁵ entre el entendimiento (facultad cognoscitiva) y la razón (o reino de la libertad). La tarea de una crítica del juicio es precisamente el análisis de la facultad de juzgar y esta crítica se materializa en las siguientes preguntas: ¿qué principios *a priori* encierra la facultad de juzgar? Y, si los tiene ¿cuál es su función y ámbito de aplicación? Ahora bien ¿cuál es la relación de las anteriores disquisiciones con respecto a la teleología? La respuesta a esta pregunta se encuentra en el mismo concepto de juicio. El juicio es la facultad de pensar lo particular como contenido en lo universal³⁶. Juzgar es lo que hace de diversas maneras el hombre ante la realidad ya sea a nivel práctico o puramente teórico. En la K.U. Kant establece una distinción entre juicio determinante y juicio reflexionante. Un juicio determinante o subsumiente es aquel que, tal como se expone en la KrV subsume lo particular en lo universal dada la regla de subsunción. Ahora bien, puede ocurrir que sólo se de lo particular, en tal caso no hay subsunción posible bajo reglas, sino sólo reflexión. De este modo la universalidad del juicio es fruto del proceso de reflexión sobre lo particular. Nos encontramos con juicios reflexionantes en diversas situaciones, particularmente en todos aquellos casos en los que la aplicación de los conceptos del entendimiento sobre los fenómenos no es directa. Estos casos son las leyes empíricas de la física, las leyes de la física no están dadas a priori, por el simple examen de las categorías presentes en el entendimiento no obtenemos leyes físicas. Pero tampoco están dadas a posteriori, pues de hecho contienen un carácter modal que no está presente en la experiencia ordinaria y permiten, además, predicciones. El juicio estético es otro caso en el que a partir de una experiencia particular deseamos predicar su validez universal. Lo bello es bello para alguien y ese alguien cree que su experiencia debe ser universal. Así, los dos juicios reflexionantes que analiza son el juicio estético, al que como es sabido está dedicada la primera parte de la K.U. y el juicio teleológico, al que dedica la segunda y última parte. Ambos juicios invocan de algún modo la finalidad, aunque sólo me detendré en estas páginas sobre el último tipo de juicios.

³⁵ Cf. Kant, K.U. Introducción, III *De la crítica del juicio como un medio de enlace de las dos partes de la filosofía como un todo.*

³⁶ Cf. Kant, K.U. Introducción, IV

El problema de la teleología natural en Kant aparece con la consideración de los juicios reflexionantes sobre la naturaleza objetiva o juicios reflexionantes objetivos materiales referidos a cosas existentes al margen del obrar humano y por oposición a los juicios reflexionantes objetivos formales (que versan sobre la matemática). El problema de estos juicios es encontrar su fundamento universal que no se halla en la experiencia sensible y que, sin embargo no está presente en los conceptos del entendimiento:

"El juicio reflexionantes, que tiene la tarea de ascender de lo particular en la naturaleza a lo general, necesita, pues, un principio que no puede sacar de la experiencia, porque ese principio justamente debe fundar la unidad de todos los principios empíricos bajo principios igualmente empíricos, pero más altos, y así la posibilidad de la subordinación sistemática de los unos a los otros. El juicio reflexionantes puede, pues, tan sólo darse a sí mismo, como ley, un principio semejante, trascendental, y no tomarlo de otra parte (pues entonces sería un Juicio determinante) ni prescribirlo a la naturaleza, porque la reflexión sobre las leyes de la naturaleza se rige según la naturaleza, y ésta no se rige según las condiciones según las cuales nosotros tratamos de adquirir de ella un concepto que, en relación a éstas, es totalmente contingente.

Ahora bien, ese principio no puede ser otro más que el siguiente: que como las leyes generales de la naturaleza tienen su base en nuestro entendimiento, el cual las prescribe a la naturaleza (aunque sólo según el concepto general de ella como naturaleza), las leyes particulares empíricas, en consideración de lo que en ellas ha quedado sin determinar por las primeras, deben ser consideradas según una unidad semejante, tal como si un entendimiento (aunque no sea el nuestro) la hubiese igualmente dado para nuestras facultades de conocimiento, para hacer posible un sistema de la experiencia según leyes particulares de la naturaleza (...). La finalidad es, pues, un particular concepto a priori que tiene su origen solamente en el Juicio reflexionante. Pues atribuir a los productos de la naturaleza algo como una relación, en ellos, de la naturaleza con fines no se puede hacer: se puede tan sólo usar este concepto para reflexionar sobre ella, refiriéndose al enlace de los fenómenos en ella que es dado según leyes empíricas."

Kant, K.U. Introducción IV

La argumentación kantiana va dirigida a explicitar cuál es el fundamento de la inteligibilidad de la naturaleza orgánica. En este sentido se trata de un argumento trascendental que intenta especificar las condiciones de posibilidad de conocimiento de la naturaleza orgánica al margen de cualquier asunción substantiva acerca de la naturaleza intrínseca de lo orgánico. Nos es dado pensar la naturaleza

bajo un principio de inteligibilidad bajo el que debemos pensarla como si de una obra atribuible a un agente se tratara *"tal como si un entendimiento (aunque no sea el nuestro) la hubiese igualmente dado para nuestras facultades de conocimiento"*. Kant aduce varios ejemplos, entre ellos la ley de parsimonia (la naturaleza toma el camino más corto), la de la continuidad en la naturaleza (la naturaleza no hace saltos) o la conocida navaja de Ockham. En la *Analítica del Juicio teleológico* (K.U. §§62-68) encontramos una clasificación de los diferentes tipos de juicios teleológicos, estos versan sobre entidades formales y naturales. Dentro de las cosas naturales cabe ver dos tipos de finalidad, una relacional o extrínseca en el sentido de que el cerdo es para la alimentación del hombre y otra interna o constitutiva. El tipo de finalidad que nos interesa es la finalidad interna que cabe ver en las cosas naturales. Este tipo de finalidad se manifiesta en juicios teleológicos absolutos y describe cómo deben concebirse las cosas naturales, como todos organizados en los que sus partes constituyentes son fines y medios:

"Ese principio, y al mismo tiempo su definición, dice. Un producto organizado de la naturaleza es aquel en el cuál todo es fin, y, recíprocamente, también medio. Nada en él es en balde, sin fin o atribuible a un ciego mecanismo natural." Kant, K.U. § 66

La finalidad interna se predica de organismos, es la única forma de teleología intrínseca en sentido clásico que cabe ver en la naturaleza; lo inorgánico está exento de consideraciones teleológicas y se explica apelando a explicaciones causales. La imagen que tenemos ahora de la naturaleza es completamente escindida. Kant apunta con esto a un importante rasgo de la ciencia moderna que perdurará hasta nuestros días. Así, delimitamos el ámbito de lo natural como un ámbito cerrado causalmente y al margen de la libertad que la circunscribimos a la esfera humana, realmente una anomalía en este mundo cerrado causalmente. Pero simultáneamente postulamos por exigencias propias de la razón dos modos de explicar la realidad objetiva, uno puramente mecánico para las entidades inertes y otro de carácter finalista para los organismos. Esto necesariamente ha de crear una situación contradictoria que Kant articula en su *Dialéctica del juicio*

teleológico (§§ 69-78). La antinomia del juicio teleológico puede formularse de la siguiente manera:

"La primera máxima de la misma es la tesis: Toda producción de cosas materiales y de sus formas debe ser juzgada como posible según leyes meramente mecánicas.

La segunda máxima es a antítesis: Algunos productos de la naturaleza material no pueden ser juzgados como posibles sólo según leyes meramente mecánicas (su juicio exige una ley de la causalidad totalmente distinta, a saber, la de las causas finales)."

Kant, K.U. §70

La solución kantiana a la antinomia³⁷ mantiene la finalidad como heurístico dentro de una concepción mecanicista del mundo, es decir, le da a las explicaciones funcionales el mismo papel que siglos más tarde les daría Hempel por razones distintas. Así, la finalidad tiene el valor de un principio regulativo que da lugar a máximas heurísticas del juicio. Kant dice cómo nos es dado pensar la naturaleza, pero no dice que la naturaleza sea intrínsecamente así. Las causas finales así quedan relegadas a un mero heurístico y su realidad resulta del todo incognoscible.

Kant es una buena síntesis de los diversos posicionamientos esbozados en las líneas precedentes. La existencia de una finalidad intrínseca a la naturaleza es, en el mejor de los casos, una tesis muy

³⁷ *"Pero, en cambio, en lo que se refiere a la máxima primeramente citada de un Juicio reflexionante, ésta no encierra en realidad, contradicción alguna, pues cuando digo: Todo suceso en la naturaleza material, y, por lo tanto, todas las formas, debo juzgarlas, en lo que toca a su posibilidad, como productos de la naturaleza según leyes meramente mecánicas, no digo con ello que son sólo de ese modo posibles (como excluyendo toda otra especie de causalidad), sino que eso quiere decir tan sólo que debo siempre reflexionar sobre aquellos sucesos según el principio del mero mecanicismo de la naturaleza, y, por tanto, desentrañar ese principio tan lejos como pueda, pues sin ponerlo en la base de la investigación no puede haber propiamente conocimiento alguno de la naturaleza. Ahora bien, esto no impide a la segunda máxima en ocasiones pasajeras, a saber, en algunas formas naturales (y, con ocasión de éstas, incluso en la naturaleza entera), reflexionar sobre ellas y buscar un principio totalmente diferente de la explicación de la naturaleza según el mecanismo, a saber, de las causas finales, pues no por eso es suprimida la reflexión según la primera máxima, sino que más bien se incita a que se prolongue todo lo que se pueda." Kant, K.U. §70*

débil que a menudo aparece justificada a través de una inferencia a la mejor explicación. Ésta pone de relieve cómo nos es dado entender los fenómenos biológicos. En este sentido es antes una exigencia de la razón que algo directamente observable. El desarrollo filosófico de la teleología parte así de tesis ontológicas cuya justificación epistemológica, aunque débil, parece irrenunciable (Platón, Aristóteles), hasta la plena asunción de una teleología inmanente al propio esquema conceptual (Kant) pero cuya realidad material queda en entredicho. La cuestión que queda pendiente será si esta teleología realmente puede tener alguna realidad más allá del ámbito de las razones. Los problemas de la teleología en la filosofía del siglo XX tras la revolución kantiana, son pues, herencia directa de la solución kantiana.

2.4. Pensamiento darwiniano y teleología

Si hacemos caso a Freud los descubrimientos de Darwin a la vez que suponen la segunda gran revolución científica después de la física del XVII, contribuye a desbancar al hombre de su lugar privilegiado. Pero la revolución darwiniana, en caso de poderla llamar así, no contribuye únicamente a derrocar al hombre de su lugar privilegiado y equipararlo con el resto de criaturas, tampoco aporta únicamente una naturalización de la antropología, sino que destierra del panorama científico los últimos remanentes teleológicos que pudieran pervivir en las ciencias biológicas. Darwin ha modificado de manera substancial nuestra cosmovisión y, como afirma Mayr (2004), probablemente su teoría haya tenido mayor impacto en el imaginario colectivo que la nueva física del siglo XX.

Mayr (2004), a propósito de Darwin y la teleología cita el siguiente texto de Huxley (1870) en el que se afirma que lo que más le impresionó *“con mayor fuerza en su primera lectura de El origen de las especies fue la convicción de que la teleología, como se entiende comúnmente, había recibido un golpe mortal a manos del señor Darwin”*³⁸

³⁸ Tomado de la edición española. 2006, p.61.

¿Qué debemos entender a propósito del "cómo se entiende comúnmente" que refiere Mayr? Es obvio que no es ni la perspectiva ofrecida por Aristóteles - opinión con la que el propio Mayr estará de acuerdo- ni la perspectiva kantiana -malinterpretada por Mayr-. Quizás la teleología *como se entiende comúnmente* es la propuesta por la denominada teología natural que nada tiene que ver en principio con la línea de argumentación aristotélica o kantiana, sino más bien con una deformación acrítica y capciosa de la primera. La teología natural se fundamenta en el argumento de la perfección del diseño para demostrar la existencia de Dios, pero no para explicar por qué los diferentes organismos son como son. En cualquier caso una explicación de este tipo estará siempre subordinada a la demostración de la existencia de Dios, por lo tanto si algún valor puede tener este tipo de argumentación, será teológico y no científico.

¿Es cierto que Darwin le da un golpe mortal a la teleología? La idea de que el diseño de los diferentes organismos responde a un plan predeterminado o la comparación sistemática entre las estructuras orgánicas y las creaciones artefactuales es muy frecuente. Incluso algunos biólogos como Williams (1966) reconocen que puede ser de cierta ayuda la comparación o el establecimiento de analogías entre ambos ámbitos. Sin embargo este tipo de estrategias están muy alejadas de las posturas extremas del creacionismo o de la teología natural. Ayala y Valentine (1983) señalan que el mérito de Darwin consiste en que proporciona una explicación natural para el diseño funcional. Así, del mismo modo que los teólogos naturales, admite que existe el diseño funcional, pero, a diferencia de estos, rechaza que dicho diseño dependa de algún plan previo e intencional. La fuente del diseño y, por ende, de la teleología que pueda querer ser ver en los seres vivos depende del proceso de selección natural. El proceso de selección natural ya no es un heurístico, sino que es una explicación completa de cómo procede la naturaleza y que para Darwin está suficientemente bien justificado³⁹.

La selección natural es una fuerza no intencional que elimina la idea de una creación propositiva del ámbito de la naturaleza, la teleología depende de un proceso mecánico y no de un proceso

³⁹ Vid. Apéndice.

intencional. Los fines naturales son producto de un mecanismo ciego y aleatorio que procede por ensayo y por error. La idea de selección natural no sería posible sin alusión al pensamiento poblacional, precisamente porque la selección natural como fuerza actúa sobre poblaciones de individuos y no sobre individuos aislados es posible explicar tanto las variaciones presentes en los organismos como su adecuación al medio en el que se desarrollan.

3. La pervivencia de la teleología en el pensamiento científico y filosófico contemporáneo

El pensamiento teleológico pervive bajo distintas apariencias en la filosofía y en la ciencia del siglo XX. Por una parte parece irrenunciable desde la perspectiva de la teoría de la acción. Por otra parte se ha insistido también desde el campo de la teoría de sistemas y la cibernética. En este sentido son claves las aportaciones de Von Bertalanffy (1968) y de Wiener (1948). Estos autores insisten en el concepto de sistema y en el de organización. Un sistema es para estos autores una clase natural a la que cabe atribuir procesos de autoorganización que se mantienen a través de un programa instanciado en el propio sistema. Explicar los fines que cabe atribuir a estos sistemas consiste en explicar cómo estos sistemas mantienen su organización pese a las alteraciones a las que los somete el medio en el que se encuentran. La homeostasis sería un ejemplo paradigmático de que los sistemas organizados tienen, en algún sentido, fines, pues están dotados de mecanismos de retroalimentación que permiten alcanzar un equilibrio óptimo para el desarrollo de sus actividades.

Siguiendo las reflexiones de Bedau (1992) a propósito del concepto de fin en la contemporánea filosofía analítica, ya más concretamente en la filosofía de la ciencia, conviene distinguir tres grandes familias donde podríamos situar las distintas aportaciones que se han ido haciendo.

En primer lugar tenemos una *aproximación mentalista* como es la mostrada por Ducasse o Woodfield. Estos autores sostienen que la genuina teleología tiene que ver con los designios de una mente humana. Así, sólo cabe esperar teleología de acciones humanas. Para explicar la acción humana recurren a la psicología de la

creencia/deseo, es decir, a la psicología popular. El argumento de todos ellos es claro; si hay que hablar de teleología sólo puede ser en un ámbito que haga referencia a las intenciones, en el resto de ámbitos no deja de ser una mera ficción con valor instrumental. Lo que estos autores no responden es a la cuestión acerca de cómo surge la intencionalidad y por tanto la posibilidad de una conducta propositiva. No obstante no será tampoco esta la línea de investigación que seguiré en este estudio.

En segundo lugar cabe distinguir una *aproximación sistémica*. Sus principales defensores son Becker, Boorse, Braithwaite o Nagel. Los partidarios de la aproximación sistémica señalan que el pensamiento teleológico se circunscribe a sistemas dirigidos a fines. Un sistema dirigido a fines en un sistema tendente a mantener un cierto estado y, por lo tanto, sus actividades están dirigidas hacia la consecución de ese equilibrio que no es otra cosa que su fin. En la naturaleza encontramos sistemas de este tipo como pueden ser los organismos.

Por último hay que señalar una *aproximación etiológica* cuyos principales artífices son Bennett, Cohen, Ch. Taylor o Wright. Para estos autores los fines que adscribimos a determinados sistemas o rasgos de estos sistemas provienen de la tendencia que estos mismos sistemas poseen para alcanzar un determinado estado. Esta tendencia se estudia en términos de la historia causal que cabe atribuir a los propios sistemas o rasgos. Dentro de esta corriente ha sido particularmente influyente la aportación de Wright que se estudia en el punto 4.1 de este mismo capítulo.

En lo que sigue, mi línea de investigación va a centrarse en la tradición iniciada por la Filosofía de la Ciencia, particularmente por Hempel y Nagel. No es que los otros enfoques destacados no sean importantes; lo son, pero la tradición analítica se ha centrado de un modo insistente en este último dejando de lado la perspectiva iniciada por los teóricos de sistemas o por aquellos autores que únicamente ceñían sus investigaciones al ámbito estrictamente intencional. Así, a partir de la polémica suscitada en torno a las explicaciones funcionales se puede dar debidamente cuenta de cuáles han sido las líneas más relevantes de los últimos remanentes de la teleología en el pensamiento filosófico y científico actual.

3.1. El problema de las explicaciones funcionales en la filosofía de la ciencia

La Filosofía de la Ciencia, desde sus inicios, ha adoptado como modelo de conocimiento científico a las ciencias físicas. Esto ha determinado en gran medida su desarrollo teórico y ha impuesto obvias restricciones para el abordaje explícito de disciplinas que, aunque científicas, quedaban más alejadas de lo que durante mucho tiempo se consideró el objeto de estudio paradigmático de nuestra disciplina.

Incluso desde una perspectiva histórica, las ciencias de la vida han corrido una suerte muy distinta de las ciencias físicas. La revolución científica tiene unos efectos más alejados en su constitución. Su “siglo de oro” no hace sino empezar a finales del XIX con la teoría de la selección natural y las consecuencias técnicas y sociales de su desarrollo se empiezan a sentir únicamente desde hace unas pocas décadas. Las ciencias de la vida a partir de la década de 1970 con la biotecnología consiguen sustituir en el imaginario social a la física de altas energías y comienzan a ser objeto de ambiciosas políticas científicas. Paralelamente, aunque no necesariamente ligado a la teoría evolutiva, podemos hablar también de una ecologización del pensamiento dentro de este mismo imaginario colectivo.

Por si lo anterior no bastara, su lugar en el conjunto del saber científico es peculiar. No hay una sola Ciencia de la Vida, sino que, bajo este rótulo, se agrupan numerosas disciplinas y subdisciplinas que, a menudo, únicamente guardan relación entre sí por dedicarse al estudio de algún aspecto de las entidades vivientes. Así, los intereses de ecólogos, bioquímicos o fisiólogos son muy dispares, al igual que los métodos o las asunciones teóricas que sirven de trasfondo a sus investigaciones. Incluso en su génesis histórica las diversas disciplinas biológicas han seguido caminos muy dispares.

Por otra parte, tampoco queda clara su relación con otras disciplinas científicas, pudiéndose hablar de dos corrientes antagónicas en lo que a este tema afecta. Para algunos teóricos la Biología es una ciencia autónoma respecto del resto de saberes científicos. Otros, en cambio, asumen lo que Rosenberg (1985) ha denominado una postura provincialista; la Biología es una provincia, eso sí, muy alejada, de la Física.

Todos estos problemas corresponden a la Filosofía de la Biología. Sin embargo, no hay que ver esta subdisciplina excesivamente alejada de las preocupaciones de la Filosofía de la Ciencia General. Es más, están mucho más próximas de lo que a primera vista pudiera parecer, incluso cabría afirmar que en ella se tratan algunos de los temas clásicos a los que los filósofos tradicionalmente se han dedicado como la cuestión de la identidad, la teleología o el realismo. Pero mis objetivos son más modestos. Una de las preocupaciones fundamentales de la Filosofía de la Ciencia ha consistido en caracterizar adecuadamente la explicación o explicaciones científicas. De hecho, éste es con toda probabilidad junto al problema del realismo en ciencia uno de los problemas más densos e importantes a los que nos dedicamos. La razón es obvia, la cuestión de la explicación científica está relacionada directamente con una de las cuestiones clásicas por antonomasia: qué es el conocimiento y cómo se justifica. Si las diversas ciencias constituyen un paradigma de racionalidad es porque nos permiten tener conocimiento de un determinado fragmento de la realidad en condiciones mucho más fiables. De otro modo: las diversas ciencias nos permiten no solo saber en condiciones, sino saber en mejores condiciones.

3.1.1. Teleología y explicación funcional

Uno de los objetivos básicos de toda ciencia empírica es su carácter explicativo. Esperamos de una buena teoría científica que sea explicativa, que dé razón de un determinado rango de fenómenos. Un saber alcanza el rango de ciencia cuando es capaz de ofrecer explicaciones, cuando puede ir más allá del descubrimiento y descripción de los fenómenos que aborda y estos son englobados en un marco sistemático y nomológico. Al margen de lo que de propio tenga cada ciencia particular, el capítulo de la explicación científica constituye el núcleo de lo que podemos denominar una Filosofía General de la Ciencia.

La etimología del término “explicar” ayuda a desentrañar algo el significado de éste. Etimológicamente *explanare* es extender, desenvolver, desplegar, como si de una alfombra se tratara. La explicación, al igual que el análisis, explicita aquello que está implícito, lo expone o lo hace inteligible. No hay que olvidar que todas estas tareas tienen un claro interés pragmático; se explica o se desenvuelven los argumentos para una audiencia determinada y conforme a ciertos criterios acerca de aquello que se considera más básico o más relevante en la comprensión de un fenómeno. A menudo se afirma que la comprensión es un proceso psicológico, mientras la explicación es lógico. No creo que la polémica deba retrotraerse a la apuntada por Von Wriqth (1971) entre comprensión y explicación circunscribiendo cada uno de los dos miembros anteriores a las Ciencias Sociales y a las Ciencias Naturales respectivamente. En cambio, creo que no debería obviarse el sesgo psicológico que toda explicación conlleva; algo no es explicativo a secas, es explicativo para alguien y bajo unas determinadas condiciones. Separar la explicación de la comprensión a un nivel conceptual puede resultar adecuado, si bien no creo que sea una separación exitosa más allá de cierto nivel de abstracción. Las explicaciones científicas, pese a su neta y clara diferenciación respecto de las explicaciones de sentido común, no dejan de estar dirigidas a una audiencia concreta ni pueden obviar la dimensión pragmática a la que apunta su finalidad. En este sentido no debiera obviarse este aspecto esencial y acentuar únicamente la caracterización de las condiciones estrictas bajo las que podemos decir que algo cuenta como una explicación científica.

Los trabajos de Hempel, Oppenheim y Nagel, además de ser pioneros en lo que al tratamiento contemporáneo de la explicación respecta, han constituido un importante hito en Filosofía de la Ciencia. Sus contribuciones han sido enormemente fecundas y han contribuido a configurar una buena porción de la disciplina tal como la entendemos hoy en día. En lo tocante al concepto de función biológica creo que cabe ver en los problemas suscitados por el modelo nomológico-deductivo y el peculiar tratamiento de las explicaciones funcionales que estos autores abordan, el origen de las diversas teorías que se han propuesto desde los diversos ámbitos. En las páginas que siguen

presento de una manera más detallada estos problemas y las soluciones que contemporáneamente se han ofrecido.

3.1.2. El modelo nomológico-deductivo de Hempel y Oppenheim y sus problemas con las explicaciones teleológicas y funcionales

El modelo nomológico-deductivo formulado por Hempel y Oppenheim (1954), y reelaborado por el propio Hempel en 1965, sostiene que explicar un fenómeno consiste en identificar las leyes generales y condiciones antecedentes conforme a las que el fenómeno se produce. De ahí que el modelo nomológico-deductivo sea también denominado modelo de cobertura legal. Un fenómeno queda explicado cuando podemos establecer una relación de inferencia deductiva entre la oración que describe ese mismo fenómeno y que denominaremos *explanandum* y un conjunto de premisas que constituyen el *explanans* u oraciones que empleamos para explicar. El *explanandum* (E), por lo tanto, se infiere deductivamente (\therefore) del *explanans*. Las oraciones que componen el *explanans* son fundamentalmente leyes científicas ($L_1, L_2, L_3, \dots, L_n$), enunciados asertivos con fuerza normativa o, de otra manera, enunciados apoyados en contrafácticos. Junto a las leyes científicas encontramos una serie de enunciados que describen las condiciones iniciales y las cláusulas *caeteris paribus* bajo las que el fenómeno aparece ($C_1, C_2, C_3, \dots, C_n$). El gráfico que a continuación se expone sintetiza bastante bien la estructura del modelo nomológico-deductivo.

$$\begin{array}{ccc} & L_1, L_2, L_3, \dots, L_n & \\ & C_1, C_2, C_3, \dots, C_n & \therefore \\ E & & \end{array}$$

Además de la relación de deducibilidad a partir del *explanans*, Hempel impone la satisfacción más precisa de una serie de requisitos o condiciones lógicas de adecuación. Una explicación cuenta como tal cuando se satisfacen las tres condiciones formales siguientes:

R.1. El *explanandum* es consecuencia lógica del *explanans*

R.2. El *explanans* contiene al menos una ley general que permite la derivación del *explanandum* a partir del *explanans*.

R.3. El *explanans* debe tener contenido empírico; el *explanans* debe ser comprobable mediante algún tipo de dispositivo experimental. Esto conlleva la condición de ser verificable y, en última instancia, significativo.

Además debe satisfacerse una condición de adecuación empírica, así:

R.4. Las oraciones que constituyen el *explanans* deben ser verdaderas.

R.4. nos sugiere que el *explanans* debe ser verdadero, ya que en caso contrario únicamente tendríamos una explicación potencial, saber que el *explanans* es verdadero nos garantiza la corrección de la explicación.

Obsérvese ante todo el carácter formal del procedimiento propuesto por Hempel y Oppenheim. El problema de la explicación científica se dirime por completo en el ámbito del lenguaje y mediante un determinado tipo de lógica, la lógica clásica. El modelo nomológico-deductivo no parte, en principio, de ningún compromiso ontológico sino de un principio de significatividad, en la medida en que las premisas tienen contenido empírico en virtud del criterio empirista del significado podemos decir que se relacionan con la realidad. No hay ninguna asunción realista que nos haga suponer que el orden de la realidad y el orden del lenguaje coinciden. En cualquier caso y en la medida en que nuestras teorías parecen funcionar suponemos que deben coincidir. La epistemología propia del neopositivismo vienés no es de base realista, no al menos si pensamos en un realismo epistemológico robusto o incluso ingenuo. En mi opinión la base epistemológica del neopositivismo hay que buscarla en el fenomenismo. A lo que un sujeto epistémico tendría acceso únicamente es a sus datos de los sentidos, todo lo que implique ir más allá de estos es puro acto de confianza, si bien puede haber razones de mayor o menor peso para confiar en que el mundo sea de una manera determinada o de otra. En principio deberíamos pensar que algún tipo de relación estrecha existirá entre los datos de nuestros sentidos y la realidad objetiva⁴⁰.

Lo dicho anteriormente tiene sentido desde el momento en que es muy fácil proceder a una interpretación realista del modelo nomológico-deductivo, fundamentalmente cuando se aborda la relación entre deducibilidad y causalidad. El propio Hempel señala que comúnmente se denomina explicación causal al tipo de explicación considerado bajo el rótulo de modelo nomológico deductivo. La explicación causal, como él mismo afirma es una variedad de tipo deductivo.

“El tipo de explicación que hemos considerado hasta aquí comúnmente se denomina explicación causal. Si E describe un hecho concreto, puede decirse entonces que las circunstancias antecedentes señaladas en las oraciones C₁, C₂, ... C_k “causan” en conjunto aquel hecho, en el sentido de que existen ciertas regularidades empíricas expresadas por las leyes L₁, L₂, ...L_r, las cuales implican que toda vez que ocurran condiciones del tipo indicado por C₁, C₂,... C_k, tendrá lugar un hecho del tipo descrito en E. Los enunciados tales como L₁, L₂, ... L_r, que expresan conexiones generales y ordinarias entre características específicas de hechos, se denominan habitualmente leyes causales o deterministas.” Hempel, (1965).

La relación de deducibilidad es un buen candidato para formalizar la causalidad, si bien creo que este último concepto es todavía demasiado basto para quedar adecuadamente recogido por la deducibilidad. Las diferencias son obvias, la causalidad se entiende habitualmente en un sentido temporal; la causa precede temporalmente al efecto mientras que las referencias temporales están fuera del alcance de la noción clásica de deducibilidad.

Lo importante para Hempel es la existencia de leyes científicas, que posibilita la explicación científica como proceso deductivo. Así, explicar consiste en realizar un tipo determinado de inferencia, una inferencia deductiva en la que el *explanandum* queda subsumido bajo el *explanans*. La explicatividad es, pues, una relación lógica en la que la conclusión se deduce de las premisas por medio de una serie de leyes. Si bien estas leyes tiene un doble carácter: por una parte poseen

⁴⁰ Un buen ejemplo de fenomenismo en teoría de la percepción aparece en Russell (1912). Russell recapitula muy bien toda la doctrina de los *sense data* y de la distinción entre cualidades primarias y secundarias del empirismo británico. Igualmente la problemática de la existencia y conocimiento del mundo externo están magistralmente expuestas en el texto citado.

una estructura lógica precisa y, por otra, están dotadas de contenido empírico.

El modelo nomológico deductivo se presenta como una teoría general de la explicación científica y, en principio, cabría suponer que debería poder ser aplicado a todas las disciplinas científicas ya que se presentaba como patrón básico de explicatividad. No obstante y pese a las modificaciones substantivas que otros autores como Salmon hicieron sobre la versión hempeliana del modelo nomológico-deductivo, éste siguió planteando no sólo problemas de aplicación general como el que claramente queda ilustrado con las explicaciones funcionales, sino también problemas de interpretación como aquellos que preocuparon a Salmon y que le llevaron a formular en modelo de la relevancia estadística o el mecánico-causal. Sin embargo, no es necesario ir a las interpretaciones de Salmon, y otros, sino que las propias reflexiones de Hempel o Nagel a propósito de la aplicabilidad de la estrategia explicativa general a ciencias especiales como la Biología, la Psicología o las Ciencias Sociales es una muestra suficientemente representativa del calado de los problemas que abordo en este capítulo.

Las explicaciones funcionales parecen jugar un papel importante en las ciencias anteriormente señaladas. Esto ya fue advertido por Hempel quien, inmediatamente tras su trabajo sobre la explicación científica general y su formulación del modelo nomológico-deductivo (1959), aborda el tema de la explicación funcional⁴¹.

El tipo de explicaciones funcionales que Hempel tiene en mente son las provistas por las contribuciones de Malinowski, Clifford Smith o Parsons, es decir, pertenecen de lleno a las Ciencias Sociales. Sólo tangencialmente aborda el papel que éstas desempeñan en Biología, si bien analiza detenidamente este tipo de casos. De cualquier modo, el uso que tales científicos puedan hacer de la teleología no tiene nada

⁴¹ Aún cuando el problema de la causación retroactiva constituya un importante escollo para la legitimidad de las explicaciones funcionales hay que señalar que lo que Hempel quiso resaltar en su reelaboración del artículo de 1945 en 1965 fue la relación de deducibilidad del *explanans* respecto del *explanandum* frente a una posible interpretación causalista de su esquema explicativo. De otra manera, no parece prestarle demasiada atención a la causalidad, entendida de un modo substantivo.

que ver con una presunta teleología universal; en este sentido no se apartan del requisito de que el *explanans* deba tener contenido empírico. Lo que de teleológico tengan estas explicaciones pertenece por entero al ámbito lingüístico, es decir, están formuladas en un lenguaje teleológico y la cuestión que se va a plantear es si el uso de este lenguaje es legítimo. La conclusión de Hempel será que tales explicaciones no son legítimas, sino que únicamente tienen valor en tanto que heurísticos. El modelo de Hempel es inaplicable a los casos que comúnmente pretenden explicar los biólogos. La lógica de la explicación científica se adapta bastante bien a las teorías físicas, pero no a las teorías y conceptos propios de la Biología⁴².

El principal problema de las explicaciones funcionales desde el punto de vista del modelo nomológico-deductivo es formal; una explicación funcional es formalmente falaz y, en consecuencia, la relación de deducibilidad entre *explanans* y *explanandum* es débil. Por otra parte, si tratamos de enmendar el argumento de manera que sea formalmente correcto nos encontramos con que o bien arroja una trivialidad irrelevante o bien se sustrae de la exigencia de tener un contenido empírico.

Podemos considerar una explicación funcional cualquiera como la ilustrada por el siguiente ejemplo: Los vertebrados son animales vascularizados, en éstos la presencia de un órgano que haga circular la sangre es condición necesaria para que el organismo funcione adecuadamente.

El objeto de análisis en este tipo de explicaciones es un determinado ítem *i* que constituye un rasgo o disposición y que ocurre en un sistema *S* de forma que $i \in S$ y *S* se encuentra en el conjunto $C = \{ C_i, C_e \}$ que es el conjunto de condiciones internas y externas del sistema *S* tal que $C(S) = \{ C_i, C_e \}$.

⁴² Éste históricamente ha sido también uno de los argumentos esgrimidos a favor de una postura autonomista de la Biología respecto de la Física. El hecho de que la lógica de la explicación científica que pretenden Hempel o Nagel no se adecue a la Biología ha sido denunciado por Mayr (2006). El problema, sostiene Mayr en una afirmación polémica y provocativa, no es que la Biología no se adapte a la lógica de la explicación científica, es que la Biología no tiene nada que ver con la Lógica.

El objetivo del análisis consiste en demostrar que *S* se encuentra en un estado de condiciones *C* tal que en *C*, *i* tiene efectos que satisfacen un requisito funcional o necesidad para que *S* siga funcionando. Podemos decir que *i* es una condición necesaria para que *S* siga funcionando o que los efectos de *i* son condición necesaria para que *S* funcione. La estrategia del modelo nomológico-deductivo procederá estableciendo la relación de dependencia deductiva entre las premisas del *explanans* y la oración del *explanandum*. De tal forma el argumento adoptaría la siguiente estructura (presento dos semi-formalizaciones equivalentes).

<p>(1) En el tiempo <i>t</i> el organismo <i>O</i> funciona adecuadamente bajo unas determinadas condiciones <i>C</i>, internas <i>C_i</i> y externas <i>C_e</i>; (2) Si esto es así es porque se cumple una determinada condición necesaria <i>N</i>, (pongamos por caso que circule la sangre, tal que podemos afirmar que sólo si circula la sangre el organismo <i>O</i> funciona adecuadamente); (3) Si el ítem funcional <i>I</i> (el corazón) está presente en <i>O</i>, entonces se cumple <i>N</i></p> <p>(4) En conclusión, <i>I</i> está presente en <i>O</i>, (el corazón está presente en el organismo).</p>	<p>(1) En <i>t</i>, <i>S</i> funciona adecuadamente en un ámbito <i>C</i></p> <p>(2) <i>S</i> funciona adecuadamente en <i>C</i> sólo si se satisface cierta condición necesaria <i>N</i></p> <p>(3) Si el rasgo <i>i</i> estuviera presente en <i>S</i>, entonces como efecto se cumpliría <i>N</i></p> <p>(4) Conclusión en <i>t</i> el rasgo <i>i</i> está presente en <i>S</i></p>
---	--

Desde un punto de vista formal no necesitamos más que la lógica proposicional para formalizar adecuadamente la anterior argumentación:

1. *p* Premisa
2. $p \rightarrow q$ Premisa
3. $r \rightarrow q$ Premisa
4. *q* *Modus Ponens* entre 1 y 2
5. *r* Conclusión

El argumento anterior es un claro ejemplo de falacia de afirmación del consecuente. Ésta reproduce la reconstrucción de Hempel de una explicación funcional desde los presupuestos del modelo nomológico-deductivo. Tiene la virtud de que distingue la condición necesaria q del ítem que la satisface, en este caso r . La razón de que esto sea así hay que buscarla en la necesidad de mantener las equivalencias funcionales. De hecho la premisa 3 podría haberse interpretado como q , es decir, únicamente constatando que, en efecto, se satisface la condición necesaria. Pero no creo que sea una formalización conveniente, pues la premisa constituye una ampliación de q y no una simple repetición o constatación. Esta objeción no se plantearía si la tercera premisa pudiera ser reemplazada por el enunciado de que el requisito de satisfacción de la condición necesaria únicamente puede ser llevado a cabo por la presencia del ítem funcional i y no por un equivalente funcional, aunque esto es empíricamente poco plausible. Volvamos al ejemplo concreto anteriormente citado. Que el corazón, el ítem funcional, esté presente en el organismo no es óbice para que circule la sangre. De otro modo, (4) no se sigue necesariamente de (3); ya que hay otros modos de satisfacer la condición necesaria, es decir, de hacer que la sangre circule. Un corazón no es una condición *sine qua non* para que la sangre circule. Bien podemos imaginar equivalentes funcionales que ejecuten la misma función que el ítem funcional que consideramos.

Para salvar el argumento podemos tratar de reformular la tercera premisa dándole un alcance mayor y afirmar, en consecuencia que, I es la clase de todas las condiciones empíricamente suficientes para llevar el requisito N en el contexto determinado por el sistema S en C (y obviamente $I \neq \emptyset$).

De una manera más intuitiva se está diciendo lo siguiente:

Sea S un sistema, I un conjunto de ítems que satisface el requisito N y N es la condición necesaria que debe ser satisfecha para que S funcione adecuadamente.

1. S
2. $S \Rightarrow N$
3. $I; I \neq \emptyset$
4. $\exists x I(x) \ \& \ I \subset S$

En 4 estamos diciendo que algunos de los ítems incluidos en la clase I están presentes en el sistema S en el momento t . el argumento es formalmente correcto pero apenas resulta explicativo. Podemos explicitar el análisis anterior todavía más si entendemos que $I = \{ i_1 \vee i_2 \vee i_3 \dots \vee i_n \}$. Este análisis es trivial salvo que tengamos información adicional acerca de los elementos de I .

Tras el análisis anteriormente presentado, Hempel concluye que el único valor que puedan tener las explicaciones funcionales es heurístico y que, a medida que el conocimiento científico progrese, éstas deberán ser sustituidas por explicaciones acordes con el modelo nomológico-deductivo. Hempel no realiza ninguna consideración a propósito de la posible autonomía de la Biología ni de las Ciencias Sociales como disciplinas científicas. En realidad no hace ningún tipo de afirmación de carácter ontológico sobre si estos ámbitos tienen alguna característica diferenciadora. Pero es que tampoco realiza ninguna afirmación con carácter ontológico sobre la Física. Lo único que Hempel señala es el tipo de inferencia que puede establecerse a partir del presupuesto de la existencia de leyes de la naturaleza, y tampoco debiera olvidarse el carácter lingüístico de éstas. Cualquier afirmación a propósito de la existencia de los objetos de estudio de las distintas disciplinas científicas no debería ir más allá de enunciados acerca de los datos de los sentidos a los que el sujeto tiene acceso y sobre los que constituye su experiencia del mundo. El resto es elaboración lógica.

3.1.3. La solución de Nagel: la inexistencia de explicaciones genuinamente teleológicas o funcionales y la traducción al modelo de cobertura legal

Nagel (1961)⁴³ analiza la misma anomalía que Hempel respecto de las explicaciones teleológicas. Aunque admite que hay diferencias obvias entre las Ciencias Físicas y la Biología, no cree que esto deba dar lugar a un pluralismo epistemológico en lo que al modelo explicativo concierne. Así, el modelo que justifica una explicación como

⁴³ Cf. Nagel, (1961) Cap. XII, secc. I. Es el punto de partida para la reflexión de Cummins (1975).

genuinamente científica es el de cobertura legal, y esto debe ser aplicable a todos los campos del saber científico, no sólo a las ciencias fisico-químicas.

Las razones para defender un modelo de explicación científica exclusivo de las ciencias de la vida hay que buscarlas en la polémica acerca de su autonomía respecto de las ciencias fisico-químicas. La diferencia entre lo vivo y lo inerte puede resultar, a los ojos de algunos, crucial. Nagel sostiene que aunque no pueda establecerse una diferencia muy precisa entre lo vivo y lo inerte, es obvio que tal diferencia existe⁴⁴. Las explicaciones teleológicas, por lo tanto, parecen tener un papel predominante en Biología a diferencia de lo que ocurre en otras disciplinas. Pero ¿debe admitirse que la Biología sea una disciplina autónoma respecto de la ciencia Física, es decir, que los principios de la lógica de la explicación propuestos para la ciencia Física como modelo de lo que es un saber científico no sean aplicables a campos como la Biología? Nagel no cree que la presencia de explicaciones teleológicas deba justificar la autonomía de la Biología, sino que en cualquier caso debe ser posible, además de deseable, reformular las explicaciones teleológicas en términos no teleológicos, es decir, adaptarlas al esquema nomológico-deductivo general. Las explicaciones teleológicas o funcionales no son más que un tipo de explicación abreviada que puede traducirse sin pérdida alguna de contenido a explicaciones que se ajusten al modelo de cobertura legal. La estrategia que sigue Nagel es una estrategia de traducción y el argumento del que Nagel se sirve para defender su tesis puede reconstruirse de la siguiente manera:

Sean *EF* y *EC* respectivamente la explicación funcional y causal (o acorde con el modelo *ND*) de un fenómeno orgánico como pueda ser la respiración celular o la fotosíntesis en las plantas verdes⁴⁵. *EF* es una forma abreviada de *EC*, de otra manera *EF* no añade nada que no esté ya presente en *EC*.

⁴⁴ Del mismo modo también es obvio que un organismo no parece ser un mero sistema aditivo, sino antes bien, es una totalidad organizada capaz de autorregulación, automantenimiento y autorreproducción. Este tipo de consideraciones acerca de la diferencia entre entidades inertes y vivas es uno de los argumentos de más peso a la hora de defender la autonomía de la biología.

⁴⁵ Este último es el ejemplo del que se sirve Nagel

Un ejemplar del tipo *EF* puede ser cualquier enunciado explicativo que contenga alguna de las locuciones “*con el fin de*”, “*para que*”, “*el propósito de*” etc. Tomaré como ejemplo ilustrativo (1).

(1) “La función de la clorofila en las plantas verdes consiste en permitir realizar la fotosíntesis a estas.”

Nagel sostiene que el mismo fenómeno al que refiere (1) puede ser expresado sin recurrir a términos típicamente teleológicos o funcionales y sin pérdida alguna de contenido explicativo y puede, en consecuencia, ser reformulado como (2).

(2) “La clorofila en las plantas verdes es una condición necesaria para el proceso de la fotosíntesis”.

Sean respectivamente *A*: una determinada sustancia o ítem cualquiera como individuo destacado del sistema *S*; *s*, tal que $s \in S$, es un miembro perteneciente a una clase de sistemas *S* con una cierta organización *C* de partes y procesos componentes. *E*: un medio determinado (interno o externo) y *P* un proceso que arroja un determinado resultado. (1) Sería un argumento del tipo mostrado en la sección dedicada a Hempel. Así, la función de *A* en *S* es permitir a *S* en el medio *E* realizar el proceso *P*. Cuando a la planta (*S*) se le suministra agua, dióxido de carbono y luz solar (*E*) elabora almidón (*P*) sólo si la planta contiene clorofila (*A*). En cambio (2) nos dice que cuando se les suministra agua, dióxido de carbono y luz solar a las plantas entonces estas elaboran almidón si tiene clorofila. Obsérvese que no se dice que la clorofila *tiene la función* de elaborar almidón o está *para* elaborar almidón. De otra manera, *A*, el ítem inicialmente funcional, es una condición necesaria para el mantenimiento de *S*.

Como premisa implícita hay que señalar que bajo las mismas circunstancias y en ausencia de clorofila, las plantas no elaboran almidón.

La formulación de cualquier enunciado como (1) quedaría del modo siguiente:

1. Premisa: Todo *S* realiza *P* en el medio *E*.

2. Premisa: Si S no tiene A entonces no realiza P
3. Conclusión: S debe tener A.

El argumento señala que la posesión del rasgo A explica por qué S realiza P sin necesidad de recurrir a ningún tipo de noción funcional o teleológica. Lo que en el análisis de Hempel aparecía como rasgo funcional aquí es una condición necesaria. No es que para Hempel fuera así, de hecho era una condición necesaria, pero sólo una entre un determinado número de alternativas o equivalentes funcionales. Para Nagel es la condición necesaria y no hay lugar para la consideración de equivalentes funcionales. De manera más clara señala:

"(L)os análisis teleológicos de la biología (o de otras ciencias en las cuales se realicen tales análisis) no son exploraciones de posibilidades puramente lógicas, sino que tratan de las funciones reales de componentes definidos en sistemas vivientes concretos." Nagel (1961) p.228 de la ed. esp.

¿Cuáles son los problemas de la propuesta de Nagel? A mi juicio hay que considerar tres grandes problemas. En primer lugar los referentes a la polémica acerca de la autonomía de la biología y las cuestiones generales que plantea el realismo científico. En segundo lugar, y de manera más específica, la estrategia que sigue Nagel en su argumentación a propósito de la eliminación de los equivalentes funcionales. De un modo más conciso esto queda concretado en la siguiente pregunta: ¿queda dañada la legitimidad de una explicación científica cuando se eliminan los equivalentes funcionales? Por último, la cuestión más importante en todo este asunto y que justifica la presencia de las diferentes teorías sobre la funcionalidad biológica en este capítulo, a saber ¿es satisfactoria la traducción que realiza Nagel entre las explicaciones funcionales y las denominadas explicaciones causales? Es obvio que no, aunque todavía no haya explicitado por qué.

En cuanto a lo primero no creo que si se asume la autonomía de la Biología deba hacerse demasiado hincapié en que sólo hay un modelo explicativo válido al que ésta deba ajustarse. Aún así es una postura burda y poco precisa; la autonomía puede entenderse al menos de dos

maneras. Podemos distinguir una autonomía metodológica o débil y una ontológica o fuerte. Si la entendemos metodológicamente es obvio que los argumentos de Hempel o Nagel son más poderosos. Podríamos preguntarnos en base a qué se supone y defiende tal autonomía metodológica. En principio cabría pensar que viene determinada por la naturaleza del propio objeto de estudio, el mundo orgánico. Pero si es así, entonces la balanza se inclina a favor de los defensores de la autonomía de la Biología y los argumentos de Nagel y Hempel pierden peso. Así las cosas, sería lícito decir que la Biología no sólo es autónoma desde un punto de vista ontológico, sino también metodológico y que no hay razón alguna para suponer un único modelo de explicación científica o decantarse por proyectos reduccionistas condenados de antemano al fracaso más estrepitoso. Sólo si se admite como postura teórica fuerte la autonomía ontológica de la Biología puede hacerse frente a los supuestos de fondo que articulan la Filosofía de la Ciencia neopositivista.

En lo que atañe a lo segundo hay que señalar que la solución de Nagel al problema es pragmática en extremo. Nagel sostiene que si los equivalentes funcionales realmente causan problemas entonces deben eliminarse. En la naturaleza generalmente no se dan tantas alternativas que no podamos considerar. Pero esta solución no es factible desde un punto de vista filosófico. ¿Por qué razón deberíamos contentarnos con asumir el modelo nomológico-deductivo? Si necesita de tantos remedios es obvio que algo no funciona bien. No parece coherente pensar en términos pragmáticos en unos momentos y en otros comportarnos como positivistas ortodoxos. Que la explicación por cobertura legal fracase en un ámbito tan amplio como en el de las Ciencias Naturales y Sociales da mucho que pensar, pues no es una porción pequeña de la totalidad de nuestro saber científico. El problema principal ha consistido en ver como anomalía lo que en realidad era la norma y el motivo la confianza en una presunta Física ideal que nunca acabó de estar muy clara.

En tercer y último lugar, la propuesta de Nagel no resulta satisfactoria. Un argumento de peso para rechazarla lo constituye la polémica anteriormente señalada a propósito de la autonomía de la Biología o el problema de los equivalentes funcionales. Estos resultados deberían cuando menos hacernos desconfiar de la

pretendida traducción de Nagel y pensar que el problema de las explicaciones funcionales no es una cuestión que pueda resolverse únicamente desde un punto de vista lingüístico.

4. De explicaciones funcionales a enunciados adscriptores de función en Biología: Los enfoques de Wright y Cummins

Los estudios filosóficos a propósito de la explicación científica deben mucho a los trabajos de Hempel y Nagel. Sin embargo, estos no son únicamente problemáticos en el terreno de las explicaciones teleológicas y funcionales, ya que como modelo acerca de qué sea una explicación científica son cuestionables. Filósofos como B. van Fraassen, ya más alejados de lo que Putnam denominó *la Concepción Heredada*, se ocuparon en la década de 1980 de desarrollar algunos aspectos que tal vez quedaran alejados del interés teórico de los representantes de la Concepción Heredada. Las críticas del modelo de cobertura legal fueron relativamente tempranas, aunque no es éste el lugar para trazar una historia general de la explicación científica a partir de Hempel. Bas van Fraassen en *The Scientific Image* (1980) argumenta a favor de lo que él denomina una pragmática de la explicación. Explicar, sostiene, es responder a un determinado porqué. En lugar de defender que la explicación científica consista en una relación diádica entre la teoría y los hechos, sostiene que hay que añadir una referencia clara al contexto en el que se formula la pregunta y en el que están presentes tanto el cuerpo de teorías a partir del que se explican los hechos como la información fáctica.

Las ideas de Van Fraassen tuvieron una gran repercusión en la literatura a propósito de la explicación científica, cualquier manual de la disciplina recoge sus aportaciones como novedosas, y lo son. No obstante, hay que señalar que el enfoque pragmático fue la primera estrategia a la que recurrieron los filósofos de la Biología como intento de superar las limitaciones impuestas por el modelo de cobertura legal casi una década antes de que Van Fraassen formulara su modelo erotético. La dimensión pragmática de los primeros intentos de los filósofos de la Biología no es deudora de la obra de Van Fraassen, sino que es el resultado de querer solucionar un problema concreto al margen de las consideraciones más generales de la Filosofía de la

Ciencia. En Filosofía de la Biología parece que los problemas, para bien o para mal, se han resuelto durante mucho tiempo en el ámbito doméstico sin una relación aparentemente muy estrecha con el resto de la Filosofía de la Ciencia.

Ya desde el primer artículo relevante de Wright (1973) del que inmediatamente me ocuparé, la estrategia pragmática está presente. En líneas muy generales esta estrategia justificatoria consigna que explicar es contestar a un determinado porqué antes que establecer una estructura nomológico-deductiva. Ese tipo de respuesta viene dada por un enunciado adscriptor de función, es decir, adscribir al ítem i la función F , o lo que es lo mismo, decir que i tiene la función F es un modo de responder relevantemente a un porqué determinado. Un enunciado funcional es, en consecuencia, suficientemente explicativo sin necesidad (aparente) del tipo de *explanans* que propone Hempel ni, como supone Nagel, parece que como *explanans* sea algún tipo de abreviatura de uno mayor que satisfaga los requisitos de Hempel. La tarea que corresponde al filósofo de la ciencia o, de, manera más concreta, de la Biología, es tratar de explicitar el conjunto de condiciones necesarias y globalmente suficientes para poder realizar dicha adscripción. La reflexión filosófica en los últimos veinticinco años no parece haberse apartado ni un ápice de esta estrategia inicial, lo que ha propiciado a mi entender, que la Filosofía de la Biología adquiriera un papel muy secundario en el conjunto de la Filosofía de la Ciencia y que se dedicara a una serie de problemas muy específicos que difícilmente podían interesar al filósofo generalista.

La estrategia de pasar de explicaciones funcionales a los diferentes análisis del concepto de función que a continuación va a exponerse, ha demostrado ser virtuosa. Sin embargo, no creo que sea la solución definitiva, si es que alguna hay, al tipo de problemas que pueden interesar a un filósofo; tampoco creo que pueda suscitar la más leve curiosidad entre los biólogos. A mi juicio el problema debería volver a los cauces de una Filosofía General de la Ciencia y enfrentarse a los problemas tradicionales que han preocupado a los filósofos, al margen de todo lo que de positivo tiene intentar hacer biología teórica.

4.1. El enfoque histórico-etiológico de L. Wright

El análisis de L. Wright (1973) constituye el punto de partida de la reflexión filosófica de las últimas décadas en torno al problema de las explicaciones funcionales y del concepto de función. Así, su análisis logra dar cuenta de la orientación hacia el futuro sin necesidad de recurrir a conceptos tan contraintuitivos como la causalidad retroactiva.

Wright parte de una concepción distinta de la explicación científica que deja un poco atrás los rigores formales de Hempel y Nagel. Considera que la legitimidad de la explicación en ciertos dominios de la ciencia como son los de la biología, la tecnología o la acción racional no vienen dados por la estructura lógica impuesta por Hempel, sino que los mismos enunciados funcionales que tienen cierta fuerza explicativa y esta se hace evidente a través de su análisis. La propuesta de Wright parte de una consideración más pragmática de lo que es una explicación científica. Precisamente, cuando adscribimos una función a un ítem determinado estamos ofreciendo un tipo de explicación, decimos para qué está donde está, cómo actúa o qué hace. Esa información que se transmite a través del enunciado funcional tiene valor explicativo. Así, la propuesta de Wright va a consistir en hacer explícito el poder explicativo de los enunciados funcionales.

Para Wright atribuir una función a un rasgo, sea este natural o artificial, consiste en decir algo sobre su etiología, sobre su historia causal. Así, consigna el siguiente esquema explicativo. Decir que la función de X es Y significa que (a) X está donde está porque hace Y, (b) Y es una consecuencia de que X esté ahí. Ejemplos que satisfacen este esquema pueden ser:

1. Decir que la función del corazón consiste en bombear la sangre significa que (a) dicho órgano está donde está porque bombea la sangre y que (b) el hecho de que bombee la sangre justifica el que esté donde está.
2. Decir que la función del interruptor es dejar pasar la corriente significa que (a) un interruptor está donde está porque deja pasar la corriente y (b) el hecho de que deje pasar la corriente justifica el que esté donde está.

(a) Exhibe la forma etiológica de las explicaciones-adscripciones funcionales, mientras que (b) describe lo que tienen de propio las etiologías funcionales respecto de las otras. El significado de (a) en principio no presenta problemas, afirma o evidencia el comportamiento causal del ítem, mientras que (b) es mucho más problemático, ya que justifica que sea éste y no otro el comportamiento del ítem. Evidencia por qué *X* ha sido seleccionado, de otra manera, cómo *X* llegó a estar donde está. De este modo deja fuera la cuestión peliaguda de que *X* estuviera donde está por accidente, ya que (b) expone la filiación del ítem que manifiesta la función, más técnicamente, su historia funcional. Por tanto explica no sólo qué hace *X*, sino que justifica también la ocurrencia de la función de *X* apelando al trasfondo causal de la existencia de *X*, el hecho de que *X* esté donde está. *Y* es una consecuencia causal de que *X* esté presente. Una de las propiedades más relevantes de la consecuencia causal es su asimetría. Que *Y* sea consecuencia causal de *X* no implica que *X* sea consecuencia causal de *Y*. Esta relación de consecuencia causal no implica que siempre que se dé *Y* deba darse *X*, puede haber una multiplicidad de ítems que den lugar a la misma función. Nos encontraríamos en este caso con equivalentes funcionales y la explicación provista carecería de adecuación material, pues no podríamos identificar el ítem que ejecuta la función, sino sólo la función misma. El caso es que si no nos fijamos solo en la función, sino que tenemos algún interés, y parece que este interés es legítimo, en saber qué ejecuta la función, debemos añadir una cláusula que especifique la filiación de la función, esta cláusula es *b*. Este aspecto de que *Y* sea una consecuencia de *X* se puede captar diciendo que se dan circunstancias en las que *X* es no redundante para *Y*.

Anteriormente he señalado que el análisis era aplicable tanto a rasgos biológicos como a entidades artificiales. En efecto, la idea subyacente al análisis que propone Wright es que la esencia de la funcionalidad hay que buscarla en cierta clase de etiología. En el caso de los organismos la etiología viene dada por los mecanismos de evolución por medio de la selección natural; en el caso de los artefactos ésta es puramente intencional y depende, en consecuencia, de los planes, deseos e intenciones del diseñador.

En suma, el análisis de Wright hace gala de satisfacer las tres condiciones siguientes: (1) Logra establecer una distinción entre los efectos accidentales de un ítem y sus funciones. Es decir, no hay lugar para una confusión entre funciones y simples efectos. (2) Es aplicable tanto a organismos como a entidades artificiales y (3) muestra como los enunciados adscriptores de función tienen de por sí una fuerza explicativa sin tener que recurrir al modelo nomológico-deductivo, es decir, la explicatividad está en el mismo concepto y no depende de una serie de inferencias entre enunciados con contenido empírico y leyes. Estas tres condiciones trazan la agenda de problemas relevantes sobre los que se ha constituido el núcleo de la discusión filosófica en torno a tópicos de las funciones en las últimas dos décadas y que, con creces, excede el ámbito de la Filosofía de la Biología o la Filosofía de la Ciencia.

4.2. El enfoque sistémico de R. Cummins

La propuesta de Cummins (1975) es la alternativa a los planteamientos histórico-etiológicos. A diferencia de Wright, Cummins no propone un enfoque etiológico, de hecho desvincula la etiología de la funcionalidad y, en su lugar, propone una teoría sistémica aplicable en cualquier ámbito en el que se requiera una explicación funcional. Por lo tanto, para Cummins la cuestión acerca de cómo llegó X a estar donde está, o a ejecutar Y, es secundaria, lo importante es cómo, de hecho, X ejecuta Y. Así, Cummins no ofrece una explicación basándose en consideraciones históricas sino en las capacidades y disposiciones actuales del ítem analizado. Por otra parte, Cummins presenta su análisis funcional como un intento de rescatar la validez que había quedado en entredicho tras el análisis de Hempel de la explicación funcional. El análisis sistémico funcional es la respuesta que da Cummins a los problemas de la explicación funcional presentados por Hempel y por Nagel. Con él pretende legitimar las explicaciones funcionales en Filosofía de la Ciencia, pues sostiene que hay fenómenos cuya naturaleza responde mejor a una caracterización genuinamente funcional. Para ello propone un modelo que sea capaz de justificar metodológicamente esa caracterización. El error de Hempel y Nagel radicaba en asumir que la explicación funcional, en

caso de ser legítima, no era esencialmente distinta de la explicación científica en general y que, por tanto, de un modo u otro, el modelo nomológico-deductivo podía justificarla. Dicho de otro modo, los fenómenos que caían bajo las supuestas explicaciones funcionales no eran esencialmente distintos de los que se ajustaban al modelo de cobertura legal y, en consecuencia, eran igualmente explicados por este último.

Cummins sostiene que decir de algo que tiene una función implica afirmar un enunciado adscriptor de función con la siguiente estructura:

“x functions as a φ in s (or: the function of x in s is to φ) relative to an analytical account A of s’s capacity to ψ just in case x is capable of φ -ing in s and A appropriately and adequately accounts for s’s capacity to ψ by, in part, appealing to the capacity of x to φ in s”. Cummins, 1975: 762.

Decir de algo que tiene una función conlleva describir un sistema y analizarlo en sus partes constituyentes de forma que podamos establecer cómo contribuye cada una de ellas en relación a la economía global del sistema. Esa contribución se entiende en términos de disposiciones o capacidades actuales. El objetivo de Cummins es, pues, ofrecer un análisis de estas disposiciones y no una subsunción de fenómenos bajo leyes generales. La búsqueda de leyes para su posterior aplicación a un modelo explicativo de cobertura legal es descartada. La propuesta de Cummins, pues, se plantea desde una estrategia analítica por oposición a una estrategia subsuntiva. La diferencia de esta última respecto de la primera consiste en que, antes que subsumir una regularidad disposicional bajo una ley no específica para el objeto dispuesto, la estrategia analítica analiza una disposición en relación a un rango de disposiciones manifestadas por el mismo objeto o por sus componentes. Este procedimiento *top-down* permite hablar de capacidades o habilidades fehacientes antes que de meras disposiciones. Así, la estrategia subsuntiva procede de abajo a arriba – dado un fenómeno, se busca las leyes que lo explican-, mientras que la analítica de arriba abajo –dado un fenómeno, se descompone en elementos más básicos-. Si suponemos que un objeto a tiene una disposición d, la regularidad disposicional asociada a a consiste en el hecho de que ciertas clases de eventos podrían hacer que a

manifestara d. Una manera de explicar esto podría consistir en averiguar algunos rasgos de a que nos permitieran representar la relación entre los eventos causantes y las manifestaciones como instancias de una o más leyes generales. Es decir, subsumimos la regularidad disposicional bajo una o varias leyes que no tienen por qué ser específicas para el objeto a, sino que se aplican generalmente a todos los objetos. De la otra manera especificamos claramente a qué disposición nos referimos o qué capacidad consideramos relevante. La estrategia subsuntiva es, pues, inespecífica para explicitar la funcionalidad de a. La estrategia analítica procede de modo inverso, en primer lugar tenemos una capacidad y, posteriormente, la analizamos en relación a un nivel determinado de ejecución que nos delimita el nivel de análisis. Así el fenómeno funcional lo descomponemos en tareas más sencillas dentro del sistema en el que éstas operan y explicitamos su contribución causal. No tenemos por qué recurrir a leyes generales e inespecíficas, sino que únicamente damos razón de cómo a contribuye causalmente en el sistema que lo alberga. Pero a la vez la explicación no resulta acientífica, pues se establece en términos de contribuciones causales. Obviamente esto da lugar a múltiples niveles de análisis y por tanto, múltiples adscripciones funcionales. Este rasgo, como se verá, resulta enormemente problemático para otros teóricos.

Este tipo de análisis son habituales en diversas áreas científicas. Podemos encontrarlo en bioquímica, fisiología, informática, electrónica, etc. La representación más habitual del mismo son los diagramas de flujo o los circuitos. Así, por ejemplo, en bioquímica podemos tener un diagrama que represente las diferentes rutas metabólicas, en fisiología un esquema que explique el funcionamiento de un sistema particular como el circulatorio, en psicología un diagrama de flujo, etc. En cualquier caso veremos que se trata, en primer lugar, de delimitar un sistema que, a su vez, puede estar relacionado con otros, y mostrar cuáles son sus elementos componentes y cómo interactúan entre sí. Si consideramos un ejemplo como el sistema circulatorio veremos que éste se subdivide en unidades como el corazón y las venas y las arterias, pero que también está relacionado con otros sistemas como el respiratorio o el nervioso autónomo. Igualmente un elemento de ese sistema puede ser tomado como un sistema en sí mismo, por ejemplo

el corazón, éste sería a su vez susceptible de un análisis sistémico similar, etc.

Las caracterizaciones funcionales tienen sentido cuando tratamos con niveles complejos de organización que fácilmente los podemos descomponer. Sin embargo en la investigación llega un momento en que la estrategia analítica empieza a zozobrar y hablar de funciones carece de sentido. Uno puede hablar de fisiología en términos funcionales, sin embargo la conductividad eléctrica de la neurona, por ejemplo, quizás resulte más adecuadamente explicada desde una estrategia subsuntiva, es decir, una estrategia basada en leyes generales como las del electromagnetismo.

Admite que el nivel de análisis venga dictado por los intereses del investigador.

En cualquier caso hay un rasgo muy importante que debe ser tenido en cuenta. Como se habrá visto cualquier enunciado adscriptor de función implica un enunciado disposicional. Atribuir una función consiste en atribuir a algo una determinada disposición. Este rasgo no debería ser tomado a la ligera, a mi juicio permite un análisis que quizás muchos de los críticos no han tenido en cuenta. Según el propio Cummins y muchos de los autores contemporáneos, tal caracterización no implica normatividad alguna. Estos últimos no dudan en considerar que el principal escollo de la teoría de Cummins sea la normatividad, ya que, como dirá Millikan, Cummins es demasiado liberal en la atribución de funciones⁴⁶

¿Puede ser normativa la teoría de Cummins? El primer paso para responder a esta cuestión consiste en enfrentarse a la arbitrariedad de los niveles de análisis que pretende. Si algo es arbitrario ¿cómo puede aportar conocimiento y cómo puede resultar, además, normativo? La cuestión epistemológica es una versión del viejo problema del realismo en filosofía de la ciencia. Tan sólo diré que desde el sano sentido común y buen hacer científico –que no defino- se soslaya cómodamente. Se trata de que entre la representación y el objeto de estudio hay una relación de adecuación que permite explicaciones satisfactorias en todas las aplicaciones de la teoría en cuestión, (la

⁴⁶La crítica de Millikan a la teoría sistémica aparece principalmente desarrollada en su artículo *An ambiguity in the notion "function"*, de 1989.

naturaleza de la adecuación de las representaciones o del mismo concepto de representación no es objeto de mi discusión). La arbitrariedad puede entenderse así: es el investigador quien determina el nivel de análisis; la elección es arbitraria en tanto que elección, pero no en lo que respecta al objeto de estudio que muy bien puede consistir en una clase natural. Por otra parte, la arbitrariedad no debe entenderse como aleatoriedad, ésta última se entiende desde el puro azar, mientras que la primera supone un acto deliberativo en el que hay un comercio de razones e intereses por parte del investigador. En un momento determinado puede resultar más exitoso un análisis micro, mientras que en otro el nivel de análisis puede ser más macro, sin que por ello se den contradicciones. En suma, cabe suponer que los intereses del investigador son legítimos y que éste no está interesado en establecer enunciados funcionales espurios como aquel que afirma que una de las funciones del corazón consiste en hacer ruido para que los médicos dictaminen su estado de salud.

El diseño del sistema puede ser visto como resultado de un proceso selectivo que orienta al conjunto de los elementos que constituyen el diseño a adquirir cierta forma. El diseño es de carácter general, dice cómo fue diseñado el objeto y para qué, pero no cómo ejecuta su función. El que ejecute la función del modo en que la ejecuta depende no sólo del diseño, sino de unas ciertas condiciones de trasfondo que acompaña al diseño.

Pero hay una manera más audaz de interpretar la normatividad que surge del mismo texto de Cummins. Éste, al caracterizar a los enunciados funcionales como enunciados que implican disposiciones, nos da la clave. Atribuir una disposición a un objeto consiste en atribuir a ese mismo objeto una regularidad que denominamos regularidad disposicional. Explicar tal regularidad disposicional implica explicar cómo se provocan las circunstancias o requisitos que la desencadenan. Es cierto que esto deja fuera de lugar a la propositividad naturalizada de la que hablará Millikan, pero también es cierto que abre la puerta a procesos que pueden restringir el grado de liberalidad de las atribuciones funcionales y que, en consecuencia, pueden garantizar cierta normatividad sin necesidad de hablar de propositividad natural. Esto se verá más adelante cuando analice los intentos conciliadores que subsumen diversas teorías histórico-causales bajo la égida de la

teoría sistémica. En cualquier caso, en mi opinión, la teoría de Cummins está formulada de una manera que no tiene por qué excluir intentos normativizadores, se trata únicamente de hacerlos presente. Por otra parte es una teoría que tiene la ventaja de compartir gran parte de nuestras asunciones de sentido común acerca de la naturaleza de la realidad física sin necesidad de recurrir a confusas nociones teleológicas.

Como se habrá podido observar hay varias similitudes entre el análisis funcional de Hempel y el de Cummins, pues ambos tratan de sistemas y contribuciones causales de sus distintos elementos. Sin embargo las diferencias son notorias. Hempel insiste en que la explicación por medio del análisis funcional requiere del concurso de leyes, mientras que esta exigencia está ausente en la propuesta de Cummins, quien únicamente hace referencia a las contribuciones causales de los distintos elementos del sistema. El análisis de Cummins renuncia en cierto modo a la presunción explicativa de Hempel ya que no busca una validez universal o una generalizabilidad máxima. En este sentido casi es legítimo decir que Hempel no está equivocado al afirmar que el valor del análisis funcional no es explicativo, al menos tal como Hempel entiende la explicación científica. Pero, por otra parte, aún considerado como lo que es, un método de análisis que, además, permite definir el concepto de función, es injusto negarle cierto valor explicativo que, a mi juicio, es obvio.

Por último hay que señalar que, veladamente, el análisis de Hempel implica la cuestión de la normatividad entendida como no liberalidad en la adscripción funcional. Esto último se hace patente al vincular el análisis funcional al concurso de leyes científicas. No es lo mismo una ley científica que la contribución causal de un ítem al sistema al que pertenece. Las diferencias son notorias, pero hay cierto aire de familia.

4.3. La herencia de Wright y Cummins

Tanto la teoría de Wright como la de Cummins ofrecen análisis relevantes sobre el concepto de función. Pero como inmediatamente mostraré no son del todo satisfactorios. La teoría de Wright tiene la ventaja de ser una teoría normativa. Explica por qué un determinado rasgo tiene la función que tiene y no otra. Es decir, es una teoría

robusta en la adscripción funcional, ya que parece impedir que se den adscripciones ambiguas. Con esto estamos recogiendo un aspecto esencial de las funciones: son normativas en el sentido que se supone que hacen algo de una determinada manera y no otra cosa. La normatividad, como muy bien ha señalado Millikan⁴⁷, tiene que ver con lo que se supone que un rasgo funcional debe hacer. Pero, si bien este es un aspecto ineludible ¿cómo podemos meter en el mundo natural el deber? ¿No estaremos cometiendo la falacia naturalista? En cualquier caso, si esto es así ha de haber alguna manera de conciliar el deber o las presuntas normas naturales y nuestra visión científica del mundo. De otra manera, hay que naturalizar el concepto de norma y así posibilitar un discurso en términos de normas naturales. Pero del mismo modo también hay que legitimar cierta forma de teleología, algo a lo que no todos están dispuestos.

La teoría de Cummins no parece correr el riesgo de ser acusada de neoteleología, sin embargo ni es normativa ni parece estar interesada en una interpretación realista de las adscripciones funcionales que a partir de ella se hagan. A su favor tiene el ser básica en extremo, argumento que algunos explotarán profusamente con posterioridad. Así, puede decirse que, aún admitiendo el enfoque contrario, toda función al principio es una función de Cummins, es decir, para que opere la selección natural y se constituya una determinada historia selectiva previamente debe caracterizarse el sistema que va a ser reproducido en términos de contribuciones causales, es decir, en términos de funciones de Cummins. En contra de esta línea argumentativa se puede decir que en tales estadios no es legítimo hablar de funciones sino sólo de meras contribuciones causales que con posterioridad y bajo ciertas constricciones específicas pueden llegar a ser consideradas como funciones.

Hasta ahora he presentado una serie de aspectos en los que ambas teorías diferían categóricamente. Sin embargo, también hay elementos comunes que, a mi modo de ver, van a canalizar gran parte de la discusión contemporánea. En primer lugar, ambos autores se apartan, de la idea de que un enunciado adscriptor de función constituye algún tipo de explicación nomológico-deductiva frustrada o encubierta. En

⁴⁷ Cf. Millikan (1989).

lugar de seguir esta vía de análisis de los enunciados adscriptores de función, elaboran una alternativa que ha demostrado ser más fructífera. Como es la de incluir la explicatividad en la propia adscripción funcional. En segundo lugar, ambas teorías tiene en común un grado máximo de generalidad. Tanto la teoría de Wright como la de Cummins pretenden dar cuenta de la adscripción funcional en cualquier ámbito, ya sea este biológico o artefactual, es decir, el alcance de ambas teorías va mucho más allá de constituir una reflexión circunscrita únicamente al dominio de la filosofía de la biología o, de manera más general, a la filosofía de la ciencia.

5. Comentarios finales sobre teleología y funcionalidad

En este capítulo, además de una caracterización general de la teleología, he presentado algunas de las teorías teleológicas más sobresalientes dentro de la tradición filosófica a través de su desarrollo histórico. Así, me he centrado principalmente en el análisis de la teleología aristotélica, su recepción medieval y moderna para finalmente señalar la importancia de Darwin en la modificación substancial de la misma. En la filosofía del siglo XX la teleología pervive de manera vigorosa bajo la cobertura del concepto de función biológica. Sin embargo, este concepto surge en filosofía como una respuesta a los problemas planteados por el modelo nomológico-deductivo frente a la explicación teleológica y funcional. Curiosamente los problemas y el diagnóstico de Hempel son prácticamente los mismos que los de Kant y, pese a la modificación substantiva que supone el pensamiento darwiniano, están presentes en las teorías contemporáneas sobre el concepto de función biológica. En el próximo capítulo se continúa con el desarrollo de esta problemática hasta presentar el estado actual de la cuestión en un importante sector del panorama filosófico actual.

II. Sobre el concepto de función biológica. El estado actual de la cuestión

CONTENIDOS

1. La polémica entre las dos teorías y el debate actual
2. La teoría disposicional de Bigelow y Pargetter
3. Los continuadores de la teoría histórico-etiológica
 3. 1. La crítica de Boorse
 3. 2. La teoría de las funciones propias de Millikan
 3. 3. La estrategia analítica de K. Neander
 3. 4. Godfrey-Smith. Explicaciones funcionales en términos de historia reciente
 3. 5. Bedau o la dimensión valorativa de la teoría de las funciones
4. Los intentos de instanciación
 4. 1. Griffiths
 4. 2. La teoría relacional de Walsh
 4. 3. La teoría débil de Buller
5. Los desarrollos recientes de la teoría sistémica: P.S. Davies
6. ¿Pluralismo o unificación?
7. El análisis de Mayr
8. Comentarios finales: Una agenda de problemas pendientes

La discusión en torno a las explicaciones funcionales y el concepto de función en las últimas décadas puede caracterizarse como una disputa abierta a propósito de cuál es el enfoque más adecuado. Si en la década de 1970 se establecen las dos principales teorías que ya he señalado, la siguiente década, a partir de 1980, constituye el momento más álgido de la discusión en el que se propondrá un abanico muy amplio de diversas opciones entre los dos polos del debate situados en las teorías de Wright y Cummins respectivamente. Finalmente, hacia el final de los ochenta (1987) Bigelow y Pargetter propondrán la teoría propensional, teoría que, pese a lo original de su enfoque, no tendrá apenas seguidores.

La polémica que intento reflejar en estas páginas puede establecerse también en términos temporales. Por una parte tenemos aquellos planteamientos que, siguiendo la estela de Wright ponen mayor énfasis en la dimensión histórica de las funciones. Para estos el pasado importa en la medida en éste se identifica con la fuente de la normatividad de aquellas. Por otra parte tenemos aquellos otros enfoques que resaltan el aspecto presente o intemporal y que parece seguirse de las consideraciones sistémicas de Cummins. Y, finalmente y de manera aislada, tenemos una tercera teoría que se analiza el concepto de función en términos de disposiciones orientadas hacia el futuro. En las líneas siguientes expongo cada uno de estos diferentes posicionamientos, así como sus derivaciones, hasta configurar el estado actual de la cuestión.

1. La teoría disposicional de Bigelow y Pargetter

La teoría de Bigelow y Pargetter (1987) supone la única alternativa a las dos grandes familias. Ambos autores caracterizan las funciones al margen de la historia pasada del ítem y en términos disposicionales o por referencia a un estado futuro del propio ítem funcional. La definición que ofrecen del concepto de función es la siguiente:

"Something has a (biological) function just when it confers a survival-enhancing propensity on a creature that possesses it." Bigelow y Pargetter (1987)

La definición anterior involucra al concepto de disposición o propensión, pero en un sentido netamente distinto del de Cummins. A diferencia de este último su teoría está pensada para organismos y, en consecuencia, se elabora con respecto a los conceptos básicos de la teoría de la evolución por selección natural. La referencia a la propensionalidad permite a estos autores caracterizar las funciones por referencia a eventos no existentes en un futuro próximo. Según estos, éste es el tipo de explicación que se requiere en Biología y es lo que ellos tratan de recoger a través de su caracterización. La justificación del requerimiento de nociones disposicionales o, de otra manera, de la necesidad de hacer una referencia explícita al futuro, viene motivada por el análisis del concepto de eficacia. Bigelow y Pargetter parten de la identificación de las funciones con las adaptaciones -tópico que discutiré en el capítulo III- y afirman que un rasgo funcional es un rasgo que resulta adaptativo. Esto último sucede cuando hay una variación en la tasa de eficacia. La teoría disposicional sostiene que las teorías etiológicas tienen un bajo poder explicativo porque explican la funcionalidad retrospectivamente y, en consecuencia, no pueden realizar predicciones. No parece legítimo afirmar que un carácter orgánico haya persistido porque sirva a una determinada función si continuamos definiendo este concepto retrospectivamente, pues la sensibilidad al medio en el que se desarrolla el organismo no está presente.

La propuesta de Bigelow y Pargetter consiste en definir la eficacia de un modo no retrospectivo, sino propensional. El recurso a propiedades subjuntivas como son las disposiciones permite dar cuenta de esta orientación al futuro. La función de cualquier carácter biológico es, pues, descrita por referencia a un evento futuro, aunque éste jamás llegue a acaecer.

Lo que confiere el estatus de función no es que el organismo sobreviva debido a la presencia de un carácter determinado a lo largo de generaciones sino que la supervivencia sea debida a las disposiciones causales que dota al organismo el rasgo en cuestión. Es el poder causal que resulta de conferir cierta propensión a la posesión o no posesión de un rasgo lo que justifica la funcionalidad de un ítem cualquiera. El rasgo puede ejecutar sus contribuciones causales en una diversidad de ambientes notable, precisamente la funcionalidad y/o

adaptabilidad tendrá que ver con la capacidad que tiene ese mismo rasgo para resultar ventajoso y, por lo tanto, eficaz en un determinado rango de ambientes. Así habrá rasgos más adaptativos o ventajosos y otros menos

La teoría de la propensión relaciona su explicación de la eficacia con el nicho ecológico del ítem funcional. Un organismo puede resultar más eficaz en un medio que en otro ya que el ítem funcional le conferirá o no una propensión mayor a sobrevivir y reproducirse, en este aspecto es un precedente de la teoría de Walsh y Ariew (1996) en el énfasis que pone en explicitar el papel del medio en la determinación de la funcionalidad.

Así, el principal competidor de la teoría propensional de Bigelow y Pargetter no es el análisis sistémico de Cummins, sino la teoría histórico-etiológica que considera que algo es una función biológica si es el resultado de un proceso de selección natural previo.

Los problemas que puede suscitar esta teoría son, a mi juicio, dos. Por una parte tenemos el propio concepto de disposición, por la otra, el problema de la normatividad. En cuanto a lo primero, podría pensarse que la idea de relacionar las funciones con eventos futuros que no han acontecido y que, por lo tanto, no existen, puede parecer problemática, pues parece mentar de alguna manera a la causación retroactiva. No obstante, es éste un falso problema. La teoría no hace referencia alguna a planes previos o diseños, sino únicamente a disposiciones. El problema filosófico está en el concepto de disposición y no en la causación retroactiva. La referencia a conceptos disposicionales o propiedades subjuntivas es, por otra parte, como quedó claro después del intento eliminativista neopositivista, irrenunciable. No parece que pueda decirse mucho más sin meterse de lleno en una polémica que excede las intenciones de mi trabajo. Lo que sí que no queda claro es el carácter modal de la teoría, puesto que el rango de posibilidades o alternativas bajo las que el ítem puede tener más o menos probabilidades de ser seleccionado necesariamente ha de ser finito, es decir, de alguna manera sólo bajo un cierto número de circunstancias ambientales el ítem tendrá una mayor propensión a ser seleccionado, pero ¿cómo puede cuantificarse? O ¿acaso la teoría disposicional no es excesivamente permisiva y, por lo tanto, indeterminada? En cuanto al segundo problema es necesario señalar que la teoría disposicional

no permite adscripción alguna de normatividad, sino más bien parece dejar en un limbo esta cuestión o, de otra manera, parece abocarnos a una situación de indeterminación. Aunque en mi opinión un análisis que así procediera sería esencialmente erróneo; es obvio que no se puede hablar de normatividad bajo los presupuestos excesivamente exigentes que requiere la perspectiva etiológica. Si se busca ese tipo de normatividad no se encontrará, aún a sabiendas de que esa misma normatividad puede colisionar fácilmente con una práctica científica saludable pasando por alto nuevas adscripciones u obviando otras. La normatividad que pretende la teoría histórico-etiológica es un mito filosófico. En su lugar, creo que la referencia al medio que ambos autores hacen desbroza el camino para interpretar la normatividad de una manera alternativa que será desarrollada en el capítulo III más extensamente.

2. Los continuadores de la teoría histórico-etiológica

El esquema de Wright es el origen del análisis del concepto de función en términos histórico-etiológicos. Aunque muchas de las explicaciones actuales del concepto de función asumen sus consideraciones histórico-etiológicas, no es un análisis definitivo, sino que abre una puerta a una serie de problemas a los que los diversos partidarios de este enfoque han tenido que enfrentarse. A continuación presento dos problemas a los que los partidarios del enfoque histórico-etiológico han tenido que dar respuesta. Las explicaciones de Millikan y Godfrey-Smith son intentos de hacer frente a estos casos incómodos que no facilitan la explicación funcional en los términos de Wright.

2.1. La crítica de Boorse y la teoría centrada en el objetivo

El trabajo de Boorse (1976) constituye una de las primeras repuestas al análisis de Wright y plantea un buen número de problemas que tratarán de ser respondidos por los autores que se circunscriben a la teoría histórico-etiológica. Wright presentaba su análisis del concepto

de función en términos de condiciones necesarias y suficientes, es decir, para que algo sea una función debe satisfacer las dos condiciones estipuladas en su análisis. Las críticas de Boorse van dirigidas a mostrar que el análisis que Wright establece en tales términos no es adecuado y que, además, no satisface los tres pretendidos desiderata. Así, es posible encontrar ejemplos de diversa índole, bien artificial, bien natural, que satisfacen las condiciones establecidas por Wright y a los que, sin embargo, no parece legítimo adscribir funcionalidad alguna. Ésta va a ser la vía argumentativa principal de Boorse que le va a permitir criticar el análisis de Wright y formular su teoría centrada en el objetivo.

En cuanto a la crítica al análisis de Wright, Boorse señala dos importantes dificultades. En primer lugar, Wright hace un uso de la Biología evolutiva panfuncionalista y supone que todo ítem funcional está donde está porque ha tenido una función en el pasado o la tiene en la actualidad; explica la presencia de cualquier rasgo apelando a una supuesta función. Esta estrategia, según el propio Boorse, conduce a una trivialización de la Biología evolutiva y ha gozado de buena aceptación entre un nutrido número de filósofos, aunque también ha sido el objetivo de críticas muy agudas. En segundo lugar, el análisis de Wright, tal como está formulado, es incompleto. De hecho, el análisis minimiza enormemente la cuestión acerca de la atribución de funciones a organismos y artefactos cuando en realidad señala que hay etiologías claramente distintas para uno u otro caso. De este modo es obvio que no es una cuestión accidental ni trivial, sino que la diferencia de etiologías debería estar presente en el mismo análisis. La situación tal como la presenta Wright es la siguiente:

Reconstrucción del análisis de Wright por Boorse

Pauta básica del análisis histórico-etiológico:

X tiene la función Z si y solo si (a) X está donde está porque hace Z y (b) Z es una consecuencia o resultado de que X está donde está.

Además, debemos señalar dos fuentes etiológicas según el ítem en cuestión sea natural o artificial:

Modelo de explicación etiológica para organismos (Modelo I): La etiología depende de la historia evolutiva del ítem; el rasgo X surge por azar y se mantiene en virtud de sus efectos sobre la economía del organismo.

Modelo de explicación etiológica para artefactos (Modelo II): La etiología de los artefactos depende de las intenciones del diseñador

Aún así, al examinar casos concretos de artefactos u organismos podemos observar que las mismas etiologías que con carácter exclusivo se atribuían respectivamente a ambos dominios, se entrecruzan. Podemos encontrar casos artefactuales en los que se satisfacen las condiciones estipuladas por Wright y, sin embargo, la fuente etiológica no se corresponde con el modelo adecuado para artefactos y, viceversa, casos naturales en los que ambas condiciones son igualmente satisfechas pero la fuente etiológica es claramente intencional. Boorse describe dos situaciones que ilustran bastante bien lo dicho con anterioridad. Entre los casos en los que encontramos etiologías diferentes del diseño o de las intenciones del diseñador y que, no obstante, se circunscriben al ámbito artefactual tenemos la siguiente situación: Un científico construye un determinado dispositivo conectado a una toma de cloro. Cuando pone en funcionamiento el dispositivo se percató de que la toma de cloro tiene pérdidas y antes de que pueda repararla, inhala el gas. Como bien es sabido, la inhalación de cloro es tóxica y va acompañada de una pérdida de consciencia. Si analizamos la misma situación desde la teoría de Wright nos encontraremos con que podemos afirmar lo siguiente, que la función de la fuga de gas en la toma consiste en dejar salir el gas, pues (1) la

salida del gas es el resultado de que la toma esté averiada y (2) la avería persiste porque sale gas por ella, pues es obvio que el científico aislado en su laboratorio e inconsciente no podrá repararla. Según Boorse nos hallamos ante un ejemplo de adscripción funcional a artefactos en los que las intenciones del diseñador no están presentes, sino al contrario, la etiología de la situación es el resultado de una situación análoga a la selección natural: el rasgo funcional surge por azar y se mantiene en virtud de lo que hace. El segundo caso en el que las etiologías se entrecruzan vemos cómo la etiología funcional en un ámbito orgánico es intencional y no depende del azar y el posterior proceso selectivo. La situación es la siguiente: en un vecindario hay un perro cuyo ladrido insistente está presente a todas horas; harto de esta situación, un vecino, aprovechando que la ausencia de los dueños del perro en su casa, salta la valla del jardín y propina una paliza al animalito rompiéndole una extremidad con la intención evidente de que el animal sufra. El resultado de la fractura, además de los posteriores alaridos, es la sensación subjetiva de dolor y la fractura ha ocurrido porque el desalmado vecino la ha propiciado. Nuevamente nos encontramos con un caso en el que el análisis de Wright resulta satisfecho: (1) X está donde está porque produce Z, X, la fractura, produce Z, el dolor (y su manifestación conductual que son los gemidos) y (2) Z es una consecuencia de que X esté donde está, es decir, el dolor es consecuencia de la fractura cuya etiología claramente es intencional.

También podemos señalar casos en los que la fuente etiológica no es satisfecha y, sin embargo, resulta plausible adscribir una función. Estos casos muestran que el análisis de Wright es defectuoso, no necesitamos la etiología para adscribir una determinada función y ni mucho menos parece plausible señalar fuentes etiológicas exclusivas para artefactos u organismos. Un ejemplo de lo primero lo encontramos en el uso de las levaduras en la fabricación de cerveza. En efecto, los procedimientos para fabricar cerveza se conocen desde muy antiguo y entre estos procedimientos está el de añadir levadura en el proceso de elaboración. Pero es obvio que hasta finales del siglo XIX no se conocieron los mecanismos por los que actuaba la levadura ni se determinó que ésta fuera un organismo vivo. Las intenciones del fabricante de cerveza no determinaron jamás la función específica de la

levadura en el proceso de obtención de esta bebida, pues únicamente aplicaba una máxima como la siguiente: "si quieres conseguir cerveza de buena calidad añade estos polvos". Es obvio que la levadura tiene una función muy precisa en el proceso de fermentación, pero no es menos obvio que ésta no depende de las intenciones del diseñador. Con este ejemplo, Boorse muestra cómo la etiología de las funciones de aquellos ítems involucrados en artefactos o procesos artificiales no viene determinada por las intenciones del diseñador. Claramente no se satisface el requisito de que la etiología dependa de las intenciones del diseñador para poder adscribir una función determinada a un ítem artificial. Sin embargo, creo que el ejemplo resultaría probablemente forzado a los ojos de Wright. En primer lugar no parece que el concepto de artificial que el propio Wright tenga en mente se corresponda con un procedimiento similar al descrito por Boorse. En segundo lugar, podemos pensar que cualquier ejemplo proveniente de la industria biotecnológica -o protoindustria- merece un tratamiento específico que no es equiparable a otros dominios artificiales como pueda ser la relojería, la producción de electricidad, o la arquitectura y la ingeniería civil.

En cuanto a aquellos casos en los que la etiología de un ítem biológico no depende de consideraciones acerca de su historia evolutiva es obvio que la historia de la ciencia es generosa en ejemplos pues, si no fuera así, nos encontraríamos con que antes de la formulación de la teoría de la selección natural no había posibilidad alguna de adscribir funciones. Quizás uno de los ejemplos más representativos sea el del corazón y la circulación de la sangre. El hecho de que la función del corazón consistiera en bombear la sangre fue demostrado por W. Harvey (1578-1657) a partir de sus investigaciones cuantitativas sobre la cantidad y velocidad de la sangre y descrito en su *De motu cordis* (1628), mucho antes de que Darwin o cualquier otro teórico evolucionista aventurase alguna teoría en contra del fijismo. El caso Harvey no constituye ninguna trivialidad, tampoco lo constituyó en su época⁴⁸, pues señala cómo es posible atribuir

⁴⁸ Las investigaciones anatómicas de Harvey constituyó un descubrimiento revolucionario, pues supone la adopción del paradigma mecanicista cartesiano en las Ciencias de la Vida y con ello asesta un importante golpe a la tradición galénica.

funciones a órganos sin consideraciones acerca de la historia evolutiva, es más, la vía experimental seguida por Harvey no necesita de ningún trasfondo histórico para ser validada como una buena adscripción funcional y es, por otra parte, el método habitual de la Fisiología, disciplina cuyo objetivo es precisamente la adscripción de funciones. El problema de las adscripciones funcionales no consiste tanto en preguntar por qué el rasgo en cuestión está donde está -pauta seguida por el propio Hempel- sino en averiguar cómo procede el sistema en el que aparece el rasgo funcional. Es decir, antes de ser una cuestión etiológica lo que nos interesa es una cuestión estrictamente teleológica. Hempel y Wright subordinan la teleología a la etiología; esta estrategia responde a un intento de naturalizar la primera mediante la segunda que no parece, a ojos de Boorse, satisfactorio. Si hemos de otorgar alguna relevancia a los contraejemplos de Boorse, la estrategia de Wright no es sólo incompleta, sino que además adscribe funciones a ítems claramente afuncionales, pues existen situaciones que claramente satisfacen las condiciones de Wright y en las que la atribución funcional es dudosa-tal y como sucedía en el ejemplo de la toma de cloro- o claramente inexistente, como ilustra una situación tan accidental como la de un tábano fastidiando a un perro en una tarde de verano. Al final el tábano acaba bajo las patas del perro. Según Boorse nada en el argumento de Wright impide afirmar que la función del tábano consistía en fastidiar al perro. En suma, el problema, según Boorse, tiene que ver con los dos modelos de explicación etiológica. Ni está claro que las intenciones jueguen un papel tan decisivo en la adscripción de funciones a artefactos ni que no se puedan adscribir funciones a organismos sin tener presente su historia evolutiva. El análisis de Wright parece introducir demasiadas complicaciones a la hora de explicar la funcionalidad que no solucionan realmente el problema en torno a las adscripciones funcionales. Por eso en su lugar Boorse no reelabora los argumentos de Wright como sí harán otros autores como Millikan, Neander o Godfrey-Smith, sino que propone una teoría teleológica no etiológica que dé cuenta de las adscripciones funcionales. Así, Boorse entiende que una función consiste en una

Harvey no hace sino contribuir con sus investigaciones a la idea cartesiana de que el organismo es una máquina.

determinada contribución a un objetivo, al margen de cualquier consideración histórica. Más exactamente, una función es una contribución a un fin dentro de un sistema organizado y dirigido a fines como pueda ser cualquier sistema biológico.

"Any goal pursued or intended by a goal-directed system may serve to generate a function statement. Functions are, purely and simply, contributions to goals." Boorse (1976; 77).

Podemos decir que una función es una contribución causal al sistema que alberga el ítem funcional y que este sistema es un sistema dirigido a fines. Esta última cláusula diferencia el análisis de Boorse de la perspectiva sistémica de Cummins. Por otra parte, es necesario matizar la finalidad que la teoría de Boorse presenta como rasgo distintivo. Antes de hablar de una sola finalidad presente en el sistema, rasgo que induciría a pensar que la teoría no es sino alguna forma velada de teleología cósmica o vitalismo, se debe señalar la importancia de los distintos niveles presentes en el sistema que precisamente posibilitan afirmar de éste que está organizado. De este modo es posible hablar de funciones en general y no de *la* función, en realidad el uso del artículo definido aquí no es tan importante, ya que lo que ocurre es que hay una contribución a distintos objetivos en diversos niveles y bajo diferentes perspectivas. El entorno del sistema es tan importante en la determinación funcional como lo pueda ser el ítem funcional mismo y el entorno seleccionado obviamente depende muchas veces de los intereses del investigador, no siendo el mismo para un ecólogo que para un fisiólogo. Así, la teleología, o la direccionalidad hacia diferentes fines, es vista por los biólogos a diferentes escalas temporales y niveles de organización. De otra manera, la determinación de los objetivos del sistema depende en gran medida del contexto de investigación, lo que no implica que necesariamente la adscripción funcional deba restringirse a criterios de carácter subjetivo o instrumental y no pueda ser considerada como una clase natural.

En cuanto a las dificultades planteadas por Boorse hay que insistir en que sus diversos contraejemplos abren la puerta a una serie de modificaciones notables o más refinadas de las ideas directrices de

Wright que serán desarrolladas por autores como Millikan, Neander o Godfrey-Smith y que mayoritariamente se circunscribirán a organismos. Se puede decir que estas contribuciones tienen bien presente el escollo que suponen los contraejemplos de Boorse y tratan de evitarlos. Estas modificaciones consistirán en un refinamiento del conjunto de condiciones necesarias y globalmente suficientes para decir de un rasgo biológico que constituye una función dentro de la línea histórico-etiológica trazada por Hempel y Wright. Los autores aludidos anteriormente son partidarios, pues, de que la única manera de naturalizar la teleología es por medio de la historia evolutiva del ítem y no acentúan las reflexiones en torno a la naturaleza peculiar que pudieran tener los organismos frente al resto de entidades. Al menos del análisis de sus principales escritos no puede establecerse la conclusión anterior. Una teoría que se base en consideraciones acerca de la naturaleza de sus objetos parecería querer retomar algunas tesis condenadas de claras resonancias vitalistas y podría pensarse que pierde crédito científico. En este sentido la teoría de Boorse es muy audaz aún a riesgo de moverse en un espacio frágil y pequeño.

Aunque quizás la crítica más relevante que deba ser hecha al análisis de Wright sea su intento aparentemente injustificado de atribuir funciones por el mismo procedimiento a artefactos y organismos. La estrategia argumentativa seguida por Boorse se basa en presentar diversos contraejemplos en los que de algún modo se ven alterados los supuestos de Wright. Estas alteraciones bien consisten en el entrecruzamiento de las etiologías que Wright propone para cada uno de los dominios, bien en la ausencia clara de las mismas o bien en la trivialidad de las condiciones supuestamente necesarias y suficientes. El problema no sólo tiene que ver con los dos dominios anteriormente referidos, sino con la ausencia de una caracterización de lo genuinamente biológico⁴⁹. La defensa que hace Boorse de la teleología no depende de argumentos etiológicos tal y como pretendían Hempel y Wright, sino de naturalezas cualitativamente distintas; no es lo mismo algo vivo que algo inerte. Lo vivo exhibe una autoorganización y una

⁴⁹ En el caso de Wright no parece haber ningún motivo de índole metafísica para mantener una teoría que adscriba funciones a artefactos y organismos por igual.

propositividad intrínseca que no aparece en las entidades inertes o artificiales.

2.2. La teoría de las funciones propias de Millikan

El interés de Millikan por el tema de las funciones está subordinado a un proyecto muy ambicioso de naturalización de los aspectos semánticos de nuestra vida mental. Para Millikan, en contra de la tesis de Brentano, la intencionalidad no es un fenómeno distinto del resto de fenómenos naturales, sino uno más que requiere una explicación en los mismos términos que el resto de fenómenos estudiados por la Biología. Así, pensamiento y lenguaje son categorías biológicas como lo pueda constituir un órgano como el corazón o un repertorio conductual estereotipado como el cortejo en diversas especies animales. En este proyecto el concepto de función biológica desempeña un papel prominente, pues es el punto de partida para la referida naturalización de la que daré cuenta en la segunda parte de este trabajo. Así, para lo que aquí nos interesa, es importante resaltar que, si bien la contribución de Millikan es relevante para la Filosofía de la Biología, no es el tratamiento que buscaría un filósofo preocupado únicamente por esta disciplina, pues, como ya he anticipado, la teoría de Millikan está condicionada por su interés en tratar el pensamiento y el lenguaje como categorías biológicas a las que adscribir funciones, objetivo principal de la teoría que no debiera ser perdido de vista en su estudio.

Millikan sigue la línea marcada por Wright y adopta una perspectiva histórico-etiológica, pero con una serie de restricciones y modificaciones importantes que, en gran medida, son de índole conceptual, pues crea una serie de nociones como "función propia", "familia establecida reproductivamente", "función propia derivada" o "explicación Normal", etc., que son conceptos de nuevo cuño ideados para fundamentar su teoría semántica. Estos conceptos no habían aparecido hasta el momento ni en el ámbito de la Biología ni de la Filosofía de la Biología.

El concepto principal es el de función propia, que es el término que acuña Millikan para referirse y acotar precisamente el término habitual y más común de "función" en Biología y, en general, de cualquier ámbito en el que adscribamos funciones. Así, Millikan no habla de

funciones biológicas, sino de funciones propias. La teoría de las funciones propias de Millikan se expone con gran detalle en los primeros capítulos de su *Language, Thought and Other Biological Categories* (LTOBC) de 1984, y en su artículo de 1989 *In defense of Proper Functions*, recogido posteriormente en su *White Queen Psychology and Other Essays for Alice* (1993), donde aparece nuevamente definido. Las definiciones que ofrece Millikan del concepto varían ligeramente de una exposición a otra. Del mismo modo que las precisiones que en torno a este concepto hace, siendo una exposición mucho más detallada y complicada la de 1984 que la de 1989.

En LTOBC (1984) define el concepto de función propia de la siguiente manera:

"Where m is a member of a reproductively established family R and R has the reproductively established or Normal character C, m has the function F as a direct proper function iff:

(1) Certain ancestors of m performed F.

(2) in part because there existed a direct causal connections between having the character C and performance of the function F in the case of these ancestors of m, C correlated positively with F over a certain set of items S which included these ancestors and other things not having C.

(3) One among the legitimate explanations that can be given of the fact that m exists makes reference to the fact that C correlated positively with F over S, either directly causing reproduction of m or explaining why R was proliferated and hence m exists." Millikan (1984: 28)

En Millikan (1989) encontramos una definición menos formal (y quizás más clara) del mismo concepto:

"The definition of "proper function" is recursive. To put things very roughly, for an item A to have a function F as a "proper function, it is necessary (and close to sufficient) that one of these two conditions should hold. (1) A originated as a "reproduction" (to give one example, as a copy, or a copy of a copy) of some prior item or items that, due in part to possession the properties reproduced, have actually performed F in the past, and A exist because (causally historically because) of this or these performances. (2) A originated as the product of some prior device that, given its circumstances, had performance of F as a proper function and that, under those circumstances, normally causes F to be performed by means of producing an item like A. Items that fall under condition (2) have "derived proper functions", functions derived from the functions of the devices that produce them." Millikan (1989; 1993: 13-14)

Ambas definiciones hacen referencia a lo mismo, a saber, que algo cuenta como función en tanto que pertenece a un conjunto sometido a una determinada historia de selección. En el caso de los ítems biológicos quizás habría que añadir que los ítems seleccionados lo son en virtud de las capacidades biológicamente relevantes que aportan al organismo portador. La primera de ellas se presenta en términos de condiciones necesarias y suficientes, mientras que la segunda es más laxa. En la primera hay una referencia explícita al concepto de familia establecida reproductivamente, mientras que en la segunda dicho concepto no se pone de manifiesto, aunque se presupone. Ambas definiciones exhiben un rasgo importante del pensamiento biológico; la referencia a categorías grupales o pensamiento poblacional, pues únicamente pueden adscribirse funciones a rasgos si estos rasgos pertenecen a un linaje o conforman, con otras palabras, una familia.

El concepto de función propia trata de dar cuenta de las nociones de diseño o de plan de ejecución o, en términos más corrientes, "propósito" en lo referente a los ítems biológicos desde presupuestos claramente naturalistas. Una función propia es aquello para lo que un ítem ha sido seleccionado y se supone que debe ejecutar como suponemos que la función del riñón es filtrar la sangre o la de los pulmones posibilitar la respiración. En este sentido es un concepto claramente normativo⁵⁰. El concepto de función propia implica una restricción importante sobre la teoría de Wright, si bien en lo esencial se puede decir que mantiene presupuestos muy similares y podemos ver la teoría de Millikan como un refinamiento de las ideas de Wright.

Una función propia es tal en virtud de su pasado, que es el que justifica la propio-funcionalidad del ítem funcional. Pero, a diferencia del análisis de Wright, algo no cuenta como función en tanto que particular,

⁵⁰ En lo que refiere al problema de la normatividad funcional sobre la que Millikan tanto insiste y que está presente en su caracterización del concepto de función biológica en términos de función propia, es patente que el énfasis que pone a propósito de la relevancia de este concepto para la Filosofía de la Biología depende de su uso en Filosofía de la Mente; es decir, la intuición subrepticia que parece guiar el trabajo de Millikan es que si las representaciones son normativas entonces es obvio que los mecanismos que las producen deben de serlo también y por eso se enfrasca en una teoría que permita adscripciones funcionales con un fuerte grado de normatividad. Como espero mostrar esta tesis no está justificada.

sino en la medida en que es un individuo perteneciente a un conjunto. Millikan en LTOBC circunscribe las funciones a lo que denomina "familias establecidas reproductivamente"⁵¹ (*reproductively established*

⁵¹ Curiosamente la referencia a las familias establecidas reproductivamente no aparece en otros escritos de Millikan como *White Queen Psychology and Other essays for Alice* (1993) en los que expone nuevamente su teoría de las funciones

Millikan distingue dos tipos de familias establecidas reproductivamente: En primer lugar las que denomina familias establecidas reproductivamente de primer orden (*first order reproductively established families*) y, en segundo lugar, aquellas que denomina de orden superior (*higher-order*). Lo que distingue unas de otras es el papel que juegan los replicadores en cada una de ellas. Las familias establecidas reproductivamente de primer orden consisten en replicadores que reproducen sus estructuras directamente al hacer copias de sí mismos. Los ítems que resultan reproducidos en dichas familias son resultado de reproducciones repetitivas a partir del mismo modelo y, en consecuencia, comparten con éste las mismas características o similares. Por ejemplo los genes que determinan el color de los ojos son miembros de este tipo de familias, del mismo modo que las numerosas reproducciones de la palabra "gato" en diversos soportes y por diversos medios o la preferencia de "miau" por diferentes gatos o por grabaciones del maullido de los gatos o porque tengamos un loro que imita prodigiosamente estos sonidos. En cualquier caso estamos ante ejemplares del mismo tipo. En cambio, para que un ítem determinado cuente como miembro de una familia de orden superior deben satisfacerse las tres condiciones siguientes: "(1) *Any set of similar items produced by members of the same reproductively established family, when it is a direct proper function of the family to produce such items and these are all produced in accordance with Normal explanations, form a higher-order reproductively established family. (...)* (2) *Any set of similar items produced by the same device, when it was one of the proper functions of this device to make later items match earlier items, and these items are alike in accordance with a Normal explanation for performance of this function, form a higher-order reproductively established family. (...)* (3) *If anything x (a) has been produced by a device a direct proper function of which is to produce a member or members of a higher-order reproductively established family R, and (b) is in some respects like Normal members of R Because (c) it has been produced in accordance with an explanation that approximates in some (undefined) degree to a Normal explanation for production of members of R, then x is a member of R.*" Millikan (1984: 24-25). En el caso de las familias establecidas reproductivamente de orden superior no nos estamos refiriendo a ítems dotados de mecanismos que producen copias de sí mismos, sino a ítems dotados de mecanismos cuya función propia consiste en producir ciertos efectos o ítems como puedan ser los distintos órganos corporales como el hígado o los pulmones, los sonidos emitidos por el sistema fonador de algunas especies de animales o las conductas aprendidas generadas por los mecanismos que las posibilitan. En cualquier caso vemos que en todos estos

family). De otra manera, un ítem tiene una función determinada si y sólo si es un miembro de una familia establecida reproductivamente. La pertenencia a éstas es lo que permite hablar de propio-funcionalidad y, en consecuencia, de normatividad.

Una función para Millikan consiste en la ejecución de unos efectos determinados por parte de un ítem perteneciente a un determinado tipo de familia establecida reproductivamente. La cuestión sobre la posesión por parte de un determinado ítem de propio-funcionalidad no es una cuestión de las disposiciones actuales que posea o no el ítem en cuestión para ejecutar una serie de efectos, sino que es una cuestión histórica, tiene que ver con el desarrollo histórico de la familia o grupo al que pertenece el mecanismo productor de la función⁵². Reconocemos un ítem como perteneciente a una categoría funcional aunque no logre ejecutar la función. Así podemos encontrar con ejemplares que no ejecutan la función que se supone que deben ejecutar y a los que, sin embargo, adscribimos una determinada funcionalidad. En este sentido la funcionalidad no sólo es normativa sino que porque es normativa posibilita establecer diversas taxonomías como puedan ser las clasificaciones de los órganos de las distintas especies, pues no es lo mismo un riñón de gato que de tigre o conejo, aunque todos ellos en virtud de la función que tienen, a saber, filtrar la sangre, son clasificados como riñones. Por otra parte, argumenta Millikan que sin referencia a propósitos no hay normatividad ni funciones. La funcionalidad va ligada a la existencia de propósitos. Según Millikan, aquellos que como Cummins y sus seguidores rechazan la existencia de propósitos en las adscripciones funcionales

ejemplos no hay una replicación directa de un ítem en otro como sí sucede en las familias establecidas reproductivamente de primer orden.

⁵² La justificación de la dimensión histórica frente a una caracterización puramente disposicional sin referencia al pasado puede hacerse a través del estudio de situaciones contrafácticas como los ejemplos aportados por Putnam o, más concretamente por el caso *Swampman* de Davidson. De hecho, ante una situación *swampman* no podríamos, según Millikan, adscribir funciones a nuestro doble. El argumento va dirigido claramente contra Boorse (1976) que considera la adscripción funcional posible al margen de las consideraciones históricas. Neander (1991a) argumentará igualmente contra la teleología no etiológica de Boorse.

no caracterizan adecuadamente la noción de función. Así, el análisis de Cummins únicamente señala que un ítem funciona de una determinada manera, pero no recoge el sentido básico del concepto de función, a saber, la propositividad.

El concepto de función propia, como ya he señalado, es el más importante y constituye la principal contribución de Millikan al debate filosófico en torno a la teoría de las funciones. Pero no es el único concepto de nuevo cuño que elabora. En Millikan (1984) encontramos una clasificación más prolija de los tipos de funciones propias que podemos encontrar como función propia directa, función propia derivada -concepto que sí aparece fugazmente en Millikan (1989)-, función propia relacional y función propia adaptada. Estas precisiones tienen sentido desde el interés de Millikan por explicar el pensamiento y el lenguaje como categorías biológicas. Es obvio que hay una diferencia considerable entre adscribir funciones a los órganos corporales o a las conductas estereotipadas y afirmar que el contenido semántico e intencional depende de las funciones biológicas de los mecanismos que producen dichas representaciones. Una diferencia notable radica en que el contenido semántico e intencional muestra una sensibilidad al contexto y variabilidad que no parece estar presente en la determinación funcional de los órganos. Así, las categorías funcionales anteriormente señaladas tienen su razón de ser en el seno de la discusión semántica, y no propiamente en un capítulo dedicado al esclarecimiento de una noción problemática en Filosofía de la Biología.

Buller (1999) considera que la teoría de las funciones propias de Millikan alcanza cuatro logros en lo que atañe al problema filosófico de la teleología: (1) En primer lugar Millikan, al igual que Wright, logra presentar una teoría claramente etiológica sin necesidad de invocar la causación retroactiva. Pero, a diferencia de aquel, precisa la primera condición de que X esté donde está. Millikan justifica esta cláusula desde el momento en que añade la restricción de que X pertenezca a una familia establecida reproductivamente. Así, X está donde está no sólo porque hace Y, sino también porque sus antepasados también hacían Y. (2) La teoría de Millikan logra evitar el obstáculo que para la teoría histórico-etiológica suponía el contraejemplo de Boorse. Las restricciones que Millikan impone para que un ítem X cuente como un ítem realmente funcional y por lo tanto normativo descartan

inmediatamente el tipo de situaciones propuestas por Boorse. Obsérvese que en la situación que ilustra el artículo de Boorse no tendría cabida en la teoría de Millikan desde el momento en que no hay historia selectiva posible que justifique el efecto de la salida de gas por el agujero de la manguera. El agujero, de otra manera, no está donde está porque algún antepasado agujero-en-la-manguera hiciera lo mismo, sino a causa de un accidente. (3) Con lo anterior y de manera más explícita se puede decir que la teoría de Millikan también logra una de las principales pretensiones del análisis de Wright como es la necesidad de distinguir los efectos funcionales de aquellos accidentales. La teoría de Millikan es normativa. (4) Por último y según el propio Buller, la teoría de Millikan también explicaría cómo la teleología puede emerger en un mundo cerrado por la causalidad eficiente. Así, la teoría de Millikan proporcionaría en el peor de los casos un camino para naturalizar la teleología y hacerla congruente con nuestra imagen científica.

En suma, Millikan define la función de un ítem en términos de su historia evolutiva. Esto supone tanto un gran paso respecto de las teorías anteriores como el origen de una serie de nuevos problemas que deberán ser dirimidos en el seno de la Filosofía de la Biología. No podemos confiar la resolución de los problemas filosóficos a la asunción de postulados científicos tomados acríticamente. Hacer lo contrario supondría robar a Pedro para pagar a Pablo. Como espero mostrar más adelante, la teoría de las funciones de Millikan se fundamenta en una determinada manera de entender la teoría evolucionista que analizada filosóficamente dista mucho de ser un la roca sólida sobre la que podamos fundamentar no sólo una teoría de las funciones, sino también aquello para lo que Millikan ha diseñado su teoría de las funciones.

2.3. La estrategia analítica de K. Neander

Las principales contribuciones de Neander al debate sobre el concepto de función biológica aparecen en dos artículos publicados en 1991 *Functions as Selected Effects: The Conceptual Analyst's Defense* y *The teleological notion of function*. Ambos desarrollan una línea de investigación centrada en dos puntos: una defensa de la teleología y

de la normatividad del concepto de función. Para Neander el interés por definir adecuadamente el concepto de función está justificado por el papel central que, a su juicio, desempeña este concepto en Biología, ya que muchas categorías biológicas no pueden ser definidas más que funcionalmente.

En *Functions as Selected Effects: The Conceptual Analyst's Defense*, justifica, en contra de Millikan, una aproximación analítica al concepto de función siguiendo la estela de Wright y suavizando el rechazo de Millikan (1989) de un análisis en términos de condiciones necesarias y suficientes⁵³. Así, el proceder de Neander es ligeramente distinto, pues si bien coincide con Millikan en lo que es una función, no en vano Neander es también una firme partidaria del enfoque histórico-etiológico, cree que debe llegarse a la definición a través de un análisis del mismo concepto tal como lo usan los propios biólogos. La tarea conseguida por Millikan y por ella misma no sería otra cosa que un análisis, un hacer explícito lo que ya está implícito en el uso científico del término.

El referido artículo elabora también una definición de función, sólo que según la propia Neander se llega a ella por vía del análisis del concepto tal como es usado por los biólogos. Así, para Neander, una función es:

"It is a/the proper function of an item (X) of an organism (O) to do that which items of X's type did to contribute to the inclusive fitness of O's ancestors and which caused the genotype, of which X is the phenotypic expression (or which may be X itself where X is the genotype) to increase proportionally in the gene pool." Neander (1991)

La definición de Neander mantiene básicamente los mismos presupuestos que la definición de Wright o Millikan en cuanto a que asume que adscribir una función a un ítem tiene que ver con hacer referencia a la historia de selección del mismo. La novedad estriba en tomar posiciones claras en torno a las unidades de selección y en

⁵³ Millikan ha rechazado siempre que su contribución a la teoría de las funciones fuera un análisis del concepto en cuestión. En su lugar elabora un nuevo concepto teórico que es el de función propia, pues no cree que sea posible dar con un análisis en términos de condiciones necesarias y suficientes a partir del uso que del concepto de función hacen los biólogos.

hacer referencia al concepto de eficacia inclusiva. Ambas precisiones no aparecía en Wright y en Millikan no parecían estar muy claras. Ambos posicionamientos son discutibles y Neander no aporta ninguna justificación de su aceptación, simplemente acepta lo que cree que es buena ciencia.

El concepto de función que desarrolla Neander es teleológico porque explica para qué está diseñado un determinado rasgo y fundamenta esta teleología en argumentos histórico-etiológicos, es decir, en términos de historias de *selección para*. La teleología etiológica que defiende Neander justifica igualmente el carácter normativo de cualquier adscripción funcional, pues para Neander, siguiendo a Millikan, la normatividad es algo intrínseco al concepto de función. Según Neander sólo un argumento que contemple la historia de selección del ítem funcional puede explicar qué se supone que ese ítem debe ejecutar aún cuando sea incapaz de ello. Lo que se supone que debe ejecutar un rasgo es aquello para lo que éste ha sido seleccionado. Por otra parte, y por las mismas razones que esgrime Millikan, una explicación en términos, ahistóricos, estadísticos o disposicionales -en el sentido desarrollado por Bigelow y Pargetter⁵⁴- no puede dar cuenta de la referida normatividad.

He señalado el uso aparentemente acrítico de los conceptos biológicos que emplea Neander en su definición, en mi opinión, aunque es un paso erróneo en su proceder, no constituye un argumento de peso para rechazar su teoría. El rechazo de la teoría de Neander no tiene que ver con el riesgo que corre al emplear acríticamente conceptos disputados en Biología sin aparente justificación alguna en el núcleo de su análisis, sino que el rechazo vendrá dado por el mismo tipo de argumentos que puedan emplearse contra la teoría histórico-etiológica en general y que desarrollaré en el capítulo III. No obstante, sí hay un argumento decisivo en contra de la formulación de Neander si

⁵⁴ El núcleo de los argumentos en contra de la teoría disposicional y a favor de la defensa de la teleología tal como es entendida por Neander aparece desarrollado en su *The Teleological Notion of "Function"*. Especialmente compárese la sección IV del referido artículo.

la consideramos como el resultado de un método. La estrategia analítica de Neander puede ser fácilmente puesta a prueba formulando las siguientes cuestiones: ¿Realmente Neander recoge un sentido unívoco del concepto de función? De otra manera, ¿es eso lo que los biólogos tienen en mente cuando realizan una adscripción funcional? Yo creo que no. En mi opinión los biólogos no tienen un concepto claro y unívoco tal como pretenden los filósofos acerca de qué sea una función biológica. En su lugar tienen un cúmulo muy borroso en el que circulan diversos parecidos de familia *à la Wittgenstein* que son aplicados en circunstancias variables y muy diversas bajo criterios distintos. No se trata de ningún cúmulo insondable, probablemente las alternativas no van mucho más allá de las dos grandes teorías que he presentado. Éstas son los dos extremos ideales en torno a los que se puede establecer un continuo que sí refleja la práctica científica real. No cabe esperar mucha más precisión en la tarea cotidiana de los biólogos ni en el resto de científicos, la exactitud de la que hacen gala es un contagio de la Filosofía.

2.4. Godfrey-Smith. Explicaciones funcionales en términos de historia reciente

Uno de los argumentos más señalados a la hora de desacreditar la teoría histórico-etiológica es que no respeta la distinción entre explicación funcional y explicación evolutiva⁵⁵ ya que pretende agotar la funcionalidad en la historia selectiva. La teoría histórico-etiológica parece suponer que entre las contribuciones causales de un ítem y la etiología de ese mismo ítem no tiene por qué haber ningún espacio vacío, sino una continuidad. Sin embargo, la anterior afirmación no parece tener el respaldo empírico que sería deseable. Ésa es una de las motivaciones que impulsaron a Bigelow y Pargetter a formular su teoría disposicional. Por otra parte, las consideraciones históricas acerca de la etiología de un rasgo funcional pueden deparar sorpresas como es el caso de los ejemplos que algunos críticos con el programa adaptacionista como Gould y Vrba han puesto de manifiesto. Los casos de sustitución o solapamiento funcional que ilustran las exadaptaciones

⁵⁵ Distinción que encontramos en Tinbergen (1963).

ponen en tela de juicio que la funcionalidad dependa de la historia selectiva.

El trabajo de Godfrey-Smith de 1994 *A Modern History Theory of Functions*, en clara continuidad con las aportaciones de Millikan y Neander, viene a aportar claridad conceptual a la teoría histórico-etiológica y tratar de solventar el problema de la vacuidad o del lapso temporal y el problema de la sustitución o solapamiento funcional. De hecho, constituye una buena formulación de lo que los partidarios del enfoque histórico-etiológico han propuesto y, como admite el propio Godfrey-Smith, no es original, sino que pretende analizar un concepto ya existente de función. La tarea es, pues, puramente analítica, desarrolla o explicita lo que ya está implícito en la teoría histórico-etiológica tal como la han formulado Millikan y Neander. Las contribuciones clarificadoras de Godfrey-Smith se centran en las constricciones temporales de la historia a las que los partidarios de la línea de Wright apelan para establecer la funcionalidad. Así el objetivo del referido artículo consiste en explicitar lo que él denomina la perspectiva de la historia moderna.

En su opinión la dimensión histórica que hay que tomar en consideración es la historia reciente del rasgo y eso es lo que la definición del concepto de función debe hacer patente. Es decir, se trata de explicitar una restricción sobre la historia de selección del ítem supuestamente funcional que evita las dificultades anteriormente señaladas, pues establece una continuidad lineal entre pasado y presente. La historia moderna de un ítem es su historia selectiva actual, con esta afirmación se evita tanto el problema del lapso temporal, pues hay una continuidad hasta el presente, como los contraejemplos que pudieran sugerirse a partir del concepto de exadaptación. Si un rasgo ha sido seleccionado en el pasado reciente porque hacía F tenderá a hacer F a día de hoy. Esta idea se corresponde con el análisis que a continuación presento y que es la definición que propone Godfrey-Smith en términos de condiciones necesarias y suficientes.

La función de un rasgo m consiste en ejecutar F si, y sólo si:

(i) m es un miembro de una familia T ,

- (ii) los miembros de la familia T son componentes de sistemas biológicos reales,
- (iii) entre las propiedades copiadas de los miembros de la familia T está la propiedad o el conjunto de propiedades C que pueden ejecutar F ,
- (iv) una de las razones por las que un miembro de T como m exista en la actualidad es el hecho de que sus miembros pasados tuvieron éxito selectivo en un pasado reciente al contribuir a la eficacia de los sistemas del tipo S y,
- (v) los miembros de T fueron seleccionados porque, al poseer la propiedad o conjunto de propiedades C , ejecutaban F .

Como puede verse la definición incorpora la referencia a las familias establecidas reproductivamente y con ello evita los contraejemplos de Boorse. Además, se aplica únicamente a entidades biológicas y no a artefactos. Por otra parte la definición explicita la noción de eficacia, aspecto que si bien había sido resaltado por Neander y por Bigelow y Pargetter desde la teoría propensional, se echaba en falta en la definición de Millikan o en la de Wright. Las constricciones realistas que constituyen el principal punto de ataque hacia el sistemismo de Cummins también están presentes; el sistema es una clase natural y no depende únicamente de los intereses subjetivos del investigador, por lo que estos deberán subordinarse al rango concreto de sistemas y subsistemas presentes en la realidad. Por último, Godfrey-Smith no pone en riesgo su definición circunscribiéndola únicamente a un mecanismo de *selección para*; no se pronuncia acerca de la dicotomía entre *selección de* y *selección para*. En este último respecto parece anticiparse a los resultados de Buller con la teoría débil.

Los problemas de la aportación de Godfrey-Smith pueden clasificarse en dos grandes grupos. Por una parte tenemos los problemas generales que afectan a la propia concepción histórico-etiológica de las funciones y que, como espero mostrar en el capítulo III, son varios. Respecto a estos admite que el recurso a la *historia moderna* no soluciona todos los problemas sobre el concepto de función y deja otras cuestiones al descubierto como las referentes al compromiso adaptacionista de la teoría histórico-etiológica o la pluralidad de sentidos con la que este término puede ser usado en Biología. Por la otra parte sí es posible cierta crítica a la contribución

positiva del autor. El principal escollo de la formulación de Godfrey Smith tiene que ver con qué debemos entender por *historia reciente*. A mi juicio falta una cláusula *caeteris paribus* referente al medio; sólo hay historia reciente si el medio en el que se opera el proceso selectivo no varía substancialmente.

2.5. Bedau o la dimensión valorativa de la teoría de las funciones

La idea de que el valor desempeña un papel central en la teleología está presente en Platón, Aristóteles, Leibniz o Kant, sin embargo, apenas ha sido insinuada hasta el momento por los autores que he presentado, claramente influidos por la Concepción Heredada. Bedau elabora lo que se conoce como teoría centrada en el valor y es, según sus propias palabras, una modificación del análisis de Wright al que se le añade una condición valorativa. En mi opinión, el análisis de Bedau lleva hasta sus últimas consecuencias la estrategia de corte pragmático que empleó Wright como respuesta a los problemas generados por el modelo nomológico-deductivo. En este sentido, si Wright tiene en cuenta elementos de carácter extralógico como que una explicación es un tipo de respuesta a una pregunta, ¿por qué no señalar el elemento valorativo como una dimensión especial de este tipo de respuesta? El problema del valor en la Concepción Heredada, y muy probablemente en la filosofía de la ciencia en general, es que, al decir del Wittgenstein del *Tractatus*, está fuera del mundo; la ciencia no puede hablar con sentido acerca del valor. Pero del mismo modo que no puede hacer esto, tampoco debería poder hacerlo sobre otros aspectos igualmente valorativos como la utilidad o la economía mental que pueda suponer un determinado recurso explicativo. Aunque en otros ámbitos más generales serán otros autores los encargados de erosionar esta concepción clásica de la explicación científica, en el caso particular de las explicaciones funcionales Wright rompe con la tendencia logicista en favor de una estrategia más pragmática. El análisis de Wright es el análisis de un concepto que se usa con fines explicativos. Así, y siguiendo esta misma línea argumentativa, Bedau decide incorporar las cuestiones valorativas como un aspecto irrenunciable de la explicación teleológica y funcional.

Su artículo de 1992 *Where is the Good in Teleology* tiene por objetivo revelar el componente valorativo de los análisis de la teleología biológica. La necesidad de la modificación de Bedau hay que buscarla en una falla del esquema de Wright. Ítems claramente no teleológicos (afuncionales) también satisfacen el esquema de Wright. Uno de los ejemplos a los que recurre es el de un madero a la deriva que al chocar con una roca queda dispuesto de tal manera que provoca una determinada estela que contribuye a que se perpetúe la posición del madero. Aunque este ejemplo satisface las dos condiciones necesarias y suficientes de la definición de Wright no hay razón para suponer que es un caso de explicación teleológica o funcional. Nadie va a ver teleología alguna en la peculiar disposición del madero en la corriente y, aunque este pudiera resultar funcional, la funcionalidad sería accidental, si por accidental debemos entender aquí una funcionalidad distinta de la que se da en organismos vivos. Bedau contrapone este caso al ejemplo clásico de la funcionalidad del corazón. En efecto, en ambos casos las condiciones estipuladas por Wright son satisfechas, pero sólo en este último hablamos de funcionalidad o de teleología. La dicotomía anteriormente citada puede parecer trivial si se tiene en cuenta que cuando Bedau escribe su artículo la teoría histórico-etiológica cuenta con recursos suficientes para evitar este tipo de situaciones como pueda ser el hecho de que la funcionalidad se adscriba a familias de ítems y no a ítems aislados. Vistas las cosas así, el ejemplo de Bedau no es sino una variación de las situaciones que exponía Boorse y mientras que el corazón depende de una familia establecida reproductivamente, el madero no. Pero para Bedau hay un elemento que ha escapado del análisis de los partidarios de la escuela histórico-etiológica. La diferencia no solamente radica en la pertenencia o no pertenencia de un ítem a una familia establecida reproductivamente, sino la contribución a un fin; mientras que la contribución causal del corazón es buena para el organismo y por eso ha sido seleccionada, la situación del madero a la deriva no contribuye a ningún bien. Bedau infiere que las alusiones al bien deben estar presentes en las explicaciones histórico-etiológicas; no basta por lo tanto con una referencia a que el ítem pertenezca o no a una familia establecida reproductivamente, aún cuando este dato constotuye una condición necesaria para que algo cuente como función biológica.

El problema del bien en las explicaciones teleológicas y funcionales es que su referencia como elemento explicativo ha sido descartado sin análisis previo alguno. La propuesta de Bedau consiste precisamente en analizar el elemento evaluativo y señala que en este se pueden distinguir tres grados. En este análisis distinguimos las tres variables siguientes: A, B y C. Donde A es un agente que hace algo o tiene una determinada propiedad B con el fin de ocasionar algún fin C. Así, tenemos los siguientes tres grados de implicación evaluativa en teleología:

(G1) Implicaciones evaluativas de primer grado: A hace B con el fin de conseguir C si y sólo si (i) A hace B (ii) Al hacer B contribuye a conseguir C y (iii) esto es bueno para A.

(G2) Implicaciones evaluativas de segundo grado: A hace B con el fin de conseguir C si y sólo si [(i) A hace B porque (ii) A al hacer B contribuye a conseguir C] y (iii) conseguir C es bueno para A.

(G3) Implicaciones evaluativas de tercer grado: A hace B con el fin de conseguir C si y sólo si (i) A hace B porque [(ii) al hacer B contribuye a conseguir C y (iii) conseguir C es bueno para A].

De una manera menos enmarañada y con el fin de aclarar las semiformalizaciones de Bedau se puede decir que en G1 aquellos items que contribuyen a la obtención de buenas consecuencias ocurren con el fin de ocasionar esas buenas consecuencias. (G1) sólo analiza las buenas consecuencias de los enunciados teleológicos en un sentido extremadamente débil, puesto que se ponen las tres variables en relación sin explicitar ninguna relación causal de tipo fuerte -i.e. sin explicitar ningún porqué como en (G2) y (G3) sí ocurre-. Así, en (G1) podríamos pensar que C es un resultado accidental, pues (G1) recoge la idea de ejecutar una función de un modo muy amplio, incluso en el sentido de Cummins. Los corchetes en (G2) y (G3) tiene la función de explicitar el porqué anteriormente referido. (G2) representa la versión débil, mientras que (G3) representa la fuerte. Bedau apunta que en (G2) las buenas consecuencias C de que A haga B son explicativas, pero no por su propia bondad. Aunque C no es accidental, sí es posible que el beneficio provocado por conseguir C lo sea. Este tipo de nivel explicativo resulta óptimo para items biológicos y puede ser ilustrado por cualquier ejemplo de los habituales como el anteriormente referido del corazón; que bombear la sangre (B) sea lo que hace el corazón (A)

con el fin de conseguir el mantenimiento del organismo (C) es un buen resultado que explica por qué (B) hace lo que hace pero no explica por qué es intrínsecamente bueno (C), es decir, la explicación que este esquema provee no dice nada acerca de la bondad de (C), pues el porqué se ciñe en exclusiva a la relación entre el hacer B por parte de A y la bondad de que A haga B para C -no dice por qué es mejor para un organismo estar vivo que muerto-. Quizás esta relación quede más clara por contraste con (G3). (G3) recoge la idea de una explicación por referencia a la bondad de las consecuencias, es decir, implica una propositividad en el sentido ordinario de este término y es también lo que ordinariamente se entiende por una explicación teleológica. Mientras que (G2) sólo implica que conseguir C reporta algún beneficio, (G3) nos dice que conseguir C es lo mejor, el bien *per se*. (G3) puede predicarse de la conducta intencional y de los artefactos, todos estos items son susceptibles de una evaluación de su finalidad intrínseca y es ésta de un modo explícito y no accidental la que da razón de por qué tienen la función o el propósito que tienen. Podría pensarse que en (G2) la supervivencia o el mantenimiento del organismo es un bien en sí mismo que explica suficientemente por qué A hace B con el fin de C. Si así fuera no habría diferencias entre (G2) y (G3). Sin embargo las hay, (G3) no es aplicable a la Biología -que no a los organismos con una conducta propositiva resultado de poseer una mente- porque la supervivencia no puede ser considerada un bien intrínseco, pues según Bedau resulta del proceso de selección natural y éste no es exclusivo del ámbito biológico. La argumentación en este punto se torna muy delicada ya que depende de una interpretación bastante forzada y extraña del concepto de selección natural. El principal problema radica en circunscribir la selección natural a organismos, algo que, según Bedau no es así, pues acontece el mismo fenómeno en los cristales de arcilla. La cuestión es que si la supervivencia debe ser considerada un bien y ésta superviene a los procesos de selección natural, la relación de transitividad es defectuosa, pues el bien natural no superviene a los procesos afectados por selección natural ya que en otros dominios donde estos acaecen no identificamos la proliferación o supervivencia como un bien. En conclusión, en Biología la teleología de tercer grado no acaece." *There are no true full-blooded teleological explanations in biology.*

Except for the teleology traceable to the mind, the conditions required for grade three explanations are never present in the natural biological world." Bedau (1992)

A pesar de traer a colación la cuestión valorativa, no parece muy problemático señalar que la teleología de segundo grado es la que opera en el ámbito biológico a diferencia de la de tercer grado. En efecto, decir que un ítem está donde está porque ejecuta una determinada función y añadir que esa ejecución contribuye a la supervivencia del organismo es decir en algún sentido extramoral que ejecutar la función es bueno para el organismo.

El elemento más sobresaliente de la contribución de Bedau es analizar una noción que hasta el momento no se había empleado mas que de una manera monolítica. Así, su tesis principal es que lo orgánico reviste de un valor que es explicativamente relevante en las adscripciones funcionales. Lo anterior es la explicación de la referida condición valorativa si bien ésta sólo ha sido hecha en el contexto de las explicaciones y del alcance que pueda tener y no ha sido caracterizada de una manera substantiva.

4. Los intentos de instanciación

El término "teorías de la instanciación" fue acuñado por Davies (2001) para referirse a aquellos teóricos que consideraban que la teoría histórico-etiológica constituía un caso particular de la teoría sistémica. Desde este punto de vista la teoría sistémica y la teoría histórico-etiológica no son ya dos polos opuestos sino dos maneras un poco distintas de ver el mismo problema que se caracterizan por acentuar unas características frente a otras. En cualquier caso lo que estos teóricos conciliadores pretenden es acercar posturas. Así, consideran que debe haber un punto medio entre la teoría sistémica y la teoría histórico-etiológica. Davies incluye en este grupo a autores como Griffiths (1993), Walsh y Ariew (1996) y Buller (1998) de los que me ocuparé inmediatamente.

4.1.Griffiths

La teoría de Griffiths (1993) pretende conjugar el análisis funcional de Cummins con las explicaciones seleccionistas propias de la Biología. Las atribuciones funcionales que realizan los biólogos son, según Griffiths, lo que en la jerga acuñada por Millikan denominamos funciones propias. Así cuando un biólogo adscribe una función lo que en realidad tiene en mente es una función propia; ese es el tipo de funciones que interesan a los biólogos. En este sentido la teoría de Griffiths depende en gran medida del análisis de Neander⁵⁶ sobre qué entiende un biólogo por función biológica; a saber, un concepto normativo que se fundamenta sobre consideraciones evolutivas.

⁵⁶ Griffiths (1993) critica las dos principales teorías histórico-etiológicas. La teoría de Millikan tiene el inconveniente de ser una teoría muy poco precisa desde el punto de vista estrictamente biológico. Es obvio que Millikan concibe una teoría de las funciones con la vista puesta en los problemas relativos a la semanticidad de nuestras representaciones mentales. En este sentido, el principal problema de su propuesta consiste en averiguar qué tipo de explicaciones biológicas sustentan sus adscripciones funcionales. Un problema derivado de este exceso de generalidad es la incapacidad que supuestamente tendría la teoría de las funciones propias de Millikan para distinguir rasgos funcionales actuales de rasgos vestigiales cuya funcionalidad no es operativa en la actualidad. Por lo que respecta a la teoría de Neander, mucho más precisa desde el punto de vista biológico, Griffiths discute la implausibilidad de su análisis. En efecto, Neander incluye en su análisis una cláusula que sostiene que el genotipo que produce el rasgo funcional X se incrementará proporcionalmente en el acervo genético de la población a la que el organismo portador pertenece. Esto es lo que se conoce como requisito de incremento proporcional. Griffiths sostiene que aunque la proporción de un rasgo funcional sea baja –es decir, no se cumpla el requisito de incremento proporcional- éste no es menos funcional. De hecho, tal como Maynard-Smith (1982) argumenta en su trabajo sobre evolución y teoría de juegos en el que menciona el conocido ejemplo entre palomas y halcones, el incremento proporcional no es una condición necesaria para poder calibrar la eficacia biológica. Un rasgo puede ser más eficaz, es decir, conferir ventaja adaptativa a su portador y en consecuencia, aumentar su probabilidad de ser seleccionado sin necesariamente estar ampliamente representado en el "pool" genético de la población. Al menos esto es lo que se deduce del ejemplo de palomas y halcones. Las reflexiones sobre eficacia biológica importan desde el punto de vista de la teoría de las funciones cuando admitimos que sólo si hay variación en eficacia un rasgo resulta adaptativo y, por lo tanto, funcional.

Para Griffiths el análisis funcional de Cummins es una estrategia muy general. Cummins no se ha centrado en la caracterización de las funciones teleológicas sino en el desarrollo de la explicación de capacidades complejas. Para Cummins hacer una adscripción funcional es proceder a un análisis sistémico en los términos que él mismo establece. Así, la función de un determinado ítem consiste en su contribución a la capacidad global del sistema al que pertenece. Cualquier capacidad compleja será susceptible de ser descompuesta en capacidades más sencillas y éstas, a su vez, serán caracterizadas en términos eminentemente causales. Con esta caracterización funcional no nos comprometemos con ningún tipo de teleología. Sin embargo, sí podemos analizar capacidades complejas como puedan ser las funciones propias. En este sentido es posible instanciar las funciones propias o, de manera más general, la teoría histórico-etiológica en el marco más amplio de la teoría sistémica.

"(T)he proper functions of a biological trait are the functions it is assigned in a Cummins-style functional explanation of the fitness of ancestral bearers of the trait."
Griffiths (1993)

"We can incorporate the etiological approach into the Cummins picture of function adscription. The proper functions of a biological trait are the functions it is ascribed in a functional analysis of the capacity to survive and reproduce (fitness) which has been displayed by animals with that feature. This means that a feature will have a proper function only if it is an adaptation for that function. The trait must have been selected because it performs that function."

Griffiths (1993)

Explicamos la eficacia biológica en términos sistémicos y al considerar también la dimensión temporal, más el hecho de que la eficacia es tomada como un valor promedio perteneciente a una población, no a un individuo tenemos, entonces, según Griffiths, una función propia. La consideración de la eficacia biológica en los términos anteriormente citados es una aplicación del pensamiento poblacional al que Mayr refiere como característico de la Biología⁵⁷.

⁵⁷ Obviamente la eficacia biológica puede ser atribuida a un individuo, pero esto es muy poco informativo.

Griffiths señala que la capacidad de sobrevivir y reproducirse, es decir, la eficacia biológica, puede ser analizada en los términos estipulados por Cummins. Griffiths procede a una lectura sistémica de los argumentos que los partidarios de la teoría histórico-etiológica han empleado para señalar el carácter teleológico y diferenciador de su postura. La teoría histórico-etiológica, o de manera más general, la perspectiva seleccionista, sostenía que un rasgo biológico -bien orgánico, bien etológico- contaba como funcional si era el resultado de un proceso de selección natural, una adaptación. En su lugar, Griffiths afirma que estos conceptos pueden ser analizados en los términos provistos por la teoría sistémica de Cummins.

En ningún momento señala que las funciones caracterizadas en los términos que propone la escuela histórico-etiológica sean explicaciones redundantes como posteriormente dirá Davies. Únicamente señala que el análisis sistémico es más básico y que, además, explica satisfactoriamente las funciones propias que constituyen la categoría funcional específica de la Biología. También trata con especial atención la problemática existente entre las atribuciones funcionales a artefactos y las atribuciones funcionales a items biológicos. Si bien se decanta claramente por el rechazo a una teoría omnicompreensiva y unificada, no discutiré aquí las razones que motivan tal elección.

Griffiths, como se ha podido ver, pone el acento sobre el enfoque sistémico antes que sobre la caracterización histórico-etiológica. La explicación en términos de historia selectiva tiene un peso mínimo en tanto que explicación autónoma; no en cambio la noción pareja y presuntamente dependiente de función propia. Griffiths pretende haber logrado la caracterización de una noción histórico-etiológica como la categoría de función propia en términos sistémicos. Así, en lugar de ser una noción primitiva o simple, la categoría de función propia es una categoría compleja y susceptible de ser descompuesta en sus elementos constituyentes. La normatividad que los defensores del enfoque histórico-etiológico pretendería defender sería, desde este punto de vista sobrevenido a la conjunción de estos elementos más simples.

4.2. La teoría relacional de Walsh

David Walsh (1996) desarrolló una teoría complementaria a las teorías histórico-etiológicas y propensionales que trataba de evitar algunos problemas que afectaban a aquellas. En un artículo del mismo año en colaboración con Ariew “*A taxonomy of functions*” señala de una manera sistemática los problemas que afectan a las teorías histórico-etiológicas y propensional respectivamente. La pretensión de Walsh no consiste en constituir una teoría alternativa si por tal teoría entendemos una más que deba eliminar al resto, sino más bien llenar un hueco explicativo que no ha sido cubierto por ninguna. De hecho, Walsh y Ariew (1996) defienden un pluralismo sistemático: hay una serie de teorías cuyas virtudes explicativas son obvias si se delimitan sus clases de referencia y se clasifican de la manera adecuada. Así, podemos clasificar las teorías vigentes en dos grandes grupos:

A) Teoría sistémica (Cummins)

B) Teorías evolutivas (o caracterizaciones funcionales evolutivas o *Evolutionary Function*)

Dentro de este segundo grupo incluiremos a la teoría histórico-etiológica, la teoría propensional y la aportación de Walsh conocida como teoría relacional. Todas ellas comparten una idea central, a saber, que la función de un rasgo, aquello que denominaremos su función evolutiva, está determinada por su contribución a la eficacia biológica con la que dota al organismo portador. Los problemas que afectan a la teoría histórico-etiológica son tres. En primer lugar, no siempre logra recoger adecuadamente el carácter utilitario que tienen las distintas adscripciones funcionales, ya que se centran únicamente en el éxito alcanzado por un rasgo determinado en la historia evolutiva pasada del mismo. Así, de la utilidad pasada de un rasgo no se sigue su utilidad presente. La teoría histórico-etiológica es insensible a las modificaciones actuales del medio, que es sobre lo que Walsh va a poner el acento en su teoría. En segundo lugar, la teoría histórico-etiológica es incapaz de establecer algún tipo de predicción, con lo que su poder explicativo parece quedar considerablemente mermado. Así,

explica por qué un determinado rasgo prevalece en la actualidad pero no dice nada acerca de su futura prevalencia. De otra manera, que la función propia de X consista en hacer Y se justifica, según la teoría histórico-etiológica, porque en el pasado X hizo Y y, además, hubo una *selección para* ese rasgo. Pero de ahí no se sigue ninguna previsión para el futuro. En este sentido la teoría histórico-etiológica es explicativamente incompleta y no precisamente porque las predicciones en biología sean imposibles, sino por la nula referencia en su formulación al papel regulador del medio. Por último también denuncian una inconsistencia en la teoría histórico-etiológica que tiene que ver con el presunto carácter hipernormativo que siempre ha sido atribuido a esta teoría. Si el carácter normativo de un ítem funcional viene dado por la contribución actual que dicho ítem realza a la eficacia del individuo poseedor y a partir de ahí se procede a una inferencia probable sobre el pasado de ese ítem, es decir, se reconstruye su historia de selección, ¿no habrá una injerencia velada del presente en el pasado evolutivo que la teoría histórico-etiológica ha obviado sistemáticamente al reivindicar siempre la historia selectiva? De otra manera, si el proceder habitual es reconstruir la historia evolutiva a partir de las contribuciones actuales de los ítems funcionales ¿cómo puede silenciarse tan rápidamente la dimensión presente para subsumirla en un pasado únicamente probable e idealizado? Así, hay un problema de relación entre el pasado y el presente que la teoría histórico-etiológica no explica.

En cuanto a la teoría propensional de Bigelow y Pargetter el problema es justo el inverso. Si bien estos autores se apresuran a defender la ahistoricidad del concepto de función y creen explicar la persistencia futura del rasgo, lo cierto es que no explican su etiología, algo que no es ni menos relevante ni importante.

La tercera teoría dentro de la clasificación de las funciones evolutivas que proponen Walsh y Ariew es la confeccionada por el propio Walsh y que está llamada a superar los problemas de las dos anteriores. La teoría relacional señala la imposibilidad de especificar la contribución que un determinado ítem pueda efectuar a la eficacia del organismo portador sin hacer referencia alguna al régimen selectivo en el que el organismo se encuentra y que, en consecuencia, determina,

la referida contribución. De otra manera, no es posible adscribir funciones a ítems sin referencia alguna al régimen selectivo. Éste determina de manera crucial la propiofuncionalidad al margen de la historia selectiva con la que puede o no coincidir. La definición que Walsh (1996) establece es la siguiente:

“The/an evolutionary function of a token of type X with respect to selective regime R is to m if and only if X’s doing m positively (and significantly) contributes to the average fitness of individuals possessing X in R.” Walsh (1996)

Y Walsh y Ariew (1996) de manera más precisa señalan también que:

“(T)he function of (trait token) x in R is just what (trait type) X is being selected for in R. This is of the utmost importance. If there is selection (in R) for a trait type X in virtue of its (token’s) capacity to do m, then the capacity to do m explains the general utility to an individual of having a trait type X. It also explains the persistence of X under natural selection.” Walsh y Ariew (1996)

La teoría relacional ofrece, en el peor de los casos, las herramientas conceptuales necesarias para llenar el lapso temporal que las dos teorías anteriores dejaban al descubierto. Tampoco es una teoría estrictamente histórica, ni presentista o propensional, sino que, en función de que el régimen selectivo sea considerado desde el punto de vista histórico o presente, las funciones adscritas adquirirán una u otra dimensión funcional, pudiendo, además, perpetuarse históricamente durante un determinado lapso temporal. En cualquier caso, la teoría relacional permite un manejo de las dimensiones temporales que determinaban respectivamente a la teoría histórico-etiológica y propensional.

La referencia explícita al régimen selectivo actual podría pensarse que estaba implícita de alguna manera en la teoría histórico-etiológica y, cómo no, en la teoría propensional. Lo cierto es que ninguno de los proponentes había declarado la importancia que el régimen selectivo actual podía tener en la determinación funcional. En cualquier caso, es mérito de Walsh y Ariew el haber destacado de manera fehaciente el papel central que tiene el régimen selectivo actual en la adscripción funcional y el haber presentado esta aportación de una manera

sistemática que se verá más claramente al exponer la taxonomía que ambos autores proponen al final de la presente sección.

Obsérvese que Walsh y Ariew aquilatan la definición de Walsh a través del concepto de *selección para* (y no de *selección de*). Además, añaden que de manera específica es la *selección para* circunscrita a un determinado régimen selectivo lo que determina la propio-funcionalidad del ítem en cuestión. Esto puede representar un problema, ya que hay rasgos persistentes que han sido seleccionados sin que haya habido *selección para* y son a su vez rasgos funcionales. En este sentido, si bien la referencia al régimen selectivo es una matización necesaria y fructífera, la referencia explícita a un proceso de *selección para* no es estrictamente necesaria.

En cuanto a la taxonomía funcional que proponen estos autores hay que recalcar que la teoría relacional tiene un papel "bisagra", ya que relaciona las dos grandes familias funcionales existentes. Tanto la teoría sistémica como la familia más numerosa de funciones evolutivas, admitirán ambos, desempeñan roles explicativos distintos, a saber explicaciones causales, como es el caso de las provistas por el enfoque de Cummins, como explicaciones teleológicas o etiológicas, en el caso de las funciones evolutivas. La teoría sistémica de Cummins, descrita en términos estrictamente causales, es la que tiene un alcance mayor y en la que se insertarían el resto de teorías que, respectivamente, suponen aditamentos explicativos sobre la generalidad del enfoque sistémico, máxime si pensamos que en las adscripciones sistémicas propiciadas por el análisis de Cummins el pensamiento poblacional está completamente ausente, algo que en modo alguno ocurre en las teorías funcionales evolutivas. Así, desde el punto de vista de la generalidad tendríamos la teoría sistémica y, como subconjunto propio de ésta, las funciones evolutivas.

4.3. La teoría débil de Buller

El debate acerca de qué tipo de teoría es más adecuada para explicar el fenómeno de la funcionalidad ha oscilado casi siempre entre posturas antagónicas y excluyentes: o se es partidario de una teoría o se es de otra, con los debidos matices. Algunos autores, en un claro intento unificador, han llegado a admitir la coexistencia de ambas teorías en un mismo dominio, como pueda ser el de la Biología, pero centradas en disciplinas distintas- por ejemplo las teorías etiológicas serían adecuadas para explicaciones en Biología evolutiva mientras que la teoría sistémica haría lo propio en el terreno de la Fisiología-. La idea de relacionar a ambas, al margen de la mutua exclusividad, es relativamente reciente. Uno de los intentos que más destaca en justificar una taxonomía que jerarquice las diversas teorías al uso estableciendo diversos niveles de complejidad es, como anteriormente he mostrado, la teoría de Walsh y Arriew (1996). D.J. Buller ha propuesto una línea similar, aunque sin empeñarse tanto en la tarea taxonómica o jerárquica. De este modo, con la teoría débil Buller (1998) pretende conseguir una unificación de las teorías funcionales vigentes en Biología.

Este autor propone distinguir entre dos formulaciones de la teoría etiológica: una fuerte, que es la que aplican los trabajos clásicos de Millikan y Neander y que puede ser considerada como la versión oficial y ortodoxa de la teoría y, alternativamente, una versión débil, que será la que él defienda.

Teorías Etiológicas -según D.J. Buller (1998)-

A. Teoría Fuerte

"A current token of a trait T in an organism O has the function of producing an effect of type E just in case past tokens of T contributed to the fitness of O's ancestors by producing E and were selected for (over alternative items) because of this contribution of the fitness of O's ancestors."

Buller (1998: 506-507)

B. Teoría Débil

"A current token of a trait *T* in an organism *O* has the function of producing an effect of type *E* just in case past tokens of *T* contributed to the fitness of *O*'s ancestors by producing *E*, and thereby causally contributed to the reproduction of *T*'s in *O*'s lineage."

Buller (1998: 507)

Lo que diferencia a una versión de la otra es que esté o no formulada en términos de *selección para* o de *selección de*; es decir, aquí la dicotomía señalada por Sober (1984) reviste de una importancia singular. Así, para Buller la estrategia de su análisis es establecer una diferenciación entre dos posibles formulaciones de la teoría histórico-etiológica y señalar que una es un caso particular de la otra. Con esto pretende unificar el concepto de función en Biología. La teoría débil es de alcance más general, pues no implica la constricción de que el proceso selectivo consista únicamente en un proceso de *selección para* a diferencia de lo que ocurre con la teoría fuerte. Así, podemos afirmar que la teoría fuerte está incluida en la teoría débil.

Que la teoría débil es más general y de mayor alcance que la teoría fuerte es algo evidente con sólo considerar las definiciones de ambas, pues las restricciones que impone la segunda son mayores que las de la primera. Sin embargo -y esta es la contribución de Buller- no hay razones para afirmar que la funcionalidad única y genuinamente radique en la formulación fuerte. Es decir, la teoría etiológica clásica nos dice tal como la formulan Millikan, Neander o Godfrey-Smith que para que un rasgo cuente como una función debe haber ocurrido un proceso de *selección para* ese mismo rasgo porque, en efecto, dotaba al organismo de mayor eficacia biológica.

La distinción entre *selección de* y *selección para*

La distinción entre *selección de* (*selection of*) y *selección para* (*selection for*) aparece en el contexto de la necesaria distinción entre eficacia global y selección, términos que a primera vista pueden parecer intercambiables. Mientras que la eficacia global es un concepto estadístico y causalmente inerte, los procesos de selección involucran genuina causalidad. Sober señala que la eficacia global, es causalmente ineficaz porque describe, por una parte, la fuerza neta de selección y no sus componentes y porque, por la otra parte, no permite la identificación de las diferencias en la significación selectiva de los rasgos que resultan globalmente eficaces. En suma, una medida de probabilidad no nos sirve para atribuir causalidad, necesitamos un análisis más fino.

La distinción entre *selección de* y *selección para* nos dice que hay una selección de objetos y una selección a causa de que un objeto posea determinadas propiedades. Mientras que el concepto de selección de objetos es bastante amplio, el concepto de selección a causa de propiedades se presta a una menor liberalidad. Por otra parte, la distinción anteriormente esbozada no aparece en el concepto de *selección a secas*. La *selección de* describe los efectos del proceso selectivo, mientras que la *selección para*, las causas. Aunque la *selección de* no implica la *selección para*. Si consideramos un rasgo como la mandíbula veremos que va siempre aparejado a poseer un mentón, sin embargo, la posesión de mentón es indiferente, la selección no actúa sobre el mentón, pues lo que ocurre es que hay un proceso selectivo a causa de la posesión de mandíbulas y no de mentones. La posesión de un mentón es neutral, no así la de mandíbulas, pues la posesión de mandíbulas implica la ventaja de poder ingerir alimentos de mayor tamaño. El desarrollo de mandíbulas significó un importante avance en la evolución de los vertebrados (Hickman et al. 2002: 572), no así el del mentón. El objeto que ha sido seleccionado es la estructura mandíbulas-mentón, pero la causa por la que ha sido seleccionado es porque la posesión de mandíbulas causa diferencias en la supervivencia de los individuos portadores y en su éxito reproductivo, es decir, supone una variación en eficacia. El anterior análisis no es posible si consideramos únicamente la eficacia global del rasgo en cuestión o si no diferenciamos las causas de los efectos en el proceso selectivo. Es decir, la distinción es necesaria porque las diferencias en eficacia entre organismos o rasgos no evidencian por qué propiedades son seleccionados. Los casos de pleiotropía o de ligamiento ilustran la distinción establecida anteriormente.

Así, para que tal selección ocurra, deben darse variantes de ese mismo rasgo que sean heredables y que se traduzcan en una eficacia diferencial respecto de los individuos portadores. En tal caso puede decirse que hay *selección para* estos rasgos en detrimento de aquellos

otros rasgos que no resultan heredados porque no son seleccionados. Sin embargo, es éste un requisito que no siempre se cumple y que, no obstante, no nos impide atribuir funciones a según qué rasgos. Buller sopesa la existencia de rasgos hereditarios que contribuyen a la eficacia de un organismo y que no resultan seleccionados, al menos en el sentido habitual de *selección para*. El conjunto de estos rasgos es bastante heterogéneo. En particular Buller refiere (i) los casos de deriva genética, (ii) los rasgos que no se han enfrentado a la competición actual desde sus distintas alternativas a través de la historia evolutiva, es decir, rasgos para los que no hay variabilidad y que, no obstante cumplen alguna función y (iii) los rasgos funcionales complejos. En el caso de los rasgos resultantes de la deriva genética al azar observamos cómo ésta, que supone un error de muestreo, es también un mecanismo reductor de variabilidad en las poblaciones que en ocasiones produce rasgos hereditarios que contribuyen a la adecuación del organismo pero para los que no ha habido selección, sino simple error de muestreo. A día de hoy parece todavía una cuestión disputada el papel real que desempeña ésta en los procesos evolutivos. Lo mismo podemos decir de los rasgos que no se han enfrentado a la competición actual desde sus distintas alternativas a través de la historia evolutiva y de todos los componentes hereditarios de rasgos funcionales complejos. Desde el momento en que rasgos de distinto orden no han variado en un ambiente selectivo común no puede haber *selección para* esos rasgos, condición necesaria para la aplicación de la teoría fuerte y, sin embargo, sí puede hablarse, según Buller, de funcionalidad. Con esta aportación Buller está poniendo en entredicho la ortodoxia de la teoría histórico-etiológica y abriendo la puerta a los tópicos más problemáticos para la concepción clásica de la Biología evolutiva.

¿Qué ventajas explicativas tiene la teoría débil frente a la teoría fuerte? La teoría débil es bastante similar a la teoría de Cummins, de hecho, Buller considera que la teoría de Cummins constituye la propuesta realmente unificadora, sólo que añade una cláusula de heredabilidad que hace que ésta última resulte aplicable al dominio biológico-evolutivo manteniéndose inmune a las críticas habituales en contra de la propuesta de Cummins. En este sentido al admitir una cláusula hereditaria la teoría débil contempla la dimensión histórica a

diferencia de la teoría sistémica y, por lo tanto, debe ser clasificada en la familia de teorías histórico-etiológicas. Es decir, la teoría débil es la teoría sistémica más una cláusula de heredabilidad.

La teoría débil, al no definir la función en términos de *selección para*, no implica que deban presentarse variaciones en el rasgo funcional, sólo requiere que el rasgo que considera funcional haya contribuido a la eficacia de sus antepasados, en eso radica su funcionalidad. Es decir, atribuye la funcionalidad a un rasgo en tanto que éste contribuye a la eficacia de sus portadores aunque jamás se hubiese dado *selección para* ese rasgo, pues no hubo alternativas posibles que se tradujeran en variaciones en eficacia. La contribución de un rasgo a la eficacia también resulta invocada por la teoría fuerte, sólo que ésta lo hace desde una perspectiva seleccionista que tiene únicamente en cuenta la eficacia relativa de los organismos que poseen variedades del rasgo funcional. La tarea de la teoría débil consiste ahora en mostrar cómo los rasgos que no han variado y que, en consecuencia, no han significado variabilidad en la eficacia de sus portadores, han contribuido, no obstante, a la eficacia del organismo por medio de la producción de ciertos efectos. El modo de analizar la contribución de un determinado rasgo a la eficacia del organismo es por medio del análisis sistémico-funcional de Cummins. Aunque Buller no reduce la funcionalidad etiológica a la sistémica. Admite que ésta última es más básica y que, para tener sentido dentro de las teorías evolutivas que explican los fenómenos naturales, debe añadirse una cláusula de heredabilidad. Por tanto, sea *T* un rasgo cualquiera perteneciente a un sistema *S*, atribuir una función a un rasgo significará para la teoría débil que: (i) *T* es heredable y (ii) *T* es una función de Cummins cuya contribución a *S* se efectúa en términos de eficacia.

La ventaja que tiene la teoría débil es que constituye un argumento unificador en contra de la idea de que en Biología hay dos conceptos de función. La teoría débil articula una explicación unificada de las funciones en el dominio biológico sin tener que diversificar tareas. En este aspecto rechaza el pluralismo dentro de la propia Biología, haciendo, por otra parte, una teoría de uso exclusivo sin necesidad de recurrir a otras de carácter más general que son aplicables a varios dominios. En su contra tiene que es difícil medir la eficacia de un rasgo

si no hay alternativas para ese mismo rasgo y, por lo tanto, poder decir que ese rasgo es eficaz porque supone una ventaja adaptativa se torna una cuestión bastante difícil. Por otra parte, si este argumento es una presa fácil para los partidarios de la versión ortodoxa es cierto que gracias a las medidas deflacionarias que impone Buller sobre la teoría permite considerar casos que no se adaptan tan bien a las explicaciones que se ciñen a un adaptacionismo riguroso y que precisamente vienen dadas por los casos señalados por el propio Buller y que tan bien han sabido explotar otros autores menos ortodoxos. Pero Buller no aporta ejemplos y todo indica que la existencia de estos casos no es algo tan habitual como para cambiar de teoría tan fácilmente. Además, el término "rasgos funcionales complejos" es bastante oscuro y depende de cómo definamos su existencia. Probablemente la cuestión quedaría más clara si en vez de hablar de rasgos complejos se hablara de pleiotropía⁵⁸ o de ligamiento genético. La complejidad nunca es un buen argumento.

5. Los desarrollos recientes de la teoría sistémica: P.S. Davies

Tanto para la teoría disposicional como las teorías históricas la funcionalidad de un rasgo se entiende en términos de contribución a la eficacia biológica. Incluso desde la perspectiva de la teoría relacional esto es así. Las contribuciones de P. S. Davies han estado enfocadas a demostrar la inviabilidad de las teorías de la familia histórico-etiológica, y son varias las ocasiones en las que Davies ha intentado rebatir la teoría histórico-etiológica⁵⁹, si bien una exposición más clara y sistemática aparece en su libro *Norms of Nature* (2001).

⁵⁸ Los genes a menudo influyen en más de un carácter, a esto lo denominamos efecto pleiotrópico o, más simplemente, pleiotropía. La idea es que una mutación en un gen puede afectar a muchos fenotipos. Así, una mutación puede ser ventajosa sobre un carácter y, sin embargo, puede alterar otras partes del organismo con consecuencias desiguales.

⁵⁹ Un claro ejemplo lo constituye su artículo de 1994 en el que argumenta contra la plusibilidad biológica de los mecanismos involucrados en el concepto de familia establecida reproductivamente. El artículo va dirigido en especial contra la formulación de Millikan, si bien en ocasiones posteriores el rechazo a la teoría histórico-etiológica es sistemático.

Para Davies únicamente existe una caracterización adecuada del fenómeno de la funcionalidad y esa caracterización es la teoría sistémica. Las funciones sistémicas son básicas ontológica y epistemológicamente. Las funciones sistémicas son ontológicamente básicas porque su existencia no depende del régimen selectivo sino que es anterior. En sistemas biológicos las funciones sistémicas pueden ser selectivamente ventajosas, desventajosas o neutrales, según los requerimientos impuestos por el régimen selectivo. En este sentido parece coincidir con Buller al señalar que lo que ocurre es que se añade una cláusula hereditaria a la formulación de Cummins. Si bien esta afirmación no está exenta de polémicas al no considerar los efectos beneficiosos en exclusiva de las contribuciones causales en la definición de función. En cuanto a la prioridad epistémica no hace sino ahondar en lo que ya expusimos a propósito de Boorse, en efecto, es posible adscribir funciones sin tener conocimiento alguno de la historia selectiva o sin suponer que se ha dado tal historia. Así, no podemos conocer la presunta función seleccionada sin conocer antes la función sistémica.

Las dos afirmaciones anteriores sobre el carácter básico de la teoría sistémica le conducen a sostener que la teoría de las funciones histórico-etiológica es redundante y que, por ende y en contra de los intentos de instanciación o cualquier otro tipo de teoría plural, sólo hay una teoría de las funciones, a saber, la teoría sistémica. Davies define el concepto de función de la siguiente manera:

"Functions are effects that contribute to more-general systemic capacities. Such functions can arise in systems affected by selection, but only insofar as they contribute to the systemic capacity to change or resist change. Functional properties are constituted by contributions to systemic capacities, not by selection. I further conclude that my formulation of this thesis is preferable to other theses extant in the literature that may appear similar in structure to mine."

Davies (2001: 71).

Si consideramos la definición que ofrece Davies veremos que su novedad consiste en restringir la concepción de Cummins a sistemas jerárquicamente organizados: una función es un efecto que desempeña un determinado rol causal en el funcionamiento de un sistema jerárquico. De una manera más clara, la contribución de Davies supone

una modificación substantiva del análisis de Cummins. Así, el análisis original quedaría reformulado de la siguiente manera:

"The theory of systemic functions ought to be reformulated to reflect the requirement that the system be hierarchial. Where "A" refers to the analysis of system S into components, and where "C" refers to the systemic capacity we wish to explicate, item I has systemic function F if and only if:

(i) I is capable of doing F,*

(ii) A appropriately and adequately accounts for S's capacity to C in terms of the organized structural or interactive capacities of components at some lower level of organization,*

(iii) I is among the lower-level components cited in A that structurally or interactively contribute to the exercise of C,*

(iv) A accounts for S's capacity to C, in part, by appealing to the capacity of I to F,*

(v) A specifies the physical mechanisms in S that instantiate the systemic capacities itemized."*

Obsérvese que la definición que ofrece Davies se formula en términos de un conjunto de condiciones necesarias y globalmente suficientes, mientras que el análisis de Cummins no definía de esta manera, sino que era mucho más informal. En realidad Davies está operando sobre una reformulación del análisis de Cummins⁶⁰, de ahí la presencia de asteriscos en la definición arriba expuesta. La definición incluye todas aquellas modificaciones substanciales a las que he aludido.

La referencia a sistemas jerárquicos es clave, pues cualquier organismo es un sistema jerárquico y probablemente eso sea lo propio del ámbito orgánico. Un sistema jerárquico es aquel en el que podemos distinguir varios niveles de un modo objetivo entre los que, además, se dan interrelaciones causales, es decir, las capacidades de un nivel pueden ejercitarse porque hay otras capacidades que se están ejecutando a un nivel más básico; hay relaciones de superveniencia. Además, con la aceptación de este requisito de organización jerárquica no hay lugar para una posible arbitrariedad en la adscripción funcional, sino que la función se adscribe conforme a un nivel y éste resulta ser una clase natural. El otro rasgo determinante es que estos sistemas

⁶⁰ Cf. Davies (2001: 27).

están afectados por procesos de selección natural; son sistemas naturales, es decir, no son producto de intencionalidad alguna. Además, el concepto de función que emplea Davies está lejos de las rigurosas restricciones seleccionistas de otros planteamientos más ortodoxos, pues la teoría permite la adscripción de funciones a rasgos que no son el producto de fuerzas selectivas. Se puede decir que la teoría formulada por Davies exhibe un holismo que está ausente en las teorías histórico-etiológicas afectadas por constricciones adaptacionistas que parecen atomizar los organismos vivos. Así, el énfasis se pone no sobre la existencia del ítem, sino sobre la capacidad del sistema que alberga al ítem y que está sometido a otras contribuciones causales. Así, hay que tomarse muy en serio el hecho de que la teoría sistémica sea un análisis sistémico y no una explicación -si por tal debemos entender una estrategia subsumptiva o de cobertura legal-. El análisis explicita lo que está implícito en el sistema y pone de relevancia cómo las distintas partes del sistema trabajan en conjunto. Esta mayor amplitud de miras tendría como principal mérito el hecho de que no se pierde ninguna información relevante que no quedara recogida en la historia del ítem y que pudiera traducirse en una mayor eficacia biológica. Por último, no hay que perder de vista que la definición es más un *desideratum* que una realidad habitual y cotidiana en la práctica científica; la cláusula (ii*) nos dice que un análisis explica de manera apropiada y adecuada cuando se especifican las capacidades subyacentes al nivel inmediatamente inferior, en eso precisamente radica su idealidad y esto no es alcanzado en todas las ocasiones por la incapacidad manifiesta que tenemos para satisfacer el requisito (v*), pues nuestro conocimiento es finito e imperfecto.

Si eliminando la teoría histórico-etiológica mediante los argumentos que esgrime Davies se consigue una unificación teórica del concepto de función, entonces es obvio que Davies tiene el mérito indiscutible de ese logro, pero a costa de eliminar también un elemento muy importante como es la normatividad. Gran parte del peso de la argumentación de Davies recae sobre la imposibilidad de la existencia de ejecuciones incorrectas de las funciones, es decir, es un ataque contra el concepto de normatividad. La teoría de Davies está lejos de dejar satisfecho a todo el mundo. Quizás uno de los problemas más

acuciantes para la teoría de Davies sea su rechazo explícito a la normatividad de las funciones, pues algunos sostienen que sin normatividad es imposible hablar de concepto de función. Pero este tópico será abordado con más detalle en el capítulo III. Baste con decir de momento que para Davies hablar de normatividad es simplemente un hábito mental, ya que no existe nada parecido a una norma natural. Los rasgos funcionales en sí mismos no son portadores de normatividad alguna sino que dicha atribución carece, en cualquier caso, de presupuestos realistas, dependiendo por entero de la perspectiva y de los intereses del investigador. De otro modo, se trata de lo que él denomina “un hábito mental”, una forma velada de neoteología o una variedad más del argumento del diseño.

6. ¿Pluralismo o unificación?

En la medida en que la solución de Davies no nos parezca convincente como teoría definitiva habrá que tener en cuenta otras alternativas teóricas. La unificación no siempre ha sido algo deseado y hay autores que como Preston (1998) la ven como un imposible. En cualquier caso unificación o pluralismo son dos opciones disponibles que tienen tanto detractores como defensores. Esto no hace sino evidenciar la magnitud del problema que, en muchas ocasiones, alcanza las aguas de la Filosofía de la tecnología como los dos autores que comentaré a continuación parecen insinuar.

El primer intento serio de unificar las teorías vigentes hasta el momento procedió de Kitcher (1993) quien en un atrevido artículo *Function and Design* propuso la idea de diseño para zanjar definitivamente el antagonismo existente entre las teorías históricas y sistémicas. Kitcher no sólo circunscribió su propuesta al ámbito de la funcionalidad biológica, sino que, en la misma línea de Wright o Cummins, pretendía también abarcar el dominio de lo artificial. Al igual que Wright, señaló que en los artefactos la fuente del diseño depende de las intenciones del diseñador, y en los organismos, de la selección natural, como ya he mencionado; al fin y al cabo no había sino diseño. Pero esto no es más que una unidad aparente; son dos tipos de diseño muy distintos. Así, es casi una propuesta nominalista. La curiosidad más notoria en este punto es que nadie ha justificado por qué debe

hablarse de la misma manera de la funcionalidad de artefactos que de la funcionalidad en Biología.

El concepto de diseño puede ser relevante en la atribución funcional de una manera directa o de un modo más indirecto. La primera vía es la explotada por Wright y sus seguidores. Ésta considera que el diseño es el resultado del proceso adaptativo a través de la selección natural. La segunda vía, más indirecta, es la seguida por Cummins. Respecto al análisis de Cummins, Kitcher sostiene que en el trasfondo persiste una fuente de diseño, si bien ésta no adquiere el papel relevante en la explicación de la funcionalidad, ya que explicamos recurriendo al análisis de las capacidades actuales del sistema considerado como un todo.

En Biología el problema del concepto de diseño como estrategia unificadora es un remedio barato; muchos son los teóricos que se muestran reticentes a hablar de “diseño” en el ámbito biológico, aunque explícitamente se identifique la fuente del diseño de los organismos con la selección natural. Puede pensarse también que los parentescos entre “diseño” y “plan” son estrechos. El diseño podría verse como una teleología soterrada. Dicho de otro modo, la propuesta de Kitcher parece aportar bien poco; incluso la idea de diseño que propone no deja de ser simplemente una metáfora de escaso valor explicativo.

Las críticas a Kitcher no se hicieron esperar, en el mismo año Godfrey-Smith publica con motivo del vigésimo aniversario de la aparición de *Function* de Wright un artículo, *Functions: Consensus Without Unity*. En este artículo, Godfrey-Smith además de presentar un estado de la cuestión, critica contundentemente la solución unificadora de Kitcher en lo referido a sus unificación teórica del concepto de función biológica. A juicio de Godfrey-Smith la apelación más general que hace Kitcher al concepto de diseño como unificación teórica entre la teoría sistémica y la teoría histórico-etiológica no respetaría la distinción establecida por Wright entre un mero efecto y una función. La crítica de Godfrey-Smith no hace sino una reafirmación de sus posiciones historicistas en la definición del concepto de función, ya que según él cualquier componente de un sistema biológico es un producto de la selección natural. La unidad que pretende Kitcher es una falsa unidad; no hay ninguna razón, según Godfrey-Smith, para recoger bajo un mismo concepto las caracterizaciones de Wright y Cummins, pues

son conceptos distintos que han tenido la mala fortuna de emplear el mismo término.

Beth Preston (1998) representa la alternativa pluralista. Como Kitcher, también ella hace alusión al problema de la funcionalidad de los artefactos. Es suyo el mérito de haberse dado cuenta de la especificidad de las atribuciones funcionales en artefactos, lo que hace que reclame una teoría específica de las funciones para estos, aunque aquí no me detendré en las razones que esgrime en favor de dicha tesis a la que dedica la segunda parte de su artículo de 1998. Las ideas de Preston a propósito del pluralismo están influidas en gran medida por el desafío al programa adaptacionista por Gould y Vrba del que trataré en el próximo capítulo. Así, Preston parte de la asimilación del concepto de función, de manera más concreta, del concepto de función propia tal como es acuñado por Millikan, al concepto de adaptación y equipara las exadaptaciones a las funciones sistémicas. Esta equiparación es bastante sensata si tenemos en cuenta que sólo las adaptaciones tienen funciones mientras que las exadaptaciones tienen efectos. En este sentido el concepto de función propia no hace sino recoger un significado importante y muy clásico del concepto de adaptación. Por otra parte, resulta sugerente admitir la legitimidad de la equiparación entre funciones sistémicas y exadaptaciones, ya que las funciones sistémicas son consideradas desde la perspectiva de Cummins como contribuciones causales por los ítems de un sistema al margen de sus respectivas historias selectivas que, en este caso, no tienen importancia alguna. La funcionalidad aquí es un efecto y para Gould y Vrba las exadaptaciones tienen propiamente efectos, no funciones. Así que, salvo la pequeña imprecisión terminológica, podríamos pensar que estamos ante el mismo fenómeno. En Biología ambos conceptos serían necesarios ya que no todo rasgo es susceptible de ser clasificado como adaptación y, como muestran Gould y Vrba, hay rasgos que en lugar de responder a los requisitos para ser considerados adaptaciones son más bien exadaptaciones. El punto débil de la postura de Preston es la equiparación anterior, que es el rasgo distintivo de su defensa del pluralismo. Si consideramos la tesis de Davies a propósito del carácter básico de las funciones sistémicas, argumento que a mi juicio es bastante sólido, veremos que la equiparación señalada por Preston entre exadaptaciones y funciones

sistémicas es espuria, pues una adaptación también sería el resultado de una función sistémica. Otra línea de ataque puede venir sugerida por el concepto de adaptación que emplea, pues, como espero mostrar en el siguiente capítulo, es muy parcial, aunque ha sido el empleado por la mayoría de los filósofos de la Biología.

7. El análisis de Mayr

El análisis que hace Mayr del concepto de función es muy sensible a la problemática abordada por los filósofos profesionales y da buena cuenta de ella.⁶¹ Si bien, su punto de vista es completamente ajeno, según el propio Mayr, a los métodos y reflexiones que han hecho hasta la fecha los propios filósofos, a su juicio siempre erróneos, ya que han elaborado sus análisis desde una filosofía de la ciencia que tomaba como modelo de saber científico a las ciencias físicas e ignoraba la autonomía ontológica y metodológica de la Biología⁶².

El tratamiento tanto de la teleología como de la funcionalidad representa para Mayr las dos caras de un mismo problema y es especialmente interesante para poner fin a las interminables disputas filosóficas acerca del concepto que aquí nos ocupa. Su análisis tiene la indudable ventaja de ser el de un biólogo que ha hecho tanto trabajo de campo como teórico y que conoce medianamente bien las teorías que sobre el concepto de función han elaborado los filósofos desde los trabajos iniciales de Hempel y Nagel. Si por teleología debemos entender la creencia en una causalidad final, teleología y selección

⁶¹ *“Muchos filósofos de la ciencia han pensado que el problema de la teleología podía ser resuelto explicando la dirección a un objetivo en forma funcional, es decir, traduciendo los enunciados teleológicos (Wimsatt 1972:1-80) como enunciados funcionales (Cummins 1975:741-765). Semejante traducción se halla también implícita en Hempel (1965), Nagel (1961) y numerosos autores posteriores.”* E. Mayr (2004) [2006: 69].

⁶² *“Durante por lo menos cincuenta años una considerable cantidad de filósofos de la ciencia ha escrito sobre la teleología basando su análisis en los métodos de la lógica y el fisicalismo “conocidos por ser los mejores” o al menos los únicos métodos confiables para esos análisis. Estos filósofos han ignorado los descubrimientos de los biólogos, aunque la teleología se relaciona en gran parte o en todo con el mundo de la vida.”* Mayr (2004) [2006:66]. Para Mayr es obvio que la biología como disciplina autónoma se sirve de nociones teleológicas no reducibles a otros tipos de causación.

natural son, para Mayr, conceptos incompatibles. Para Mayr la teleología no es un fenómeno unitario, antes bien, es partidario de distinguir al menos cinco usos del término "teleológico" que se corresponderían respectivamente con diferentes fenómenos naturales. Podremos hablar legítimamente de teleología cuando nos refiramos a cuatro de los cinco procesos, aunque será siempre preferible el uso del término correspondiente que no da lugar a equívocos o confusiones. Los procesos naturales para los que habitualmente se ha empleado el término teleológico son: (1) los procesos teleomáticos, (2) los teleonómicos, (3) la conducta deliberada y (4) los rasgos adaptativos. Sólo hay un concepto que deba defenestrarse como es la teleología cósmica. Los procesos teleomáticos son aquellos que alcanzan su fin de manera automática, regulados por leyes físicas, cuando su potencial se agota. Tampoco están circunscritos necesariamente al ámbito biológico y buen ejemplo de ellos lo constituyen fenómenos físicos como el cambio de temperatura. Los fenómenos teleonómicos son aquellos que están guiados hacia un fin a través de un programa determinado. En este caso sí podemos decir que estos procesos son exclusivos del ámbito orgánico en el que el programa es producto de la selección natural. La presencia de este tipo de procesos sería, según Mayr, un argumento de peso en la defensa de la autonomía de la Biología. Es obvio que la conducta deliberada está orientada a fines, en este sentido Mayr únicamente asiente ante el hecho de que la presencia de este tipo de conductas están presentes en el reino animal, si bien no resulta menos problemático explicar en qué consiste la intencionalidad. Por último, los rasgos adaptativos no pueden denominarse ni teleológicos ni funcionales, si aparecen mencionados aquí es porque pueden llevar a cabo actividades teleonómicas. Respecto a este último punto Mayr señala que un rasgo funcional no es un rasgo adaptativo. También sostiene que deberíamos distinguir las actividades funcionales genuinas de las adaptaciones, el tema de la adaptación constituiría el aspecto histórico de una determinada actividad funcional, si bien ésta podría ser caracterizada de alguna manera con independencia de este trasfondo.

Para Mayr el concepto de función tiene dos significados muy distintos en Biología. Por una parte hablamos de funcionalidad en un sentido estrictamente fisiológico y no teleológico. Por la otra, sostiene

que el concepto se ha extrapolado ilegítimamente al ámbito de la Biología evolutiva y se ha asimilado al de fin. El concepto de función sería para Mayr un concepto casi estrictamente fisiológico o relativo al papel que un rasgo determinado pueda adquirir en el ciclo vital de un organismo⁶³. Toda acepción que vaya más allá de estas lindes supone una intromisión en el ámbito de la explicación teleológica que para el propio Mayr también tiene su legitimidad. De lo que Mayr no se percata es de la dimensión normativa a la que la perspectiva adaptacionista contribuye y que, por otra parte, resulta ineludible en cualquier adscripción funcional. En mi opinión, la aportación de Mayr resulta muy poco sensible a la problemática filosófica del concepto de función, además de que parece obviar gran parte la literatura al uso. La postura de Mayr es la de un pluralista que asume que no hay una única definición del concepto en cuestión, sino varias. Su proceder es instrumental o pragmático, antes de definir a priori qué sea una función o qué deba entenderse por tal concepto recoge y explicita lo que cree que los biólogos entienden en su trabajo cotidiano. De otra manera, ¿qué entiende la mayoría de las veces un biólogo cuando emplea el concepto de función? Hasta aquí Mayr procede de una manera aparentemente descriptiva.

Las críticas a las teorías y métodos de los filósofos se centran principalmente en dos aspectos: los filósofos de la ciencia no han entendido adecuadamente ni el objeto ni el método propios de la Biología. En cuanto al objeto, el principal problema es que el proceder habitual de nuestra disciplina ha entendido a los organismos como si fueran entidades inanimadas como las que aborda la Física o la Química y, en lo tocante al método, ha procedido como si la Biología en realidad fuera una extensión de estas dos ciencias.

No creo que le falte razón a Mayr al señalar que la Biología tiene un objeto y método propios. Es ésta una cuestión que ya he debatido. Sin embargo, y pese a las críticas vertidas contra la Filosofía de la ciencia

⁶³ "(L)a palabra función se emplea en biología con dos significados muy diferentes, que deben ser cuidadosamente distinguidos en cualquier análisis teleológico (...) función se usa a veces para procesos fisiológicos y a veces para el papel biológico de un rasgo en el ciclo vital del organismo." Mayr (2006:69). El problema de la postura de Mayr es que no aclara cómo se relacionan ambos tipos de procesos funcionales cuando es obvio que alguna relación tienen.

de los filósofos académicos, hay que decir que la propuesta de Mayr no tiene más virtud que la experiencia aportada por un buen biólogo que hace un guiño a la Filosofía de la Biología con notable éxito. Quizás sea bueno distinguir entre la labor de un teórico y un filósofo de la Biología. Mayr, a mi juicio, pertenece a la primera categoría pero no a la segunda. No creo que Mayr haya entendido adecuadamente cuál es el propósito que guía al filósofo de la biología en su labor científica.

La postura de Mayr tiene una serie de consecuencias relevantes para la Filosofía de la ciencia. En primer lugar, y de manera manifiesta, supone un compromiso con el pluralismo, ya que como he indicado, Mayr entiende que el concepto se emplea de dos maneras diferentes que se corresponden a dos clases de referencia bien distintas. Esto no constituye ningún problema siempre y cuando se distingan los usos. Aunque quizás fuera mejor reservar un término específico que refiriera a cada una de las clases de referencia si es que realmente tal diferenciación es tan obvia como supone Mayr. De hecho, el problema del pluralismo es que sea un pluralismo instrumentalista que no individualiza adecuadamente la referencia del término. Éste ha sido el problema que ha preocupado a los teóricos partidarios de una perspectiva unificacionista o monista a propósito del concepto de función. Bien se defendiera una u otra postura lo que se quería dejar claro es que sólo había una clase de referencia que se correspondiese con el término. Mi duda en este punto es si Mayr quiere reflejar con su pluralismo dos clases de referencia distintas o es simplemente una estrategia pragmática. La impresión que da es que como científico está comprometido con cierto realismo pero que en su práctica cotidiana no deja de ser instrumentalista. Cuando menos no creo que sea una cuestión que contribuya a esclarecer. La razón que me induce a formular este juicio es la facilidad con la que pasa de una postura puramente descriptivista a una prescriptiva en exclusiva. Pasar de la descripción del uso del término a la formulación normativa del uso y de la referencia del mismo no es un buen modo zanzar la disputa filosófica. Si admitimos que el término se usa en dos ámbitos o niveles explicativos diferentes estamos asumiendo una postura instrumentalista pero no explicamos qué relación hay entre esos dos niveles que, sin embargo, aparentemente, están relacionados.

8. Comentarios finales: Una agenda de problemas pendientes

Lo visto a lo largo de este capítulo puede sugerir no sólo cuál es la complejidad de un debate en el que hay numerosos interlocutores, sino también cuál es la magnitud de este debate. En mi opinión las cuestiones suscitadas por el mismo pueden dirimirse en dos áreas muy distintas. Por una parte tenemos aquellas cuestiones que pertenecen de lleno a la Filosofía de la ciencia y, de manera más particular, a la Filosofía de la Biología. Por otra parte están las cuestiones que afectan bien a la Filosofía de la técnica bien a la Filosofía de la mente. Así, el concepto de función es lo suficientemente rico como para afectar a todas ellas y una caracterización exhaustiva del mismo debería tratar de abarcarlas en la medida de lo posible.

Respecto de las cuestiones del primer grupo creo que el tópico involucrado en la discusión más general en Filosofía de la Ciencia es el problema de la explicación funcional y teleológica, del que he dado cuenta en este capítulo. En cuanto a los tópicos que pertenecen de lleno a la Filosofía de la Biología creo que la exposición presentada aquí apunta claramente a dos: la cuestión de la equiparación entre funciones y adaptaciones y el problema de la normatividad. En mi opinión se hace necesario aclarar filosóficamente ambos conceptos que, de manera constante, aparecen en todas las teorías analizadas hasta el momento. Ambos son analizados a continuación en el capítulo III.

En cuanto al segundo grupo de cuestiones hay dos aspectos que habría que mencionar. Por una parte, la cuestión de los artefactos, por la otra la teleosemántica. En lo que afecta a los artefactos es importante no olvidar que desde las propuestas de Wright y Cummins hay un interés manifiesto en que la teoría que justifica la funcionalidad de rasgos orgánicos sea también aplicable al ámbito de los artefactos. ¿Por qué insisten algunos teóricos en este punto? ¿Por qué debería ser deseable una teoría no solamente unificada en Biología sino que diera también razón de la funcionalidad de los artefactos? Puedo aventurar una razón epistemológica: si sabes cómo hacerlo, entonces sabes cómo funciona. Esto no es más que una paráfrasis de una idea de Dretske aparecida en su célebre *If You Can't make One, You Don't Know How It Works* (1994). Sin embargo, es esta una línea de

investigación muy diferente a la mía y no la explotaré en este trabajo. En cuanto a la teleosemántica, razón de ser de esta investigación, hay que señalar que en la medida en que su caracterización del contenido depende del concepto de función es muy importante tener claro qué puede ser una función biológica y qué implicaciones tiene cada definición para la teleosemántica.

III. Adaptación y normatividad

CONTENIDOS

1. Adaptación y función biológica
 - 1.1. Adaptación y Selección Natural
 - 1.2. Exadaptaciones y Adaptacionismo
 - 1.3. La relación entre las adaptaciones y las funciones
 - 1.4. Conclusiones sobre las relaciones entre los conceptos de función y adaptación
2. Funciones y normatividad
 - 2.1. El rechazo de la normatividad dentro de la teoría sistémica
 - 2.1.1. La crítica de Davies a la neoteleología: la imposibilidad de las ejecuciones erróneas de las funciones
 - 2.1.2. Crítica de Cummins a la neoteleología: ni adscripciones funcionales ni normatividad
 - 2.1.3. Consecuencias del rechazo de las nociones normativas
 - 2.2. Nomologicidad y normatividad
 - 2.2.1. Sobre el concepto de ley científica en el neopositivismo
 - 2.2.2. La crítica de Dretske a la concepción neopositivista de las leyes de la naturaleza
 - 2.3. ¿Por qué es necesaria una teoría normativa de las funciones?
 - 2.4. Funciones sistémicas normativizadas

Este capítulo aborda dos tópicos. En la primera sección analizo el problema de la equiparación entre adaptaciones y funciones, un problema que se deja entrever muy a menudo en la discusión contemporánea pero que no ha sido suficientemente tratado. Mi

propuesta va a consistir en equiparar el concepto de función con el de adaptación, pero desde una clarificación conceptual suficiente del concepto de adaptación biológica. La segunda sección de este capítulo está dedicada al estudio de la normatividad que considero de gran importancia. Mi aportación personal a la discusión actual expuesta al final de la misma analiza la normatividad sin recurrir a la teoría seleccionista y ubica la fuente de la normatividad en el contexto en el que se localiza el ítem funcional. Ambos tópicos son de importancia considerable para un análisis adecuado de la teleosemántica y volverán a aparecer en el capítulo VI.

1. Adaptación y función biológica

Una estrategia habitualmente seguida por los autores familiarizados con el desarrollo de la Biología evolutiva ha consistido en asimilar el concepto de función biológica al de adaptación. La estrategia no es nueva, fue inaugurada por Tinbergen (1963)⁶⁴ y continuada por etólogos y ecólogos conductuales. Williams (1966), desde otros presupuestos y con otros intereses, igualmente relaciona ambos conceptos de manera casi automática. Otros autores como Sober (1993), desde una perspectiva mucho más general y acorde con las aportaciones de Williams, no muestran interés alguno en centrar su definición en el ámbito etológico. En su lugar y, de modo mucho más general, establece la misma relación entre los conceptos de función y adaptación; así, su definición de adaptación tiene claros tintes funcionales. En cualquier caso, la estrategia general de los autores anteriormente citados consiste en igualar el concepto de función con el de adaptación. Todos ellos sostienen de una manera más o menos implícita que decir de un rasgo que es funcional es decir que es

⁶⁴ Tinbergen (1963) en su célebre artículo "On the Aims and Methods of Ethology" señala que las cuatro principales áreas de estudio de la etología son: (1) La evolución, (2) la causación, (3) el estudio de la adaptación o función biológica que supone la conducta y (4) el desarrollo. Obsérvese que el proyecto teleosemántico entronca directamente con la escuela etológica. De manera más concreta se afirma que la conducta resulta adaptativa o no. Una conducta adaptativa es una conducta funcional.

adaptativo. Aún así, esta estrategia tan simplificadora es susceptible de diversas interpretaciones.

Los filósofos hemos equiparado las funciones con las adaptaciones guiados por una presunta confianza en que estábamos haciendo buena ciencia. De hecho, teníamos razones para proceder así si científicos como Tinbergen también actuaban de la misma manera. Por otra parte, la revisión de la literatura científica estándar, por ejemplo, de aquellos manuales⁶⁵, algunos ya clásicos, empleados por los estudiantes y profesores de Biología parecen apoyar la idea. Sin embargo, no hay que fiarse en exceso del rigor conceptual que algunos atribuyen a los autores de manuales. Éste no es igualmente entendido por un científico que por un filósofo de la ciencia. La razón es obvia; hay intereses claramente divergentes. No es que los científicos sean poco rigurosos, probablemente si quisieran ser tan presuntamente rigurosos como pretenden ser algunos filósofos no harían ciencia. Las precisiones que interesan a los filósofos no interesan a los científicos. Así, lo que ha sucedido en el debate actual a propósito de la teoría de las funciones es que el apoyo científico se ha tomado a conveniencia para apoyar una u otra vía argumentativa. No es que esto sea una mala práctica, en modo alguno. Pero sí es una práctica que hasta cierto punto se ha realizado de manera acrítica; los filósofos en muchos

⁶⁵ Por ejemplo el ya clásico de Dobzhansky, Ayala, Stebbins y Valentine (1983) ampliamente usado en las universidades españolas, si bien deja muy claro en qué consiste la evolución y cuáles son sus mecanismos, incluso finaliza con un capítulo dedicado a la Filosofía de la Biología, no aclara en modo alguno ninguna de estas cuestiones. Otro tanto puede decirse de Ayala y Valentine (1983). O de Soler (Ed) (2003) que, si bien no es estrictamente un manual, sino una recopilación de trabajos introductorios y algunos más específicos, combina excelentemente precisión conceptual con estudios de casos. Por otra parte, Fontdevila y Moya (2003) en su manual tampoco son muy prolijos al definir el término, si bien su definición no se hace depender de consideraciones históricas "*Cualquier carácter, morfológico, fisiológico, de conducta, o de desarrollo que incremente la supervivencia y/o el éxito reproductivo de un organismo.*" p.415. El problema no es que cada uno de ellos defina el concepto de una manera u otra, es la nula discusión en torno a los conceptos, ninguno de ellos señala que haya definiciones alternativas o que la definición no es de alcance universal. Es obvio, por otra parte y en su defensa, que se trata de manuales. Si bien, no hay que olvidar que el conocimiento que tengan en el futuro los biólogos dependerá en gran medida de sus lecturas de formación.

casos han pretendido disolver un problema filosófico tratando de relacionarlo o reducirlo al dominio especial de un determinado ámbito del saber científico y para ello se han servido de los conceptos propios de ese mismo ámbito. Este trabajo analiza los intentos de reducción de la intencionalidad a la funcionalidad biológica. Sin embargo, nos encontramos con que tal reducción no puede ser acometida de un modo directo y sin salvar importantes dificultades. Lo mismo ocurre con otros conceptos. Así, no parece que haya conceptos científicos que silencien definitivamente los debates filosóficos. Las páginas que viene a continuación tratan de analizar la legitimidad de la equiparación entre funciones y adaptaciones. Mi proceder va a ser el siguiente: En primer lugar voy a presentar el problema en torno a la definición del concepto de adaptación. En segundo lugar, discutiré lo que se ha conocido como programa adaptacionista, la polémica noción de exadaptación y qué implicaciones tiene para el debate en torno a las funciones. A continuación se exponen a aquellos autores que de manera explícita han relacionado los conceptos de adaptación y función. Finalmente, presento una serie de conclusiones que exponen mi punto de vista sobre este problema.

1.1. Adaptación y Selección Natural

El concepto de adaptación, junto con los conceptos de selección natural, o eficacia biológica, es uno de los conceptos centrales de la Biología evolutiva. Quizás sea el concepto más importante después de la idea de evolución, ya que no en vano distinguimos el concepto de evolución como mero cambio, de la evolución adaptativa. En los estudios de microevolución⁶⁶ admitimos igualmente que algunos

⁶⁶ Distinguiré microevolución de macroevolución. Los estudios sobre microevolución se ocupan de los cambios genotípicos y fenotípicos que acaecen en breves lapsos de tiempo en el seno de las distintas biopoblaciones, siendo este cambio bajo determinadas circunstancias adaptativo. Los estudios de macroevolución, por otra parte, se ocupan del proceso de especiación y manejan unidades temporales mucho más amplias (millones y billones de años) que los estudios de microevolución. Obviamente un trabajo como éste no puede estar sino circunscrito al estudio de la microevolución.

cambios resultan más adaptativos que otros dentro de una determinada población, mientras que otras variaciones son simplemente neutrales.

Sin embargo, el concepto de adaptación no es patrimonio exclusivo de los biólogos evolutivos, ni es un concepto acuñado por el propio Darwin, sino que tiene una larga existencia en el seno de las diversas corrientes dentro de las Ciencias de la Vida. Para un teólogo natural como Paley una adaptación no era sino el producto de un diseño preestablecido por Dios. La perspectiva darwiniana supone un giro bastante considerable respecto de esta opinión ampliamente generalizada. Una de las grandes aportaciones de Darwin es el concepto de selección natural. Éste tiene un papel central en la explicación naturalista del proceso evolutivo, ya que explica cómo se originan las distintas especies.

Evolución y Selección Natural según la doctrina darwinista y la síntesis neodarwinista

A. Teoría Clásica

En *On the Origins of Species by Means of Natural Selection*, publicado en 1859, Darwin aporta numerosos ejemplos para explicar la idea de selección natural. El razonamiento de Darwin para justificar la evolución por medio de la selección natural se basa en tres premisas y dos conclusiones (Fontdevila y Moya, 2003).

Las premisas del argumento son:

- (1) Crecimiento geométrico de las biopoblaciones.
- (2) Progresión aritmética en el aumento de los recursos naturales.
- (3) Existencia de variabilidad interindividual heredable.

Las consecuencias son dos: C1) En las poblaciones se genera una lucha por la supervivencia que propicia que exista una selección de los más aptos. (C2) Los individuos seleccionados por su mayor aptitud dejarán mayor descendencia, lo que propiciará que gradualmente cambie la composición de la población originaria pudiendo llegar a mostrar unos caracteres morfológicos y fisiológicos muy distintos de la población ancestral. Darwin no dispuso jamás de una teoría de la herencia, por lo que no pudo explicar cómo las variaciones observadas en la composición de las diferentes biopoblaciones eran transmitidas a la siguiente generación. Fue con el surgimiento y desarrollo de la genética cuando pudo comprobarse la existencia de la selección natural.

B. Teoría sintética

Desde la síntesis neodarwinista se señala que hay selección natural si, y sólo si, concurren las siguientes tres condiciones: (1) Existencia de variabilidad interindividual, (2) Existencia de una relación entre variabilidad en eficacia biológica y variabilidad fenotípica y (3) Existencia de una relación entre las características fenotípicas y el genotipo del individuo.

La selección natural hace innecesaria cualquier invocación a la existencia de un Dios diseñador, pues es un proceso mecánico y no intencional⁶⁷. La teoría de la selección natural tiene por objeto explicar, sin el concurso de causas externas a la propia naturaleza, el origen de las distintas especies que pueblan o han poblado el planeta. Pero, en la medida en que Darwin asume que la evolución interespecífica es gradual, también explica las diferentes modificaciones fenotípicas en relación con la adaptación de la misma especie a su medio. Es decir, no sólo explica la evolución interespecífica (cladogénesis), sino también la evolución intraespecífica (anagénesis).

El concepto de adaptación biológica dista mucho de ser unívoco. Desde un punto de vista bastante intuitivo puede decirse que un rasgo constituye una adaptación en la medida en que aparentemente parece estar diseñado para un determinado fin. Eso es lo que se desprende de la etimología del término, si es que ésta tiene algo que decir. Así "adaptación", ateniéndonos a su etimología, resulta de "ad" y "aptus", apto para un determinado fin o propósito. La noción misma ya tiene un sesgo teleológico difícilmente eludible. Decimos que las uñas de un gato son adecuadas para cazar ratones o que el hecho de que las diferentes orquídeas tengan un fidelísimo parecido con las gónadas de algunos insectos es una adaptación y tiene por objeto o función facilitar la polinización de estas plantas. Si analizamos la morfología de los picos de las aves, ya Darwin se percató de ello en su visita a las Galápagos, veremos que las formas varían desde un pico muy generalizado propio de un eurifago como el pico del cuervo común, *Corvus corax*, hasta formas más precisas como la que posee una

⁶⁷ Vid. Apéndice.

estenófaga como el jilguero, *Carduelis carduelis*, cuyo pico está adaptado para recoger y partir pequeñas semillas. Otras aves acuáticas como las pertenecientes al orden de las pelecaniformes son ictiófagas y tienen el pico adaptado para pescar. Las definiciones más generales nos dicen que una adaptación es algo que resulta útil⁶⁸, y se apresuran a relacionar el concepto de adaptación con otros conceptos como función, diseño o eficacia biológica, conceptos con los que es obvio que el concepto está relacionado de alguna manera. Sin embargo, estas definiciones son poco rigurosas, recogen más bien el uso cotidiano, informal, del término y, no menos, el uso habitual y acrítico que cualquier biólogo no especializado atribuiría a un rasgo cualquiera que resultara *útil* para algo.

También usamos el mismo término tanto para referirnos a un proceso como a un resultado⁶⁹; el proceso de adaptación, o una adaptación en curso, y el hecho de que un rasgo determinado haya culminado el citado proceso. Tampoco parece ser el mismo uso aquel que puedan hacer un biólogo evolutivo o un fisiólogo. Así, no es lo mismo decir de un organismo particular que está adaptado al esfuerzo físico en virtud de su entrenamiento deportivo previo, o referirnos también a cualquier fenómeno fisiológico en términos de adaptabilidad como sucede cuando estudiamos el tejido muscular y señalamos los cambios producidos por una estimulación repetida o la respuesta de acomodación del cristalino ante los estímulos visuales, que señalar que el mecanismo piloerector es una adaptación para mantener la homeostasis en los mamíferos o mencionar la forma del pico de los pinzones como adaptaciones al medio. En este sentido puede resultar útil distinguir entre adaptaciones ontogenéticas y adaptaciones evolutivas, sólo estas últimas van a ser abordadas en la presente sección.

En Biología evolutiva “adaptación” puede definirse de varias maneras. Autores como Gould y Vrba (1982), Gans (1988) o Mayr (1988) señalan la ambigüedad existente en el uso científico del término. La mayoría coincide en señalar que el término adaptación bien se emplea en un sentido eminentemente histórico o, por el contrario,

⁶⁸ Cf. Ayala y Valentine (1983: 163)

⁶⁹ Cf. Sober (1984: 196)

exclusivamente presentista y centrado en la contribución en términos de eficacia que hace un determinado rasgo al organismo portador. La conveniencia de decantarse por una u otra tendrá repercusiones más que notables en las investigaciones que en este capítulo se presentan acerca del concepto de función.

Entre las definiciones clásicas en términos históricos podemos encontrar la de Williams (1966) en su ya clásico *Adaptation and Natural Selection*, un libro "escrito en prosa inglesa, sin una sola ecuación a la vista", en palabras de Sober (1993). Williams se toma muy en serio la tarea de definir el concepto de adaptación y calibrar el alcance de la selección natural. A propósito del concepto de adaptación señala que:

"Evolutionary adaptation is a special and onerous concept that should not be used unnecessarily, and an effect should not be called a function unless it is clearly produced by design and not by chance. When recognized, adaptation should be attributed to no higher a level of organization than is demanded by the evidence. Natural selection is the only acceptable explanation for the genesis and maintenance of adaptation"

Williams (1966: v)

Williams tiene muy claro cuáles son los peligros de no tomarse en serio el concepto de adaptación y el problema que entraña interpretarlo con excesiva liberalidad. Por otra parte, la relación que establece Williams entre funciones y adaptaciones es automática, un rasgo del que decimos que es una adaptación (y no meramente un rasgo adaptativo) es un rasgo funcional. En cuanto al origen de las adaptaciones, Williams señala el proceso de selección natural, lo que implica admitir sin lugar a dudas que sólo un proceso de selección natural puede dar lugar a adaptaciones⁷⁰. Así, un rasgo puede ser

⁷⁰ Obsérvese que Williams (1966) critica la idea de una selección de grupo, de ahí que afirme en el texto citado que *adaptation should be attributed to no higher a level of organization than is demanded by the evidence*. Un nivel superior sería el grupo, a su juicio es una idea difícil de asimilar y propone seguir admitiendo el individuo como unidad de selección, si bien tampoco considera implausible que pueda haber selección de grupo. La idea de que así podía suceder es debida a Wynne-Edwards y tuvo su momento de esplendor hacia la década de 1960. La selección de grupo se refiere a la existencia de poblaciones en competición dentro de una misma especie y se planteó como explicación del origen y proliferación de rasgos como las

adaptativo, pero sólo será una adaptación si es resultado de un proceso de selección natural. En este punto Williams es categórico:

"Natural selection would produce or maintain adaptation as a matter of definition."
Williams (1966: 25)

La idea que Williams tiene del concepto de adaptación es claramente histórica⁷¹, no es que sólo la selección natural produzca adaptaciones, sino que, además, en tanto que proceso que se desarrolla en el tiempo, sólo cuando ha habido un lapso de tiempo suficientemente largo en el cual ha ocurrido la selección del rasgo en cuestión, podemos hablar de adaptaciones. La noción de adaptación es, así un concepto histórico que debe diferenciarse de aquellos rasgos que si bien son útiles en el desempeño de alguna función actual que aumenta la eficiencia del organismo, no constituyen genuinas adaptaciones. En esto consiste la onerosidad del concepto que Williams reclama.

Williams intenta dar respuesta a las dificultades que se presentan para determinar si un determinado rasgo es adaptativo o es una adaptación. Es necesario un trabajo teórico que aporte cierta claridad sobre estos puntos oscuros y él lo lleva a cabo. Pero las relaciones entre la historia evolutiva del rasgo que consideramos como adaptación y la eficacia o utilidad actual del mismo no es un problema que solucione Williams, quien claramente se decanta por mantener la interpretación historicista del concepto de adaptación. El problema principal consiste en articular adecuadamente las relaciones que pueden establecerse entre ambas dimensiones temporales, el pasado, sugerido por el concepto de diseño, y el presente, exigido por la referencia ineludible a la eficacia o aptitud del rasgo en cuestión.

proporciones aberrantes de individuos de un sólo sexo en las que una explicación gradualista centrada en el individuo parecía implausible.

⁷¹ Williams (1966: 6) sostiene que *"We attribute the origin and perfection of this design to a long period of selection for effectiveness in this particular role."* La dimensión histórica es patente, sólo podemos decir de un rasgo que es una adaptación si ese rasgo ha sido seleccionado en virtud de la funcionalidad que entraña y de la eficacia con que dota al organismo durante un largo periodo temporal.

Lewens (2007) señala lo que él denomina tres estilos de definición de la adaptación. En primer lugar tenemos una definición -o más bien una pseudodefinition- que se inferiría de algunos pocos ejemplos de lo que consideramos adaptaciones. Obviamente este tipo de "definiciones" son muy poco informativas. En segundo lugar tenemos definiciones más informales con claras resonancias de la teología natural. En este sentido Lewens no duda en incluir en esta categoría a un autor como Williams. La definición de Williams es problemática según Lewens porque en realidad, y pese a los esfuerzos del propio Williams por alejarse de cualquier interpretación teológico-natural o creacionista, es susceptible de una interpretación en los mismos términos que Williams rechaza⁷². El problema nuevamente es la noción de diseño, si se hace depender la funcionalidad, o la propiedad de ser una adaptación, de la de estar bien diseñado. ¿Qué impide a un teólogo natural apoyar sus reflexiones en este tipo de razonamientos principalmente cuando a pesar de todos los reparos posibles acabamos acudiendo a nociones intencionales y valorativas como "buen diseño"? Para Lewens las alternativas sólo pueden ser dos, o bien nos deshacemos del concepto de adaptación o bien proponemos una buena definición. Puesto que lo primero no parece viable, habrá que ensayar la segunda estrategia, una definición formal del término "adaptación". Las definiciones formales que constituyen el tercer tipo de definiciones consideradas por Lewens igualmente reflejan la polémica entre las dos principales familias señaladas anteriormente, a saber, definiciones históricas frente a definiciones no históricas. Nuevamente el problema de la dimensión temporal reaparece.

El ejemplo paradigmático de definición formal en términos históricos lo aporta Sober (1984):

⁷² Williams (1966: 6-7) explícitamente rechaza cualquier interpretación creacionista: "*This book is based on the assumption that the laws of physical science plus natural selection can furnish a complete explanation for any biological phenomenon, and these principles can explain adaptation in general and in the abstract and any particular example of an adaptation. This is a common but not a universal belief among biologist.*"

"A is an adaptation for task T in population P if and only if A became prevalent in P because there was selection for A, where the selective advantage of A was due to the fact that A helped perform task T."

La definición de Sober claramente está inspirada en la de Williams, pues al igual que éste no duda en asumir que una adaptación es el resultado de un proceso de selección natural. Sin embargo, Williams no asume la inversa de la proposición anterior, es decir, no dudará en negar que la prevalencia de un rasgo sea debida sólo porque haya habido selección para ese mismo rasgo y, en consecuencia, pueda contar como adaptación. De otra manera, la existencia de un proceso selectivo no garantiza suficientemente el surgimiento y mantenimiento de adaptaciones, debemos distinguir las verdaderas adaptaciones de las meras degeneraciones y para tal fin no es suficiente apoyarse en un proceso selectivo, sino sólo necesario. De hecho, Williams argumenta la insuficiencia de la selección natural en el surgimiento de adaptaciones a partir de un análisis muy crítico sobre los experimentos de Waddington⁷³ en *Drosophila melanogaster*. Según Williams, para que algo cuente como adaptación no debe únicamente ocurrir que sea el resultado de un proceso de selección natural sino que además debe

⁷³ Waddington experimentó los efectos del éter sobre las huevas de *Drosophila*. Para ello sometió a las huevas a una dosis subletal de vapores de éter y observó que mientras muchos de los supervivientes nacían normalmente, otros pocos desarrollaban un bitórax. Estos ejemplares anormales fueron seleccionados como progenitores de la siguiente camada y las huevas que produjeron fueron igualmente sometidas a la misma exposición experimental de vapores de éter observándose resultados análogos. Sin embargo, a partir de una determinada generación la aparición del bitórax es generada sin necesidad de exposición alguna a los vapores de éter. Waddington denomina a este proceso *asimilación genética* y señala que el proceso de selección natural debe ser suplementado por la asimilación genética en el desarrollo de las características fenotípicas. Waddington no es lamarckista, pese a lo que imprudentemente pudiera pensarse acerca de su dispositivo experimental. La idea que quiere poner de manifiesto es que el genotipo es canalizado por las condiciones ambientales, dicho de otra manera, el genotipo asimila las condiciones ambientales y responde en consecuencia con un fenotipo adaptado. Si se desarrolla un bitórax no es de la nada, sino porque existen alelos generadores de bitórax en el genotipo de *Drosophila*. La presencia o no de éter modifica la expresión de estos alelos y, en consecuencia afecta al incremento de su frecuencia en las poblaciones experimentales de *Drosophila*.

concurrir que el rasgo en cuestión constituya una respuesta adaptativa y no un mero resultado. Y para que algo constituya una respuesta adaptativa deben estar presentes los mecanismos sensoriales que posibilitarían al organismo percibir la situación ambiental y disponer los mecanismos efectores que nos capacitan a subvertirla. En última instancia lo anteriormente dicho depende del diseño del organismo como resultado del desarrollo de un programa determinado de información genética bajo un supuesto principio de economía de la información genética que Williams postula al inicio de su libro.

Las ideas de Williams son discutibles desde diversos puntos de vista. Para Sober hay suficiente evidencia empírica para no admitir ese supuesto principio de economía genética como pueda ser la existencia de DNA basura, dato que invalida automáticamente el referido postulado. A un nivel más conceptual debe recordarse el excesivo énfasis que pone Williams en el diseño, este argumento, siguiendo las interpretaciones de Lewens, fácilmente lo pueden utilizar los creacionistas y los partidarios de la teología natural.

En definitiva, la definición de Sober sí admite la conversa de la propuesta de Williams. Las diferencias que el primero añade son suficientemente relevantes y suponen un claro avance respecto de los planteamientos de Williams. En primer lugar, obsérvese que Sober no formula su definición en términos de condiciones necesarias, sino de condiciones necesarias y globalmente suficientes. En segundo lugar no se apoya en un concepto desnudo de selección, sino que lo matiza en términos de *selección para*. No sólo si hay selección hay, en consecuencia, adaptaciones, sino que también ocurre que un rasgo prevalece porque ha habido *selección para* ese mismo rasgo y así cuenta como adaptación. Por esta razón se justifica la existencia de un bicondicional en la definición de Sober.

Por las razones expuestas anteriormente podemos decir que la definición de Sober es histórica y formal, mientras que la historicidad de la definición de Williams era mucho más lábil y no había ningún cuidado en la corrección formal de la misma.

La historicidad de la definición de Sober no puede ser puesta en duda, incluso el propio Sober manifiestamente refiere a ella:

"Adaptation is a historical concept. To call a characteristic an adaptation is to say something about its origin. It is for this reason that adaptations stand in contrast to fortuitous benefits"
Sober (1984: 199).

Decir de algo que es una adaptación implica, según este autor, señalar algo sobre su origen y sobre su prevalencia a través de los mecanismos selectivos. Sin embargo, no todos están dispuestos a aceptar tales constricciones históricas. La definición del concepto de adaptación en términos historicistas afirma que un rasgo contará como una adaptación si ha habido un proceso de *selección para* ese mismo rasgo en algún momento de la historia evolutiva del mismo. Este mismo rasgo ha sido seleccionado porque resultó adaptativo o funcional en, al menos, un momento determinado. Pero de ahí no se sigue que el rasgo ejecute la función para la que fue seleccionado y que lo convierte en adaptación. La interpretación historicista del concepto de adaptación es incompleta; puede haber sucedido así, pero también puede suceder que el rasgo considerado como adaptación sea a día de hoy afuncional o que un rasgo para el que no haya habido selección alguna resulte en el presente adaptativo. Enfatizar el rol actual del rasgo considerado como adaptación tiene problemas muy similares; un rasgo puede resultar adaptativo y no haber sido el resultado de un proceso selectivo. La definición de Sober, si bien ha gozado de relativa aceptación entre el reducido núcleo de los filósofos de la Biología, no ha levantado ninguna pasión entre los biólogos que parecen más inclinados a decantarse por interpretaciones del concepto de adaptación en términos de eficiencia actual del rasgo.

Las razones que pueden esgrimirse en una crítica a las definiciones historicistas de la adaptación son varias, desde (A) la negativa a abandonar el concepto de diseño, pese a la utilización sistemática de este concepto como recurso argumentativo por parte de los creacionistas, (B) la insistencia en que el concepto de adaptación va ligado al de eficiencia y, en consecuencia, debe dar razón de la eficiencia o adaptabilidad del organismo en el momento actual al margen de cualquier consideración histórica, hasta (C) las dificultades metodológicas y epistemológicas con las que podemos encontrarnos al tratar de reconstruir la historia evolutiva de un determinado rasgo o de

una especie. Una narrativa histórica, que es lo único que puede proporcionarnos dicha reconstrucción, no atribuye causalidad alguna de manera inequívoca, sino que señala tan sólo un posible curso de los acontecimientos. Como podrá ver el lector de estas líneas el problema no difiere en gran medida de lo apuntado respecto al concepto de función en el anterior capítulo.

Respecto de la primera de las objeciones es patente que hay biólogos que prefieren entender las adaptaciones como rasgos que han sido modelados o modificados por la selección natural y no únicamente rasgos cuya frecuencia ha progresado a causa del proceso selectivo. En este sentido la idea de diseño juega un papel muy importante que quedaba reducido a su mínima expresión en la definición de Sober. Biólogos como Gould no sólo han empleado sin prejuicio alguno el concepto de diseño en su caracterización de la adaptación, sino que además subordinan la existencia de aquellas a las constricciones impuestas por el desarrollo morfológico del organismo. Para estos la selección natural es una fuerza de diseño.

La defensa de estos frente a los creacionistas consistirá en la reivindicación de la cientificidad del concepto de diseño alejándose de cualquier uso vernáculo y preteórico del término. Esta estrategia es muy respetable a mi juicio como *desideratum* pero difícilmente aplicable.

En cuanto al segundo aspecto, las razones para defender una definición de adaptación en términos no históricos son igualmente poderosas. Algunos como Bock y Von Wahlert (1965) y Bock (1980) señalan que una adaptación es cualquier rasgo que incrementa la aptitud de su portador, al margen de las consideraciones que pudieran hacerse a propósito de la historia evolutiva de los distintos rasgos. Definen el concepto de la siguiente manera:

"Evolutionary adaptation, the process, is defined as any evolutionary change which reduces the amount of energy required to maintain successfully a synerg, or the niche as the case may be, toward the minimum possible amount."
Bock y Von Wahlert (1965: 287).

Obsérvese que (1) los autores sitúan en primer plano la relación actualmente existente entre el rasgo y el medio, (2) el balance

energético que ésta supone y que indica el grado de adaptabilidad o en qué medida es eficiente y (3) la nula referencia explícita a cualquier consideración histórica. Respecto a este último punto no niegan en ningún momento el papel causal que tenga el proceso selectivo como no podía ser, por otra parte, de otra manera, sólo que este papel es irrelevante en la definición de adaptación, pues es obvio que algún tipo de selección se ha producido para que el rasgo esté donde está y desempeñe la función que desempeña.

Bock y Von Wahlert no sólo se ciñen a la tendencia general entre los biólogos de considerar como relevante el incremento de la aptitud, sino que, además, incorpora el elemento de duda metodológica que supone trazar la historia evolutiva y que he señalado en tercer lugar en la clasificación de las dificultades que afectan a las definiciones historicistas.

Mayr (1988) sigue en gran medida los pasos de Bock y Von Wahlert en la definición del concepto de adaptación. Así, considera que una adaptación es un resultado del proceso de selección natural que opera sobre el fenotipo de los individuos. Y define el concepto en términos del incremento de la eficacia fisiológica y ecológica en la mayoría o en el promedio de los individuos que componen una determinada población gracias a la presencia de uno o más rasgos que resultan más útiles en la lucha por la supervivencia. La definición de Mayr obvia, por los mismos motivos anteriormente citados, las referencias históricas en su definición.

Más contemporáneamente autores como Reeve y Sherman (1993) insisten en acentuar la contribución de un rasgo a la eficiencia del organismo al margen de cualquier consideración histórica. Su definición se centra pues en el concepto de eficacia y no en el de *selección para* como proponía Sober.

"An adaptation is a phenotypic variant that results in the highest fitness among a specified set of variants in a given environment."
Reeve y Sherman (1993: 9).

Una adaptación es la variante fenotípica más eficaz⁷⁴ de entre todas las presentes en una determinada población. La definición hace referencia al proceso de selección; no en vano está generalizada la idea de que lo que realmente se selecciona es el genotipo⁷⁵, pero, a diferencia de las definiciones historicistas, Reeve y Sherman no hacen referencia alguna a la historia pasada del rasgo.

Los partidarios de las definiciones no historicistas estarían dispuestos a admitir que incluso cuando hay selección de un rasgo podemos hablar de adaptaciones sin necesidad de que haya explícitamente un proceso de *selección para*, bien porque no acaezca, bien porque resulte metodológicamente imposible establecer la historia evolutiva del ítem en cuestión. En este aspecto las similitudes con la teoría débil de Buller (1998) que expuse en el capítulo anterior son claras. Al igual que sucede en la teoría de las funciones, la polémica entre las dos interpretaciones del concepto de adaptación muestra que la historia evolutiva puede importar o no y, las razones para que sea tenida en cuenta son suficientemente poderosas en ambos casos.

Por último, no habría que olvidar la relación existente entre el concepto de adaptación y el de eficacia biológica. Al igual que sucede con el primero la definición de éste último no es menos problemática. Así, habría que señalar que en la medida en que el concepto de

⁷⁴ Obsérvese que cuando afirmamos que se trata de la variante más eficaz estamos en realidad afirmando que es la más eficiente, es decir, la que mejor optimiza los recursos de los que dispone el organismo. La eficacia no es evaluativa, mientras que la eficiencia sí lo es, para la primera sólo cuenta la consecución de los resultados, mientras que para la segunda importa también el ajuste entre costes y beneficio. Como afirma Quintanilla: "*Intuitivamente consideramos que una acción es ineficiente no sólo cuando no logra los objetivos previstos sino también cuando los logra derrochando recursos.*" Quintanilla (2005: 131). Si bien hay que matizar que Quintanilla se refiere a la evaluación tecnológica y no a la valuación de las funciones de los órganos de los seres vivos.

⁷⁵ Cf. Mayr (2004). Mayr defiende que la selección actúa sobre el organismo y, de manera más concreta, sobre el fenotipo. Mayr incluye en el fenotipo tanto el repertorio conductual de una especie como sus productos. Al admitir que lo que resulta seleccionado es el fenotipo, Mayr está rechazando la idea de que la unidad de selección pueda ser el gen (idea mantenida por el propio Williams) o el grupo (idea en contra de la que argumenta Williams en su libro de 1966). Mayr considera que el gen aislado, sin referencia alguna al medio, no es visible para la selección y, por lo tanto, defiende que la unidad de selección es el fenotipo.

adaptación se relaciona con el de función también en éste último tiene algo que decir el debate sobre la eficacia, si bien de manera más tangencial a diferencia de lo que ocurre entre adaptaciones y funciones. Principalmente si entendemos la adaptación como un proceso y no como un producto. Sober (1984) señala que decir de un organismo que está adaptado no es sino decir que es eficaz en cierta medida; "adaptado" es otro nombre de "eficaz". Sin embargo, ya he señalado que la definición de Sober presenta problemas al considerar la eficacia presente de un rasgo y su carácter en sentido estricto de adaptación, más aún cuando insiste en distinguir procesos de resultados, pues un rasgo puede resultar adaptativo sin ser una adaptación; que sea o no sea una adaptación, conforme a la definición de Sober, será cuestión de tiempo.

No en vano decimos también que una población evoluciona por selección natural cuando ésta posee una variación en la eficacia heredable⁷⁶.

1.2. Exadaptaciones y Adaptacionismo

A menudo se ha señalado la importancia de distinguir una adaptación del beneficio fortuito que puede suponer un determinado rasgo para un organismo. Incluso en la teoría de las funciones se señalaba la distinción de la función propia de un rasgo de los efectos accidentales que éste pudiera tener como un aspecto que debía ser considerado.

El término exadaptación, debido a Gould y Vrba (1982), surge para dar cuenta de la divergencia funcional de los caracteres adaptativos. El concepto constituye una herramienta para distinguir entre la génesis histórica de un rasgo y su funcionalidad actual. Así, un rasgo puede haber tenido inicialmente una función determinada y, sin embargo, su funcionalidad actual puede diferir al haber adquirido el carácter una nueva función. Para Gould y Vrba una adaptación es un rasgo diseñado por la selección natural para algún uso. De una determinada

⁷⁶ Cf. Sober (1993) Cap. 3. la idea de que hay selección natural si, y sólo si, hay variación heredable en eficacia es original de Lewontin (1970).

adaptación decimos que tiene una función. Sólo las adaptaciones tienen funciones, mientras que las exadaptaciones tienen efectos. Una exadaptación resulta eficaz para la ejecución de su rol actual. En este sentido, decimos que tiene la propiedad de ser apta para ese fin actual o, lo que es lo mismo, es eficaz en función de su forma. Pero su forma no ha sido inicialmente seleccionada para cumplir con ese rol; no ha habido una *selección para* ese rol. En consecuencia, no puede hablarse de adaptación al menos en el sentido que Williams o Sober lo hacen y que Gould y Vrba toman por adecuado.

El ejemplo que habitualmente se trae a colación es el surgimiento de las plumas de las aves. Actualmente consideramos que la función de las plumas consiste en capacitar a las aves para el vuelo, y decimos que las plumas constituyen una adaptación para el vuelo. Sin embargo, el análisis del registro fósil a partir de *Archaeopteryx lithographica*, descubierto en 1861, que evidencia las relaciones filogenéticas entre aves y dinosaurios terópodos, muestra que las plumas surgen como una respuesta adaptativa para el mantenimiento de la temperatura corporal, siendo la capacitación para el vuelo mucho más tardía. Algunos neornites no voladores como el avestruz *Struthio camelus* o el kiwi *Apteryx hastii* o cualquiera de las cuatro especies restantes de kiwis. Todos tienen plumas y ninguno de ellos vuela. ¿Para qué sirven? Las aves inicialmente no volaban y de hecho la pérdida de la capacidad para el vuelo ha evolucionado en aquellos lugares que, como las islas, existen pocos depredadores. Así, decimos que las plumas son una exadaptación para el vuelo y no una adaptación.

El concepto de exadaptación está íntimamente relacionado con la noción de diseño. Una exadaptación tiene que ver con aquellas posibilidades inherentes al diseño de un organismo que le permiten contribuir a su propia evolución.

La cuestión de las exadaptaciones está también muy ligada a la crítica sistemática de Gould hacia lo que se ha denominado "programa adaptacionista". El adaptacionismo se define a menudo como una tesis acerca de la potencia de la selección natural.⁷⁷ Así, un adaptacionista extremo tenderá a ver cualquier rasgo fenotípico como adaptativo y, en la inmensa mayoría de los casos, como un producto de la selección

⁷⁷ Cf. Sober (1993).

natural. La polémica sobre el adaptacionismo fue iniciada por S.J. Gould y Lewontin (1979) en el ámbito de un debate mucho más amplio sobre la forma. Gould y Lewontin discuten acerca de lo que consideran el modo generalizado de interpretar la morfología de los seres vivos que tiende a considerar la forma como resultado exclusivo de un proceso de adaptación previo. A este modo de considerar la estructura de los diferentes organismos lo denominan adaptacionismo. En su lugar proponen que no todo rasgo estructural es resultado de un proceso de adaptación sino que en gran medida es resultado de este mismo proceso de adaptación bajo ciertas constricciones arquitectónicas que determinan en gran medida el resultado final. Así, puede decirse que opera un plan previo de construcción con sus propias restricciones o que Gould y Lewontin denominan con la palabra alemana "*Bauplan*". La crítica al adaptacionismo se basa en la excesiva liberalidad con la que los biólogos a menudo atribuyen el carácter de adaptaciones a diversos rasgos fenotípicos. Por otra parte, a menudo se dice que un rasgo es una adaptación sin haberlo considerado en relación al resto de rasgos que constituyen el organismo. Gould y Lewontin reclaman un tratamiento holístico en el estudio de la morfología de los distintos organismos. Sería absurdo atomizarlo en partes más y más pequeñas a la vez que postulamos una función/adaptación determinada para cada una de ellas. En este aspecto el adaptacionismo dista mucho de tener una visión global sobre el diseño del organismo y atribuye adaptaciones sin una base sólida. La existencia de exadaptaciones se justifica porque, pese a que haya habido una divergencia funcional, estos rasgos han llegado a formar parte integrante del sistema y, en virtud de esta razón, no son eliminados. La noción de exaptación a la que anteriormente he hecho referencia es, pues, clara deudora de este debate.

Las críticas de Gould y Lewontin fueron contestadas por Mayr (1983) en un célebre artículo titulado *How To Carry Out the Adaptationist Programm?* Mayr, en líneas generales, considera legítimas las críticas de Gould y Lewontin sólo que considera que estos hacen una lectura excesivamente panglosiana del adaptacionismo y, en este aspecto, disiente. No es que la selección natural produzca organismo perfectos, sino que de los que produce resultan

seleccionados los más perfectos y esto ya no es panglosianismo, sino fidelidad a los hechos.

En cuanto a la crítica de Gould y Lewontin a propósito de la consideración holística de los diferentes rasgos que constituyen el fenotipo de un organismo a los que anteriormente hacía referencia hay que señalar que, aunque Mayr alaba este enfoque propugnado por Gould y Lewontin, tampoco condena taxativamente la disección del fenotipo en caracteres aislados. De otra manera, cabe preguntarse si los rasgos aislados son clases naturales o bien responden en exclusiva a los intereses del investigador. La polémica en torno a la individuación de los rasgos adaptativos es muy similar a las críticas vertidas a propósito de la teoría sistémica de Cummins por los partidarios del enfoque contrario como Millikan. Sin embargo, los científicos poco tienen que decir al respecto, ya que es un problema exclusivo de Filosofía de la Ciencia.

¿Qué alternativa proponen en realidad Gould y Lewontin? A juicio de Mayr ninguna, pues considera que sus críticas no constituyen sino una versión modificada de aquello que critican. La estrategia holística, la consideración del *Bauplan*, debería ser el último cartucho. Recurrimos a la estrategia sugerida por Gould y Lewontin cuando los criterios de significación adaptativa⁷⁸ no son aplicables. Según Mayr, la crítica que efectúan Gould y Lewontin sólo sería legítima si, en primer lugar, se definiera el adaptacionismo de manera estrecha y reductivista fragmentando al organismo en rasgos unitarios y proponiendo una historia adaptativa para cada uno de ellos y, en segundo lugar, si los biólogos asumieran que la selección natural es un mecanismo que produce rasgos perfectos. Pero los científicos a los que Gould y Lewontin califican con cierto sobresalto de "adaptacionistas" hacen ninguna de las dos cosas. El adaptacionismo al que Gould y Lewontin

⁷⁸ Un criterio de significación adaptativa nos permite reconocer a una adaptación como tal, es decir, nos permite dirimir la cuestión acerca de si un determinado rasgo es una adaptación o, en su lugar, no es el resultado de ninguna fuerza selectiva y no puede ser calificado como tal. Mayr (1983) propone considerar los siguientes dos criterios: (1) Que pueda establecerse una correlación tentativa entre un rasgo y una característica del medio y (2) Que sea posible analizar el punto anterior en el marco de un estudio comparativo de otros organismos expuestos a la misma característica del medio para tratar de averiguar si han adquirido también el rasgo especializado.

se refieren es un adaptacionismo que no tiene cabida en la mente de los biólogos evolutivos que, en su lugar y a juicio de Mayr, aplican una estrategia heurística. El adaptacionismo tiene un valor heurístico, tratar de averiguar cuál pueda ser la función de un órgano o estructura es una buena pregunta para empezar a investigar.

La versión que ofrecen Gould y Lewontin es, aunque sugerente, quizás excesivamente hiperbólica, incluso parece estar dirigida a una cierta audiencia ávida por desmontar el conocimiento científico animados por la idea de que se esconde tras él una gran conspiración. El adaptacionismo que describen Gould y Lewontin también podría ser visto no sólo desde la óptica panglosiana, sino también panfuncionalista. Decir que algo es una adaptación implica decir también que ese algo tiene o ha tenido una función necesariamente. El biólogo busca la funcionalidad de los rasgos que describe. Desde el momento en que señalamos que la dimensión histórica puede jugar un papel relevante en el debate en torno a la teoría de las funciones, éste fácilmente puede estar viciado de un adaptacionismo muy pernicioso. Sin embargo, las diversas acusaciones de adaptacionismo que puedan hacerse a las distintas teorías de las funciones pueden minimizarse al considerar la dimensión heurística presente en las diferentes teorías científicas. En gran medida lo que importa a los científicos es usar buenas hipótesis de trabajo, métodos suficientemente confiables que puedan arrojar resultados interesantes. Decir qué sea una función ha ocupado mucho tiempo a los filósofos de la Biología, pero el concepto de función biológica no es un concepto filosófico, sino científico y como tal pertenece a la teoría de la evolución, si bien hay que recordar que no es un concepto central dentro de esta teoría y que muchas exposiciones de la misma no lo tienen presente. Que los filósofos hayan discutido acerca de la idoneidad de una serie de definiciones no implica que hayan desarrollado un trabajo de primera magnitud filosófica, sino más bien que han sacado a la luz algunos de los problemas filosóficos que ese mismo concepto científico podía generar a partir de un trabajo de reflexión propio de la Biología teórica. En este sentido, traer a colación el adaptacionismo puede arrojar una serie de resultados interesantes desde el momento en que se señala la dependencia de los conceptos de función y adaptación -aspecto que inmediatamente abordaré en la siguiente sección-. La consideración

que los propios científicos tienen acerca de las adaptaciones y de cómo se adscriben a los diferentes rasgos debería constituir una guía para la elaboración de las distintas teorías de las funciones. La discusión en torno al adaptacionismo contribuye en gran medida a aclarar no sólo el concepto de adaptación, sino también la propia naturaleza de las Ciencias de la Vida.

1.3. La relación entre las adaptaciones y las funciones

Las relaciones entre las adaptaciones y las funciones han sido abordadas por varios autores. Si bien dicha relación nunca ha sido puesta de manifiesto de manera explícita. Sober (1993) es de los pocos autores que abordan específicamente la relación de las adaptaciones con las funciones abogando por su identificación.

"Tal vez debemos contemplar el concepto de adaptación, tal como lo hemos definido aquí, como la única roca firme en todo este movedizo océano semántico. Si se entiende que la función significa adaptación queda bastante claro lo que significa el concepto. Si un científico o filósofo utiliza el concepto de función en algún otro sentido, debemos pedirle que nos aclare el concepto. El de función no es un concepto que pueda tomarse en su sentido literal inmediato.

El término "función" está con frecuencia en boca de los biólogos. Pero eso no quiere decir que sea un término teórico de alguna teoría científica. "Función" no es como "coeficiente de selección" o "deriva genética al azar". Se utiliza para hablar sobre teorías, pero no aparece de modo ineliminable en ninguna teoría."

Sober (1993) [1996: 149]

Las razones que ofrece Sober para sostener la equiparación entre funciones y adaptaciones son poderosas y, a mi entender, es lo más cabal que se ha dicho a propósito de este largo debate. De hecho, coincido con Sober en afirmar que realmente dicha equiparación es la única roca firme a la que sujetarse. Sin embargo, creo haber expuesto suficientemente bien las razones para rechazar la definición de Sober del concepto de adaptación. Así, si se quiere establecer una relación entre ambos conceptos tendrá que ser a partir de otra definición de adaptación que constituya una roca mucho más firme que la de Sober y que evite los problemas señalados anteriormente. Por lo demás, es obvio que para un biólogo la funcionalidad dependerá de las

adaptaciones y, en consecuencia, cualquier otro uso del término requerirá de la justificación pertinente. El filósofo en este sentido no puede sino ir de la mano del científico.

En cuanto al rango que tenga el concepto de función biológica dentro de la teoría evolutiva es patente que, como he mostrado, el propio Sober duda de que el concepto de función sea un concepto teórico en el sentido de que no es ningún concepto teórico que aparezca como esencial o ineliminable en alguna teoría. Argumenta que análogamente no necesitamos el concepto de función del mismo modo que sí necesitamos los conceptos de gen o de selección. No le falta razón, no hay ninguna teoría que dependa del concepto de función en su formulación más elemental. En este sentido podemos pensar que, a diferencia de lo que sucede con otros conceptos plenamente científicos, el concepto de función está más cerca del lenguaje cotidiano que del científico. Sólo el uso en el sentido anteriormente definido, función como adaptación, es legítimamente científico. El punto flojo de este argumento es que no se ajusta a lo que constituye la práctica científica. Un biólogo habla de funcionalidad aunque no siempre sepa a qué se está refiriendo o no sepa si su uso responde más al sentido científico definido con anterioridad o al uso cotidiano. En realidad no está empleando un uso cotidiano, sino un término vago, o está empleando el término en varios sentidos diferentes. En cualquier caso y aunque no lo necesite como concepto central en sus teorías, si es obvio que caracteriza ciertos elementos como funcionales en un sentido u otro o que realiza adscripciones funcionales. Y las realiza porque lo cree necesario. Decir, además, que quizás no sea un concepto científico es una solución poco valiente, con esta afirmación Sober abandona el problema demasiado pronto. Si no es un concepto científico y pertenece en gran medida al lenguaje cotidiano, ¿por qué tantas décadas dedicadas al análisis del concepto? ¿Acaso los filósofos de la ciencia no tienen buen ojo clínico para seleccionar los problemas importantes? El criterio para determinar si un concepto cuenta como concepto genuinamente científico parece estar muy claro para Sober, un concepto es un concepto científico o un término teórico si está presente en alguna teoría científica como término ineliminable. Es cierto que un término teórico obviamente es aquel que pertenece a

una teoría de modo ineliminable, es decir si $t \in T$ y T es una teoría, entonces t es un término teórico.

Soler (2003), un biólogo evolucionista, no duda en relacionar las adaptaciones biológicas con la funcionalidad. Así, afirma que:

"Basándose en que las características de los organismos deben de estar próximas a un óptimo, es posible predecir que muchas de las características de los seres vivos han sido seleccionadas por realizar una función en el ambiente que explotan."

Soler (2003: 128).

Y más adelante de manera explícita sobre el concepto de función señala que:

"Cualquier carácter, o conjunto de caracteres, que permite a los individuos que lo poseen superar con éxito uno o más agentes de selección es lo que genéricamente se denomina adaptación. Por tanto, una adaptación podría ser defendida como un carácter de un organismo cuya forma es el resultado de procesos de selección en un contexto funcional particular. En consecuencia, el proceso adaptativo es la modificación evolutiva de un carácter bajo selección y que implica una mayor eficiencia o ventajas funcionales (eficacia biológica) en un ambiente determinado con respecto a poblaciones ancestrales."

Soler (2003:148).

Como se habrá notado, las definiciones de Soler son bastante laxas, centrándose en una ocasión sobre la contribución del rasgo a la eficacia biológica del organismo o, en la otra, acentuando su dimensión histórica. Sin embargo, el problema principal de esta definición es la falta de precisión. La definición de Soler está inspirada en la de West-Eberhard (1992)⁷⁹. No obstante, Soler no recoge las restricciones que

⁷⁹ West-Eberhard (1993:13) define el concepto de adaptación de la siguiente manera: *"According to strict usage in evolutionary biology, it is correct to consider a character an "adaptation" for a particular task only if there is some evidence that it has evolved (been modified during its evolutionary history) in specific ways to make it more effective in the performance of that task, and that the change has occurred due to the increased fitness that results. Incidental ability to perform a task effectively is not sufficient; nor is mere existence of a good fit between organism and environment. To be considered an adaptation a trait must be shown to be a consequence of selection*

West-Eberhard impone en su definición, principalmente porque West-Eberhard, siguiendo a Sober, la estipula en términos de *selección para*, mientras que Soler no establece tal distinción.

En realidad tal laxitud tiene la ventaja de proporcionarnos una definición bastante integradora que trata de articular las dos dimensiones temporales a las que me he estado refiriendo. Por una parte, para que algo cuente como adaptación debe haber sido seleccionado de alguna manera dentro de un conjunto de rasgos poseídos por algunos de los individuos que integraban la población ancestral de la cual es descendiente el o los individuos portadores del rasgo considerado como adaptación. Por la otra parte, el rasgo debe ser funcional, es decir, debe incrementar la eficacia biológica de sus portadores.

La definición de Soler representa lo que podría considerarse una visión estándar del concepto de adaptación y función biológica y que muy bien podría coincidir con la teoría presentada por Godfrey-Smith. Estas teorías delimitan lo que cabe esperar tanto de una adaptación como de su efecto funcional y en buena parte de los casos considerados tanto como adaptaciones como funciones resultan aplicables. Los problemas ya los he presentado; sumariamente, corremos el riesgo de no distinguir una adaptación genuina de un rasgo vestigial o de no reconocer como adaptación un rasgo para el que no ha habido selección y que, no obstante, ha sido seleccionado o, por último no reconocer un rasgo actual cuya contribución a la eficacia del organismo es indubitable como adaptación.

1.4. Conclusiones sobre las relaciones entre los conceptos de función y adaptación: Las funciones son adaptaciones

Los problemas filosóficos acerca del concepto de función biológica surgen, como he señalado en el capítulo anterior, en el contexto del modelo nomológico-deductivo para irse progresivamente independizando como concepto explicativo. Por otra parte algunos de

for that trait, whether natural selection or sexual and social selection- whether the selective context involves what Darwin called "the struggle for existence", or competitive interactions with conspecifics."

los teóricos de las funciones, principalmente aquellos partidarios de una perspectiva histórica, han apoyado sus teorías de un modo más o menos explícito en el concepto de adaptación. No en vano han apuntado al proceso de selección natural como productora y mantenedora de rasgos funcionales. Sin embargo, hasta el momento nadie ha señalado la dependencia estricta existente entre las funciones biológicas y las adaptaciones. En mi opinión el debate que mayoritariamente ha preocupado a los biólogos y a algunos filósofos de la Biología acerca de la conveniencia de definir el concepto de adaptación de una u otra manera es en buena medida análogo al debate suscitado principalmente entre los filósofos a propósito del concepto de función. Coincido con Sober (1993) en querer redirigir el debate al terreno de las adaptaciones y en minimizar, por las mismas razones, la importancia que los filósofos han querido otorgar al concepto de función biológica. Una función no es sino, en la mayoría de los casos, la ejecución o lo que hace una adaptación. Bien se defina este último concepto en términos históricos o en términos de eficiencia, es obvio el paralelismo existente entre las teorías histórico-etiológicas, en el primer caso, y las teorías sistémicas, en el segundo. Para los biólogos resulta claro que un rasgo adaptativo es en gran medida un rasgo funcional, si bien, acorde con algunas propuestas definitorias del concepto de adaptación como la defendida por Sober, podemos encontrarnos con adaptaciones que en la actualidad son afuncionales.

En mi opinión los filósofos hemos sido poco sensibles a la relación existente entre los dos conceptos, probablemente se hubiera evitado mucha polémica inútil si hubiéramos prestado más atención al uso que hacen los biólogos no sólo del concepto de función, algo que no se nos podrá reprochar, sino a los conceptos adyacentes al uso del mismo término, a saber, adaptación, eficiencia y selección. No podemos pretender solventar los problemas filosóficos que el concepto de función entraña subsumiéndolo en un concepto aparentemente apromblemático como pudiera parecer a primera vista que era el de adaptación. Al inicio de este capítulo ya critiqué esta estrategia. En su lugar creo haber mostrado suficientemente que el concepto de adaptación, lejos de exhibir una univocidad patente que permitiera dirimir la polémica funcional en el terreno de la Biología Evolutiva no hace sino arrojar grandes lagunas. Vistas las cosas así, nos

encontramos con que los biólogos prefieren definiciones ahistóricas del concepto de adaptación. Podemos relacionar sin grandes problemas sus preferencias con las teorías sistémicas, tan criticadas por muchos filósofos. Así, en la medida en que una adaptación no es sino la aportación causal con la que un determinado ítem contribuye a la eficiencia de un organismo en un determinado régimen selectivo, podemos entender que adaptación y funcionalidad van de la mano, siendo las dos caras de una misma moneda. El rasgo, en definitiva, es una adaptación, y lo es porque produce adaptabilidad al ser funcional. No hay, en consecuencia, funciones al margen de las adaptaciones, o al menos no como la mayoría de los filósofos como Millikan, Neander o Godfrey-Smith piensan.

El seleccionismo defendido por aquellos ha sido mal asimilado por la influencia perniciosa de querer insertar a toda costa la normatividad señalando que una función era un concepto normativo y pretendiendo fundamentar dicha normatividad en el proceso histórico de una *selección para* determinados rasgos. Las diferentes teorías que se han presentado bajo el título de intentos de instanciación han suavizado más los requisitos considerando la existencia de *selección de* en los procesos que dan lugar a adaptaciones. Una definición del concepto de adaptación en términos no históricos debería fácilmente acoplarse con lo señalado por estas teorías y más aún una teoría establecida en términos puramente sistémicos como pretende Davies. De hecho, si la réplica habitual que condena al fracaso estos intentos por la inmensa mayoría de la comunidad filosófica es el argumento de que el concepto de función es normativo lo que debe hacerse es tratar de entender en qué medida lo es también el de adaptación y no elaborar quimeras filosóficas forzosamente normativas y alejadas del uso habitual de los biólogos, a riesgo de acabar haciendo mala ciencia y traicionando el sano espíritu científico del que los abanderados del naturalismo hacen gala. En mi opinión la exigencia de normatividad es una exigencia que compete más a la Filosofía de la Mente que a la Filosofía de la Biología. En la siguiente sección abordaré este importante tópico.

2. Funciones y normatividad

De las adscripciones funcionales se dice que son normativas y uno de los rasgos más explícitos de muchas de las teorías al uso es que intentan preservar esa misma normatividad, ya Wright insistió en preservar la distinción entre los efectos propios y accidentales de un ítem. También importa señalar esta distinción en el terreno de las adaptaciones, presumiblemente menos polucionado por la influencia de los filósofos, de hecho esta misma distinción se menciona a propósito de las exadaptaciones y la necesidad de distinguir entre una genuina adaptación y el beneficio fortuito que puede conllevar la posesión de un rasgo determinado.

Una de las principales disputas en torno a la teoría de las funciones es, pues, el supuesto carácter normativo de cualquier adscripción funcional. Así, afirmar de un ítem que tiene una función, implica seleccionar un efecto de ese ítem entre otros posibles efectos como su función. Distinguir entre lo propio y lo accidental es, pues, uno de los objetivos que, según muchos teóricos de las funciones, debería satisfacer una buena teoría, ya que admiten que una función no es un efecto cualquiera⁸⁰.

La insistencia en los aspectos normativos puede resultar un tanto extraña en un ámbito inicialmente alejado de la Filosofía Moral o de la Filosofía del Derecho, incluso puede pensarse que hablar de algo así como "normas naturales" está muy alejado de nuestra imagen científica actual. Sin embargo, a nadie le sorprende la existencia de leyes naturales, un término no menos jurídico que el "norma".⁸¹

Los partidarios de las teorías histórico-etiológicas tienen muy claro que los conceptos funcionales son conceptos normativos. Millikan (1984, 1989) señala que el concepto de función propia es un concepto normativo, un ítem funcional es una función propia cuando suponemos

⁸⁰ Vg. Wright (1973).

⁸¹ Según Popper lo propio de las leyes es vetar hechos al mundo, al igual que las leyes civiles. "*(L)as leyes naturales pueden compararse a "vetos" o "prohibiciones". No afirman que exista algo, o que se dé un caso determinado, sino que lo niegan. Insisten en que no existen ciertas cosas o situaciones, como si las vedaran o las prohibieran: las excluyen.*" Popper (1934: § 15) [Ed. española de 1962, pág. 66.].

que debe hacer algo de una manera determinada y no de otra. Neander (1991) igualmente señala este rasgo de las funciones y también Godfrey-Smith (1993, 1994).

La idea general que tienen los partidarios de la teoría histórico-etiológica es que la normatividad de las funciones sólo puede sustentarse en su historia de selección. Un ítem ejecuta la función que ejecuta porque ha sido seleccionado históricamente en virtud de que hacía lo que hacía y, en consecuencia, incrementaba la eficacia del organismo portador. Estos autores no dudan en equiparar las funciones con las adaptaciones. Pero su caracterización del concepto de adaptación es histórica y, por lo tanto, corren el riesgo de atribuir funciones a rasgos afuncionales en la actualidad o de considerar como afuncionales y anormativos rasgos plenamente funcionales. Los problemas pues, son idénticos a los problemas señalados a propósito de las consideraciones historicistas del concepto de adaptación.

Por el contrario, autores como Cummins (2002) han insistido en que cualquier intento de hablar de funciones normativas debería ser entendido como una forma de neoteleología sin el menor respaldo científico. Otros como Davies (2001) rechazan de plano cualquier referencia a conceptos normativos. Ambos presentan líneas de argumentación muy distintas, si bien coinciden en mantenerse fieles a la teoría sistémica.

2.1. El rechazo de la normatividad dentro de la teoría sistémica

Tanto Davies (2001) como Cummins (2002) presentan dos interesantes líneas de argumentación en contra de la pretendida normatividad de las atribuciones funcionales. El argumento de Davies descansa en la imposibilidad de las ejecuciones erróneas de los ítems funcionales, mientras que Cummins se centra en una crítica más general a la teoría seleccionista que él califica de neoteleológica.

Tanto uno como otro ofrecen interesantes aportaciones al análisis de la funcionalidad, si bien, como más adelante argumentaré, creo que la crítica a las nociones normativas es inconsistente.

2.1.1. La crítica de Davies a la neoteleología: la imposibilidad de las ejecuciones erróneas de las funciones

Davies (2001) argumenta en contra de la inexistencia de cualquier tipo de normatividad en las adscripciones funcionales a partir de su rechazo de la existencia de ejecuciones erróneas de funciones. No hay normatividad porque no hay posibilidad de malfuncionamiento. El rechazo del concepto de "*malfunction*" es completo, ya que asume que ni la teoría histórico-etiológica, ni los denominados intentos de instanciación ofrecen los suficientes recursos explicativos para justificar la normatividad de las funciones o la posibilidad de que se den ejecuciones erróneas o apartadas de la norma. Así, la consecuencia inmediata del rechazo del concepto de "*malfunction*" es la inexistencia de cualquier tipo de normatividad natural.

Los rasgos funcionales pertenecen a tipos funcionales y no a ejemplares. Un ejemplar ejecuta una función determinada en la medida en que pertenece a un tipo funcional que a su vez resulta una variante exitosa o más eficaz perteneciente a un determinado tipo genérico. Así, un tipo genérico es, por ejemplo, poseer un rasgo como el pico en la familia de los pinzones y un rasgo funcional es poseer la misma estructura adaptada a las constricciones del medio tal como sucede con las diferentes variedades de pinzones que Darwin observó en las Islas Galápagos. Sin embargo, un pico perteneciente a un sólo ejemplar que resulta ser superadaptado no es más que una rareza si la posesión de ese rasgo no está generalizada o en condiciones de generalizarse. No podemos adscribir funciones aisladas a los ejemplares, sino que el concepto de función es dependiente del concepto de biopoblación⁸². Davies argumenta que la ejecución incorrecta de una función, es decir los casos de "*malfunction*" es imposible del modo siguiente:

⁸² En este sentido Millikan ha argumentado satisfactoriamente la necesidad de adscribir funciones a lo que ella denomina familias establecidas reproductivamente que no es sino una aplicación del concepto más general de biopoblación que Mayr reclama como rasgo diferencial de la Biología.

(1) Un ejemplar de un rasgo cualquiera puede no haber adquirido la propiedad para la que el rasgo ha sido seleccionado, bien como consecuencia de un error genético, bien por algún otro tipo de daño irreparable y permanente.

(2) El ejemplar que no posee el rasgo para el que ha habido selección queda descalificado como miembro de la tipología funcional. Así, un ejemplar que no reúne las características que definen al tipo funcional no resulta funcional aunque sea descendiente de organismos que sí son ejemplares pertenecientes al tipo funcional.

(3) En conclusión, en la medida en que los ejemplares incapaces de ejecutar la función son miembros del tipo genérico pero no del tipo funcional es obvio que no poseen el rasgo funcional para el que ha habido selección y en consecuencia no pueden poseer la presunta malfunción seleccionada.

Como ya he señalado, las funciones se adscriben considerando tipos funcionales que resultan más exitosos que los tipos genéricos. Los ejemplares, en caso de no poder ejecutar la función adscrita al tipo al que pertenecen, dejan simplemente de funcionar, pero no son malfuncionantes. Para este autor uno de los principales problemas de la teoría histórico-etiológica es que pretende mantener la normatividad del ítem aún cuando la capacidad física para ejecutar la función está ausente. Para ilustrar su argumento Davies refiere a un hipotético caso en el que realizamos un trasplante de corazón de un cuerpo a otro y la intervención resulta un fracaso. Cuando esto sucede así no decimos que el corazón trasplantado funciona mal, sino simplemente que ha dejado de funcionar.

¿Simple cuestión de palabras? A mi juicio sí. El argumento que presenta Davies no es concluyente ni defenestra la normatividad tal como pretende. Creo que hay al menos tres razones que sustentan mi afirmación. En primer lugar, hay un concepto estadístico de normalidad que, aunque rechazado por Millikan, sí puede decirse que tiene algún papel en la normatividad que queremos presuponer a las funciones biológicas. Un rasgo es funcional en la medida en que está distribuido normalmente en una determinada biopoblación. Cualquier individuo perteneciente a esa biopoblación que no posee el rasgo o que lo posee imperfectamente es estadísticamente anormal. El ejemplar afuncional o

malfuncional lo es respecto de sus progenitores y congéneres. En segundo lugar, si debemos recurrir al análisis del lenguaje cotidiano tal como hace Davies que no duda en señalar un uso extraño o anómalo de “malfunción” no es menos cierto que existen categorías patológicas que se definen en términos de malfuncionamiento y que no son menos empleadas. Así, un individuo padece arritmia cardíaca si su corazón no bombea la sangre con una frecuencia rítmica y acompasada. En tal caso, decimos que el corazón funciona mal y no que no pertenece a la categoría funcional de los corazones. Admitir la inexistencia de malfunciones implicaría admitir la no existencia de patologías. De este modo decir que no hay normas en la naturaleza es como decir que no hay enfermedades. En tercer lugar, la posesión de un rasgo funcional no es una cuestión categórica tal como parece desprenderse de la lectura de Davies, pues la patología nos enseña que algo puede ejecutar en mayor o menor medida la función para la que hubo selección. En este punto es necesario señalar que el concepto de selección que parece tener en mente Davies es panglosiano; no es que se seleccionen los óptimos, sino los menos malos.⁸³

2.1.2. Crítica de Cummins a la neoteleología: ni adscripciones funcionales ni normatividad

La teoría sistémica es reacia a cualquier concepción normativa en la adscripción de funciones, pues no en vano ha supuesto la gran alternativa a la concepción histórico-etiológica de las funciones. Sin embargo, Cummins (2002) nada dice a propósito de la normatividad sino que directamente carga contra la teoría histórico-etiológica y sus derivaciones a las que acusa de formas veladas de teleología o, más claramente, de neoteleológicas. Lo que justifica el comentario de Cummins (2002) en esta sección dedicada al concepto de normatividad es el rechazo tajante a la postura contraria en la que la normatividad es omnipresente y, de modo más concreto, a la fuente de esa misma normatividad. En consecuencia, al despojar al concepto de función de su presunto fundamento en la teoría evolutiva tal como propone Cummins también estamos negando la fuente de la normatividad del

⁸³ Cf. Mayr (2004) cap. 8.

mismo y, a su vez, rechazando de plano la misma idea de normatividad en la ejecución funcional concreta.

Para Cummins la explicación teleológica o el proceder habitual de los partidarios de la teoría histórico-etiológica y versiones anejas trata de explicar un rasgo biológico, estructura o conducta por medio de su misma función biológica. Por el contrario, el análisis funcional del que es partidario es un análisis de las capacidades del sistema que contiene el objeto o la adscripción funcional concretas. No se pregunta por la funcionalidad del rasgo *T* en términos de "¿para qué es *X*?" sino en términos de las capacidades en que *T* contribuye al sistema. Y trata de responder especificando el diseño del sistema. De hecho, el concepto de diseño es para Cummins previo y más importante que el concepto de función. Sólo el diseño del sistema nos posibilita el análisis funcional. Así, Cummins presenta una visión holística de la situación. Los componentes de un determinado sistema son identificados funcionalmente en relación al sistema mismo y no intrínsecamente. El contexto tiene, pues, una importancia determinante. Las funciones tienen sentido únicamente como elementos del análisis funcional y no como razones que expliquen por qué están donde están los rasgos a los que las adscribimos.

La teleología, o la versión neoteleológica de la teoría de las funciones, requiere de una historia de trasfondo, ésa es la idea que dirige los análisis en la línea de la escuela histórico-etiológica. En la explicación teleológica de los rasgos y procesos biológicos, así como también de la conducta, el proceso fundacional sobre el que se yergue la adscripción funcional es la teoría de la selección natural; la idea de que los diferentes rasgos están donde están porque ejecutan una función determinada para la que han sido seleccionados. De otra manera, si tenemos corazones es porque los corazones bombean la sangre, es decir, por su propósito. Éste es el principal remanente de teleología en la moderna Biología.

El argumento de Cummins en contra de la neoteleología sostiene que la selección natural no produce rasgos en un sentido que pueda resultar relevante para los partidarios de aquella. El proceso que produce rasgos es insensible a la función de los mismos, por lo tanto la funcionalidad no es una explicación suficiente para dar cuenta de por qué el rasgo en cuestión está donde está. Si bien, esto es algo que

aceptan los partidarios de la corriente neoteleológica, añaden que, aunque los rasgos carecen de toda funcionalidad hasta el momento en que son producidos, los procesos que hacen que los diferentes rasgos proliferen y se preserven sí son sensibles a las funciones de estos mismos rasgos. Sin embargo, la tesis fuerte con la que Cummins pretende invalidar la neoteleología es que estos no proliferan a causa de las funciones que les adscribimos. El argumento principal contra las tesis seleccionistas fuertes es que un rasgo sólo puede ser seleccionado por su contribución funcional únicamente si éste implica una variación adaptativa en la población. Es decir, a juicio de Cummins debe haber al menos dos subpoblaciones que posean el mismo rasgo y que cada una de ellas difiera en eficacia. Así, si admitimos que los corazones han sido seleccionados porque bombean la sangre, entonces debemos admitir también que existe una subpoblación de estructuras cardíacas que, en efecto, bombea la sangre y otra que no.

Cummins añade que el hecho de que un rasgo *t* tenga o no una función es algo relativamente independiente de que sea seleccionado o de que se propague. En este sentido el resultado de su argumento es aplicable a las teorías más débiles o los denominados intentos de instanciación en la escuela neoteleológica⁸⁴. Así, en contra del pansleccionismo atomista que implícitamente parece subyacer a las corrientes neoteleológicas, Cummins defiende una postura holista centrada en el concepto de diseño, muy en la línea de Gould y la idea de *Bauplan*. El argumento de Cummins descansa sobre cómo interpretar el seleccionismo de que hacen gala los neoteleólogos. De este modo, la variación que observamos pertenece de lleno a la ejecución de las funciones -una función se ejecuta mejor o peor, con mayor o menor eficiencia- y no a las propias estructuras funcionales. En cualquier caso, lo que es obvio para Cummins es que la funcionalidad no puede dirigir en modo alguno el proceso selectivo. Así, Cummins desliga el concepto de función de los conceptos de selección

⁸⁴ De manera más clara: "(A) *Although the connection between adaptiveness and selection is more indirect than is sometimes appreciated. What is uncontroversial is that a trait is more fit than the competition. Exactly the same thing can be said with equal truth about every trait of that host. Every trait of the winning host spreads, regardless of how adaptative it is -regardless, indeed, of whether it is adaptative (ot has a function) at all.*" Cummins (2002:166). (El subrayado es mío).

y adaptación y a partir de este posicionamiento conceptual respecto de las funciones justifica su crítica a la neoteleología. La estrategia de equiparar funciones y adaptaciones es también sopesada. Pero el diagnóstico de Cummins es que la neoteleología no puede salvarse apelando a la referida equiparación, pues mientras que la adaptación es una cuestión de grado, la funcionalidad, a su juicio, no lo es, sino que es una propiedad categórica. El análisis funcional se basa, en consecuencia, fundamentalmente en el concepto de diseño bajo una perspectiva global.

Cummins insiste en la necesidad de desligar el concepto de función de cualquier tinte teleológico, así desaparece cualquier referencia a fines o a procesos orientados finalísticamente. La pretensión última de Cummins es naturalizar el concepto de función biológica, es decir, hacer compatible un concepto eminentemente teleológico con nuestra visión científica actual.

En mi opinión hay dos grandes escollos en el proyecto de Cummins. En primer lugar su argumento en contra de la equiparación entre los conceptos de adaptación y función creo que es erróneo aún insistiendo en la idea de diseño. A mi juicio y en contra del posicionamiento de Cummins, sostengo que la funcionalidad no es una propiedad categórica, sino que puede ser gradual. Algo es más o menos funcional y no completamente funcional o, por el contrario, afuncional. La patología nos aporta numerosos ejemplos que creo pueden invalidar la tesis anterior. Así un paciente aquejado de mordida en tijera -una patología maxilofacial- no posee una mandíbula completamente funcional, pero tampoco afuncional. De hecho como rasgo anatómico puede estar sin diagnosticar debido a que no curse dolor o molestia alguna y sea relativamente funcional. La mandíbula de dicho paciente no responde a un buen diseño, es decir, no es completamente funcional, pero tampoco es afuncional, pues le permite desenvolverse con soltura en su vida cotidiana. Quizás ante requerimientos extremos, por ejemplo, que el paciente fuera un cantante o se viera sometido a una dieta rigurosa en la que predominaran en exclusiva alimentos durísimos que obligaran a un esfuerzo extra de su musculatura mandibular y al consecuente desgaste condilar, resultaría progresivamente afuncional.

En segundo lugar, y en lo que afecta al proyecto naturalizador que pretende proponer Cummins, creo que no ofrece los recursos conceptuales suficientes para acometerlo. Su análisis se presenta como eminentemente instrumental, pues la identificación del sistema al que sobrevienen las funciones no es una clase natural, sino que depende de los intereses del investigador. Si bien esto es lo que habitualmente ocurre en la práctica científica y el análisis disposicional de Cummins tiene el mérito de reflejarlo, no es lo que cabría esperar de una buena naturalización. Un científico espera en la mayoría de los casos que sus teorías sean objetivas. Si la naturalización consiste en adoptar un punto de vista indistinguible del científico, es obvio que con el sesgo subjetivista de los intereses del investigador en la determinación del sistema que es objeto de estudio no podemos acometerla. En su lugar creo que serían necesarios una serie de aditamentos ontológicos como que el sistema estipulado se corresponda con una clase natural, es decir, sea objetivo, para acometer satisfactoriamente la referida naturalización en caso de que sea plausible.

2.1.3. Consecuencias del rechazo de las nociones normativas

Las consecuencias del rechazo de la normatividad son muy importantes, ya que es en esta presunta normatividad sobre la que descansan los fundamentos de la teoría moral, la filosofía de la mente o la epistemología naturalizadas. Si la propia teoría de las funciones es incapaz de establecer firmemente la normatividad desde presupuestos estrictamente naturalistas ¿qué cabe esperar de disciplinas que como la teoría moral, la filosofía de la mente o la epistemología emplean conceptos estrictamente normativos?

El problema de los autores que acabo de examinar es que en principio ya no quieren entrar en el juego de la normatividad y, por lo tanto, para ellos ésta no es ningún problema, sino un pseudoproblema. La idea que parecen defender ambos es que los rasgos funcionales en sí mismos no son portadores de normatividad alguna, sino que dicha atribución carece de los más mínimos presupuestos realistas y depende por entero de la perspectiva y de los intereses del investigador que realiza la adscripción funcional.

En mi opinión sí hay mucho de rescatable en las ideas de estos sólo que bajo un marco normativo. Así, en primer lugar voy a argumentar a favor de la normatividad en la teoría de las funciones y, en segundo lugar, voy a presentar un escenario en el que es lícito hablar de normas naturales al menos en cierto sentido.

2.2. Nomologicidad y normatividad

Anteriormente he señalado que el concepto de "ley científica" no resulta extraño a nuestra imagen científica, nadie piensa que sea ésta una intromisión ilegítima del derecho o la moral en el ámbito de la ciencia natural. La idea de que existen leyes científicas o leyes de la naturaleza parece estar bien asentada en la mente de los científicos e incluso es motivo de orgullo para ellos. En principio, cabe pensar que la existencia de leyes naturales no parece suponer ninguna intromisión ilegítima de un concepto jurídico en el terreno de la ciencia natural porque las leyes que ésta postula nada tienen que ver con las leyes jurídicas o con los imperativos morales, salvo el que se llamen de la misma manera o prohíban. Pero, como espero mostrar, esto no es así. La única diferencia existente es la referida a los objetos acerca de los que postulamos leyes y no a las leyes en sí mismas. Con esto pretendo argumentar a favor de la normatividad de las funciones que atribuimos a las entidades biológicas. La normatividad no es un concepto extraño para la ciencia natural ni supone ninguna intromisión o caso alguno de falacia naturalista.

2.2.1. Sobre el concepto de ley científica en el neopositivismo

Es generalizada la idea de que los diversos saberes científicos tratan de averiguar las leyes de la naturaleza. El concepto de ley científica juega un importante papel en la explicación científica según la Filosofía de la Ciencia desarrollada por los neopositivistas. Así, puede pensarse que si las leyes tienen alguna importancia es porque constituyen el elemento determinante en el proceso de explicación científica según el modelo nomológico-deductivo.

El concepto de ley científica o de ley de la naturaleza ha sido sistemáticamente objeto de análisis filosófico y, como sucede con cualquier otro concepto filosófico, las teorías disponibles son muy diversas, desde aquellas que afirman su existencia absolutamente indubitable y necesaria, hasta las que lo niegan taxativamente. Sin embargo, no parece que a día de hoy los científicos se tomen muy en serio las indicaciones de Cartwright, Van Fraassen o Giere y quieran prescindir de sus leyes universales. Mientras ellos no lo hagan, los filósofos podemos seguir disfrutando del goce intelectual que nos propicia la discusión sobre el referido tópico. En el peor de los casos la existencia de leyes es una ficción muy útil que no justifica que prescindamos de ella, del mismo modo que nadie duda de la existencia de las leyes jurídicas a menos de que alguien dude de que realmente ha sido incapacitado o encarcelado por desobedecerlas.

Los neopositivistas como Hempel o Nagel, además de clasificar los diferentes tipos de leyes como puedan ser las leyes deterministas o las leyes probabilistas, han insistido también en la importancia de caracterizar la forma lógica de todas ellas. Así, una ley tiene la forma de un enunciado condicional verdadero de alcance universal del tipo:

$$\forall x (Fx \rightarrow Gx)$$

¿Por qué un enunciado universal? La respuesta debería parecerse obvia y entender que esta formulación se corresponde con lo que cabría esperar de una ley científica, a saber, un enunciado, pues no en vano el saber científico se expresa lingüísticamente, que establece una verdad de alcance universal o un enunciado cuya validez es universal. Así, desde este punto de vista lo que nos interesa es que las leyes puedan aplicarse a un rango potencialmente infinito de objetos o fenómenos de un determinado tipo. La formalización anterior debería ser, además, un buen esquema de leyes científicas como las que a continuación se muestran, ya que en todas ellas estamos ante enunciados universales que, expresados en forma de ecuaciones matemáticas, expresan regularidades sobre una serie de fenómenos o propiedades.

1. Todo cuerpo permanece en estado de reposo o de movimiento en línea recta con velocidad constante si no actúa ninguna fuerza sobre él. (Primera ley de Newton).
2. A temperatura constante, el volumen V de una muestra de gas es inversamente proporcional a la presión P , tal que se verifica la siguiente ecuación: $P \cdot V = \text{constante} \propto ; V \propto 1/P$ (Ley de Boyle)
3. La fuerza eléctrica entre dos partículas con cargas Q y q está dirigida a lo largo de la línea que las une y es directamente proporcional a sus cargas eléctricas e inversamente proporcional a la separación entre ellas. $F_{el} = K_e \cdot (Qq/r^2)$ (Ley de Coulomb)

Aunque esto es a todas luces una caracterización insuficiente, pues hay multitud de ejemplos que satisfacen la fórmula $\forall x (Fx \rightarrow Gx)$ y no son leyes, aunque toda ley satisfaga la referida fórmula. Decir, por ejemplo que *Todos los pájaros son aves* o que *Todos los perros tienen cuatro patas* son enunciados que también satisfacen la fórmula anterior.

Los neopositivistas como Nagel⁸⁵ (1961) o Hempel⁸⁶ (1966) destacan que el rasgo que principalmente distingue las leyes científicas respecto de otros condicionales generalizados que igualmente expresan regularidades de la naturaleza es que las primeras tienen apoyo en contrafácticos⁸⁷, es decir, podemos identificar una serie de condiciones que implican en un sentido más fuerte del habitual que lo

⁸⁵ Nagel (1961) señala que una ley científica tiene la forma lógica de un condicional generalizado que expresa una regularidad en la naturaleza. Las leyes científicas deben ser distinguidas de las generalizaciones accidentales, tal que no toda regularidad factual es una ley científica, pero cualquier ley científica implica una regularidad factual. Nagel, si bien no establece unas condiciones necesarias y suficientes para decir de un enunciado que es una ley, sí señala una serie de rasgos que distinguen a las leyes científicas de otros condicionales universales, tales como que estos enunciados -leyes- sean universalmente irrestrictos, no sean vacuamente verdaderos, sean explicativos y sirvan para realizar predicciones o sean sistemáticos.

⁸⁶ Hempel (1966) destaca como las características que distinguen una ley científica de un mero enunciado condicional universal verdadero la explicatividad y el apoyo en contrafácticos.

⁸⁷ El verdadero artífice de esta concepción de las leyes científicas fue Goodman (1947)

establecido por la ley es el caso y, en caso contrario, la ley no se cumple. Una ley, por lo tanto, no se limita a constatar hechos, sino a encuadrarlos en un determinado marco modal. Otra cuestión es cómo deba de interpretarse esa modalidad y qué problemas filosóficos pueda entrañar.

El análisis positivista de las leyes científicas no es, en principio, un análisis realista -no está comprometido con tesis realistas, sino fenomenistas-. La nomicidad se entiende como una característica de los enunciados nomológicos y no como una propiedad de las cosas. Algunos positivistas defienden una interpretación de la necesidad como fundada en nuestros hábitos de pensamiento y no en las cosas mismas. La necesidad es lógica y no ontológica, aunque también es obvio, como más adelante argumentaré, que la necesidad implicativa que tienen en mente no puede ser, muy a su pesar, la que muestra el condicional material, sino algún otro tipo de implicación. Las soluciones a este tipo de problemas son las que a menudo han puesto en serios aprietos a las caracterizaciones positivistas. Como buenos empiristas, los positivistas abrazan la crítica humeana al principio de causalidad y desdeñan la idea de que exista alguna conexión necesaria que ligue entre sí los distintos objetos y acontecimientos. Las regularidades observadas no se fundamentan en ningún tipo de necesidad física. En este respecto Nagel (1961) es muy claro cuando señala que las leyes naturales son lógicamente necesarias y descarta una necesidad física o real como reformulaciones oscurantistas del problema. Hempel (1965), por su parte, analiza la noción de ley desde un punto de vista principalmente lingüístico, es decir, el análisis del concepto de ley es el análisis de las oraciones legales. Hempel caracteriza las leyes como regularidades sin necesidad de recurrir a una presunta necesidad en la naturaleza.

Una de las principales objeciones que puede ser hecha al regulativismo encarnado por los positivistas es su falta de compromiso realista u objetivo. Para un regulativista no realista o humeano como puedan ser Hempel o Nagel, las leyes no tienen por qué ser objetivas más allá de los límites que impone nuestro sistema de conocimiento. Así, si bien admiten que las leyes no son el resultado de nuestra libre invención, tampoco explican por qué no son objetivas en la medida en que la necesidad que éstas exhiben depende del sistema de

conocimiento al que pertenecen y no del mismo mundo que tratan de explicar. Un empirista lógico no puede ser jamás realista porque es fenomenista, obviamente para ninguno de ellos la falta de objetividad denunciada aquí constituye un problema. Esta perspectiva no es sensible al avance científico, pues si la nomicidad depende de nuestras diferentes teorías y no del mundo es obvio que, en la medida en que hay cambio científico o existen diversos sistemas teóricos, la diferencia entre regularidades y leyes variará, salvo que postulemos una ciencia ideal a la que pertenezcan igualmente nuestras leyes ideales⁸⁸.

Uno de los problemas de la propuesta positivista es que no establece las condiciones necesarias y globalmente suficientes para decir de un enunciado que es una ley científica. Aunque el apoyo en contrafácticos es un buen indicador de la legalidad, no constituye una condición necesaria y suficiente para que un enunciado cuente como una ley, pues un enunciado analítico cualquiera como " *Todas las zorras son raposas*" justifica su correspondiente contrafáctico, "*si X fuera una zorra, sería una raposa*"⁸⁹. El tipo de justificación que buscamos debe de ser distinto.

2.2.2. La crítica de Dretske a la concepción neopositivista de las leyes de la naturaleza

Un buen análisis de los problemas que afectan al concepto de ley científica propugnado por los neopositivistas y que resulta muy clarificador es el de Dretske (1977), autor que aparecerá más adelante a propósito de la naturalización del contenido mental. La propuesta de Dretske tiene la virtud de que puede servir para argumentar a favor de la normatividad de las funciones, de ahí que haga mención al trabajo de este autor y no a otras concepciones quizás más relevantes o de alcance más general a propósito de la polémica en torno al concepto de ley científica.

Este autor, a diferencia de los neopositivistas, presentó un análisis realista del concepto de ley científica y dudó de que la forma lógica de éstas fuera la de un enunciado universal. En su lugar, propuso que las

⁸⁸ Cf. Díez y Moulines (1997: 169).

⁸⁹ Cf. Diéguez (2005:63).

leyes científicas consistían en enunciados singulares acerca de propiedades universales y señaló que podíamos tener un conocimiento objetivo de las mismas.

Las razones de la crítica de Dretske hay que buscarlas en el análisis que éste hace de la caracterización neopositivista del concepto de ley. El diagnóstico de Dretske es el siguiente: no hay explicación suficiente alguna que nos permita dirimir entre una ley y una verdad universal al menos desde un punto de vista intrínseco. Además, ocurre que, a diferencia de las verdades universales, las leyes exhiben una opacidad semántica que no está presente en las primeras. Pero dicha opacidad no cabe buscarla en el condicional, pues en ningún momento se ha querido señalar la posibilidad de que éste consistiera en algún tipo de implicación diferente a la implicación material. Si esto es así, entonces cabe pensar que la nomologicidad depende de factores extrínsecos al enunciado condicional. Pero el análisis de tales factores extrínsecos no aporta la claridad esperada y se muestra insuficiente en la justificación de la nomologicidad. Por lo tanto, la idea de que una ley es un enunciado universal debe ser rechazada.

El punto de partida de Dretske es que nuestra concepción de ley científica difiere claramente de nuestra concepción de verdad universal. Una ley científica no puede ser una verdad universal sin más, en esto es obvio que también coinciden los neopositivistas. La discrepancia entre ambas facciones tiene que ver con las razones que se esgrimen para señalar la diferencia entre leyes y verdades universales sin más. Dretske afirma que la diferencia entre ambos tipos de enunciado es intrínseca, mientras que para los neopositivistas esta diferencia depende de factores extrínsecos. Los positivistas habían señalado que la nomologicidad de las leyes de la naturaleza se basaba en una conexión nómica que las diferenciaba de las simples generalizaciones universales. Las leyes científicas, arguyen, si bien son formalmente idénticas a cualquier generalización universal, exhiben una conexión mucho más fuerte entre el antecedente y el consecuente. Pero ¿en qué consiste la naturaleza de esa conexión? En principio cabe pensar que en algún tipo de modalidad, entre alguna forma de necesidad entre los *F* y los *G*. Ahora bien, esta necesidad introduce la formulación del concepto de ley en un contexto opaco. Tanto una ley científica como

una generalización universal tienen la forma lógica siguiente: $\forall x (Fx \rightarrow Gx)$. Si añadimos otro predicado K coextensivo de F tal que $Fx \leftrightarrow Kx$ tenemos que si $\forall x (Fx \rightarrow Gx)$ es una verdad universal entonces también lo es, en virtud de la antedicha coextensividad, el enunciado $\forall x (Kx \rightarrow Gx)$. Sin embargo, este tipo de sustituciones coextensionales no pueden realizarse en enunciados legaliformes, luego estos son claramente intensionales. Así, si tomamos una ley científica como “*Todos los diamantes tienen un índice de refracción de 2,419*” y la propiedad “*ser un diamante*” es coextensiva con la propiedad “*encontrarse en acumulaciones de kimberlita*” no podemos inferir que el enunciado “*cualquier cosa encontrada en una acumulación de kimberlita tiene un índice de refracción de 2,419*” sea una ley de la naturaleza. En los enunciados que describen leyes de la naturaleza el valor de verdad no se mantiene necesariamente bajo cualquier sustitución coextensiva y, por consiguiente, constituye un contexto opaco, algo que no ocurre en una generalización universal. Según Dretske, la opacidad viene dada por el carácter especial de necesidad que atribuimos a las leyes de la naturaleza y éste aparece como un aditamento a la generalización universal tal que no esperamos de una ley científica que sea una simple generalización, sino una generalización con una serie de propiedades que son las que introducen la opacidad en la formulación de la ley.

Dretske señala que para los neopositivistas la diferencia entre una generalización universal y una ley científica es extrínseca. La reconstrucción que hace del concepto de ley empleado por los neopositivistas es el siguiente:

Ley científica =_{def.} verdad universal + X

donde X se entiende que es el conjunto de rasgos que establecen que la verdad universal sea, además, una ley científica. De este modo, el carácter legal no viene dado por ninguna propiedad misteriosa del condicional, sino por el aditamento que supone X . Como miembros de X a menudo se incluye que (1) la verdad universal candidata a ser ley tenga un alto grado de confirmación, (2) sea ampliamente aceptada, (3) tenga un gran potencial explicativo, (4) forme parte de un sistema

deductivo o (5) pueda ser utilizada para realizar predicciones. Esta formulación tiene la virtud de ser enormemente simple. Sin embargo, esta estrategia no está exenta de graves problemas, pues cada uno de estos aditamentos no aclara en absoluto el carácter nomológico de las leyes que debiera de distinguirlos de las generalizaciones universales. Los argumentos de Dretske van dirigidos a mostrar que la perspectiva neopositivista que aborda la nomologitud como algo extrínseco al propio enunciado legal es errónea, pero la vía de argumentación no va a consistir en asumir sin más que una verdad universal es una ley y que el condicional en cuestión debe ser interpretado de una manera alternativa. Sino que la caracterización de las leyes como verdades universales es incorrecta; no son verdades universales, sino enunciados particulares que expresan propiedades universales. Veamos cuál es la línea argumentativa que emplea

En la revisión de los rasgos que habitualmente decimos que componen X en la formulación tradicional encontramos que los dos primeros casos constituyen ejemplos de condiciones epistémicas; ser ampliamente aceptado o tener un alto grado de confirmación son propiedades que sobrevienen al sujeto epistémico y no al objeto de estudio. En tanto que características epistémicas no contribuyen a explicar esa misma opacidad que exhiben las leyes científicas. En cuanto a la tercera propiedad, tener un gran potencial explicativo, es ésta una propiedad falaz si se quiere predicar de una generalización universal, puesto que el argumento supone una petición de principio. Fundamentamos la explicatividad de las leyes en su confirmación, al menos Dretske así interpreta la explicatividad tal como la entienden los neopositivistas, pero la confirmación no es sino una generalización universal y la generalización universal carece de poderes explicativos. De que el lector y yo seamos mortales y de que todos los seres humanos sean mortales no obtendremos ninguna explicación acerca de en qué consiste ser mortal. El cuarto rasgo, la integración de una ley en un sistema deductivo, no puede contar como factor que permita distinguir una ley de una verdad universal; únicamente pospone el problema al suponer que la ley es verdadera y es deducible de otro teorema o hipótesis de nivel superior. Pero nada se dice de la opacidad que ésta pueda exhibir, por lo que no soluciona el problema de la opacidad, sino que lo pospone. Por último, Dretske sostiene que una

verdad universal no puede servir para hacer ningún tipo de predicción sin el concurso de enunciados suplementarios o leyes de alcance no universal. Aún cuando están confirmadas todas las instancias de una verdad general en el caso de que ésta tenga un alcance finito, o una parte suficientemente considerable de las mismas, no estamos justificados para poder calificar de ley a esa misma verdad general. Un enunciado universal puede ser verdadero y estar suficientemente confirmado y no por ello ser una ley. La confirmación de un enunciado es insuficiente para decir que estamos ante una ley científica⁹⁰.

Anteriormente he señalado la opacidad que cree ver Dretske en los enunciados legales a través del argumento de la imposibilidad de las sustituciones de términos coextensivos. En este punto Dretske señala que la opacidad de las leyes no viene dada por los aditamentos anteriormente señalados, sino porque hay un cambio en la referencia de los términos, ya que al afirmar “es *una ley científica*” estamos

⁹⁰ Dretske (1977: 256-258) ilustra la idea anterior mediante un ejemplo de probabilidad. Supongamos que disponemos de una moneda absolutamente normal con su cara y su cruz. El que lanza la moneda conjetura que en diez lanzamientos obtendrá 10 caras, hipótesis que nadie en su sano juicio admitiría aún sin haber estudiado rudimentos de teoría de la probabilidad. Como que nuestro lanzador de monedas es muy osado se dispone a lanzar diez veces su moneda, es decir, procede a la confirmación de su hipótesis y observamos que en el momento del noveno lanzamiento sólo ha obtenido caras, tal como él mismo había predicho. Es obvio que la probabilidad de que la hipótesis de nuestro audaz lanzador de monedas sea verdadera es del 90%. Ala vista de los resultados no me queda sino conjeturar que la moneda está trucada, es decir, establezco una hipótesis auxiliar que me ayuda en la justificación de los hechos, pero en modo alguno rechazo mi hipótesis inicial que considero como ley, a saber, que si lanzo aleatoriamente una moneda la probabilidad de obtener cara o cruz es de 0.5. pero en modo alguno creo que los lanzamientos hayan contribuido a confirmar la hipótesis descabellada del lanzador de monedas. La probabilidad de obtener caras o cruces es la misma antes que después del experimento, el experimento en cuestión no aporta evidencia alguna. Las conclusiones a las que Dretske quiere llegar por medio de este experimento es que una verdad universal por sí sola no tiene confirmación posible si no es mediante el concurso de leyes o asunciones auxiliares que expliquen las regularidades manifestadas en los casos estudiados. Sin embargo, los neopositivistas a menudo mencionan el concurso de estas hipótesis auxiliares que contribuyen a la nomologización de la verdad universal de tapadillo. De ahí que una verdad universal no pueda ser confirmada sin más y que la confirmación de la misma no la califique de ley.

introduciendo un operador modal, mientras que en ausencia del mismo, la referencia de los términos es translúcida. Ahora lo que interesa es salir del contexto opaco y reformular el concepto de ley, para ello Dretske propone la siguiente formulación:

“Law-like statements are singular statements of fact describing a relationship between properties or magnitudes. Laws are the relationships that are asserted to exist by true law-like statements. According to this view, then, there is an intrinsic difference between laws and universal truths. Laws imply universal truths, but universal truths do not imply laws. Laws are (expressed by) singular statements describing the relationships that exist between universal qualities and quantities; they are not universal statements about the particular objects and situations that exemplify these qualities and quantities. Universal truths are not transformed into laws by acquiring some of the extrinsic properties of laws, by being used in explanation or prediction, by being made to support counterfactuals, or by becoming well established.” Dretske (1977:253-254).

La propuesta de Dretske acerca de las leyes caracteriza a éstas como enunciados singulares. Un enunciado legal expresa una relación entre magnitudes y no entre objetos. Así, ahora en lugar de una expresión como $\forall x (Fx \rightarrow Gx)$ tenemos que una ley viene enunciada no por el anterior tipo de expresión, sino por uno del tipo *F-idad* \rightarrow *G-idad*, donde ambos predicados expresan cualidades y cantidades universales. Consideremos un ejemplo como la ley de Arquímedes. Ésta afirma que cualquier cuerpo sumergido en un fluido experimenta un empuje hacia arriba igual al peso del fluido que desaloja, no debe ser entendida como un enunciado universal acerca de todo objeto o cuerpo sumergido en un fluido, sino como un enunciado singular acerca de la propiedad de ser un cuerpo y de estar sumergido en un fluido. La diferencia entre una ley y una verdad universal es intrínseca, no extrínseca, es decir, una verdad universal no adquiere el rango de ley por los factores anteriormente señalados. Las leyes no se aplican a los individuos sino a lo que los individuos hacen o al rol que desempeñan, a sus propiedades. Denominamos ascenso ontológico al hecho de hablar de propiedades de objetos como cualidades y cantidades. De este modo sólo tratamos las intensiones que los diferentes objetos ejemplifican y no los objetos mismos.

Tanto Dretske como los neopositivistas se enfrentan al problema de la opacidad de distinta manera. Para los segundos, en el caso de reconocerla, sería algo que quedaría solventado con la fórmula anteriormente señalada que deja el condicional material intacto y añade los aditamentos anteriormente señalados. Sin embargo, ya he señalado que esta solución no resulta convincente, pues no establece claramente en qué consiste una ley científica, sino que sólo señala una serie de rasgos destacados de las mismas. El posicionamiento de Dretske asume peculiarmente la opacidad y elabora una estrategia de ascenso ontológico que permite un tratamiento extensional de la misma. Los enunciados legales no son enunciados acerca de las extensiones de sus términos constituyentes, sino sobre sus intensiones, como pude mostrar con el ejemplo acerca de la imposibilidad de realizar sustituciones *salva veritate* con términos coextensivos. Sin embargo, que esto ocurra así no supone para Dretske un impedimento para la caracterización de las leyes científicas, pues la caracterización que hace de las mismas es extensional. La formulación alternativa que propone, a saber, $F\text{-idad} \rightarrow G\text{-idad}$, es una relación extensional entre las propiedades ejemplificadas por los términos del condicional. Cualquier término que se refiera a la misma cantidad o cualidad como *F-idad* puede ser sustituido por *F-idad* en la anterior fórmula.

Podemos preguntarnos si no sería más fácil introducir conceptos modales en la caracterización del concepto de ley. No obstante no parece que Dretske esté inclinado a este tipo de caracterizaciones. Pues es una tónica en su trabajo el rechazo a cualquier tipo de opacidad. Como se verá en el capítulo V su tratamiento naturalista del contenido se basa en los mismos presupuestos, Dretske intenta naturalizar la intencionalidad en términos puramente informacionales, y, en consecuencia, extensionales. Otro asunto es el éxito de esta estrategia.

El carácter modal o necesitativista de las leyes es, para Dretske, sobrevenido. La necesidad sobreviene a los objetos mismos y se genera a partir de las relaciones existentes entre los mismos. Podemos encontrar ejemplos que ilustran fácilmente esta tesis en la organización de los grupos. Las relaciones nómicas se predicán del grupo y no del

individuo, que, sin embargo las acata. Así, el rector de la universidad no es rector por la posesión de unas características esenciales y radicalmente distintas del resto de catedráticos de la universidad, sino por una serie de relaciones de naturaleza social que le confieren el cargo de rector y que, a su vez y de manera intrínseca, conlleva una serie de obligaciones y derechos. Del mismo modo podemos decir que la necesidad de las leyes naturales no es menos sobrevenida:

"Natural laws may be thought of as a set of relationships that exists between the various "offices" that objects sometimes occupy. Once an object occupies such an office, its activities are constrained by the set of relations connecting that office to other offices and agencies; it must do some things, and it cannot do other things. In both the legal and the natural context the modality at level n is generated by the set of relationships existing between the entities at level $n+1$. Without this set of higher order relationships there is nothing to support the attribution of constraints to the entities at a lower level." Dretske (1977: 265)

Lo que Dretske está tratando de explicar es como surge la normatividad. Como puede verse ésta es intrínseca a las distintas relaciones que se dan entre los objetos que relevantemente relacionamos en el sistema a partir del que queremos establecer leyes. Es discutible que las leyes puedan extrapolarse más allá de los sistemas o modelos a partir de las que se formulan⁹¹. Sin embargo, ahora no voy a explotar esta línea de investigación. En cualquier caso lo importante es que la normatividad o el carácter legal de los enunciados que expresan leyes no es algo extrínseco o que dependa de la perspectiva del espectador, sino que tiene un cierto carácter objetivo que, a su vez, constituye el problema principal de la propuesta de Dretske, a saber, su compromiso con el realismo de universales.

La explicación ofrecida a propósito de la legalidad debería hacernos dudar de que la normatividad constituya una injerencia en el terreno de la Ciencia natural. Tanto normatividad como legalidad se explican del mismo modo. O, de otra manera, la normatividad, al menos en un sentido muy básico, no es una noción menos válida que la noción de ley y tiene pleno derecho a cohabitar entre el resto de conceptos

⁹¹ Cf. Cartwright (1999).

plenamente científicos. No supone ninguna veleidad neoteleológica ni nada similar.

En Biología las regularidades derivan de las leyes de la evolución y en este sentido podemos hablar de normatividad, si bien los fenómenos regulares que exhiben algún tipo de normatividad no deben ser considerados como leyes dada su singularidad y dependencia estricta de las leyes de la evolución biológica. La teoría de la evolución por selección natural claramente tiene un carácter nomológico pese a que como teoría no exhiba el mismo grado de sistematicidad que algunas teorías físicas⁹².

No es necesario buscar una similitud perfecta entre las leyes propias de la Física, que tampoco se presentan de manera monolítica, y las leyes o enunciados de carácter legaliforme que componen la teoría evolutiva. Si se pretende buscar, no se encontrará, ya que el criterio con toda seguridad consistirá en aproximar los enunciados centrales de la teoría de la evolución a unas supuestas leyes ideales de la Física. En su lugar, coincido con Sober (1993) en que existen enunciados con carácter legaliforme y, además, añadiré que satisfacen el requisito de apoyo en contrafácticos⁹³ en el que los neopositivistas

⁹² Ruse (1973) destaca que, si bien la estructura de la teoría de la evolución no es comparable en perfección deductiva a, por ejemplo, la mecánica clásica, sí podemos decir que al menos participa de muchas de las características que exigimos a las teorías propias de las ciencias físicas. Así, en la medida en que la teoría de la evolución según el propio Ruse tiene como núcleo a la genética de poblaciones, teoría que sí está axiomatizada, comparte cierto rigor deductivo con otras disciplinas científicas.

⁹³ Es fácil ver cómo esto es así, si señalamos las tres condiciones que deben satisfacerse para que ocurra evolución por selección natural como (1) que haya variación fenotípica, (2) que ésta suponga una variación en la eficacia del organismo portador y (3) que las características sean heredables, veremos que todas ellas tienen apoyo en contrafácticos. Tomemos como ejemplo las colas de los gatos. No todos los gatos domésticos son iguales, los hay con colas más cortas y más largas, otros como los Manx, originarios de la Isla de Man, carecen de cola debido a un alelo dominante que en caso de presentarse en homocigotos resulta letal y hace inviable el feto. Una cola muy larga puede perjudicar a su portador, del mismo modo que una cola muy corta. La cola, entre otras cosas, es usada por los gatos para mantener el equilibrio y es obvio que este hecho tiene un claro valor adaptativo en la economía global de los gatos, así es posible pensar que ha habido un proceso de selección en el tamaño medio de las colas. La primera condición tiene el siguiente apoyo

tanto insistieron si con ello se quiere señalar un indicio sólido de su nomologicidad. Si bien la teoría evolutiva considerada en su conjunto es mucho más compleja y pueden ser señaladas otras posibles causas de evolución, hay que advertir que muchos de los enunciados que la componen tienen claramente el rango de ley y no son meras generalizaciones. Por otra parte, es obvio que en manos de un filósofo pocas teorías científicas satisfacen nuestros requisitos, ya que tampoco tenemos muy claro qué es una ley científica o qué pueda ser una buena teoría científica. Lo anterior no supone ninguna lacra o motivo de envidia hacia los saberes positivos, sino al contrario, es el estado habitual de cualquier saber vivo.

Volviendo al tema de la normatividad natural, creo que no hay razones para desprenderse de ella. En la medida en que hay normas puede haber leyes en Biología. La nomologicidad de la teoría de la evolución sobreviene a la existencia de normas de ejecución funcional. Incluso históricamente es así, pues con anterioridad a la formulación de la teoría de la evolución los fisiólogos y anatomistas adscribían funciones a los distintos rasgos orgánicos. Así, si la normatividad puede ser puesta en cuestión con igual razón podemos dudar de la nomologicidad de las leyes en Biología.

contrafáctico, si no hubiera distintos tamaños la selección natural no podría actuar sobre las colas. En el segundo caso igualmente podemos pensar que la variación supone una variación en eficacia, no hay demasiados gatos con colas muy largas o muy cortas, si el tamaño de la cola no tuviera efecto alguno sobre las diferentes tasas de supervivencia de sus portadores entonces la selección natural no favorecería el tamaño medio (difícilmente un gato persa sobrevive fuera del estado doméstico). Por último, si la posesión del rasgo no fuera heredable entonces no se explicaría el cambio progresivo en los fenotipos "cola media" de las siguientes generaciones de gatos.

2.3. ¿Por qué es necesaria una teoría normativa de las funciones?

Del mismo modo que los positivistas como Hempel o Nagel rechazaron hacer cualquier tipo de afirmación sobre la existencia extralógica de leyes de la naturaleza podríamos pensar en principio que dicha cautela ontológica puede ser aplicada también a la teoría de las funciones. Desde este punto de vista nuestras adscripciones funcionales sí pueden tener un valor normativo en virtud de su forma lógica que no debería ser más problemático que la diferencia entre una generalización accidental y una genuina ley científica. Sin embargo, al igual que las leyes en el contexto de las discusiones en torno al neopositivismo, debemos garantizar de alguna manera la objetividad de nuestras adscripciones funcionales. Las razones que presenté en la discusión a propósito de los argumentos de Davies en contra de la existencia de normas de ejecución y de malfunciones me parecen suficientes para poder afirmar que hay razones objetivas, al menos más objetivas que la simple estructura lógica de nuestras afirmaciones a propósito de la realidad orgánica, para mantener los conceptos normativos. Tampoco hay por qué pensar que ambos requisitos, estructura lógica y objetividad no sean compatibles. De hecho prefiero pensar que constituyen diferentes apoyos al concepto de normatividad.

2.4. Funciones sistémicas normativizadas

Observar que durante un determinado lapso de tiempo las cosas han ocurrido de una determinada manera debido a una constancia en las condiciones ambientales a las que ha sido sometido el organismo no da pie a pensar que todo deba ocurrir tal como ocurrió. Identificar un rasgo como seleccionado para acometer un determinado rol causal o como ejecutando un determinado rol causal en virtud de otras constricciones ajenas a las fuerzas de selección directas tampoco implica que cualquier ítem deba actuar de esta manera bajo otras condiciones. Al menos no en un sentido tan fuerte de "deba" como el que esperamos de un imperativo moral. Lo único que se afirma es que *caeteris paribus* a partir de cierto momento en la historia reproductiva

del organismo y hasta un determinado momento temporal las cosas han ocurrido de esta manera. Si esa es toda la normatividad a la que podemos aspirar, pues adelante, estaremos trasladando un término moral a un ámbito estrictamente natural, pero es obvio que no es esa la normatividad que uno esperaría encontrar. En este sentido es tan legítimo mantener esta normatividad *sui jure* como negar su existencia y apelar a la perspectiva del observador como, por otra parte, no podía ser menos. Así estaremos expresando en nuestro peculiar lenguaje una relación causal mantenida durante un tiempo más o menos estables bajo una concatenación de circunstancias igualmente estables y diremos que es normativa porque mientras las circunstancias se mantengan constantes dentro del lapso temporal acotado será extraño observar fenómenos diferentes en la ejecución de la función adscrita.

¿Cómo hace frente esta propuesta a los contraejemplos que a menudo traen a colación los partidarios de la perspectiva histórico-etiológica como la distinción entre lo propio y lo accidental? Que la función del corazón, el ítem favorito de muchos filósofos, consiste en bombear la sangre y no en producir sonidos rítmicos y acompasados se justifica en virtud de que el organismo es una estructura organizada en la que la contribución de producir sonidos rítmicos y acompasados no representa ningún beneficio intrínseco a su portador. El reconocimiento de que existe una organización en los organismos vivos es una exigencia de la razón, de lo contrario cualquier posibilidad de discurso mínimamente científico estaría condenada al fracaso y, por si fuera poco, observamos que dicha estrategia basada en admitir como legítima una ontología que considera como substancias o entidades a los organismos da resultado. Así coincido con Davies en que es necesario asumir que las funciones tal como las describe Cummins sólo pueden resultar exitosas como teoría explicativa dentro del marco de la Biología admitiendo la existencia de organismos como estructuras jerárquicas y organizadas. Un organismo es un particular ontológico, una categoría propia dentro de nuestra descripción del mundo.

Gran parte del problema de las normas naturales ha radicado en considerar únicamente al ítem en cuestión y su propagación en las generaciones futuras dejando de lado el poder regimentador del medio. Walsh y Ariew (1996) han reclamado una mayor atención para el medio en la determinación de la funcionalidad del ítem. Desde mi punto de

vista la aportación de ambos es muy pertinente y creo que puede servir, tal como espero haber mostrado, como fundamento de la normatividad que esperamos encontrar en cualquier adscripción funcional. Como puede verse la solución que ofrezco es kantiana; admitimos que la normatividad es un *como sí* que nos permite hablar del mundo orgánico, a sabiendas que en él no hay más procesos que aquellos que son descritos mecánicamente. Así, hay razones suficientes para hablar de normas naturales sin necesidad de postulados vitalistas o teleológicos substantivos. Éstas únicamente describen lo que de hecho ocurre dado un determinado regimen ambiental.

IV. Una Caracterización del contenido mental

CONTENIDOS

Introducción

I. Esbozo histórico del problema: La caracterización tradicional del contenido mental

II. Presentación sistemática: Una presentación sistemática del contenido y sus dimensiones

1. Contenido mental, contenido lingüístico y lenguaje del pensamiento

1.1. La caracterización de nuestras actitudes

1.2. La teoría representacional de la mente y el lenguaje del pensamiento

1.3. Pensamiento y lenguaje

2. La individuación del contenido mental

2.1. La teoría ideacional del significado y la caracterización internista tradicional del contenido

2.2. Los argumentos externistas

2.3. Internismo y externismo a debate

3. Normatividad

4. La caracterización informacional del contenido y la dicotomía entre lo conceptual y lo no conceptual

5. Sistemática, composicionalidad y holismo semántico

6. Recapitulación

Introducción

La concepción intuitiva o de sentido común que tenemos de nuestra mente es desde la conciencia. La conciencia, o la reflexión sobre nuestros estados conscientes, es la vía de acceso al territorio de lo mental más inmediata.

Si tuviéramos que trazar un mapa o una caracterización de lo mental podríamos distinguir principalmente dos tipos de estados. Por una parte señalaríamos los estados intencionales, mientras que por la otra haríamos referencia a los estados fenomenológicos. Ambos constituyen contenidos de conciencia, si bien de modo muy distinto. Los primeros se caracterizan por tener un contenido determinado ante el que un sujeto adopta una cierta actitud. Mientras que los segundos destacan el carácter cualitativo o fenomenológico de la experiencia que, en principio, no es reducible a contenido alguno y es eminentemente subjetivo. Este capítulo se centra únicamente en el estudio y caracterización del primer tipo al que denominaré contenido mental (intencional) y que a menudo se ejemplifica mediante las actitudes proposicionales. Este contenido intencional es independiente de los estados conscientes: nuestro acceso epistémico al mismo es desde la reflexión consciente, pero también puede ser adscrito con mejor o peor fortuna a organismos carentes de conciencia reflexiva. A este contenido le atribuimos comúnmente propiedades como la posibilidad de ser semánticamente evaluado o el ser causa de nuestras acciones⁹⁴. El primer caso es estudiado principalmente por la filosofía de la mente y es sobre lo que versan las denominadas teorías del contenido, mientras que el segundo, si bien es obvio que no puede separarse de la filosofía de la mente, es principalmente analizado por la teoría de la acción. De este modo mi objeto de estudio va a ser la semánticidad de las actitudes proposicionales: dicho de otra manera, la cuestión de la inexistencia intencional que poseen las representaciones

⁹⁴ Cf. Fodor (1987:10). Fodor señala que las actitudes proposicionales están formuladas desde la psicología popular. Si no he considerado antes esta posibilidad es porque estoy tratando de ofrecer una caracterización muy general del contenido y considero que el tema de la psicología popular constituye un tópico específico.

y su relación con lo representado. A este importante tópico dedico este capítulo y los dos siguientes.

I. La caracterización tradicional del contenido mental

No es mi pretensión elaborar en estas páginas una *Historia del Concepto de Contenido Mental*, si bien es cierto que ésta, muy a menudo confundida en los manuales de teoría del conocimiento con la propia disciplina epistemológica, reviste de gran interés. Cualquier reflexión clásica sobre los problemas del conocimiento ha generado en mayor o menor medida una teoría sobre el formato o modo de presentación de ese conocimiento. La razón es obvia, pues es necesario un mediador entre el mundo conocido y el sujeto cognoscente. El mediador y el papel relevante que éste pueda tener en el proceso cognoscitivo ha ocupado buena parte de las páginas de los filósofos desde los mismos inicios de la Modernidad.

A mi juicio la reflexión epistemológica se circunscribe en gran medida a una teoría de la justificación epistémica, es decir, a una teoría que permita dirimir entre los casos de genuino conocimiento y los espúrios en línea con la definición de conocimiento analizada por Gettier (1963). Sin embargo, en muchos momentos de la Modernidad y, desgraciadamente, en numerosas exposiciones sistemáticas de la propia disciplina, se han confundido inconscientemente Epistemología y Filosofía de la Mente. Los problemas que genera el contenido mental afectan a varias esferas de la reflexión filosófica moderna y contemporánea como puedan ser la cuestión general sobre la intencionalidad sobre la que se abundó de manera sistemática a finales del XIX y en la primera mitad del siglo XX por parte de la Fenomenología y sus derivaciones existencialistas o, más concretamente y ceñido a una perspectiva eminentemente analítica, los problemas acerca de la determinación o fijación del contenido, los tipos de contenido, la normatividad del mismo, su poder causal o la congruencia de las caracterizaciones de la intencionalidad con nuestra visión científica -naturalización del contenido-.

En cualquier caso con la idea de un *tertium quid* que constituye el medio o formato del contenido y que según el caso puede ser más o menos relevante desde un punto de vista epistemológico, se han

iniciado los diferentes aspectos de la reflexión epistemológica en la Modernidad. Es, pues, a partir de este momento donde hay que situar la reflexión sobre el contenido mental. En la Modernidad tanto el problema del conocimiento como la cuestión sobre el propio medio del conocimiento cobra un importante papel que hace que frente a posicionamientos de carácter más metafísico se acentúen cada vez más los planteamientos de corte epistemológico. Las diversas teorías sobre la ideación y de manera más concreta, sobre el concepto de representación, constituyen tanto el núcleo embrionario de la epistemología como de la Filosofía de la Mente modernas.

Aristóteles consideraba que el conocimiento procedía de los sentidos y que estos recibían las formas sensibles de los objetos sin su materia, según cada modalidad sensorial específica⁹⁵ y con el concurso del sensorio común. Posteriormente el entendimiento agente⁹⁶ procedía de alguna manera sobre estas formas sensibles para llegar a un conocimiento más abstracto y general, puesto que sólo de lo general hay ciencia⁹⁷. En cualquier caso nos encontramos con un proceso de abstracción en que a partir de la percepción de particulares se llega al conocimiento que, por definición, debe tener un carácter general. Así, los contenidos mentales del entendimiento son para Aristóteles ideas⁹⁸ que tienen su origen en la actividad del alma sensitiva. En Aristóteles encontramos algunos problemas muy relevantes acerca del contenido mental que han sido abordados prolíficamente por algunos autores contemporáneos como el carácter general de los conceptos y el problema del error. El primero de ellos ha sido subrayado en nuestros días por G. Evans (1982) a través de la restricción de generalidad de la que trataré en la sección 4 de este mismo capítulo. El segundo ha sido el problema del error que ha

⁹⁵ "El sentido es la facultad de recibir las formas sensibles sin la materia al modo en que la cera recibe la marca del anillo sin el hierro ni el oro: y es que recibe la marca de oro o de bronce pero no en tanto que es de oro o de bronce." *De Anima* II, 12.

⁹⁶ Cf. Aristóteles, *De Anima* III, 5.

⁹⁷ Cf. Aristóteles, *De Anima*, II, 5.

⁹⁸ "Por lo tanto, dicen bien los que dicen que el alma es el lugar de las formas, si exceptuamos que no lo es toda ella, sino sólo la intelectiva y que no es las formas en acto, sino en potencia." Aristóteles, *De Anima* III, 4.

preocupado a diversos autores como Fodor, Dretske o a los partidarios del enfoque teleosemántico que se exponen en el siguiente capítulo.

En la Modernidad, los empiristas clásicos retoman la doctrina de la abstracción en un formato representacional. El caso de Hume es paradigmático. Para este autor las ideas se forman a partir de impresiones. Los contenidos mentales o son impresiones o son ideas, pero en cualquier caso son datos de experiencia y no ideas innatas. Lo que distingue a una impresión de una idea es su fuerza, así las ideas tienen una existencia dependiente de las impresiones. A partir de la generalización de dos o más impresiones semejantes obtenemos conceptos. De este modo, la capacidad básica de la mente humana es la de detectar semejanzas. La complejidad de los contenidos conceptualizados se hace depender los distintos tipos de posibles combinaciones que acaecen en la mente humana. A mi juicio el empirismo constituye una buena filosofía de la mente por cuanto trata de describir el mobiliario mental y se centra en la cuestión del origen de las ideas. Desde el punto de vista epistemológico no constituye una buena teoría de la justificación epistémica pues no ofrece los recursos necesarios y suficientes para distinguir el conocimiento genuino de la mera opinión. Con el racionalismo sucede lo contrario, es una teoría de la justificación relativamente sólida y constituye el paradigma del fundacionalismo. No obstante, como teoría acerca del contenido mental es pobre, explica qué cuenta como contenido mental pero no explica qué relación tiene con la realidad que representa. Ambas, no obstante, se caracterizan por ser teorías representacionistas que admiten que los datos del conocimiento no se presentan de manera bruta o desnuda, sino bajo cierta apariencia que puede ser más o menos relevante y más o menos confiable. El concepto de una representación dotada de cierto contenido es pues fundamental en el debate epistemológico de la Modernidad más allá de las divergencias que puedan existir entre empirismo y racionalismo.

La tradición analítica describe los pensamientos en términos de las denominadas actitudes proposicionales, en eso consistiría una primera aproximación de sentido común al estudio del pensamiento. Los verbos actitudinales como creer, saber, desear, pensar, y todos aquellos que de un modo u otro refieren al pensamiento van seguidos de un complemento. Quien piensa, piensa algo, quien cree, igualmente cree

algo y así. Lo creído, lo pensado o lo deseado son los contenidos de nuestros pensamientos y deseos, es decir, el contenido de nuestras actitudes hacia el que tenemos una cierta "actitud" psicológica. Una actitud proposicional nos indicaría, pues, el estado psicológico en el que un determinado sujeto se encontraría respecto de un cierto contenido. Este estado estaría indicado por el verbo actitudinal que exhibiría el tipo de relación concreta que se da entre el sujeto y el contenido. Así S(ujeto) puede creer/ desear/ saber/ dudar/ que p.

Desde el punto de vista clásico hay dos maneras de entender el contenido proposicional de nuestras actitudes. Así podemos entender las proposiciones en un sentido fregeano o russelliano. La noción fregeana de "pensamiento" es un claro antecesor de la idea más contemporánea de contenido⁹⁹. Para Frege (1879) una proposición expresa un sentido o pensamiento. La caracterización del pensamiento se dirime en la dicotomía existente entre sentido y referencia. Para Frege no sólo tienen sentido y referencia los nombres propios, sino que también las oraciones. Una oración expresa un pensamiento y el pensamiento expresado por la oración constituye el sentido de la misma -es decir, es un constituyente proposicional- y es, por lo tanto, el responsable de fijar los valores de verdad o la referencia de esa oración. Así, la única vía de acceso a un análisis de los pensamientos es a través del lenguaje que es la explicitación objetiva y pública del pensamiento. Por esa razón los problemas clásicos de la filosofía de la mente revierten, a partir de Frege, en un estudio sistemático sobre las expresiones lingüísticas y, de modo más específico, en el estudio de las actitudes proposicionales. Conceptualizar el pensamiento en los términos anteriormente descritos tiene muchas ventajas a la hora de establecer distinciones. Russell rechaza la idea fregeana de sentido, para Russell el significado de una expresión lingüística es lo designado por ella y no un sentido; el significado se agota en la referencia. Así, las proposiciones son un complejo estructurado cuyos constituyentes proposicionales son los individuos y propiedades designados hacia los que el sujeto manifiesta una determinada actitud.

La caracterización en términos de actitudes proposicionales puede ser vista como un refinamiento de la tradición filosófica que señalaba

⁹⁹ Cf. Farkas (2008:159).

que el acto del pensar se daba en el juicio, es decir, pensar consistía en juzgar. La filosofía kantiana es un buen ejemplo de este tipo de conceptualizaciones. La tabla de los juicios en la *Crítica de la razón pura* (KrV) muestra qué tipos de operaciones formales pueden ser hechas sobre el contenido aportado por la sensibilidad. Pero, además, y en esto se aproxima a la caracterización esbozada anteriormente de las actitudes proposicionales, señala Kant la necesidad de la apercepción trascendental subyacente a todo juicio, es decir, si el juicio es un ejemplar de un tipo de pensamiento es obvio que ese pensamiento es instanciado en un sujeto.

Tanto la caracterización del pensamiento tradicional en términos de juicios como la versión refinada que se ofrece a través de las actitudes proposicionales implican hablar de contenidos intencionales. Si bien pueden hacerse muchas más precisiones y puede discutirse la idoneidad de este tipo de caracterización clásica, lo cierto es que es la manera más tradicional y más próxima a la psicología del sentido común que puede encontrarse, por eso creo que constituye un buen punto de partida aunque más adelante discuta su idoneidad.

A mi juicio el estudio de las actitudes proposicionales ha seguido dos vías en la tradición analítica. Por una parte no han faltado aquellos estudios que dentro del más puro estilo analítico y en continuidad con la filosofía lingüística del lenguaje se han centrado en lo que podríamos ver como una caracterización lógica del contenido. Por otra parte, creo que podemos distinguir un ámbito de reflexión más centrado en lo que Kenny (1989) denominaría la metafísica de la mente. Aquí podemos encontrar teorías muy generales que tratan de explicar la relación entre lo físico y lo mental como los diferentes monismos, dualismos o epifenomenalismos. Aunque también desde esta perspectiva lo que ha importado ha sido caracterizar el vehículo de las actitudes proposicionales proponiendo diversas teorías como los diversos tipos de funcionalismo, la autonomía de lo mental o las aproximaciones neurocomputacionales a las actitudes proposicionales, posturas que dominan prácticamente el debate actual en Filosofía de la mente.

II. Una caracterización contemporánea del contenido mental y sus dimensiones

Tras la caracterización informal o intuitiva con la que se iniciaba este capítulo parece obligado decir algo más concreto acerca del contenido mental tal como se entiende en la literatura especializada actual. Si bien, en lo básico no creo que difiera mucho de lo esbozado desde un punto de vista histórico.

El contenido tiene que ver con nuestras representaciones acerca de los diferentes estados de cosas; es lo que la tradición moderna ha denominado "ideas". Aunque el problema actual no tiene tanto que ver con las ideas como con el alcance del término "representación" que es sobre el que se ha configurado gran parte de la contemporánea psicología cognitiva.

Con las precisiones históricas esbozadas anteriormente poco camino puede recorrerse. Bien dan por supuesto ciertos aspectos del contenido que resultan manifiestamente obviados, bien no fueron tematizados en su momento. Entre estos aspectos creo imprescindible apuntar una serie de cuestiones que actualmente ninguna teoría del contenido puede pasar por alto.

Una teoría del contenido mental tendría que dar cuenta de diversos aspectos. He aquí un conjunto de seis desiderata para una teoría del contenido:

(I) En primer lugar una teoría del contenido debe dar razón de su carácter selectivo. Es decir, el contenido mental de nuestras actitudes no parece reducible a la clase de referencia sobre la que en principio cabe suponer que versa. De ahí que no sea fácil caracterizarlo en términos meramente causales o informacionales, ni que cualquier otro tipo de recurso naturalista constituya un recurso idóneo para su caracterización. El contenido versa sobre el mundo pero es bastante independiente respecto de él. Así, fijando la referencia no fijamos el contenido.

(II) En segundo lugar sería deseable que una teoría del contenido explicara las contribuciones que hacen a éste tanto el entorno, como la historia o el contexto socio lingüístico. Es decir, una teoría del

contenido debe decantarse bien por un planteamiento internista bien por uno externista.

(III) El contenido es fuertemente normativo; debe poderse tener de modo erróneo. Una teoría del contenido debe por lo tanto explicar satisfactoriamente la normatividad que exhibe el mismo y localizarla en algún lugar.

(IV) El contenido involucra aspectos tanto conceptuales como no conceptuales o cualitativos. Una teoría del contenido debería explicar qué papel epistemológico desempeñan los aspectos cualitativos de la posesión de estados intencionales.

(V) El contenido mental no es lo mismo que el contenido lingüístico, pero de alguna forma ambos tipos de contenido están interrelacionados. Hablar no es lo mismo que pensar, pero ambos fenómenos están intrínsecamente relacionados. Así, del mismo modo que observamos propiedades sistemáticas y holísticas en las distintas lenguas naturales también el contenido exhibe tales propiedades.

(VI) El contenido mental tiene una dimensión práctica; es decir, no sólo tiene funciones representacionales sino que apelamos a él para explicar la conducta de otros agentes. La dimensión práctica del contenido implica una articulación racional del mismo en la explicación de la conducta. Porque tenemos actitudes podemos atribuir a los demás otras actitudes y deseos para justificar racionalmente por qué hacen lo que hacen: explicar la conducta de alguien supone atribuirle las actitudes y deseos que la desencadenan.

En las páginas siguientes desarrollaré cinco de los anteriores seis *desiderata*, excluyendo la dimensión práctica del contenido, pues pertenece por completo a la teoría de la acción. Si bien asumo que el contenido mental es causa de la acción y que para la explicación de determinadas conductas es necesario postular la existencia de estados mentales representacionales que tienen una función directiva. Así parto de presupuestos mentalistas. Este procedimiento está en consonancia con la investigación contemporánea en psicología cognitiva y es el habitual.

En primer lugar me centraré en la cuestión de la caracterización de nuestras actitudes. Tradicionalmente se ha admitido que el pensamiento tiene una estructura lingüística. No obstante esta

caracterización, si bien es esencialmente correcta desde el punto de vista de la tradición filosófica, no hace justicia a la actividad cognitiva de organismos alingüísticos y se enfrenta a diversas dificultades relacionadas con la confusión de niveles explicativos en la filosofía de la mente y en la ciencia cognitiva.

En segundo lugar abordaré la cuestión de la individuación del contenido, es decir, el segundo de los *desiderata* mencionados más arriba.

En tercer lugar una buena teoría del contenido mental es una teoría capaz de explicar la normatividad de aquel. Por lo tanto no sólo esperamos que identifique la base sobreviniente o que se pronuncie acerca de si lo debemos abordar desde una perspectiva internista o externista, sino que también esperamos que de cuenta de por qué nuestros contenidos mentales intencionales son contenidos acerca de un determinado tipo de objetos y no de otros, o por qué permiten discriminar adecuadamente entre la diversidad de objetos intencionales que son causa de estos. Así, es éste un punto importante no sólo para una caracterización general del contenido mental, sino de manera más específica para el conjunto de teorías que abordo en este trabajo, ya que con frecuencia las críticas han procedido de la incapacidad de la teleosemántica para dar cuenta de la normatividad a través de los casos de disyunción.

A continuación caracterizo la distinción entre contenido conceptual y contenido no conceptual. Mi interés por dicha dicotomía viene motivado por la influencia que ha tenido la idea no menos poderosa de que pensar consistía en conceptualizar; hablar tanto de pensamientos como de juicios implica en cierto modo hacer referencia a los elementos más básicos del pensamiento, los conceptos. El núcleo del debate radica en la idoneidad de una caracterización exclusivamente conceptual de nuestra vida mental, tal como se viene insistiendo desde la tradición kantiana. No obstante, junto a esta idea principal se discuten una serie de tópicos de relevancia manifiesta para un estudio sobre el contenido mental. Entre estos cabe señalar la naturaleza de los conceptos, las condiciones de posesión de los mismos o la relación, desde presupuestos informacionales como los de G. Evans, con lo no conceptual y con el resto de nuestra vida mental. Finalmente esbozo las cuestiones relativas a la sistematicidad, la composicionalidad y el

holismo como rasgos que exhiben los contenidos mentales y que deberían tenerse en cuenta en una caracterización completa del tópico que investigo.

1. Contenido mental, contenido lingüístico y lenguaje del pensamiento

1.1. La caracterización de nuestras actitudes

Tradicionalmente el estudio de los contenidos mentales apenas se ha dissociado del estudio de las oraciones que los expresan. De hecho, no tenemos otra vía de acceso a los contenidos mentales de una persona normal y adulta que no sea a través del lenguaje. Desde Aristóteles¹⁰⁰ la tradición ha entendido el lenguaje como expresión del pensamiento y, aunque esta idea así formulada es muy general, persiste con substanciales modificaciones en la filosofía contemporánea. En este sentido el caso de Frege es muy ilustrativo, pues emplea el análisis del lenguaje como un medio de análisis de los contenidos mentales. De hecho, el proyecto fregeano bien puede recibir el nombre de una de sus obras *Begriffsschrift*, una conceptografía. Frege emplea el término *Gedanken* o pensamiento en el sentido de contenido mental de naturaleza puramente conceptual. Un pensamiento es una entidad abstracta -no psicológica- que consiste en lo pensado o el contenido de pensamiento que es comunicado a través del lenguaje y que refiere a alguna realidad. En este último sentido los pensamientos constituyen los sentidos de los enunciados, mientras que los objetos intencionales a los que apunta constituyen su referencia o también valor de verdad.

Del mismo modo buena parte de la filosofía analítica del siglo XX no duda en relacionar los contenidos mentales con el lenguaje llegando más allá de las tesis clásicas y afirmando que el propio pensamiento tiene una estructura lingüística tal como hace Sellars (1956:§47). Sellars al señalar que los episodios mentales son episodios lingüísticos afirma que el contenido mental es lingüístico. Aunque no puede decirse

¹⁰⁰ Cf. Aristóteles, *De Interp.* 16a.

que este posicionamiento sea el único a partir de la segunda mitad del siglo XX. Otros autores como Grice no mantienen una postura tan extremadamente lingüística a propósito de la estructura del contenido. Desde el punto de vista de este autor la frontera entre el contenido lingüístico y el contenido mental parece desdibujarse, pues el análisis del primero depende de la existencia del segundo; el significado de lo dicho depende de las intenciones de un determinado hablante y éstas, a diferencia de lo que ocurre en otras semánticas no mentalistas de corte conductista y positivista, tienen un papel decisivo en el estudio del significado lingüístico; significar es querer decir tal como sugiere cierta interpretación del término inglés *meaning*.

Como puede observarse disponemos de muy diversos posicionamientos filosóficos en torno a las relaciones entre pensamiento y lenguaje que, además, se han ido enriqueciendo con las aportaciones procedentes de la psicología. Los posicionamientos filosóficos van desde planteamientos que asumen que el lenguaje es únicamente expresión del pensamiento hasta la idea más sofisticada de que el pensamiento tiene de alguna manera una estructura lingüística.

La Filosofía de la mente contemporánea en gran medida ha seguido la idea clásica de los filósofos analíticos que categorizaban el contenido en términos actitudinales y asumían que éste tenía una estructura claramente lingüística. Estudiar el contenido se convertía así en el estudio de las actitudes proposicionales y, ya desde presupuestos más substantivos, se admitía que éstas eran eficazmente causales y, por lo tanto, reales. Es decir, el recurso a las actitudes como explicación del contenido no era en principio ninguna estrategia instrumental para explicar la conducta o atribuir contenidos que se pusiera en funcionamiento a falta de una mejor ciencia natural, sino que sostenía que las actitudes tienen plena realidad. A esta concepción que defiende que las actitudes tienen una existencia genuina y que se hallan involucradas causalmente en la génesis de la conducta intencional la denominamos realismo intencional. Aunque otras posturas también son concebibles como puedan ser las defendidas por los partidarios del antirrealismo intencional. Estos rechazan la caracterización del contenido en términos de actitudes proposicionales por considerar que al estar basada en la psicología popular es una conceptualización

principalmente errónea¹⁰¹. Los partidarios del antirrealismo intencional pretenden ofrecer una visión "científica" del contenido y para ello eliminan el vocabulario intencional, dado que entienden que éste está fuera de lugar en nuestra imagen científica. Sin embargo, el recurso a la psicología popular no implica ausencia de científicidad, al menos así parecen entenderlo la gran mayoría de los autores partidarios del realismo intencional que, en mayor o menor medida, tratan de naturalizar sus teorías sin renunciar a la jerga actitudinal. Dentro del realismo intencional podemos encontrar una gran variedad de posicionamientos. Aunque todos ellos aceptan que los estados mentales son estados con un contenido, pueden discrepar a propósito de la manera adecuada de caracterizar el contenido. Así mientras unos conceptualizan las actitudes proposicionales en términos representacionales, otros son partidarios de caracterizarlas como disposiciones globales. El debate se enriquece además con diversos posicionamientos en torno a la individuación del contenido o a la caracterización de otras propiedades semánticas relevantes tal como más adelante mostraré. En lo que sigue me centraré únicamente en el realismo intencional y, de modo más específico, en aquellos autores que defienden de alguna manera la teoría representacional de la mente, dado que estos posicionamientos establecen las líneas generales dentro de las que cabe situar a la teleosemántica, objeto de investigación de este trabajo.

1.2. La Teoría Representacional de la Mente y el lenguaje del pensamiento

Tal como mostré en la primera parte de este capítulo, la idea de representación es un tópico común en filosofía al menos desde la Modernidad. Una representación es, en principio, una presentación en formato mental de una realidad externa al sujeto y constituye, por lo tanto, grosso modo, el contenido mental. Tener un contenido mental determinado pasa por tener ciertos tipos de representaciones. Lo menos que podemos decir de las representaciones mentales es que

¹⁰¹ En esta posición encontraríamos las teorías de materialismo eliminativo de los Churchland, los trabajos de Stich o los posicionamientos que defiende Dennet.

son entidades intencionales, es decir, los estados mentales están asociados a objetos intencionales que constituyen las condiciones de verdad de las representaciones. No obstante, el concepto de representación dista mucho de ser un concepto claro. Más bien todo lo contrario, es un concepto que ha sido dotado de una pluralidad de sentidos. Como he señalado anteriormente, Aristóteles propuso la idea de forma desmaterializada que se hacía presente a los sentidos y el empirismo británico incidió en la relación de semejanza que se daba entre los representado y la representación. Con esta relación deberíamos supuestamente explicar la semántica de nuestras representaciones, es decir, en virtud de qué las representaciones representan. En cualquier caso, dicha relación no puede ser diádica, sino triádica. Algo representa a otra cosa en virtud de un criterio de representación. Pero explicar en qué consiste el criterio de representación no explica por qué la representación representa. Esta estrategia corre el riesgo de ser acusada de emplear el argumento del tercer hombre y, por tanto, de ser falaz. Wittgenstein (1953: §454) argumenta en contra de la relación de similitud para explicar el significado de las representaciones mentales; ninguna imagen es capaz de estar dotada de significado en virtud de sus propiedades intrínsecas de semejanza natural porque las propiedades intrínsecas de las imágenes son claramente insuficientes en la determinación de las propiedades semánticas que pudieran poseer. Una imagen por sí misma no significa nada¹⁰². El argumento de Wittgenstein en contra de las imágenes mentales como determinantes de las propiedades semánticas de las representaciones supuso la quiebra definitiva de las teorías basadas en la similitud.

Fodor (1990: 33) señala de una manera más sistemática que las teorías de la semejanza se encuentran con dos problemas. En primer lugar, la relación de representación no es una relación simétrica mientras que la relación de semejanza sí lo es. En segundo lugar, aunque admitamos que la relación de semejanza es una relación

¹⁰² Poseer una determinada propiedad semántica depende para Wittgenstein del seguimiento de reglas. No obstante en los §§ 85 y 86 de las *Investigaciones Filosóficas* presenta argumentos en contra de que podamos dar una justificación racional de por qué seguimos una regla, simplemente la seguimos; es una práctica.

triádica y no diádica, esto no explica la generalización que manifiestan nuestros conceptos. Así, mi concepto GATO es aplicable a todos los gatos y no sólo a Leo, Atina y Ciro, a pesar de que las presentaciones de gatos sean limitadas a lo largo de mi vida. Es decir, puedo aplicar un concepto general a individuos particulares, algo que no podría hacer si el concepto GATO procediera por semejanza con las presentaciones particulares y limitadas de los gatos, luego la semejanza es insuficiente. Una manera alternativa a la de conceptualizar las representaciones en términos de semejanza es recurrir al concepto de causalidad tal como hacen el propio Fodor (1990) y Dretske (1981) o Stampe (1977). Sin embargo, sobre esta teoría incidiré en el siguiente capítulo. Fodor (1987:145) considera que las teorías que emplean el concepto de representación adscriben a estas la función de especificar el contenido mental de los diferentes estados intencionales o actitudes y, por otra parte, constituyen el dominio sobre el que los procesos cognitivos, en tanto que caracterizados funcionalmente, se definen.

La teoría representacional de la mente (TRM) tal como la entendemos hoy en día es una familia de teorías mucho más sofisticada que aquellas teorías clásicas que simplemente proponen que los contenidos mentales son representaciones. La TRM afirma en líneas generales que los estados mentales son estados representacionales, ahora bien, hay muchas maneras posibles de entender en qué consiste esa misma representacionalidad. Fodor (1998), uno de los autores más significativos dentro de esta teoría, describe lo que él entiende como TRM a partir de las cinco siguientes tesis substantivas: (1) La explicación psicológica es típicamente nomológica, es decir, apela a leyes causales que expresan relaciones entre estados mentales descritos intencionalmente. (2) Los portadores del contenido intencional son las representaciones mentales. (3) El pensamiento consiste en computar símbolos-representaciones mentales. (4) El significado es información (5) El contenido está individualizado no relacionamente, es decir, está en la cabeza. Lo presentado anteriormente constituye la TRM para Fodor, no obstante, dentro de la misma hay divergencias. Es decir, no todos los autores partidarios del representacionalismo estarían de acuerdo con las cinco tesis substantivas de Fodor. Las matizaciones que dentro de la TRM introducen cierta pluralidad de pareceres tienen que ver con la

individuación del contenido o la polémica a propósito de su carácter holístico o atomista que ya abordaré más adelante. Un aspecto importante de la TRM es que está comprometida en mayor o menor medida, según los autores, con la psicología popular. La TRM apela a creencias y deseos, ambos tipos de estados actitudinales están por estados de cosas o necesidades/aspiraciones de los individuos que las albergan. Creencias y deseos se manifiestan en forma de actitudes proposicionales y gracias a ellas explicamos la conducta de otros organismos al atribuirles un contenido intencional. Por otra parte, este compromiso con la psicología popular no es visto como un menoscabo de la cientificidad de las distintas teorías¹⁰³, sino como algo irrenunciable para explicar nomológicamente la conducta; la psicología popular es indispensable. No podemos explicar nomológicamente mi deseo de defender una tesis apelando a los niveles de serotonina presentes en mi sistema nervioso, sólo podemos apelar a otras razones, no a eventos físicos. Lo que no quiere decir que las razones no puedan naturalizarse de algún modo o hacerse compatibles con nuestra imagen científica sin necesidad de eliminarlas.

La existencia de representaciones mentales no deja de ser un tópico comprometido con el realismo intencional. Existen varias razones para pensar que el recurso a las representaciones mentales no es una simple fábula. Como ya he señalado, la tesis (3) afirma que el pensamiento tiene que ver con la computación o con la manipulación de los símbolos mentales. Algunas interpretaciones de la TRM como las de Fodor (1975, 1987) sostienen que hay una sintaxis de las representaciones que alberga una mente y que se configura en un lenguaje del pensamiento (LOT) o *mentalés* distinto de las lenguas naturales. Estos planteamientos enunciativistas entienden que, del mismo modo que podemos establecer una sintaxis para las lenguas naturales en las que definimos un conjunto de reglas de formación sobre un conjunto de símbolos o términos lingüísticos, también podemos entender el pensamiento como un conjunto de reglas que se

¹⁰³ Por ejemplo, Fodor (1987:16) señala que "We have no reason to doubt -indeed, we have substantial reason to believe- that it is possible to have a scientific psychology that vindicates commonsense belief/desire explanation." Para una defensa del realismo intencional Cf. también Acero (2007).

aplican sobre un dominio de símbolos mentales o conceptos¹⁰⁴. Así, sólo los símbolos o representaciones mentales son entidades susceptibles de estar dotadas de una sintaxis¹⁰⁵. De este modo, el contenido mental puede entenderse como frases - fórmulas bien formadas o secuencias gramaticales de símbolos mentales- en el lenguaje del pensamiento. La hipótesis del lenguaje del pensamiento sostiene que los estados mentales típicamente tienen una determinada estructura constituyente¹⁰⁶ que funciona como una sintaxis para un lenguaje dado. El objetivo de dicha hipótesis es explicar cómo pueden nuestros estados mentales representarse el mundo y causar nuestra conducta¹⁰⁷.

La hipótesis del lenguaje del pensamiento ha ejercido un poderoso influjo tanto en la ciencia cognitiva como en la filosofía de la mente. Es particularmente apta para explicar ciertos aspectos de los denominados procesos superiores, aunque también ha generado importantes críticas. Fodor (1987) presenta tres argumentos a favor de la hipótesis del lenguaje del pensamiento. En primer lugar disponemos de un argumento metodológico. Fodor argumenta que si la conducta (incluyendo la conducta verbal) tiene una estructura segmentada, es decir, pueden identificarse constituyentes que son, además, necesarios para su explicación, entonces lo mismo debe ocurrir con las actitudes que supuestamente son causa de la conducta. La etiología de la conducta es intencional y, por ende, también la conducta depende de la estructura de la intencionalidad. Si yo quiero mover mi brazo izquierdo y de hecho lo muevo, también puedo querer levantarme de la silla y hacerlo, y también puedo optar por querer hacer ambas cosas y, de

¹⁰⁴ Para Fodor (1975) los conceptos son innatos.

¹⁰⁵ Fodor (1987:19-20) plantea su existencia como una inferencia a la mejor explicación, "*There must be mental symbols because, in a nutshell, only symbols have syntax, and our best available theory of mental processes -indeed, the only available theory of mental processes that isn't known to be false -needs the picture of the mind as a syntax-driven machine.*"

¹⁰⁶ Cf. Fodor (1987:136).

¹⁰⁷ No obstante, la causación de la conducta no depende de las propiedades intencionales de los contenidos, en este sentido Fodor muestra un claro rechazo a la semántica del rol conceptual. No existe algo así como una intencionalidad irreducible en los contenidos, sino que la intencionalidad que estos exhiben es fruto de la sintaxis.

hecho, realizarlas. Del pensamiento complejo "Querer levantar mi brazo izquierdo y ponerme derecho" que es causa de mi conducta también puedo tener los pensamientos constituyentes que se obtienen por la eliminación de la conjunción. No podemos obviar que el pensamiento tiene una estructura constitutiva si admitimos que las acciones exhiben dicha estructura. Es decir, debe darse de alguna forma una relación isomórfica. Como puede verse el argumento es subsidiario del realismo intencional y procede por inferencia a la mejor explicación. En segundo lugar, Fodor argumenta que diferentes procesos psicológicos como la comprensión lingüística o la percepción sólo pueden ser explicados recurriendo a la idea de que los estados cognitivos tienen una estructura constituyente formulada en mentalés. Por ejemplo, cuando alguien profiere una determinada oración nuestra comprensión de este ejemplar depende de la representación que nos podamos hacer de su estructura constituyente. Así, existe un paralelismo semántico entre pensamientos y enunciados. Lo mismo ocurre con las escenas perceptivas: interpretar una determinada escena pasa por reconocer los principales elementos que la constituyen y que pudieran aparecer en otros contextos. Finalmente, la hipótesis LOT constituye una buena explicación de la productividad y sistematicidad tanto del pensamiento como del lenguaje¹⁰⁸. En suma, si consideramos el pensamiento adulto normal desde el realismo intencional, es decir, el pensamiento como un evento mental discursivo en el que los elementos sobre los que se realizan transformaciones e inferencias son las proposiciones y los conceptos que las constituyen, entonces la hipótesis del lenguaje del pensamiento es una buena candidata a explicar el contenido de nuestras actitudes, ya que estas se caracterizan desde la psicología popular mediante estructuras lingüísticas. Así las cosas no estaríamos defendiendo sino un refinamiento de la teoría clásica del contenido de las actitudes. Además, en tanto que las actitudes se caracterizan como eventos

¹⁰⁸ Cuando en la década de 1970 se planteó la hipótesis LOT supuso un gran avance respecto del conductismo dominante. Hay que recordar que incluso otros movimientos como la psicología genética de Piaget o las propuestas de Vigotski no explican la actividad cognitiva superior tan finamente como lo hace la propuesta de Fodor, si bien sus aportaciones no son nada desdeñables.

lingüísticos, la hipótesis del lenguaje del pensamiento constituiría una buena explicación de la sistematicidad, composicionalidad y productividad del pensamiento del mismo modo que se explican estas propiedades para cualquier lenguaje recurriendo a una sintaxis. Es decir, si admitimos que el pensamiento tiene una determinada estructura constitutiva entonces la hipótesis LOT debe ser verdadera.

No todos los autores que defienden de algún modo el representacionalismo son partidarios de la hipótesis del lenguaje del pensamiento. Se puede ser partidario de la TRM e indiferente o contrario respecto de la hipótesis del lenguaje del pensamiento. La hipótesis LOT es idónea para explicar aspectos como los antedichos. Pero muy bien puede ser un estorbo para explicar otros tipos de procesos cognitivos más básicos como por ejemplo los de aquellos organismos que no están dotados de lenguaje verbal o cuya competencia en la resolución de problemas es más limitada y, sin embargo, poseen de alguna manera macroconceptos o categorías y aplican algunos modelos de racionalidad. Por otra parte, la hipótesis LOT no explicaría aquellos aspectos fenomenológicos de nuestra vida mental, pues claramente son asintácticos -no hay una sintaxis de los *qualia*-. Así, las explicaciones del contenido que recurren a LOT para explicar los aspectos cualitativos y no conceptuales de nuestra vida mental se encuentran con importantes dificultades. Lo no conceptual, por oposición a lo conceptual es asintáctico, no es el tipo de cosas que puedan entrar a formar parte de una proposición y ser por tanto el objeto de las actitudes. La relevancia de lo no conceptual no radica solamente en ser un aspecto más de nuestra vida mental, sino que es un tópico que necesariamente ha de ser explotado en la caracterización de la vida mental de otros organismos tal como mostraré en la sección 4 del presente capítulo.

Desde posicionamientos muy próximos a la investigación en ciencia cognitiva se ha argumentado en contra de la hipótesis LOT tal como Fodor la presenta. Los ataques han procedido principalmente de posturas conexionistas que son la alternativa a la explicación computacional clásica de la mente. Según Clark (1994) LOT tiene los tres siguientes inconvenientes: (1) LOT, tal como lo formula Fodor (1975) se enfrenta a los problemas del innatismo: si los conceptos son innatos como señala Fodor ¿cómo es posible la ampliación de nuestro

bagaje conceptual? (2) Por otra parte, el atomismo representacional que Fodor defiende es implausible desde el punto de vista de la ciencia cognitiva -al menos tal como la entiende Clark-. Es decir, la hipótesis LOT necesita de una determinada arquitectura mental que no es la que de hecho poseen los organismos. Sin embargo, hay que decir que también las arquitecturas conexionistas han recibido importantes críticas. (3) Por último, y quizás el argumento más fuerte, LOT no logra explicar los procesos centrales, es decir, aquellos procesos que no son modulares como los procesos que acaecen en los sistemas de entrada y de salida. Los procesos que tienen lugar en los sistemas centrales exhiben propiedades isotrópicas y quineanas, es decir, tienen una estructura fuertemente holista que se alimenta con información procedente de cualquier sistema encapsulado. Tanto la isotropía como el quineanismo representan un obstáculo en la comprensión de estos sistemas. Clark no niega que LOT pueda tener cierta plausibilidad, sin embargo, tiende a minimizar enormemente la importancia que le atribuye Fodor para la ciencia cognitiva.

Otros autores mantienen las tesis representacionistas a la vez que son neutrales respecto de LOT. Esta postura es la principalmente adoptada por los autores partidarios de la teleosemántica que estudiaré en el siguiente capítulo. Así, autores como Dretske o Millikan no precisan de la hipótesis del lenguaje del pensamiento para explicar las capacidades representacionales de los organismos.

En suma, puede decirse que si bien la TRM cuenta con un amplio apoyo, la hipótesis del LOT es más controvertida. Los argumentos en contra tratan de poner de manifiesto la inadecuación de dicha hipótesis para una caracterización del contenido de muchos organismos. Si bien creo que esencialmente es una buena caracterización de una porción importante del contenido de los humanos adultos normales.

1.3. Pensamiento y lenguaje

Hasta ahora he incidido principalmente sobre la teoría representacional de la mente y el lenguaje del pensamiento. Tal como he mostrado, las relaciones entre pensamiento y lenguaje son muy íntimas, pero no creo que sean idénticas. Respecto de las relaciones entre contenidos mentales y contenidos lingüísticos creo que hay

razones suficientes para afirmar que hay una separación no muy clara entre nuestras actitudes y el lenguaje que usamos; pensar no es lo mismo que hablar con independencia de que el significado de nuestros términos tenga mucho que ver con nuestras intenciones, al menos para los humanos adultos normales. En esta línea ha argumentado precisamente Evans (1982) quien no duda en separar pensamiento y lenguaje.¹⁰⁹

Algunas extrapolaciones del contenido lingüístico al contenido intencional son discutibles. De hecho los argumentos empleados para caracterizar el significado de nuestros términos pueden no ser adecuados para dar cuenta del significado de nuestras representaciones¹¹⁰. En mi opinión, el pensamiento es hasta cierto punto independiente del lenguaje público por lo que respecta a sus propiedades semánticas. Así, puedo usar una palabra sin ser consciente de todo lo que implica esa palabra; mi uso de esa palabra responderá a mis intenciones y no necesariamente a las acepciones más habituales del diccionario de uso del español. Los usos raros o particulares no son sinónimos de lenguajes privados, sino de lenguajes más reducidos.

Otra razón importante para considerar la dicotomía entre contenido lingüístico y contenido mental radica en la posibilidad de adscribir vida mental a seres prelingüísticos y alingüísticos. Este enfoque pretende ofrecer una imagen coherente con planteamientos naturalistas respecto de la vida mental. Bermúdez (2003) ofrece una explicación acerca de cómo es posible atribuir pensamiento a seres alingüísticos sin necesidad de identificar ninguna arquitectura cognitiva en particular, es decir, el objetivo de Bermúdez está al margen de las cuestiones sobre arquitectura cognitiva. Para este autor una teoría que explique el pensamiento no lingüístico debe satisfacer los siguientes cuatro

¹⁰⁹ "One must be able to think in terms of the possession of semantic properties by expressions of the language independently of any particular occasion of use, as well as in terms of significance of particular utterances. There immediately opens up the possibility of a gap between what a speaker means to say by uttering certain words - what he strictly and literally says, according to the conventional meanings of the words he utters, on the other." Evans (1982: 67).

¹¹⁰ Tal es el caso de la argumentación de Crane (1991) en contra de Putnam y Burge.

*desiderata*¹¹¹: (1) Debe identificar el vehículo o la ontología que manifiesta dicha modalidad de pensamiento. (2) Debe dar razón de las propiedades semánticas del pensamiento no lingüístico. (3) También debe proveernos de algún dispositivo metodológico que nos permita identificar su contenido. (4) Debe permitir explicaciones intencionales de la conducta exhibida por los organismos alingüísticos. Según este autor, el pensamiento de seres alingüísticos no puede caracterizarse únicamente en términos de LOT porque al carecer de lenguaje no tenemos modo alguno de acceder a los enunciados en mentalés que determinarían los contenidos de las actitudes de estos organismos. Por lo tanto, si no se posee un lenguaje verbal ¿qué razones hay para inferir que exista un lenguaje del pensamiento con una estructura similar al lenguaje hablado? Es decir, no se satisface el criterio (3). Aunque esto es un problema metodológico relevante no implica que se desestime LOT, de hecho no implica ninguna asunción acerca de LOT. Más bien sólo se afirma que no es una teoría de aplicación máximamente general en la explicación del pensamiento. Además, en seres alingüísticos se nos hace difícil tanto atribuir actitudes, tal como propone Fodor (y la psicología popular), como disponer de algún medio para identificar el mentalés, pues al carecer de lenguaje no podemos inferir la sintaxis de su código interno ya que no se evidencia ni sistematicidad ni composicionalidad. Sin embargo, todo ello no implica que podamos afirmar que los organismos alingüísticos no albergan pensamientos. Es factible pensar que estos organismos se representan el mundo y, en algunos casos, como los córvidos o los chimpancés, resuelven problemas. Bermúdez es partidario de atribuir pensamiento a organismos no lingüísticos para posibilitar algún tipo de explicación psicológica. Para ello es necesario establecer tanto una ontología como un procedimiento metodológico adecuado. Con el establecimiento de una ontología podemos identificar el rango de objetos sobre los que un organismo alingüístico puede pensar. El procedimiento metodológico tiene por objeto establecer las atribuciones de pensamientos a estos organismos. En este último respecto Bermúdez se decanta por recurrir a la semántica del éxito (*success semantics*) cuyos objetivos son más modestos que otras semánticas disponibles, ya que no pretende

¹¹¹ Cf. Bermúdez (2003: 12).

caracterizar el contenido completo de un determinado estado mental. La semántica del éxito nos debe permitir atribuir pensamientos estructurados a una criatura alingüística sin necesidad de postular un lenguaje del pensamiento, es decir, sin necesidad de identificar el vehículo de ese pensamiento. Para este tipo de semánticas el contenido de una creencia viene determinado por la utilidad que implica poseer esa misma creencia, mientras que el contenido de un deseo es su condición de satisfacción. La semántica del éxito puede resumirse afirmando que las creencias verdaderas causan acciones que satisfacen los deseos. Esta semántica no es, según Bermúdez, de aplicación general, sino sólo apta para aquellos organismos cuya economía cognitiva está encaminada a la satisfacción de sus deseos. El deseo aquí no es algo caprichoso, sino algo que tiene que ver directamente con la supervivencia, es decir, deseos entendidos como apetencias básicas. En este sentido, la posesión de creencias útiles es más importante que la posesión de creencias verdaderas. Podemos determinar, al menos parcialmente, el contenido de una creencia si asumimos que ésta, tal como sugiere la semántica del éxito, es causada por un deseo y para ello identificar el estado de cosas que causan la satisfacción del deseo.¹¹² De este modo, identificando las clase de cosas que causan la satisfacción del deseo identificamos al menos parcialmente el contenido de una creencia albergada en la mente de un organismo no lingüístico. El principal problema al que ahora nos enfrentaremos es el de determinar la estructura conceptual del organismo alingüístico, es decir, la cuestión de identificar los conceptos que pudiera emplear tal organismos está afectada por los problemas de la traducción radical y por un importante sesgo del investigador que puede dudar de hasta qué punto no está extrapolando sus propios conceptos a la mente del organismo. Así el problema es en realidad el de identificar la ontología del organismo y esta cuestión está afectada de indeterminación. Un rasgo muy importante de nuestras ontologías es que se centran en la cuestión acerca de qué cuenta como un objeto como aspecto básico de nuestro pensamiento¹¹³. De

¹¹² Cf. Bermúdez (2003: 66).

¹¹³ Tanto Aristóteles como Kant insisten en este importante tópico, la categoría central sobre la que gira tanto su conceptualización de la realidad como en el caso concreto

hecho, si un concepto ha sido importante a lo largo de toda la historia de la Filosofía ése es el de objeto¹¹⁴ o entidad. No obstante esto no tiene por qué ser así para otras especies. Las investigaciones resultantes de la aplicación del paradigma de deshabitación muestran que los organismos alingüísticos son capaces de percibir un mundo estructurado. Según Bermúdez (2003: 82) podemos comprender la ontología de un organismo en la medida en que somos capaces de identificar el conjunto de propiedades físicas que el organismo admite que constituyen el objeto. Pero ¿cómo determinar los constituyentes de las actitudes de los organismos alingüísticos? Bermúdez propone una estrategia basada en la semántica del éxito. En primer lugar hay que determinar cuál es el objetivo de la conducta del animal. En segundo lugar es necesario identificar la condición de utilidad relevante, es decir, qué estados de cosas podría satisfacer su estado motivacional. Para este fin evidentemente también proponemos un conjunto de enunciados candidato a identificar la condición de utilidad. Así, a partir de esta estrategia proyectamos los constituyentes de los enunciados a los contenidos que alberga la mente del organismo y así identificamos sus constituyentes proposicionales. No obstante, con el segundo paso de esta estrategia nos encontramos en una situación de indeterminación en la que podemos identificar varios candidatos a ser los constituyentes proposicionales de la actitud. Es decir, si bien la clase de referencia identificada por la semántica del éxito puede estar relativamente clara, no ocurre lo mismo con el sentido. Por ejemplo, si

de Kant, la caracterización de los conceptos del entendimiento es la de substancia. Más contemporáneamente Strawson ha vuelto a incidir sobre el mismo tópico y desde la perspectiva de la psicología del desarrollo centrada en las explicaciones de la psicogénesis. Por ejemplo Piaget (1937) insiste de manera determinante en el papel que desempeña el desarrollo del concepto de objeto: *"El primer problema que conviene plantearse, para comprender cómo la inteligencia naciente construye el mundo exterior, es saber si durante los primeros meses el niño construye y percibe las cosas, como lo hacemos nosotros mismos, bajo la forma de objetos sustanciales, permanentes y de dimensiones constantes. Suponiendo que no ocurra nada de esto, sería necesario entonces explicar cómo se constituye la noción de objeto."* Piaget (1937: 15 de la versión española).

¹¹⁴ Entenderé que un objeto se determina no sólo por constituirse en un agregado de propiedades que son percibidas como un todo, sino que, además, ese agregado debe tener constancia o permanencia temporal.

ponemos a una rata en una situación experimental como la diseñada por Tolman, un laberinto en forma de cruz en el que ponemos comida en uno de sus brazos, podemos decir que aquello que elicitaba la conducta motora de la rata es la comida. Es decir, el objetivo de localizar la comida cuenta como el objetivo que guía la conducta de la rata. Pero, respecto de la representación que alberga la rata, podemos tener varios candidatos que no se determinan por la mera consideración de la referencia de la conducta exploratoria, es decir, la representación o pensamiento que alberga la rata tiene un sentido y nuestra atribución puede no ser unívoca desde que tenemos varios candidatos a ser el sentido de la representación de la rata como que la comida se localice en unas determinadas coordenadas dentro de un espacio egocéntrico o que ocurra lo mismo pero en el espacio del laberinto entre otras opciones disponibles. Así, si bien tenemos cierta seguridad acerca de qué cuenta como la referencia del pensamiento no verbal de la rata, la cuestión del sentido y, por lo tanto, la de identificar los constituyentes proposicionales y delimitar su estructura permanece indeterminada por la mera aplicación de la semántica del éxito. Podremos determinarla sometiendo a test más rigurosos la situación experimental para finalmente descartar aquellos candidatos a sentido que resultan ser espúrios. Estos pensamientos tendrían componentes distinguibles que podrían especificarse a través de la correspondiente oración actitudinal. Según este autor, aunque no podamos dilucidar claramente cuál es el vehículo del pensamiento de estos organismos, sí podemos atribuirles un contenido estructurado y ofrecer así una explicación intencional de su conducta. De hecho sólo podemos proceder a una explicación psicológica si postulamos que estos organismos poseen pensamientos estructurados. El problema que conlleva admitir la posibilidad de una explicación psicológica de este tipo de organismos alingüísticos es la caracterización de modelos de racionalidad. En estos se hace difícil ejemplificar las reglas de inferencia que pudieran emplear respecto de las caracterizaciones plenamente lingüísticas y sintácticas que empleamos los humanos. Sin embargo Bermúdez señala que la racionalidad no es algo categórico, sino gradativo¹¹⁵. Estos pensamientos alingüísticos estructurados

¹¹⁵ Bermúdez (2003: 109 y ss.) clasifica tres niveles de racionalidad progresivos para

servirían para hacer determinadas inferencias conforme a una protológica. Los límites del pensamiento sin lenguaje son claros. Entre ellos quizás el más importante es que sin un lenguaje público es imposible tener pensamientos de segundo orden, es decir, pensamientos sobre pensamientos. De modo que el pensamiento reflexivo o las capacidades metarrepresentacionales son imposibles para los organismos carentes de lenguaje.

La propuesta de Bermúdez tiene claras ventajas frente a enfoques minimalistas del pensamiento de las criaturas alingüísticas. Una contribución clara es que logra separar el pensamiento del lenguaje de una manera bastante habil y plausible, identificando, además, las limitaciones de dos modos esencialmente distintos de pensar: pensar sin palabras y pensar con ellas. Pero también se enfrenta a dificultades. He aquí dos: (1) Es cierto que en algunos casos la indeterminación persiste; la estrategia de la semántica del éxito permite fijar la referencia del contenido, pero no tan fácilmente su sentido. (2) No está claro a qué animales alingüísticos podemos atribuir estados representacionales y a cuáles no: no es lo mismo un gato que una oruga. En ese punto hace falta una mayor precisión. Si bien creo que la estrategia es adecuada para mamíferos y aves. Aunque en cualquier caso se trata de un dato puramente empírico.

los organismos alingüísticos. El nivel 0 estaría formado por el repertorio de respuestas estereotipadas o que responden al condicionamiento clásico, es decir, aquellas respuestas que ni responden a un procedimiento de decisión entre alternativas ni se aplican a conductas particulares sino que se elicitan automáticamente en virtud de alguna tendencia o disposición. El nivel 1 aunque no implica propiamente un procedimiento de decisión entre diversas alternativas sí conlleva una aplicación a conductas muy concretas, es decir, a ejemplares de un tipo de conducta como sucede por ejemplo cuando un animal se encuentra amenazado y esta misma situación elicit o bien la conducta de huida o la agresividad sin posibilidad de elección alguna. El animal elicit una conducta u otra en virtud de la circunstancia concreta: si está acorralado, peleará, si no, huirá. Finalmente lo que Bermúdez denomina nivel 2 implica tanto la aplicación a instancias particulares de las conductas como procedimientos de decisión. Ejemplos de este último nivel lo constituyen el desarrollo de útiles líticos por parte del *Homo habilis*, homínido carente de lenguaje. La explicación psicológica carecería de sentido en el nivel 0, sería cuestionable en el 1 y perfectamente asumible en el 2.

Como he mostrado hasta ahora el pensamiento no es únicamente habla interiorizada ni puede explicarse en su totalidad apelando a LOT. Las objeciones que pueden presentarse a LOT son poderosas, sin embargo lenguaje y pensamiento van de algún modo ligados. El lenguaje, por ejemplo, es imprescindible para la resolución de según qué tareas cognitivamente complejas tal como evidencian los estudios sobre habla subvocal que disponemos desde Vigotski o Piaget. La conciencia también tiene un soporte lingüístico, soy consciente en la medida en que puedo representarme proposicionalmente mis propios estados. Por lo tanto, la idea original de Sellars no es tan descabellada, pues tenemos acceso consciente a nuestro pensamiento en un formato lingüístico que coincide con nuestra lengua natural. Carruthers ha investigado esta cuestión y sus conclusiones sobre este tópico son bastante razonables. La cuestión que trata de resolver Carruthers (2009) es el papel que desempeña el lenguaje en la cognición humana. Carruthers defiende un posicionamiento mixto en el que por una parte si bien acepta la plausibilidad limitada de LOT, también cree que el mentalés es en realidad el lenguaje público interiorizado. Además, Carruthers circunscribe el uso de LOT únicamente al pensamiento proposicional. Aquellos pensamientos no proposicionales no se ven afectados por LOT, por ejemplo, aquellas tareas que resolvemos manipulando únicamente imágenes mentales. Es decir, las representaciones en formato lingüístico son claves para entender las interacciones entre un buen número de sistemas centrales cuasi modulares¹¹⁶ y a que en buena parte la facultad del lenguaje implica un control consciente de las ejecuciones del propio sujeto. Carruthers se apoya en una inferencia a la mejor explicación para sostener que el lenguaje del pensamiento es lenguaje natural interiorizada. Así Carruthers parte de la consideración del acceso introspectivo a nuestros estados mentales. Cuando nos referimos a ellos o cuando empleamos el habla subvocal como método de control de nuestras acciones e inferencias lo hacemos en nuestra lengua vernácula y no en mentalés. El problema de la hipótesis de Carruthers es que parece confundir lo que es una explicación a nivel subpersonal que tiene que ver tanto con el formato de nuestras representaciones mentales como

¹¹⁶ Cf. Carruthers (1998).

con la arquitectura mental con una explicación a un nivel personal y formulado en términos de psicología popular.

Todo lo anterior no evita que nos sea dado pensar acerca del contenido en términos lingüísticos; la caracterización por excelencia del contenido no puede ser sino lingüística y todas las demás variedades se dicen por analogía a la caracterización lingüística del mismo. No hay otra vía de análisis posible, los casos patológicos o anormales como el niño de Aveyron o las representaciones mentales que puedan tener los sordomudos de nacimiento no son más que eso, casos anormales que no permiten la constitución de ciencia alguna. Sin embargo, los problemas surgen cuando se quiere compatibilizar el realismo intencional con la ciencia cognitiva. Como he mostrado, la relación no es nada fácil, pues a menudo ocurre que ambas perspectivas comparten conceptos y supuestos comunes. Por ejemplo, es frecuente la confusión de distintos niveles de análisis: así lo que ocurre a nivel doxástico de lo que ocurre a nivel subdoxástico, lo que puede acaecer a un nivel personal de lo que acaece a un nivel subpersonal.

2. La individuación del contenido mental

La caracterización o individuación del contenido mental puede hacerse básicamente desde dos perspectivas contrapuestas, bien se es internista (o individualista) respecto de la identidad de los estados mentales, bien externista¹¹⁷. Los partidarios del internismo intencional

¹¹⁷ La polémica entre internismo y externismo es un tópico actual en Filosofía del Lenguaje, Filosofía de la Mente y Epistemología. En cada disciplina adquiere matices distintos aunque sí es fácil identificar una raíz común. Como ya se ha señalado básicamente tiene que ver con el problema de la individuación. En el caso de la Filosofía del Lenguaje estamos tratando de individualizar significados o, de otra manera, nos preguntamos por las condiciones de individualización de los significados de los términos lingüísticos. El alcance de la Filosofía de la Mente es mucho más general, principalmente porque la cuestión de la lingüisticidad no desempeña un papel primordial y porque, en cierto modo y en continuación con ciertas asunciones propias de la tradición, se acepta que el lenguaje es la expresión del pensamiento, además de que se consideran también como objeto de estudios los estados mentales de seres alingüísticos o aquellos estados en los que los conceptos pueden no desempeñar ningún papel relevante (contenido no conceptual). Así, esta polémica

sostienen que el medio externo no desempeña ningún rol en la fijación del contenido de los estados intencionales. Por el contrario, los partidarios del externismo consideran que los estados psicológicos intrínsecos de un individuo son insuficientes para determinar la identidad de sus contenidos mentales y que, en esta tarea deben involucrarse necesariamente elementos externos al propio sujeto¹¹⁸. Cada uno de estos posicionamientos tiene sus propias implicaciones y ambos delimitan los dos polos de un interesante debate en torno a nuestras actitudes.

2.1. La teoría ideacional del significado y la caracterización internista tradicional del contenido mental

Habitualmente se señala que la perspectiva tradicional en la caracterización del contenido intencional es el internismo. Sin embargo, creo que, a la luz de los argumentos externistas no sólo es difícil de sostener sino que incluso el mismo concepto de internismo intencional ha quedado muy oscurecido llegando a parecer contraintuitivo.

Las caracterizaciones internistas o individualistas del contenido mental son conceptualizadas principalmente por dos autores modernos: J. Locke y R. Descartes y son asumidas por autores más recientes dentro de la tradición analítica como Frege o Russell¹¹⁹.

recae, desde la Filosofía de la Mente en la individualización del contenido de los estados mentales.

¹¹⁸ Farkas (2008:73) señala que son varias las maneras posibles de caracterizar el externismo, así (1) El contenido mental depende de hechos externos al propio sujeto. (2) El contenido no sobreviene a los estados internos del sujeto. (3) Un sujeto que alberga determinados pensamientos presupone la existencia o la naturaleza particular de cosas que son externas al propio sujeto. (4) El contenido de un sujeto no está determinado por sus propiedades intrínsecas. (5) Las propiedades de determinados contenidos son relacionales. Cada una de estas maneras de definir el externismo puede tener sus consecuencias a la hora de criticarlo.

¹¹⁹ Evans (1982) caracteriza el posicionamiento de Russell acerca del lenguaje como "*not an intrinsically social phenomenon*." (p.67) y recalca el progresivo abandono de esta perspectiva por parte de la contemporánea filosofía analítica. Las ideas de Russell sobre el lenguaje son extrapolables a su caracterización del contenido, no es

Locke señala la dependencia del lenguaje respecto de los contenidos mentales del sujeto, sin embargo este posicionamiento no es exclusivo de Locke, pues no hace sino explicitar la postura comúnmente asumida por la tradición. El caso de Descartes es más complejo y tiene que ver con la autoridad de primera persona en lo que respecta al acceso epistémico a nuestros contenidos. Así lo que cuenta como contenido para Descartes es aquello que puede ser examinado en el espacio mental con independencia de su procedencia. En efecto, cómo se nos aparecen las cosas es algo determinable inmediatamente por el yo cartesiano y que, además, puede ser considerado al margen de cómo sea de hecho el mundo. Sin embargo, la simple consideración de los posicionamientos de ambos autores no parece tan pernicioso como suponen Putnam o Burge. ¿Qué es lo que convierte al internismo en problemático?

El partidario del internismo intencional defiende que las condiciones de identidad del contenido son exclusivamente subjetivas y ahistóricas, es decir, cómo sea el medio circundante y la historia del sujeto es irrelevante en la determinación del contenido¹²⁰. Por el contrario, aquellos filósofos que defienden el externismo intencional sostienen que entre las condiciones de identidad del contenido no cabe contar únicamente las intrasubjetivas, sino que el medio y la historia juegan un determinado papel. De una manera más precisa el internista está comprometido con al menos dos tesis: en primer lugar sostiene que el contenido intencional viene determinado exclusivamente por factores internos al individuo. En segundo lugar, afirma que el contenido depende constitutivamente de esos mismos factores. La primera de las tesis es propiamente una tesis acerca de la identidad del contenido, nos dice qué es lo que determina o a partir de qué superviene el

vano el estudio de este importante tópico es a través de las denominadas actitudes proposicionales.

¹²⁰ Burge define el internismo de la siguiente manera: "*Individualism is a view about how kinds are correctly individuated, how their natures are fixed. (...) According to individualism about the mind, the mental natures of all a person's or animal's mental states (and events) are such that there is no necessary or deep individuating relation between the individual's being in states of those kinds and the nature of the individual's physical or social environments.*" Burge (1986: 3-4).

contenido¹²¹. La segunda tesis es mucho más refinada, ya que explicita directamente el tipo de dependencia que se da entre el contenido intencional y su base de superveniencia. Respecto de la primera tesis obsérvese que queda a nuestro libre arbitrio la cuestión acerca de cómo debemos entender esos factores internos al individuo. De otra manera, nos dice que hay una ontología, pero no precisa cuál. Un mentalista únicamente señalará que esos factores internos lo son con respecto a su mente, sea lo que fuere dicha entidad. Un materialista, en cambio entenderá que los factores a los que se refiere no son sino factores internos a su organismo.

En lo que atañe a la segunda tesis, el asunto más delicado estriba en la necesidad de esclarecer la propia relación de dependencia. En líneas generales podemos decir que Y depende de X si Y necesita de algún modo a X para cumplir con cierto cometido o propósito. Sin embargo, hablar de *un cierto cometido o propósito* no es más que señalar una vaguedad camuflada bajo el armazón de una definición. Así, puede ser adecuado distinguir entre al menos dos tipos de dependencia, a saber, una dependencia causal y una dependencia constitutiva. Una relación de dependencia causal es la que se da por ejemplo entre el hecho de accionar el interruptor de la luz y el encendido instantáneo de aquella. Decimos que la luz se ha encendido porque alguien ha accionado el interruptor o, de otra manera, decimos que ese mismo gesto es la causa principal que, *caeteris paribus*, se encienda la luz. En este caso es obvio que seguimos la definición tradicional de causalidad y aceptamos que la causa precede al efecto en el tiempo. Sin embargo, no decimos que el interruptor constituya la luz o que no pueda haber luz sin interruptores. La relación de dependencia constitutiva es intrínseca a la propia entidad, es decir,

¹²¹ Preguntarse por su base superviniente ya es un modo de responder a la pregunta a propósito de sus condiciones de identidad. Este tipo de caracterizaciones las hallamos en aquellos autores que mantienen un realismo a propósito de las actitudes proposicionales como Fodor (1987) dado que no se pretende ningún tipo de reducción de los diferentes estados mentales a términos no intencionales. Sin embargo podría procederse de una manera que evitara los términos intencionales y caracterizar la vida mental -considerando el propio concepto de "vida mental" como psicología popular- en términos de estímulos, conducta, asambleas neuronales, estructuras neurológicas, etc. Aunque todo esto forma parte de otro debate.

forma parte de las condiciones de identidad de manera necesaria o esencial, no hay prioridad temporal de ningún tipo. La constitutividad no es algo con lo que contingentemente puedan establecerse intercambios causales, no es cuestión de ninguna interacción contingente, sino de necesidad ontológica. Por ejemplo, un corazón está constituido necesariamente de material orgánico, en caso contrario es cualquier otra cosa; incluso cuando se dé el caso de que sea un equivalente funcional como pueda ser una prótesis, pero no es un corazón en sentido estricto. En este sentido, al igual que Tomás de Aquino, debemos ver únicamente el principio de individuación en la materia. La relación de constitutividad nos permite afirmar que bien el contenido es sobrevenido exclusivamente a los factores internos del sujeto o bien que no es únicamente sobrevenido a estos y concursan también los externos.

Traslademos ahora estos distinguos al caso concreto del contenido intencional. Quizás el modo más sencillo, y por otra parte, más relevante para el análisis de la teleosemántica, sea un ejemplo basado en la experiencia perceptiva. Para comenzar el análisis bastará con un ejemplar cualquiera que satisfaga un esquema de enunciado como "S percibe O. Así, al decir que *"veo a mi gata sentada en el sofá"* básicamente estoy afirmando desde presupuestos realistas que hay una situación tal que causa en mi la creencia perceptiva de que, en efecto, mi gata está sentada en el sofá. En ningún momento he negado que el contenido intencional de mi pensamiento no estuviera causado por factores externos, aunque también podría darse un caso de alucinación. Pero en modo alguno he afirmado que el medio, es decir, el hecho de que mi gata estuviera sentada en el sofá, pudiera tener algún papel en la constitución del contenido. Sólo he señalado a ese hecho extramental (y extralingüístico) como la referencia de mi pensamiento. De hecho, podría darse el caso de que bajo unas determinadas circunstancias cambiara esa misma realidad y el contenido de mi creencia perceptiva se mantuviera. En tal caso albergaría una creencia errónea provocada por el mal funcionamiento de mis sentidos o por una alucinación. De otra manera, sólo si aquellos factores que cuentan como internos como puedan ser mis datos de los sentidos o un supuesto objeto interno de aprehensión directa cambiaran, entonces se vería alterado el contenido de mi percepción.

2.2. Los argumentos externistas

La teoría ideacional del significado afectaba tanto a cuestiones lingüísticas como semánticas o de contenido. Los planteamientos clásicos de Frege y Russell son claros herederos de este tipo de concepciones acerca de la naturaleza del significado y del contenido mental. No obstante, esta imagen internista se va erosionando a lo largo del siglo pasado a medida que se van resaltando los factores contextuales o extrasubjetivos de nuestras actitudes y preferencias. El externismo, en contra de los posicionamientos internistas señala que los contenidos no se individualizan ateniéndose únicamente a las propiedades intrínsecas del sujeto, sino que tienen un componente claramente "externo".

Putnam (1973, 1975), Burge (1979, 1986) y Davidson (1987) respectivamente elaboran tres conocidos argumentos basados en situaciones contrafácticas cuyo objetivo es la defensa del externismo. Cada uno de ellos se centra en una dimensión específica.

Putnam desarrolla una argumentación en favor del externismo semántico partiendo del estudio de los términos de género natural en situaciones contrafácticas¹²². Su objetivo es mostrar que la teoría tradicional del significado descansa sobre dos supuestos incompatibles: (1) Conocer el significado de un término consiste en hallarse en un determinado estado psicológico. (2) El significado determina la referencia. Para ello idea el conocido experimento mental de la tierra gemela: Imaginemos que existe un planeta exactamente igual al nuestro en todos los aspectos salvo en que el término "agua" que emplean sus hablantes no refiere a H₂O sino a una sustancia XYZ. H₂O y XYZ son únicamente distinguibles empleando métodos muy sofisticados, pues en apariencia son sustancias idénticas. Así, si un terrícola viajara a la tierra gemela se encontraría con que hay agua hasta el momento en que tuviera posibilidad de analizar el líquido y averiguara que es en realidad XYZ. Por lo tanto, también averiguaría

¹²² Putnam es partidario de la teoría de la referencia directa. En sus artículos de 1973 y 1975 no hace sino extrapolar las ideas de Kripke (1971) sobre los nombres propios como designadores rígidos.

que el término "agua" en esa Tierra gemela no refiere a H₂O. Si por ejemplo fuera imposible realizar un análisis químico, pongamos que porque nos encontramos en una época en la que la ciencia química no ha avanzado lo suficiente no podríamos averiguar la composición exacta del líquido presente en ambos planetas, pero, arguye Putnam el término "agua" sólo referiría a lo que en la Tierra identificamos como H₂O y no XYZ y viceversa. Supongamos que un habitante de la Tierra, Oscar cree que el agua calma la sed, lo creído por Oscar es creído a propósito de la substancia que en la Tierra tiene la composición química H₂O, mientras que lo creído por Bioscar, el habitante de la Tierra gemela, refiere únicamente a XYZ. En ambas situaciones las condiciones de verdad de la creencia son diferentes y, arguye Putnam, también los contenidos de las creencias de Oscar y Bioscar. Por lo tanto, la cuestión es si Oscar y Bioscar, al proferir el término "agua", significan lo mismo. Para Putnam ambos significados dependen de que en cada uno de los planetas haya de hecho agua terrícola o biterrícola. Así, Putnam concluye que los estados psicológicos no determinan la referencia de los términos de género natural y que, por lo tanto, para que los significados puedan determinar la referencia de aquellos no pueden ser individuados a partir de las propiedades intrínsecas del sujeto que los posee: los significados no están en la cabeza.

El argumento de Burge (1979, 1986) es más idóneo para una caracterización del contenido ya que apela directamente a nuestras creencias y no a los términos de género natural. Burge trata de poner de manifiesto la importancia de los factores sociales o extrasubjetivos en la determinación del contenido intencional. Burge cree que podemos atribuir diferentes contenidos a una misma actitud. Así, lo que pretende mostrar es que el contenido que albergue un sujeto puede diferir conforme al contexto en el que se mueva el sujeto permaneciendo la fisiología o las actitudes del sujeto iguales¹²³. Las diferencias únicamente están afuera. Para ilustrar esta concepción acerca de la individuación de los contenidos mentales idea la siguiente situación

¹²³ De manera más concreta Burge señala que: "A person's intentional states and events could (counterfactually) vary, even as the individual's physical, functional (and perhaps phenomenological) history, specified non-intentionally and individually, is held constant." Burge (1986: 4).

contrafáctica: (1) Imaginemos que las actitudes de un determinado individuo contienen el concepto "artritis", así S cree que tiene artritis en sus rodillas, en sus muñecas, en las falanges, etc... y... ¡también cree que tiene artritis en el muslo! Nuestro individuo hace un uso equivocado del término "artritis" porque tiene una comprensión parcial del significado del mismo. (2) Ahora imaginemos la situación contrafáctica en la que el mismo individuo (cuya identidad es exactamente la misma que la del anterior) se mueve en un ámbito en el que todo es igual salvo el contexto social, en el referido nuevo contexto el término "artritis" se aplica a cualquier padecimiento reumático. Ahora ya no hay un uso incorrecto del término tal como sucedía en (1). En el contexto contrafáctico "artritis" no significa *artritis*, no podemos decir que el hablante tiene una comprensión parcial del concepto. Los contenidos presentes en la situación normal y la situación contrafáctica son distintos; no podemos realizar sustituciones *salva veritate* entre ambas situaciones. Sin embargo lo único que ha cambiado de una situación a otra es el contexto social, pues el individuo sigue siendo el mismo. Por lo tanto, concluye Burge, las diferencias de contenido son atribuibles únicamente a las diferencias presentes en el medio social. La defensa del externismo o antiindividualismo de Burge no se sustenta únicamente en la consideración de situaciones contrafácticas como la presentada anteriormente, según Burge algunas teorías psicológicas sólidas como la teoría de la percepción visual de Marr (1982) nos ofrecen básicamente caracterizaciones antindividualistas del contenido¹²⁴.

¹²⁴ Según Burge (1986) la teoría de Marr (1982) no es individualista porque sostendría que el contenido intencional de nuestras representaciones visuales se individualiza ateniéndose a los antecedentes causales sobre los que versa la representación. Así, lo que varía en nuestra percepción son esos mismos referentes y nos las representaciones. De manera más explícita: "*The individuation of the intentional features of numerous representations depends on a variety of physical constraints that our knowledge of the external world gives us. Thus the individuation of intentional content of representational types, presupposes the veridicality of perception. Not only have the explanations, but the intentional kinds of the theory presupposed contingent facts about the subject's physical environment.*" Burge (1986: 32-33). Este artículo de Burge fue respondido por Segal (1991). Para este autor la teoría de Marr no es antiindividualista, sino todo lo contrario. Así, considera que el análisis de Burge es erróneo.

Por último, el argumento del hombre del pantano de Davidson explora la dimensión histórica en la constitución de nuestros contenidos mentales. El caso "Swampman" u hombre del pantano, el famoso experimento mental de Davidson aparecido en su *Knowing One's Own Mind* (1987) aborda la cuestión temporal como fuente de externismo semántico. En este artículo, Davidson presenta una vía de argumentación distinta a la de Putnam y Burge en la defensa del externismo semántico y resalta la dimensión temporal como una fuente en la determinación del contenido más allá de la superveniencia respecto de una misma base física. La situación imaginaria que propone Davidson es la siguiente: Imaginemos una situación fantasmagórica en la que existe una réplica molécula por molécula del propio D. Davidson a la que denominaremos "hombre del pantano" porque nos la hemos encontrado por la noche en un pantano junto a un árbol caído. Dicha réplica se mueve exactamente de la misma manera que Davidson y es reconocida (o confundida) por sus amigos como el propio Davidson. Según Davidson hay diferencias entre él y su réplica. Su réplica no puede reconocer a sus amigos ni puede recordar cuál es su casa ni puede escribir artículos sobre la interpretación radical. Los contenidos que pueda albergar el hombre del pantano no son los contenidos que posee la mente de Davidson. El hombre del pantano, pese a ser un duplicado molécula por molécula carece de contenidos mentales porque carece del trasfondo histórico. La constitutividad física es así insuficiente para la determinación del contenido y, en consecuencia el contenido no superviene a los propios estados físicos del individuo sino que tiene componentes externos, en este caso temporales.

Los dos primeros casos son bien conocidos y constituyen un tópico habitual en las discusiones generales sobre contenido mental. Pero el argumento de Davidson, como mostraré en el capítulo VI, al centrarse en la dimensión temporal, ha tenido una importancia decisiva para las reflexiones en torno a la teleosemántica y las teorías acerca del concepto de función biológica. En mi opinión no debe tratarse de rebatir los tres argumentos en bloque, aunque todos partan de una situación contrafáctica como recurso experimental, creo que las dimensiones que exploran son suficientemente distintas como para merecer un análisis pormenorizado.

2.3. Internismo y externismo a debate

La filosofía contemporánea de la mente está plagada de estudios con argumentos a favor y en contra del externismo. Son muchos los autores de primera línea que han dedicado considerables esfuerzos a delimitar posiciones en torno a ambas tesis y, aún hoy el debate sigue generando un considerable volumen de artículos en revistas especializadas, si bien se empieza a percibir cierta saturación de posiciones externistas¹²⁵.

Disponemos de un amplio abanico de propuestas que van desde aquellas que asumen un individualismo extremo (Bach, Crane, o Farkas), individualistas moderados y partidarios de teorías duales como Fodor (1987) o Chalmers (2002) o, finalmente, externistas exacerbados como Dretske, Millikan o Clark & Chalmers (1998)¹²⁶.

Fodor (1987) señala que las caracterizaciones amplias del contenido no son adecuadas desde el punto de vista de la ciencia cognitiva, pues la caracterización amplia de nuestras actitudes en tanto que estados psicológicos entra en conflicto con la idea de que el contenido es causalmente eficaz, ya que las propiedades relacionales no son determinantes para la individuación de los estados cerebrales y no pueden ejercer causalidad alguna. El contenido estrecho sería para Fodor lo que compartimos tanto mi gemelo como yo y sólo apelando a este tipo de contenido podríamos explicar satisfactoriamente nuestras actitudes como oraciones en mentalés. Sin embargo el contenido estrecho es, según Fodor, inexpresable (Fodor 1987:50) porque es únicamente un contenido en potencia, ya que sólo es contenido cuando es el caso que puede ser semánticamente evaluado. De esto se sigue que el contenido estrecho tiene existencia al margen de sus condiciones de verdad. Lo que comparten ambos gemelos es un estado mental cuya evaluación semántica es relativa a un contexto.

¹²⁵ I.e. Farkas (2008).

¹²⁶ El caso de Clark y Chalmers (1998) es de un externismo extremo o, en sus propias palabras, un externismo activo basado en la consideración de las interacciones activas con el medio en los procesos cognitivos. Lo que hacen estos autores es ir más allá del planteamiento externista de Putnam y señalar que el contenido no sólo se individúa relacionalmente sino que, además, el medio tiene un papel activo en los procesos cognitivos.

Las representaciones o los conceptos, en tanto que unidades mínimas del pensamiento, deben ser caracterizados en sentido estrecho para poder ser combinados conforme a los principios de composicionalidad y generatividad. Así, lo que puede ser combinado ha de ser discreto y concreto, por lo tanto ausente de propiedades relacionales como las que se suponen desde el externismo¹²⁷. No podemos atribuir causalidad a aquello que no está delimitado. Frente al argumento de Putnam, Fodor señala que la identidad de los poderes causales que tiene un determinado ítem debe ser evaluada a través de los distintos contextos en los que puede aparecer y no desde dentro de esos mismos contextos, es decir, dado un determinado ejemplar de representación o preferencia lingüística sus poderes causales se mantienen intactos a través de todo mundo posible. Así ante la preferencia "tráeme un vaso de agua" la respuesta del interlocutor será la misma en este mundo o en el otro¹²⁸. En cualquier mundo posible en el que el referido enunciado fuera proferido hubiera tenido los mismos efectos. Es lícito decir "sírvenme un vaso de agua" al margen de que el agua servida sea H₂O o XYZ. La teoría fodoriana es una teoría que

¹²⁷ Una manera de ilustrar esta idea podría partir de la consideración del ejemplo que propone Fodor (1987:33) sobre los poderes causales de una moneda en una máquina expendedora de bebidas: que dispongamos de una moneda verdadera o falsa es indiferente para evaluar sus poderes causales. Cualquier dispositivo análogo ejemplifica la misma situación. Así, los carritos de la compra que encontramos en los supermercados tienen una ranura por la que hay que introducir o bien una moneda de un euro o bien de cincuenta céntimos. Si introducimos un disco de metal con las mismas características físicas el carrito se suelta de la hilera y puede ser utilizado. El carrito no distingue entre una moneda de un euro, cincuenta céntimos o la pieza de metal. De hecho, para explicar por qué el carrito se suelta de la hilera no necesitamos recurrir al valor de la moneda, sino sólo a las propiedades físicas del objeto que llamamos "moneda". Lo mismo ocurre con los estados mentales. Si asumimos una postura funcionalista en la línea de Fodor nuestros estados mentales consisten en conjuntos de representaciones sometidas a reglas de composicionalidad que regulan sus posibles combinaciones y transformaciones.

¹²⁸ *"It's true that when I say "water" I get water and when my Twin says "water" he gets XYZ. But that's irrelevant to the question about identity of causal powers, because these utterances (thoughts) are being imagined to occur in different contexts (mine occur in a context in which the local potable is H₂O, his occur in a context in which the local potable is XYZ)." Fodor (1987: 35).*

considera dos dimensiones respecto del contenido: Existe un contenido que se individualiza al margen de las propiedades relacionales y otra dimensión que tiene que ver con la evaluación semántica respecto del contexto en el que se sitúa el organismo pensante. La que es relevante para la ciencia cognitiva es sólo la primera¹²⁹.

K. Bach (1988) sostiene una postura individualista extrema en contra de las tesis de Burge. Según Bach los argumentos de Burge están mal planteados porque si observamos cómo son usados los términos relevantes en cada uno de los contextos particulares de los hablantes no encontraremos ninguna diferencia radical en las actitudes proposicionales que albergan. Así, (1) la gente puede tener actitudes cuyos contenidos sean objetos -contenidos de primer orden- sin que posean un término adecuado para referirse a ellos. (2) Aunque Burge insista en el carácter social de la normatividad lingüística ésta no constituye un factor individualizador de las actitudes. Bach parte de una caracterización cartesiana de la mente en la que distingue el contenido mental del lenguaje que se emplea para expresarlo. Así uno puede tener contenidos correctos caracterizados de modo individual y emplear mal el vocabulario cuya normatividad es social. Bach insiste en señalar las diferencias entre pensamiento y lenguaje y creo que esto es positivo ya que permite ver la estrategia de Burge como un exceso. Se podría decir que la dificultad con la que podría encontrarse Bach es la de explicar el contenido sin apelar a un lenguaje. Así ¿cómo tener contenidos si estos no pueden ser conceptualizados adecuadamente? Pero este argumento se desarma fácilmente diciendo que no es cuestión de tener o no tener un lenguaje con el que conceptualizar, sino que uno puede poseer el lenguaje con el que conceptualiza -su habla interna o vocabulario interno- sin tener que albergar el diccionario normativo de su lengua vernácula; las imperfecciones son la norma.

¹²⁹ También hay otras teorías duales del contenido como la de Chalmers (2002). Estas afirman que los contenidos mentales son estados compuestos por elementos caracterizables individualísticamente y elementos amplios; sólo los primeros serían responsables de la causación de la conducta. Así, Chalmers distingue un contenido estrecho que califica de epistémico y uno amplio o de carácter subjuntivo.

T. Crane (1991) discute los argumentos de Putnam y Burge a favor del externismo¹³⁰ y no el externimo en sí mismo. Según Crane los argumentos de Putnam y Burge no nos compelen a decantarnos por un posicionamiento u otro porque en realidad evidencian un pseudoproblema ya que no existen situaciones contrafácticas como la que ambos autores plantean¹³¹. Considera este autor que la idea de un contenido estrecho es un concepto *ad hoc* para solventar las inconsistencias generadas por las teorías de Putnam y Burge. La idea de un contenido estrecho es, pues, un concepto que tiene sentido por contraste. Así, su revisión no propone ninguna teoría substantiva en favor del individualismo, sino únicamente muestra la insuficiencia de las argumentaciones esgrimidas a favor del externismo; las situaciones contrafácticas presentadas por Putnam y Burge son en realidad irrelevantes. Respecto de Putnam Crane arguye que términos como "agua" no son en realidad términos de clase natural, sino términos pertenecientes al lenguaje ordinario, un término de clase natural genuino sería por ejemplo el nombre de cualquier elemento de la tabla periódica. Incluso el concepto de clase natural no es fácil de definir. De un pseudotérmino de clase natural como "agua" no es esperable que fije unívocamente su referencia. El caso del agua es tanto para Oscar como para Bioscar un caso en el que la referencia es distinta pero nada impide que ambos compartan los mismos conceptos

Según Crane deberíamos distinguir entre el significado convencional de un término y el concepto que un hablante trata de expresar. Los dos individuos que ilustran el ejemplo de la tierra gemela

¹³⁰ Crane (1991: 9) sostiene que las tesis de Putnam respecto del carácter social del lenguaje son esencialmente correctas; lo que rechaza es la extrapolación a los contenidos mentales, es decir, el paso que da Putnam desde los significados a los contenidos mentales es ilegítimo.

¹³¹ "Despite its influence, I think that reflection shows that it is inadequate, and that his positive claims are either implausible or largely irrelevant to the problem of intentionality and mental representation." Crane (1991:9). Ya en Burge (1986:22) se evalúa la posible acusación de irrelevancia de los experimentos mentales "The empirical implausibility of the thought experiments is irrelevant to their philosophical point -which concerns possibility, not plausibility." No creo que Crane argumente en contra de la importancia de las situaciones contrafácticas en filosofía, sino en contra del diseño experimental de Putnam o Burge.

pueden muy bien compartir sus conceptos sólo que ambos pueden no ser todo lo precisos que desearíamos. Puede darse el caso de que los conceptos no sean capaces de discriminar entre dos clases naturales como H₂O y XYZ, en tal caso el concepto incluye a ambas puesto que se aplica a ambas. El significado de un término como "agua" no puede depender, tal como supone Putnam, de la microestructura del agua, puesto que la microestructura del agua (sea H₂O o XYZ) no causa intrínsecamente nuestro contenido AGUA, tanto Oscar como Bioscar tienen el mismo concepto. Así, la idea de Putnam de que los términos de género natural funcionan como indécicos es, según Crane, falsa, el mundo o el contexto son incapaces de fijar por sí mismos el significado de estos términos, no al menos como sucede con indécicos como "ahora" o "esto". Tampoco funciona la tesis del bautismo inicial por parte del experto que preconizan las teorías referenciales directas como la de Kripke. En última instancia podríamos considerar que el argumento de Putnam se basa en presupuestos esencialistas y epistemológicos. La idea subyacente a la defensa del externismo es que nuestro conocimiento del significado de los términos de género natural es incompleto desde una caracterización estrecha de los mismos y, en consecuencia, insuficiente en la determinación de la referencia precisa de esos mismos términos -aquí es donde entra en escena el esencialismo-. Sin embargo, para Crane un significado incompleto es ya algún tipo de significado, al menos en lo que refiere a nuestros conceptos.

Respecto de Burge, Crane señala que este autor olvida que a menudo la expresión de creencias en un lenguaje involucra creencias acerca de la conveniencia de los términos usados, de otra manera, hablar implica manejar creencias de primer y de segundo orden, creencias sobre hechos y creencias sobre otras creencias. Así no deberíamos entender la conducta lingüística únicamente como expresión de creencias como "tengo artritis en mi muslo" sino como una conducta en la que se imbrican también creencias de segundo orden como "la frase anterior es la frase correcta para expresar mi creencia". No es lo mismo el significado de los enunciados en un lenguaje público que el contenido de nuestras creencias. La inadecuación de la oración "tengo artritis en mi muslo" es resultado de una creencias de segundo orden falsa. Según el análisis de Crane el

sujeto posee una creencia verdadera y una metacreencia falsa que invalida la primera. De este modo las relaciones entre el lenguaje público y la conducta individual no estarían mediadas sino por propiedades intrínsecas al sujeto. Nuestro sujeto albergaría muchas creencias verdaderas sobre la artritis, podría aplicar el concepto de manera adecuada en muchas ocasiones, pero no en todas. El acierto de sus aplicaciones estaría relacionado con las creencias de segundo orden y no con ningún elemento externo al propio sujeto.

De este modo los casos similares al de la Tierra gemela o a los relatados por Burge en realidad no tienen lugar; el aspecto referencial o veritativo no es un constituyente proposicional. Las diferencias no están más que afuera, el contenido es exactamente el mismo para ambos sujetos.

Farkas (2008) parte de un posicionamiento típicamente cartesiano de nuestra mente. Lo mental para esta autora es "aquello que se individualiza" desde el punto de vista del sujeto". Para Farkas los argumentos en contra del internismo no son tan definitivos como parece suponerse. A su juicio los externistas se enfrentan a tres tipos de argumentos que no pueden responder satisfactoriamente como son la imposibilidad que manifiestan los sujetos para discriminar positivamente en situaciones como la relatada por Putnam (1975) entre XYZ y H₂O, la confrontación con la idea de acceso privilegiado de primera persona y la interpretación errónea de la idea de que el sentido determina la referencia. Ni la indiscriminabilidad, ni el acceso privilegiado son tratados adecuadamente por el externismo, mientras que la teoría a propósito de la fijación de la referencia que los externistas atribuyen a los internistas es errónea. Así, las creencias se individualizan a través de sus contenidos y no de sus condiciones de verdad; diferentes contenidos indican diferentes creencias. El sentido no determina por sí sólo la referencia con respecto al estado del mundo actual de forma que cambios en el mundo no necesariamente implican cambios en el contenido de una proposición. Siguiendo la misma línea de Crane, las diferencias sólo están en el mundo.

En mi opinión defender una postura cartesiana tal como hacen Bach y Farkas tiene claras virtudes en humanos adultos. Pero es difícil de mantener si se extrapolan esas mismas conclusiones a otros organismos más básicos. Por lo tanto, una postura que insiste en la

conciencia como rasgo de lo mental no es fácil de asumir desde presupuestos naturalistas. Por otra parte, respecto de los tres argumentos externistas clásicos creo, al igual que sus críticos, que exhiben diversas debilidades. En lo que respecta a Putnam su argumento parte de un supuesto erróneo como es la omnisciencia que pudiera suponer un presunto conocimiento absoluto de los sujetos acerca de lo que significan sus representaciones. Así, que los sujetos tengan acceso privilegiado a sus contenidos -idea cartesiana por excelencia- no significa que estos contenidos sean perfectos. Los contenidos pueden ser sesgados, sin embargo el sujeto los conoce tal como son para él y se maneja relativamente bien con ellos. El sujeto no debe ser omnisciente acerca de todo aquello que resulta subsumido bajo el concepto o el término correspondiente. El argumento de Putnam procede inversamente, es como si se esperara una omnisciencia imposible por parte del sujeto. Por otra parte, aunque Putnam señale que los significados no están en la cabeza y con esta célebre frase pretenda eliminar el supuesto de que conocer la intensión -es decir hallarse en un determinado estado psicológico es condición suficiente para determinar la referencia- cabe preguntarse si acaso los conceptos pueden estar en otro lugar sino "en la cabeza". Por otra parte, como bien argumenta Farkas, el argumento de Putnam se encuentra afectado por dificultades considerables acerca de qué cuenta como interno o externo a los sujetos, o en qué sentido es lícito admitir que los estados fenoménicos de ambos sujetos son realmente indiscernibles. Sin embargo creo, al igual que ella, que el argumento más poderoso es el de la imposibilidad de compatibilizar acceso privilegiado y externismo.

La argumentación de Burge tiene diversas lecturas. En primer lugar creo que pone de relieve la normatividad intrínseca al empleo del lenguaje. En este respecto no podemos decir que sea partidario de ningún cartesianismo lingüístico si es que alguna vez existió tal posicionamiento. El problema creo que hay que verlo en la extrapolación de la constitutividad normativa del lenguaje al contenido. Que los hablantes emplean mal o parcialmente los términos es un fenómeno conocido, el uso regimenta el significado de las palabras y éste se constituye en un ámbito comunitario. Pero es una cuestión disputada decir lo mismo acerca de la constitutividad de los contenidos.

Según Crane (1991) una de las ideas que subyace a los experimentos de Burge es que los estados intencionales pueden verse afectados causalmente sin que varíe su fisiología, es decir, tal como sucede con el lenguaje, el medio tendría efectos causales constitutivos sobre el contenido. Por otra parte en lo que refiere al contenido no ya de nuestros términos, sino de nuestros perceptos, Burge parece asumir que la percepción es un proceso que bajo las condiciones apropiadas arroja representaciones verdaderas y, por esta razón, puede afirmarse que el contenido de esas mismas representaciones no se fija de manera individualista, ya que puede haber un cambio en la referencia sin que varíe la representación. Pero esto no es más que decir lo que ya sabemos convirtiendo las condiciones de verdad en constituyentes del contenido. Que las percepciones puedan ser erróneas no implica que lo percibido sea un constituyente del contenido. Si así fuera ¡todos seríamos externistas desde los inicios de la Filosofía!

Finalmente, por lo que respecta a Davidson es, en mi opinión, el argumento de mayor peso y el menos estudiado. Creo que éste no se plantea en base al error de contenidos o la inexactitud de las clases naturales, sino a la posibilidad categórica de que el hombre del pantano tenga contenidos. El argumento de Davidson es el argumento realmente peligroso porque al afectar a la historia del individuo pone en cuestión que realmente tengamos otro individuo idéntico, es decir, lo que el argumento explota es la noción de identidad individual. En mi opinión el externismo defendido por Davidson ha sido el más influyente para la teleosemántica.

3. Normatividad

La normatividad de lo mental es un tópico clásico. Las entidades mentales no sólo son intencionales sino que, además, son normativas. Algo es normativo cuando es conforme a algún tipo de regla, es decir, cuando se impone una regimentación que tiene apoyo contrafáctico. Esto hace igualmente normativas tanto a las leyes científicas que versan sobre fenómenos empíricos como al contenido mental. Sin embargo, el contenido aunque es legaliforme, no depende de leyes causales en el mismo sentido que muchos fenómenos tal como son descritos por algunas ciencias. El contenido exhibe un grado de

complejidad y singularidad que difícilmente puede ser obviado en su consideración. No ocurre lo mismo en algunos fenómenos naturales como la dilatación de un sólido, el comportamiento de los gases o la aceleración de un grave, etc. Por si fuera poco, las leyes científicas no abundan en biología, disciplina más próxima para una reducción científica. Aunque sería ridículo negar complejidad a los fenómenos descritos a través de leyes por ciencias como la física o la química, lo cierto es que muchos son susceptibles de una explicación razonable apelando a unas pocas leyes. También su clasificación es menos prolija y no exhibe dobleces intencionales tal como ocurre con los estados psicológicos con un contenido intencional. Sin embargo, la generalidad exhibida por los fenómenos naturales se difumina rápidamente cuando consideramos la acción racional, es decir, la acción guiada por creencias y deseos: una misma conducta puede ser desencadenada por diversas razones y deseos, no hay tal cosa como una tipicidad conductual. Explicar una conducta particular requiere un complejo sistema de creencias y deseos al margen de la descripción física del organismo que la ejecuta. Por ejemplo, la explicación biomecánica de un determinado movimiento corporal como una patada es exactamente la misma para un entrenamiento de karate que para un partido de fútbol o que para una pelea callejera; las razones que estructuran el mismo fenómeno físico, no. Esa es precisamente la diferencia existente entre actuar por causas y actuar por razones. Las acciones que no son estereotipadas o que no responden a una pauta de acción fija, involucran creencias y deseos que constituyen en algún sentido razones para actuar más allá de simples causas. El caso extremo de racionalidad es el humano. En estos la acción es plenamente racional porque implica una determinación de la voluntad, es decir, un acto reflexivo y consciente.

Aunque ya he señalado un paralelismo entre la nomologicidad y la normatividad (cap. III) es preciso apuntar que la normatividad mental no es traducible a nomologicidad física, es decir, no es reducible. Identificar un fenómeno mental como normativo no es lo mismo que equipararlo con algún tipo de evento físico caracterizable nomológicamente. Davidson (2001: 253) señala que una descripción psicológica no puede ser reducida a una caracterización física, biológica o fisiológica. Si bien el monismo anómalo de Davidson es

discutible, lo cierto es que cuando hablamos de normatividad de lo mental no estamos hablando de relaciones causales ni estamos constatando una regularidad, sino que imponemos un deber ser. El ámbito de la normatividad mental pertenece al espacio de las razones, no de las causas. Como se verá en el próximo capítulo, la teleosemántica es un intento de naturalizar el contenido, de aproximar las razones a las causas naturales que suponemos en todo fenómeno natural.

La normatividad se dice también de muchas maneras, así, cuando se habla de la normatividad de lo mental puede entenderse que (1) se hace referencia al hecho de que los conceptos designen lo que designan y no otra cosa, (2) también al hecho de que las relaciones entre conceptos están sujetas a reglas tal que se puedan formar enunciados o proposiciones y, por último, (3) a las relaciones de orden inferencial que se dan entre los enunciados en la mente de los sujetos y que se manifiestan en la conducta lingüística de muy diversos modos. De otro modo, los tres elementos clásicos de la lógica; el concepto, el juicio y la inferencia exhiben normatividad.

En el primer caso podemos entender que a lo que nos estamos refiriendo es a la cuestión de la fijación del contenido de los conceptos; un concepto designa lo que designa y no otra cosa. Los conceptos subsumen representaciones de los objetos y así estos últimos pueden ser pensados por el entendimiento. Pero lo que es subsumido no es heterogéneo; cabe suponer que se subsume un conjunto de ejemplares del mismo tipo de representación conforme a la norma del concepto que se presenta como función. Como Kant señala:

"Todas las intuiciones, en cuanto sensibles, se basan en afecciones, mientras que los conceptos lo hacen en funciones. Entiendo por función la unidad del acto de ordenar diversas representaciones bajo una sola común." Kant, KrV A68/B93.

Además, Kant añade que el orden de estas representaciones subsumidas bajo el concepto -a mi entender la norma subsuntoria- procede de la espontaneidad del sujeto, es decir, se funda en la propia naturaleza del entendimiento y no en la sensibilidad; de otra manera, no responde a causas, sino a razones. Así, si pensamos lo que pensamos es porque nos es dado pensarlo así según la propia

naturaleza del entendimiento. De este modo encontramos no sólo conceptos empíricos que en un momento dado pudiera pensarse que se corresponden con clases naturales sino también conceptos puros que se constituyen en condición de posibilidad de los anteriores. El planteamiento de Kant es obviamente un planteamiento trascendental. A diferencia de otros planteamientos no propone que los conceptos necesariamente tengan que identificarse con la esencia de las cosas subsumidas, sino que se centra en desarrollar las condiciones de posibilidad para la experiencia y, de modo más particular, para una experiencia cognoscitiva. Otros planteamientos han sido posibles a lo largo de la historia del pensamiento filosófico. Las teorías filosóficas acerca de los conceptos son harto conocidas y al margen del estatuto ontológico que pueda atribuirse a un concepto desde el punto de vista filosófico hasta bien entrado el siglo XX se admite que un concepto debe estipular en términos de condiciones necesarias y globalmente suficientes que entidades o rango de fenómenos o *sense data* deben ser subsumidos. Las teorías psicológicas acerca de qué sea un concepto (i.e. teorías probabilistas, teorías de ejemplares, etc.) no nos importan desde el punto de vista de la pura teorización, sino únicamente para dar cuenta de la actuación de los sujetos particulares. Por esa razón no las contemplo.

Un concepto tiene un significado determinado y es usado o aplicado a aquellos objetos que deben ser subsumidos por el mismo. Los conceptos en tanto que tienen contenido exhiben normatividad e igualmente, como las palabras, pueden ser usados de manera correcta o incorrecta. Uno puede subsumir bajo el concepto PERRO a canes domésticos y chacales, de hecho, esta circunstancia no es extraña en los primeros estadios del desarrollo psicoevolutivo, pero el concepto PERRO sólo corresponde a la clase biológica de los perros (*Canis lupus familiaris*) y no chacales (*Canis adustus*) o lobos (*Canis lupus*). Cuando aplicamos un concepto no lo hacemos para un sólo caso, sino que tácitamente suponemos que será extrapolable en el futuro a otros casos, es decir, al usar un concepto estamos haciendo suposiciones acerca de sus futuras aplicaciones. En este sentido, decimos que los conceptos son normativos.

En el segundo caso nos referimos a relaciones entre conceptos es decir, juicios. En los juicios identificamos a un individuo como tal y le

adscribimos un predicado o concepto tal que enunciamos *Fa*. Si bien y de manera más limitada, la formulación tradicional nos indica que adscribimos un predicado a un sujeto. Un juicio involucra así al menos dos conceptos; un concepto individual -que de por sí es discutible que sea un concepto en algunos casos cuando lo que se persigue es denotar a un individuo particular- y un concepto general o clase. Estas relaciones también son normativas en algún sentido, pues no todo concepto es combinable con otro concepto. En el caso de los juicios analíticos el requisito mínimo es que se respete el principio de no contradicción. Como dice el propio Kant

"La ausencia de contradicción interna constituye la condición universal -aunque sea únicamente negativa- de todos nuestros juicios, sea cual sea el contenido del conocimiento y el modo según el cual se refiera al objeto. De no cumplir tal condición, esos juicios no significan nada en sí mismos. (...) La proposición "A ninguna cosa le es adecuado un predicado que la contradiga" recibe el nombre de principio de contradicción y constituye un criterio universal, aunque meramente negativo, de toda verdad." Kant, KrV A150-151.

En el caso de los juicios sintéticos nos encontramos con que ha de ser posible imaginar de algún modo la adscripción predicativa o la síntesis entre los dos conceptos que mínimamente componen el juicio y, por lo tanto, necesitamos de una adecuación material con algún estado posible de cosas en el mundo¹³². Es decir, un juicio sintético relaciona dos conceptos si describe o se corresponde con una experiencia posible. Así, a la luz de la experiencia actual acumulada no parece posible que podemos tener una experiencia de un "elefante volador" pero sí la de un "gato con cuatro patas". Ser elefante y poseer la propiedad de volar no parece plausible; son conceptos no relacionables en un juicio, mientras que no ocurre lo mismo con la propiedad de ser gato y poseer cuatro extremidades.

¹³² Kant señala que el elemento que sintetiza los dos conceptos que componen un juicio sintético es el todo formado por el sentido interno -el tiempo- la imaginación y la apercepción trascendental o yo que acompaña a todas las representaciones: *"Es pues en el sentido interno, la imaginación y en la apercepción donde hay que buscar la posibilidad de juicios sintéticos y, dado, que estas tres facultades contienen las fuentes de las representaciones a priori, habrá que buscar igualmente en ellas la posibilidad de juicios sintéticos puros."* KrV A155.

Finalmente es obvio que las inferencias exhiben normatividad, no sólo material, es decir, no solo una normatividad derivada de los juicios y conceptos sobre los que se constituyen las inferencias sino además una normatividad intrínseca a la propia inferencialidad: hay inferencias correctas -ajustadas a las diversas normas lógicas- e inferencias incorrectas -o inferencias falaces-. No en vano pensar es conocer mediante conceptos¹³³ y el lugar propio de los conceptos no es la sensibilidad, sino el entendimiento, lo que para Kant es la facultad de las reglas.¹³⁴ No es necesario plantear esta caracterización en consonancia con el carácter apriorístico de las mismas, si bien para Kant es obvio que lo es en tanto que la normatividad impuesta por los conceptos puros del entendimiento es *a priori*.

Otra manera de explicar la normatividad del contenido es vinculándolo directamente con la noción de verdad. Si bien es esencialmente la misma estrategia seguida para el estudio de los juicios sintéticos. Dado que no somos seres perfectos ni omniscientes y que no siempre acertamos, pues el error es una constante en nuestras vidas e incluso en términos evolutivos tiene un precio muy elevado, éste merece una consideración especial. Una característica esencial del contenido mental es que éste ha de poder ser tenido erróneamente. Así, puedo creer que lo que está delante de la estufa es un gato y, como ya he mostrado, estar equivocado, también puedo creer que los champiñones que encontré ayer en el bosque son comestibles cuando en realidad no eran champiñones sino amanitas (*Amanita muscaria*). Así el contenido es normativo en la medida en que puede ser evaluado respecto de lo que es o no es el caso, es decir, en tanto que la relación entre los conceptos que componen el juicio o enunciado se adecua o no respecto de una realidad dada. Por lo tanto, decir del contenido que es normativo implica decir también que éste es susceptible de una evaluación semántica. Podemos creer que p y nuestra creencia de que p será verdadera si y sólo si es el caso de que p , y falsa en caso contrario. Si creo que mi gata Atina está delante de la estufa mi creencia es verdadera si es el caso que ella está delante de la estufa y

¹³³ Cf. Kant, KrV B94.

¹³⁴ Cf. Kant, KrV A126.

falsa si, por ejemplo, es una rata grande y negra que he confundido con mi gata en la oscuridad ya que de noche todos los gatos son pardos.

En Filosofía del lenguaje también es un tópico común resaltar la normatividad del mismo, la idea de que existe una íntima relación entre el significado de una palabra y su uso, idea que fue explorada por Wittgenstein en sus *Investigaciones Filosóficas* y que ha sido el germen de importantes reflexiones sobre este tema. Wittgenstein conceptualiza la normatividad en términos de seguimiento de reglas; decir del lenguaje que es normativo no es sino decir que es una práctica guiada por reglas. Nuestras creencias, al igual que el lenguaje que las manifiesta, también son normativas.

El aspecto más básico de la normatividad es el de la fijación del contenido de los conceptos. Sobre éste se incide de manera sistemática en las teorías que presento en el siguiente capítulo y constituye el paradigma de fijación del contenido. Si bien los otros dos no son descuidados, hay que señalar que en un enfoque naturalista como es el teleosemántico dependen en gran medida del primero.

2.4. La caracterización informacional del contenido y la dicotomía entre lo conceptual y lo no conceptual

En la sección anterior introduje dos conceptos kantianos que han tenido una gran repercusión: intuiciones y conceptos. Respecto a la necesaria unión o síntesis de ambos para posibilitar una experiencia cognoscitiva o simplemente para posibilitar el juicio da buena cuenta el famoso dictum kantiano:

"Los pensamientos sin contenido son vacíos; las intuiciones sin conceptos son ciegas. Por ello es tan necesario hacer sensibles los conceptos (es decir, añadirles el objeto en la intuición) como hacer inteligibles las intuiciones (es decir, someterlas a conceptos). Las dos facultades o capacidades no pueden intercambiar sus funciones. Ni el entendimiento puede intuir nada, ni los sentidos pueden pensar nada." Kant, KrV A51.

Kant afirma que hay razones para mantener la distinción entre sensibilidad y entendimiento, entre *sentiencia* y *sapiencia*. Las intuiciones y conceptos kantianos perviven en el discurso filosófico actual. Kant conceptualiza un importante tópico moderno que ha sido

contemporáneamente explotado por filósofos como McDowell (1994), aunque no de modo exclusivo, pues en la filosofía analítica ha sido siempre un tópico importante, especialmente en lo referido a la teoría de la percepción. Una manera de abordar la distinción kantiana entre conceptos e intuiciones es la que más contemporáneamente se ha establecido entre lo conceptual y lo no conceptual (Evans, 1982). Sin embargo, antes de analizar esta nueva distinción es conveniente decir algo sobre el punto de partida de filósofos como Evans.

Ya se ha mencionado la importancia que tenía el proyecto naturalizador de la intencionalidad en este trabajo. Si la intencionalidad ha de ser objeto de estudio de alguna disciplina científica no puede ser una propiedad tan misteriosa, debemos tratar de eliminar lo que de mágico o misterioso pueda tener, de otra manera, lo intencional aún a riesgo de constituir una anomalía, debería poder ser susceptible de un tratamiento que casase con nuestra visión científica del mundo. Muchos filósofos con intenciones naturalistas inspirados por la revolución cognitiva han pensado que una buena manera de proceder a un tratamiento científico de la intencionalidad pasaba por el análisis del concepto de información. En psicología el uso de la teoría de la información no es nada nuevo, al menos en el estudio de la percepción y más concretamente de la percepción visual. de hecho todos estos autores de una manera u otra se basan en el trabajo de J.J. Gibson (1968) *"The Senses Considered as Perceptual Systems"*. El recurso a las caracterizaciones informacionales es la estrategia seguida por la semántica del indicador causal de Dretske (1981) o Stampe y por los partidarios de la teleosemántica. Una teoría del contenido cercana a los planteamientos psicológicos ha de emplear de algún modo el concepto de información.

Hablar de relaciones informacionales es un primer paso para una naturalización del contenido; la noción de información no es una noción intencional, sino causal. Lo que se espera de una teoría naturalista del contenido es que lo explique sin recurrir a nociones intencionales. Una naturalización de la intencionalidad se consigue cuando ésta puede ser abordada con los mismos conceptos que el resto de disciplinas científicas. En este sentido es obvio que el vocabulario causal sería aparentemente el propio y deseable de cualquier disciplina científica. Si

bien esto es una simplificación excesiva y mitificada de la actividad científica.

Desde un punto de vista estrictamente filosófico el papel de la obra de G. Evans (1982) es fundamental en la caracterización del contenido mental en términos informacionales. Evans rechaza una caracterización del contenido en términos puramente doxásticos o actitudinales. En su lugar se decanta por una caracterización informacional. Así en lugar de decir que S cree/desea que p considera más acertado decir que S se encuentra en un estado informacional determinado con un cierto contenido. Las razones de Evans al proponer una caracterización informacional no responden a un simple capricho filosófico, sino que se justifica desde el presupuesto de una neta diferenciación entre lo conceptual y lo no conceptual sugerida por la independencia doxástica que se da en algunos estados mentales. No obstante hay que advertir que la noción de contenido no conceptual no está suficientemente clarificada por el propio Evans.

Al margen de la importancia que pueda tener el concepto de información en la caracterización del contenido –este tópico será discutido más adelante en el lugar que le corresponde- lo que sí es cierto es que la dicotomía conceptual/ no conceptual es muy rica a la hora de establecer una caracterización del contenido mental y lo es por el gran número de precisiones que permite establecer.

Hablamos de contenido conceptual como de aquel contenido propio de nuestras creencias o juicios que, constituidos por conceptos son, además, evaluables en términos de condiciones de verdad. Así, los conceptos tradicionalmente son vistos como los constituyentes de las actitudes proposicionales, la posesión de conceptos es, en consecuencia, una condición previa para poder tener un pensamiento.

La cuestión realmente interesante en este debate será la que nos diga en qué situaciones podemos decir que un contenido determinado es un contenido conceptual. Tratar de decidir esta cuestión tendrá que ver con la posesión o no posesión de conceptos por parte del sujeto. Pero ¿cuándo podemos decir que se poseen conceptos o, de otra manera, cuáles son las condiciones para la posesión de conceptos? En principio puede decirse que conceptualizar implica discriminar, un concepto es un discriminante que delimita los objetos que caen bajo su dominio separándolos de aquellos que quedan fuera. Una

discriminación es algo dicotómico que separa entre elementos pertenecientes a la clase o elementos no pertenecientes, por ejemplo, lo comestible y lo no comestible para un determinado animal o lo que cuenta como un sauce y aquellos individuos que, pertenecientes a la clase de los árboles, no pueden ser clasificados como sauces. Así, una de las tareas básicas de los conceptos es la aplicación del mismo a una determinada realidad. Pero esta caracterización es excesivamente vaga si lo que queremos es señalar qué de propio tiene lo conceptual. La discriminación de por sí no es un elemento muy clarificador en la explicación de nuestras actitudes.

G. Evans (1982) desarrolla una peculiar aproximación al problema, uno de sus objetivos consiste en explicar cómo están estructurados los pensamientos en términos del ejercicio de habilidades conceptuales. Así, para Evans, la posesión de conceptos que sería lo que caracterizaría a las diferentes actitudes tiene que ver con el ejercicio de habilidades conceptuales. Conceptualizar nos debería permitir pensar sobre lo conceptualizado, no sólo formar pensamientos sino proceder a inferencias; el poder establecer inferencias es algo intrínseco al pensar. En esto consistirían los contenidos conceptuales de nuestras actitudes. G. Evans hace referencia a un requisito de generalidad como condición básica para poder atribuir la posesión de conceptos a un sujeto determinado, es lo que se conoce como *Generality Constraint*. Su requisito de generalidad afirma que:

"(I) If a subject can be credited with the thought that a is F, then he must have the conceptual resources for entertaining the thought that a is G, for every property of being G of which he has a conception." Evans (1982: 104).

El requisito de generalidad afirma es que un concepto es algo que puede ser aplicado en numerosas ocasiones, no sólo una vez. Si uno es capaz de reconocer a este árbol como sauce entonces debe ser capaz de reconocer a otros muchos ejemplares como sauces y estos conceptos deben poder aparecer no sólo en un enunciado sino en muchos otros con diferente estructura, es decir, debemos ser capaces de producir nuevos pensamientos a partir del conjunto de conceptos que un determinado sujeto posee. En este sentido decimos del pensamiento que es articulado.

Por otra parte, los conceptos conforman proposiciones, son los ladrillos del pensamiento, y permiten articular razones. El ámbito de lo conceptual es el ámbito de las razones. Y con las razones es con quienes tenemos que vérnoslas cuando tratamos de justificar nuestras creencias. Lo conceptual constituye nuestra fuente justificatoria pero ¿cómo se relaciona con la realidad?

Evans (1982: 122) señala que las personas constituyen un tipo particular de sistemas informacionales. Un sistema informacional es un sistema en el que las transacciones entre sus diversos componentes se caracterizan en términos informacionales; un sistema que recoge, almacena y transmite información, entendiéndose aquí que el concepto de información se toma como primitivo. En el caso de los humanos, las funciones anteriormente citadas están representadas por la percepción, la memoria y la comunicación, capacidades que ejercitamos respectivamente cuando obtenemos información del medio a través de nuestros sentidos, cuando la almacenamos o cuando la recibimos en el proceso comunicativo.

La ventaja de caracterizar el contenido en términos informacionales antes que en términos doxásticos hay que buscarla precisamente en la necesidad de establecer límites entre lo epistémico y lo no epistémico, entre lo que pertenece al ámbito de las intuiciones y lo que entra dentro del espacio de las razones, el entendimiento. Sin embargo, Evans no desarrolla muy ampliamente la noción de contenido no conceptual; únicamente afirma que el contenido de nuestra experiencia perceptiva es no conceptual¹³⁵.

¹³⁵ "The informational states which a subject acquires through perception are non-conceptual, or non-conceptualized. Judgements based upon such states necessarily involve conceptualization: in moving from perceptual experience to a judgement about the world (usually expressible in some verbal form), one will be exercising basic conceptual skills. But this formulation (in terms of moving from a experience to a judgement) must not be allowed to obscure the general picture. Although the subject's judgements are based upon his experience (i.e. upon the unconceptualized information available to him), his judgements are not about the informational state. The process of conceptualization or judgement takes the subject from his being in one kind of informational state (with a content of a certain kind, namely, non-conceptual content) to his being in another kind of state (with a content of a different kind, namely, conceptual content.)" G. Evans (1982: 227).

El término “contenido no conceptual”¹³⁶ adquiere su significado por oposición con el de contenido conceptual. Se trata, por tanto, de una noción que resulta de un contraste. Evans sugiere que el contenido de nuestros estados perceptivos es de una clase distinta al de nuestros estados cognitivos. Es decir, por una parte tenemos los estados perceptivos cuyo contenido sí es representacional pero no conceptual o no conceptualizado y, por la otra, los estados de creencia caracterizados por la posesión de conceptos. Estar en un determinado estado doxástico implica ejercer habilidades conceptuales, pues las creencias están formadas por conceptos. En la sección anterior apunté qué implicaba la posesión de conceptos y cómo ésta se caracterizaba a través del requisito de generalidad, ejercer cualquier tipo de habilidad conceptual dependerá, en consecuencia, de la referida posesión de conceptos.

Creo que hay dos vías posibles para enfocar de un modo más preciso la cuestión del contenido no conceptual. En primer lugar deberíamos poder afirmar que la posesión de contenido no conceptual implica la no satisfacción de las condiciones para la posesión de conceptos. De un modo más preciso, implica la incapacidad que tienen las representaciones perceptivas para poder entrar a formar parte de

¹³⁶ Hay que señalar que la noción de “contenido no conceptual” se aplica al menos a tres ámbitos diferentes y es obvio que cada uno de ellos merece de por sí un tratamiento diferenciado, de ahí que solamente me circunscriba en este ensayo al estudio del tercero de ellos que afecta en exclusiva a la filosofía de la percepción. El contenido no conceptual lo podemos encontrar en animales o en niños en el estado preoperatorio, también en estados representacionales a un nivel subpersonal o subdoxástico y, finalmente, en estados perceptivos. Las dos primeras variedades pueden suscitar la pregunta acerca de qué acceso tienen al mundo las criaturas carentes de un aparato conceptual complejo y, por supuesto, de lenguaje. La idea es que en algunos casos como los aludidos se puede atribuir contenido sin atribuir necesariamente la posesión de conceptos. Por lo tanto lo que se quiere recalcar es que puede haber estados representacionales sin necesidad de que haya conceptualización. De todo esto se sigue que no es necesario que para atribuir pensamiento a un individuo sea necesario atribuirle la posesión de conceptos. Con el tercero de los ámbitos señalados, el referido al contenido no conceptual de la percepción sucede que ésta es un proceso en el que este tipo de contenido parece estar claramente presente en el *quale* de mi percepción particular que se resiste a ser conceptualizado. Es lo que se conoce como el argumento de la riqueza de la percepción

una estructura proposicional en tanto no estén conceptualizadas. No obstante, hay que dejar claro que la no satisfacción parece ser entendida por Evans de manera categórica; no se satisfacen las condiciones para la posesión de conceptos porque no pueden ser satisfechas de ninguna manera sin dejar de ser no conceptuales. No es algo defectivo, sino que, en cualquier caso, se trata de dos naturalezas distintas de contenido.

En segundo lugar y, a mi juicio, de modo complementario y de una manera más específicamente centrada en el estudio de la percepción, podemos decir que el carácter fenomenológico de ésta¹³⁷ constituye el contenido no conceptual de la misma. Con esto estamos señalando que aquello que tiene de propio la percepción es una riqueza empírica no subsumible en el ámbito conceptual o, de manera más directa, no conceptualizable. Es lo que se conoce como el argumento de la riqueza perceptiva y que se basa en la incapacidad que tenemos para dar cuenta conceptualmente de la pluralidad de matices que constituyen nuestra experiencia perceptiva. La idea de fondo que subyace a este argumento es que la conceptualización es incapaz de agotar la riqueza de la percepción. Así, pese a una posterior conceptualización de la información recibida por los sentidos, siempre quedaría un remanente no conceptualizado o no conceptualizable.

Cuando señalamos el carácter no conceptual de la percepción y, por lo tanto, no epistémico, debe ser hecha una precisión: experiencia y

¹³⁷ A propósito de la fenomenología de la percepción y dentro del debate entre contenido conceptual y no conceptual, Bermúdez (2003) señala tres rasgos de lo que él denomina una fenomenología de la percepción y que deberían servirnos para trazar las diferencias entre un contenido perceptivo y uno proposicional o conceptual; de otra manera ¿en qué podemos fundamentarnos para decir que la percepción se aparta en gran medida de la pura creencia como señalaba Armstrong? Bermúdez alude al hecho de que la percepción se presenta como análoga y no digital, a diferencia de las actitudes proposicionales. También que la percepción está libre de una metrización o de aplicación de conceptos métricos ya que sólo las actitudes proposicionales pueden expresar magnitudes. En tercer lugar, señala que el contenido perceptivo es más detallado que el contenido proposicional. Así, puedo percibir un sinnúmero de matices de gris pero difícilmente podré conceptualizar todos esos matices de gris que percibo. Esta última distinción vuelve a ser el argumento de la riqueza de la perceptiva.

percepción son, para Evans, cosas distintas. El concepto de percepción es estrictamente informacional, en este sentido Evans se aparta tanto de la concepción tradicional de la percepción considerada desde un punto de vista psicológico como filosófico. Una experiencia perceptiva no es lo mismo que un estado informacional perceptivo. En el primero necesariamente está involucrada la conciencia, mientras que en el segundo no. La experiencia es conceptual o conceptualizable, es decir, una experiencia perceptiva es el resultado de tratar de aplicar nuestras habilidades conceptuales a lo dispuesto por el sistema informacional aún cuando pueda quedar el mencionado remanente de no conceptualidad.

Tanto el argumento de la incapacidad de satisfacer las condiciones de posesión de conceptos por parte de las representaciones perceptivas como la caracterización fenomenológica de la percepción constituirían dos buenas aclaraciones acerca de qué sea el contenido no conceptual. Aunque falta la cuestión más importante, ¿cómo se relacionan?

Lo que está en juego en la discusión a propósito del contenido no conceptual y el contenido conceptual en lo que atañe a la teoría de la percepción es el carácter epistémico o no epistémico que ésta pueda tener. La cuestión es si la percepción pertenece o no pertenece al denominado espacio de las razones. Evans sostiene que la experiencia perceptiva es un estado del sistema informacional y, por lo tanto, su contenido no puede ser conceptual. Con esto está negando su valor epistémico ya que, al no intervenir conceptos en la percepción, ésta queda fuera del espacio de las razones. Así percepción y creencia son dos cosas distintas. Las creencias se formarían sobre la base no conceptual de la percepción en tanto que ésta es un elemento más del sistema informacional. Sólo en la medida en que el contenido perceptivo pueda conceptualizarse puede entrar a formar parte del espacio de las razones. La percepción únicamente aporta los materiales sobre los que se ejercen nuestras habilidades conceptuales. Desde este punto de vista se asume que entre las metas del contenido no conceptual está la de servir de fundamento al contenido conceptual.

Por su parte McDowell (1994) responde a Evans arguyendo que la experiencia perceptiva debe entrar dentro del espacio de las razones. McDowell ensaya lo que podríamos denominar una salida kantiana en

el problema de determinar el lugar que el contenido empírico ocupa en nuestro conocimiento. Si tendemos hacia un punto de vista coherentista como el que pudiera ofrecernos una teoría de la percepción como creencia pura, es decir, una teoría que enfatizara el aspecto eminentemente doxástico y epistémico de nuestras percepciones, nos encontraríamos con dificultades a la hora de caracterizar lo que de propio tuviera el contenido empírico ya que éste se difuminaría fácilmente en el espacio de las razones. Si, por el contrario, ensayáramos una aproximación fundacionalista como la propuesta por ciertos empirismos fácilmente caeríamos en el mito de lo dado, ya que no habría lugar para las razones. Si la virtud está en el medio es ésta una situación en el que los extremos se alcanzan con mayor facilidad. McDowell piensa que el conocimiento empírico resulta de la cooperación entre sensibilidad y entendimiento. Para él la experiencia sensible, aquello que usando una terminología empirista podríamos denominar “impresiones” –y que es el término que el propio McDowell emplea- tiene de por sí contenido conceptual, ya que comparte el ejercicio de las mismas capacidades puestas en escena en el juicio; es decir, participa de la espontaneidad del sujeto o, lo que es lo mismo, del control racional y autónomo de nuestros pensamientos. La percepción es, por lo tanto, una fuente más de justificación epistémica. Para McDowell, la postura defendida por Evans de que el contenido de la experiencia perceptiva es no conceptual es una versión modificada del mito de lo dado y es insuficiente, en consecuencia, como fuente de justificación epistémica.

Es obvio que ambos tienen una concepción distinta de la percepción que es ilustrativa de la dicotomía que he intentado caracterizar. Para Evans la percepción es un proceso informacional sin más, para McDowell la percepción tiene una dimensión epistémica que no está recogida en la caracterización más estrecha de Evans. Al incorporar esta dimensión epistémica McDowell está volviendo de algún modo a una cierta variedad de la teoría de la percepción como creencia o estado doxástico. En cualquier caso, ambos presentan el tipo de dilema al que nos enfrentamos cuando queremos abordar el estudio filosófico de la percepción como fuente de muchos de nuestros contenidos mentales. Así las cosas, podemos considerar dos dimensiones muy distintas de la percepción, una buena teoría del

contenido deberá articular de algún modo estos dos importantes componentes.

2.5. Sistemática, composicionalidad y holismo semántico

Los contenidos mentales no se tienen aisladamente, uno no tiene una sola idea en la cabeza. Aunque tampoco la complejidad de los estados mentales es tal que apenas podamos decir algo más allá de que "son complejos" o que todo depende del contexto, si bien es obvio que el contexto, lejos de ser un término vago, desempeña algún papel en la cuestión del contenido. Albergar un determinado contenido implica, entre otras cosas, la posibilidad de poder establecer relaciones con otros contenidos. El contenido tiene una estructura fuertemente holista y sistemática. Así, sistemática, composicionalidad y holismo son conceptos que deben aquilatarse suficientemente bien en un estudio acerca del contenido.

La propiedad común y más básica que tienen los conceptos de composicionalidad y holismo es la sistemática. De manera general decimos de algo que es sistemático cuando es un elemento que forma parte de una estructura mayor -sistema- y mantiene diversas relaciones o roles en esa misma estructura. Muchos objetos abstractos como los fonemas, las palabras o los conceptos tienen propiedades sistemáticas. La propiedad de ser sistemático consiste pues en ser parte de un sistema. De una manera más específica y centrada en la reflexión propia de la filosofía de la mente decimos que un símbolo cualquiera es una parte de un sistema más amplio de significados y no una entidad aislada que no pueda relacionarse con nada más. Son muchos los filósofos que han insistido en las propiedades sistemáticas tanto del lenguaje como del pensamiento como Frege¹³⁸ o Wittgenstein¹³⁹. Los símbolos -lingüísticos o mentales- no se tienen aisladamente y únicamente poseen significado en un contexto más amplio.

¹³⁸ Es el caso del denominado principio de contexto.

¹³⁹ Cf. Wittgenstein (1953: §199) "*Entender una oración significa entender un lenguaje.*"

Evans (1982) recoge esta idea aplicada a las creencias bajo el concepto de "Generality Constraint"¹⁴⁰. La restricción de generalidad da cuenta de cómo está estructurado el contenido en términos del ejercicio de habilidades cognitivas diferentes. Cuando pienso que mi gata Atina come también puedo pensar que Leo come o que duerme. En cada caso he ejercido una habilidad conceptual que consiste en la aplicación del mismo concepto "comer" tanto para Atina como para Leo. Así el predicado F puede aplicarse a diversos objetos como a , b , c , etc. Y estos mismos objetos pueden aparecer con diversos predicados F , G , H , etc.

Fodor y Lepore (1992), de manera más concreta y aplicada al lenguaje, señalan que un lenguaje natural es sistemático si cuando puede expresar una determinada proposición p también es capaz de expresar proposiciones semánticamente cercanas a p . Lo mismo es aplicable al contenido, si podemos atribuir a un sujeto el contenido "A está a la derecha de B" también podemos atribuirle el contenido "B está a la derecha de A". Por otra parte, que nuestro contenido sea sistemático es algo que está determinado por las propias condiciones para la posesión de conceptos como la mencionada restricción de generalidad. Finalmente, del hecho de que tanto el pensamiento como el lenguaje posean propiedades sistemáticas se deduce que podemos entender y producir un número potencialmente infinito de oraciones.

Otro aspecto fundamental del contenido que ya fue desarrollado por Frege (1879) es la composicionalidad del mismo. En líneas generales se puede caracterizar la composicionalidad, ya sea de un lenguaje verbal o del mismo lenguaje del pensamiento, diciendo que el significado de sus enunciados dependen de del significado de las expresiones que los componen¹⁴¹. Hay diversas maneras de interpretar este principio. Por ejemplo, para Frege (1879) la composicionalidad es

¹⁴⁰ "In discussing the nature of our conceivings we have little enough to go on, but there is one fundamental constraint that must be observed in all our reflections: I shall call it "The Generality Constraint".

It seems to me that there must be a sense in which thoughts are structured. The thought that John is happy has something in common with the thought that Harry is happy, and the thought that John is happy has something in common with the thought that John is sad. Evans (1982:100).

¹⁴¹ Cf. Fodor y Lepore (1992:258).

referida tanto al sentido como a la referencia de las expresiones complejas. El sentido de un enunciado así, está en función de los significados particulares de cada una de las expresiones y lo mismo sucede con la referencia. De este modo tenemos tanto el principio de composicionalidad del sentido como el principio de composicionalidad de la referencia. Pero no todos están dispuestos a admitir que debemos distinguir un sentido y una referencia en cualquier expresión lingüística y, cabe suponer, también en cualquier representación. Así proceden aquellos que se adscriben a corrientes neorussellianas. Se sea neofregeano o neorusselliano, en cualquier caso la cuestión de la composicionalidad puede ser abordada al margen de este tipo de disputas. Así, otra manera de entender esto es señalando que el significado de nuestras representaciones, en tanto que de algún modo éstas son eventos poseedores de alguna lingüisticidad, es también composicional. Las expresiones que componen un lenguaje –como pueda ser el mentalés- se combinan conforme a ciertas reglas de modo que a partir de un conjunto de símbolos primitivos o léxico se pueden obtener potencialmente infinitas secuencias de fórmulas o enunciados significativos en ese lenguaje. A esta propiedad la denominamos también generatividad. Dentro de la hipótesis del lenguaje del pensamiento podemos entender que el conjunto de conceptos es ciertamente finito, pero podemos pensar muy diversos pensamientos aplicando el mismo procedimiento: los conceptos se someten a la acción de las reglas de formación de secuentes en mentalés y dan lugar a nuevos pensamientos. El pensamiento o la actividad representacional también exhibe rasgos sistemáticos y su riqueza se explica a través de la composicionalidad. Decir que el pensamiento es composicional implica para algunos como Fodor la afirmación de que pensar consiste en computar símbolos mentales conforme a ciertos algoritmos. Otra cuestión es si los conceptos son innatos (i.e. Fodor 1975) o son adquiridos. Aunque este último tópico no lo abordaré.

Finalmente, el contenido mental parece exhibir propiedades holísticas, es decir, los contenidos mentales parecen estar fuertemente ligados unos de otros. Así, las teorías holistas del contenido mental, en líneas generales, sostienen que los contenidos de nuestras actitudes dependen constitutivamente de las relaciones inferenciales que estos

mismos mantienen con el resto de contenidos que alberga la mente de un determinado individuo¹⁴². El holismo semántico es una tesis que procede del campo de la epistemología. En epistemología se refiere al proceso de confirmación de las distintas teorías científicas. Así, según Quine (1953) la evaluación semántica de una teoría científica depende de toda la teoría y no se procede a la confrontación individual de los enunciados con la realidad. Es decir, no hay una correspondencia uno a uno entre los enunciados que componen la teoría y la clase de referencia de esa misma teoría. Del mismo modo en filosofía de la mente entendemos también que el proceso de fijación de las creencias es global; una creencia queda individualizada a partir del sistema global de creencias al que pertenece. De otra manera, el sistema de creencias es la unidad mínima del contenido intencional.¹⁴³

Al holismo semántico se opone el atomismo. El atomismo puede entenderse de diversas maneras. En lo que afecta a las representaciones mentales uno es atomista si cree que la referencia de una representación se determina autónomamente respecto de las referencias de otras representaciones mentales¹⁴⁴. Dicho de otro modo, implica asumir que todos nuestros conceptos son simples.

Las razones para decantarse por posicionamientos holistas son principalmente de orden epistemológico, es decir, relativas a la justificación de nuestras creencias. Sin embargo el problema principal del holismo semántico es que parece perder de vista la referencia de sus términos. Si lo que individualiza un término o una representación mental es su relación con otros términos o representaciones ¿cómo pueden referir a una realidad externa? ¿qué papel desempeña la referencia? La tensión entre atomismo y holismo no está resuelta. Admitir posicionamientos atomistas tiene la virtud de que permite explicar el contenido de las actitudes recurriendo únicamente a la sistematicidad y composicionalidad. De este modo el pensamiento podría entenderse como un juego *lego* en el que se ensamblan unas piezas con otras. Por otra parte, los posicionamientos holistas dan cuenta de la complejidad de algunas de las actitudes que pueden

¹⁴² Cf. Peacocke (1997: 227).

¹⁴³ Cf. Fodor (1987: 63).

¹⁴⁴ Cf. Neander (2006: 376).

albergar los organismos inteligentes; en estos se observa que el contenido de algunas creencias depende en gran medida de otras creencias relacionadas o que el significado de un concepto varía conforme al contexto en el que aparece. Con esto queda reflejado el primero de los *desiderata* que se contemplan en este capítulo: una teoría del contenido debe dar razón del carácter selectivo del mismo, el contenido de nuestras actitudes no parece reducible a la clase de referencia sobre la que en principio cabe suponer que versa. Así, fijando la referencia no fijamos el contenido.

Recapitulación

Al inicio de este capítulo presentaba los siguientes seis *desiderata* para una teoría del contenido:

D.1. Una teoría del contenido debe dar razón del carácter selectivo del contenido

D.2. Debe ser capaz de individualizarlo adecuadamente, es decir, debe dictaminar si el contenido se individualiza en base a propiedades relacionales o si, por el contrario, responde a una caracterización internista.

D.3. Debe dar razón de su dimensión normativa enfrentándose eficazmente a puzzles como el de la disyunción.

D.4. El contenido involucra aspectos tanto conceptuales como no conceptuales o cualitativos. Una teoría del contenido debería explicar que papel epistemológico desempeñan los aspectos cualitativos de la posesión de estados intencionales.

D.5. Debe explicar qué relación se establece entre contenido mental y contenido lingüístico

D.6. Debe dar cuenta de su dimensión práctica

Estos configuran así una agenda de problemas a los que debería enfrentarse una teoría del contenido. Para estos problemas no hay una única solución. Así, la pericia con la que una posible teoría del contenido enfrente o logre evitar esos problemas puede constituir en principio un buen síntoma de la idoneidad de esa teoría. A lo largo del capítulo he ido abordando de un modo u otro algunas de las diversas

teorías que se han vertido acerca de estos distintos *desiderata*. De manera más concreta he examinado las cuestiones relativas a la teoría representacional de la mente, las relaciones entre pensamiento y lenguaje como ejemplo de la caracterización de la estructura de las actitudes, el problema de la individuación del contenido, la normatividad o la dicotomía entre lo conceptual y lo no conceptual. Finalmente he hecho una breve alusión al problema de la sistematicidad, composicionalidad y holismo. El estudio de todos estos tópicos responden a cinco de los anteriores seis *desiderata*, puesto que la dimensión práctica, si bien es presupuesta, no se aborda por pertenecer a la teoría de la acción. De este modo, el objetivo de este capítulo era introductorio: presentar una caracterización general del contenido que permita una evaluación de las distintas teorías que sobre él existen.

V. La caracterización teleológica del contenido mental: Dretske, Millikan, Neander y Papineau sobre la fijación del contenido

Contenidos

- I. Introducción: la caracterización teleológica del contenido
- II. Estudio de cuatro teorías teleosemánticas
 - 1. La semántica causal o del indicador: Dretske
 - 1.1. La primera teoría causal de Dretske: cuando Dretske no era teleosemántico
 - 1.1.1. Información y conocimiento
 - 1.1.2. Información y contenido. La definición del contenido informacional de una señal
 - 1.1.3. Contenido informacional y actitudes proposicionales
 - 1.2. La caracterización teleológica del contenido en Dretske
 - 1.2.1. Tipos de sistemas representacionales
 - 1.2.2. La explicación del error
 - 1.2.3. El papel de los procesos de aprendizaje
 - 1.2.4. Cómo se fija el contenido
 - 2. Semánticas basadas en el consumidor *down up*
 - 2.1. La teleosemántica de Millikan
 - 2.1.2. Signos, intencionalidad y representaciones
 - 2.1.3. ¿Cómo fijar el contenido? Propiofuncionalidad y contenido mental
 - 2.2. Neander: *Low Church* frente a *High Church*
 - 2.2.1. La teleología *Low Church*
 - 2.2.2. La determinación del contenido
 - 2.2.3. La explicación del error
 - 2.2.4. Comentarios finales
 - 3. Papineau: el enfoque *top down*
 - 3.1. La alternativa *top down*
 - 3.2. La determinación del contenido
 - 4. Comentarios finales: Una comparación de las teorías expuestas en este capítulo

I. Introducción: La caracterización teleológica del contenido

En el capítulo anterior presenté una caracterización general sobre el contenido mental y clasifiqué los principales problemas que una teoría acerca del contenido debería abordar. Como es de suponer son diversas las teorías existentes a propósito del contenido¹⁴⁵. En este capítulo me centro en desarrollar aquellas propuestas que, de manera general, se engloban bajo la etiqueta de "teleosemántica". Denominaré de modo genérico "teorías teleosemánticas del contenido" a aquellas teorías naturalistas del contenido que de algún modo lo relacionan con el concepto de función biológica. Así el rasgo que diferencia a las teorías teleosemánticas de otras teorías acerca del contenido mental es el uso que hacen del concepto de función biológica. La teleosemántica no pretende explicar el contenido de un estado mental basándose en sus causas típicas, sino que apunta al diseño biológico o función que una determinada representación puede tener. Es decir, los contenidos representacionales de los estados mentales se explican en términos teleológicos, a saber, en términos de las funciones que dichos contenidos tienen dentro de la economía de un organismo vivo sometido a presión evolutiva. De otra manera, un estado mental tiene la función que tiene en virtud de una determinada historia selectiva.

"The central thesis of the teleological theory (I shall take 'of mental representation' as read from now) is that the representational contents of mental states can be explained in terms of the biological functions of those states. More specifically, this theory equates the contents of informational states (and in particular, therefore, beliefs) with those conditions in which those states are biologically supposed to be present. It further equates the contents of affective states (and in particular, therefore, of desires) with the conditions which those states are biologically supposed to produce." Papineau, (1998: 1).

Para las teorías teleosemánticas el contenido de las representaciones es fijado por las funciones biológicas que tienen los sistemas que las producen (Dretske) o las consumen (Millikan, Neander, Papineau). Para estas teorías el concepto de función

¹⁴⁵ Cf. Acero (1995).

biológica es fundamental porque permite dar cuenta de la naturaleza normativa del contenido. Así, según estos autores, la normatividad semántica depende de la normatividad funcional. La teleosemántica pues, se dirige de manera específica a abordar un problema: el de la normatividad del contenido y quizás, tal como señala Neander (2004) o Millikan (2004), sea entendida mejor antes como una estrategia general para dar cuenta de la normatividad del contenido que como una teoría alternativa del contenido mental en general. Por otra parte, hay que señalar que las teorías teleosemánticas son teorías informacionales. El recurso a los flujos informacionales es, en mayor o menor medida, el punto de partida para la naturalización de la intencionalidad junto con una serie de consideraciones de carácter ecológico que encuadran estos flujos informacionales.

Las teorías teleosemánticas del contenido forman una familia bastante amplia que, además de emplear el concepto de función biológica para fijar el contenido de las actitudes, comparten diversos rasgos:

- (1) Se presentan como teorías naturalistas; su objetivo es naturalizar la intencionalidad.
- (2) Son realistas intencionales.
- (3) Son partidarias de la teoría representacional de la mente pero no se comprometen con la hipótesis del lenguaje del pensamiento.
- (4) Defienden posicionamientos atomistas respecto de la estructura de las actitudes, es decir, no son partidarios del holismo semántico.
- (5) Están comprometidas con el externismo.

A su vez esta familia de teorías puede subdividirse en subfamilias. Neander (2004) establece la siguiente clasificación en cuatro subfamilias:

- (I) Teorías informacionales o semánticas basadas principalmente en el concepto de información.
- (II) Semánticas del consumidor o centradas en los mecanismos que consumen representaciones.
- (III) Teorías *top down* como la de Papineau.

(IV) Teorías combinatorias modestas o centradas únicamente en el contenido innato.

Históricamente las teorías teleológicas del contenido se han desarrollado a partir de varias fuentes, si bien la principal es la teoría causal-informacional del contenido (Stampe 1977; Dretske 1981) que, a su vez, está muy ligada a las teorías de la referencia directa o teorías histórico-causales de la referencia¹⁴⁶ (Kripke, Putnam). Éstas recurren a mecanismos causales para explicar la referencia de nombres propios y términos de clase natural sin apelar a la mediación de sentidos fregeanos. Estas teorías tal como fueron presentadas en su día por Kripke y Putnam no eran teorías naturalistas. Sin embargo, con el recurso al concepto de causalidad son susceptibles de ser naturalizadas, pues el concepto de causa es, en apariencia, un concepto científico básico. La teleosemántica adoptaría de la teoría de la referencia directa ciertos elementos antifregeanos señalando la importancia de la referencia al margen de los sentidos para la cuestión de la semántica de las representaciones mentales. Si bien, en lo que atañe a la naturaleza del contenido, aboga por una perspectiva naturalista e informacional. Las teorías teleológicas del contenido mental también son teorías naturalistas que tratan de rastrear la intencionalidad en el mundo natural rechazando la tesis de Brentano de que la intencionalidad es irreducible a fenómenos naturales. En este sentido, las teorías teleosemánticas entienden la intencionalidad y el significado como un elemento más de la naturaleza que debería ser explicado a través de las mismas teorías que empleamos para dar cuenta de los fenómenos naturales. Así, la teleosemántica sería una teoría acerca de la fijación del contenido; un intento de evitar problemas como el de la disyunción mediante el recurso a la noción de función biológica¹⁴⁷.

¹⁴⁶ Cf. Neander (2006: 377).

¹⁴⁷ La teoría de la dependencia asimétrica de Fodor (1984; 1990) es la gran rival de las teorías teleosemánticas pues tanto su objetivo como sus presupuestos informacionales y causales son los mismos que los de las teorías teleológicas.

En las páginas que siguen voy a proceder de la siguiente manera: En primer lugar expondré la semántica causal de Dretske y, en segundo lugar, las diversas teorías que componen la familia teleosemántica y hacia las que se dirigen las principales críticas que se abordarán en el capítulo VI. Mi exposición, próxima a la clasificación de Neander (2004), se centrará sobre cuatro autores: Dretske (en su faceta teleológica), Millikan, Neander y Papineau. He considerado que estos eran los más importantes. Sus aportaciones han configurado el panorama actual del enfoque teleológico del contenido mental. Si bien éstas han sido diversas y se han vertido sobre distintos aspectos de la Filosofía de la mente, mi exposición se centrará casi de manera exclusiva en la cuestión de la fijación del contenido.

II Estudio de cuatro teorías teleosemánticas

1. La semántica causal o del indicador: Dretske

Quizás uno de los primeros intentos de naturalizar el contenido sea el empirismo británico de Locke o de Hume. Si todo conocimiento procede de los sentidos y en la mente no hay nada innato, entonces el mundo exterior juega un papel determinante en la constitución de nuestra vida mental en general y, de manera más precisa, en la determinación de nuestros contenidos mentales. Aunque las obras de estos autores se centran en cuestiones epistemológicas, también constituyen un primer intento naturalista de explicar el contenido mental. Otro elemento que hay que destacar es el papel que juega la causalidad. Son las cosas -sea lo que sea una "cosa"- aquello que causa las impresiones de los sentidos y, por lo tanto, aquello de que depende el contenido de nuestras ideas. Aunque la semanticidad de nuestras representaciones no depende de conexiones causales, sino de relaciones de semejanza, para los empiristas, en última instancia, la semanticidad del contenido se justificaría en base a una relación de semejanza entre las representaciones y el mundo representado.

La idea original de los empiristas perduró durante mucho tiempo¹⁴⁸ como una teoría posible para explicar las relaciones entre la mente y el mundo. No obstante, esta sencilla explicación conlleva una serie de problemas que no la hacen adecuada. En líneas generales puede decirse que la idea de parecido es insuficiente para la caracterización de nuestro contenido en parte porque no todo nuestro contenido mental es de carácter imaginativo¹⁴⁹.

En el siglo XX ha habido intentos de explicar la semántica de nuestras representaciones que recurrían a las relaciones causales sin pasar por el concepto de similitud entre lo representado y la representación¹⁵⁰. La causalidad tiene a su favor que no parece a primera vista una noción muy problemática, sino todo lo contrario, parece que apelar a la misma dota de cientificidad al discurso filosófico, pues supuestamente las ciencias maduras explican basándose en relaciones causales. Por otra parte, el concepto de causalidad no es intencional, hay causas al margen de las intenciones; las causas naturales, dicen algunos, son objetivas.

Uno de los primeros filósofos en señalar que el contenido intencional podía explicarse en términos causales fue Dennis Stampe. El artículo de Stampe *Toward a Causal Theory of Linguistic Representation* aparecido en 1977 en los *Midwest Studies in Philosophy* de la Universidad de Minnesota, puede considerarse como el primer intento serio de desarrollar una semántica naturalista que se aparta de las caracterizaciones intencionales del contenido. La contribución de Stampe, como su propio título indica, se centra antes en las representaciones lingüísticas que en las representaciones mentales, es decir, es un texto más propio de la Filosofía del lenguaje que de la Filosofía de la mente, si bien en los momentos en que se escribe las fronteras entre ambas disciplinas empiezan ya a estar más bien difuminadas. Por otra parte el mismo texto es heredero de las reflexiones elaboradas en torno a las teorías de la referencia directa. La teoría causal de Stampe señala que las representaciones de un

¹⁴⁸ Mill y James son claros partidarios de este enfoque.

¹⁴⁹ Cf. el capítulo anterior a propósito de la crítica de Wittgenstein sobre la relación de representación.

¹⁵⁰ Cf. Fodor (1990: 33).

determinado tipo de referencia refieren a las causas de los ejemplares de ese mismo tipo.

Los problemas a los que se enfrenta la teoría causal de Stampe son varios¹⁵¹. Entre ellos el más importante es el conocido problema del error: si el concepto GATO sólo puede ser causado por los ejemplares de gatos ¿cómo es posible que exista el error? La teoría causal no deja lugar para el error y, sin embargo, el error es una constante en la vida de los organismos. Otro problema importante es el del contenido distal; la causación de nuestros ejemplares representacionales se produce a través de un continuo estimular en el que aparecen involucradas diversas causas. Por ejemplo, un ejemplar de GATO es posible porque de alguna manera los rayos lumínicos inciden sobre un gato particular que al impactar sobre los órganos de los sentidos elicitan un proceso fisiológico en el que se opera una transducción del estímulo hasta llegar a activar un percepto o concepto perceptivo que resulta ser el ejemplar GATO. Sin embargo, esta explicación no da cuenta de por qué lo representado es un gato particular y no los rayos lumínicos o la transducción nerviosa. Aunque parezca una perogrullada lo cierto es que desde un punto de vista causal no tenemos suficientes razones para aislar en el continuo estimular la causa determinante de nuestras representaciones. En tercer lugar nos encontramos con un problema ya clásico para las teorías del significado como la de Russell (1905) como es el problema de los términos (o representaciones) arreferenciales. ¿Cómo es posible instanciar representaciones de cosas que no existen y que por lo tanto son ineficaces causalmente? Por último, y como señalé en el anterior capítulo, el contenido es de grano mucho más fino que la referencia pero si debemos admitir que el contenido es causado por la referencia ¿cómo se explica ese aspecto tan específico? Las teorías causales y teleosemánticas explican la referencia de nuestras representaciones pero difícilmente, pueden dar cuenta del sentido, es decir, identifican lo que causa la activación de un determinado concepto pero ¿explican su modo de presentación? ¿Dan cuenta de la ocurrencia de tales conceptos en contextos oblicuos?. En cualquier caso cabe ver a las relaciones causales como insuficientes por sí solas para fijar adecuadamente el contenido.

¹⁵¹ Cf. Neander (2006: 378).

El trabajo filosófico de Dretske es amplio y variado, habiendo abordado diversos aspectos de la Filosofía como la teoría de la percepción, la teoría de la acción, o algunos aspectos de la Filosofía de la ciencia como pueda ser el tema de la explicación científica. Si bien, todos estos temas confluyen en un interés por la Filosofía de la mente y la Teoría del conocimiento, tópicos que aborda desde una perspectiva naturalista junto con una enconada defensa del externismo.

Su trayectoria filosófica responde a un plan muy bien trazado que tiene por objeto desarrollar una teoría del conocimiento de corte naturalista que le permite desarrollar paralelamente una Filosofía de la mente muy sólida. Su investigación epistemológica empieza con un texto sobre teoría de la percepción *Seeing and Knowing* (1969) en el que desarrolla una distinción ahora ya clásica entre ver no epistémicamente y ver epistémicamente. Las reflexiones de Dretske sobre la Teoría de la percepción constituyen una primera aproximación a la epistemología que se inserta en la discusión sobre este tópico de la tradición empirista y de autores más contemporáneos como Russell (1911) o Armstrong (1961) y que puede verse como una manera de plantear la cuestión de la no conceptualidad en el contenido perceptivo que sigue a día de hoy.

Dretske desarrolla un enfoque causal-informacional sobre el contenido mental que dirime la cuestión de la determinación del contenido en aquello que lo causa. No obstante, su filosofía de la mente depende de su teoría del conocimiento y desarrolla la primera como preludio para la segunda. Así, la motivación de sus investigaciones acerca del contenido mental hay que buscarla en el intento de desarrollar una teoría del conocimiento basada en el concepto de información, proyecto que acomete en *Knowledge and the Flow of Information* (1981). Por lo tanto, sus reflexiones sobre la caracterización del contenido mental en términos informacionales responden a un interés epistemológico y no exclusivamente a un tratamiento desde la Filosofía de la mente. Sus hallazgos en Filosofía de la mente también tienen una aplicación determinada a la teoría de la acción. Dretske investiga la dimensión causal del contenido en su *Explaining Behaviour* (1988). Finalmente, sus últimas contribuciones, en un claro intento de caracterizar la vida mental al completo, afectan al

tópico de la conciencia y aparecen desarrolladas en su *Naturalizing the Mind* (1995).

El valor de la obra de Dretske para la caracterización teleológica del contenido mental es determinante. De hecho, las teorías teleosemánticas en general pueden ser vistas como un intento igualmente naturalista de hacer frente al problema de la disyunción que, cabría suponer, afecta de manera decisiva a la propuesta de Dretske a propósito del contenido. Así las cosas, puede verse la teleosemántica como un refinamiento de la naturalización del contenido acometida por Dretske que, en principio, apelaría al concepto de función biológica para encarar el problema de la disyunción.

1.1. La primera teoría causal de Dretske: cuando Dretske no era teleosemántico

1.1.1. Información y conocimiento

Dretske desarrolla en *Knowing and the Flow of Information* (1981) una teoría naturalista del conocimiento basada en el concepto de información. La información, a diferencia de otros conceptos, no es intencional, sino objetiva, puede medirse y es, por lo tanto, la candidata óptima para acometer una naturalización de la epistemología; el conocimiento, a fin de cuentas, consiste en ciertos flujos de información. Así, el recurso a la noción de información viene motivado por su interés en elaborar una teoría naturalista del conocimiento y en ofrecer una explicación de la noción de justificación epistémica que tan problemática resulta tras los contraejemplos de Gettier (1963) para la definición tripartita tradicional de conocimiento. El recurso a la noción de información no es sólo original porque provea una peculiar solución al problema de Gettier, sino también por la naturalización de la epistemología que conlleva¹⁵². Sin embargo, su propuesta no hay que

¹⁵² En este sentido es un intento muy original si tenemos en cuenta qué naturalizaciones de la epistemología tenemos en el mercado filosófico. Podemos considerar, entre otros casos, la epistemología genética de Piaget o la teoría evolutiva del conocimiento de K. Lorenz (1941) o, más próxima en el ámbito analítico, la propuesta quineana dominada por una interesante reelaboración del pragmatismo americano desde el punto de vista del conductismo.

situarla únicamente en una naturalización de la epistemología, sino también en una naturalización de la intencionalidad a la que nos vemos llevados inmediatamente por la caracterización de las actitudes proposicionales en términos informacionales.

Los planteamientos de Dretske están inspirados en la teoría de la información desarrollada por Shannon (1948) y Shannon y Weaver (1949) cuya influencia en las ciencias cognitivas ha sido considerable¹⁵³. Dretske aplica la teoría de la información desarrollada por Shannon y Weaver a determinados problemas filosóficos como la justificación epistémica y el análisis de las actitudes proposicionales en términos informacionales. Estos dos tópicos centrales son desarrollados a partir de una teoría semántica de la información y a partir de estos dos podemos dar cuenta de otros aspectos más prácticos del contenido como es la explicación de la conducta.

La teoría matemática de la comunicación o teoría de la información tiene que ver con la cuantificación de información que es transmitida a través de un canal en un determinado sistema informacional. Desde el punto de vista de esta teoría la información se define como la reducción de incertidumbre respecto a un estado de cosas dado; a mayor información, menor incertidumbre. Las relaciones informacionales que median en los distintos sistemas informacionales son cuantificables y susceptibles de ser tratadas matemáticamente.

1.1.2. Información y contenido. La definición del contenido informacional de una señal

La teoría de Shannon no contempla ni la definición del concepto de información, pues aborda a ésta como una magnitud, ni consideración alguna sobre el contenido, ya que es innecesaria su consideración desde el punto de vista del tratamiento matemático de la información. Si bien Dretske no se centra tanto en el canal de información como en

¹⁵³ A partir de la formulación de la teoría de la información muchos psicólogos descontentos con el paradigma conductista dominante empezaron a estudiar la actividad cognitiva humana en términos informacionales -i.e. Broadbent (1958) y su modelo atencional-, consideraron que los sentidos son receptores de información y el material sobre el que se ejecutan los diversos procesos psicológicos es la información.

las señales consideradas aisladamente. De ahí que insista en la importancia de la definición del contenido informacional de una señal. Este rasgo desde el punto de vista de la teoría de la información podría parecer una anomalía, pues la teoría no estudia las señales individuales, sino el flujo de las mismas a través del canal. En el caso de Dretske esta estrategia se justifica porque sólo así podemos abordar el tema del contenido, ya que con muchas señales circulando por el canal no podemos saber nada sobre qué contienen estas.

La definición de contenido informacional que ofrece Dretske es la siguiente: Sea r una señal cualquiera, y k el conocimiento previo del receptor sobre las probabilidades de que s sea o no F , diremos que:

Informational content: A signal r carries the information that s is F = The conditional probability of s 's being F , given r (and k), is 1 (but, given k alone, less than 1).

Dretske (1981: 65).

Un aspecto de la definición de contenido informacional sobre el que merece la pena detenerse es la forma en la que se expresa este mismo contenido, s es F , puesto que la definición expresa el contenido informacional *de re* de la señal y no *de dicto*. La razón es que desde un punto de vista cognitivo el primero tiene un carácter más básico que el segundo. Un contenido *de re* no necesita de la existencia de otros seres racionales para determinarse ni requiere de habilidades más complejas como la comunicación o el lenguaje. Así la simplicidad cognitiva es obvia si consideramos que s es un elemento indéxico o demostrativo que denota un particular en la fuente informacional y al que le adscribe un determinado predicado F cuya misión es reflejar el contenido informacional de la señal. Es decir, sin el predicado F la señal no entraría a formar parte de lo conceptual y no podría ser un elemento constitutivo de una determinada actitud proposicional. En este sentido el contenido informacional *de re* recoge la idea de que una determinada representación lleva la información de que s es F . Como se puede suponer no nos referimos sino al individuo mismo nombrado por el indéxico y no a nuestra manera de describirlo. El acceso epistémico que podamos tener hacia s está determinado

perceptivamente; s es un objeto perceptivo y no un concepto¹⁵⁴. Dicho de otra manera, s es F nos dice que un determinado individuo satisface un predicado F o, lo que es lo mismo que un concepto es adscrito a un particular o , en otras palabras, que éste es subsumido bajo un concepto. Sólo así, siguiendo el *dictum* kantiano, cabría la posibilidad de un pensamiento sobre la realidad. El contenido de la proposición s es F se determina tanto en base a la existencia del particular denotado como por la información que s acarrea sobre el particular y que permite satisfacer el predicado F . Pero ¿de qué depende F y dónde está? Un concepto que quizás pueda ser útil para responder a esta cuestión es el requisito que exige que el sujeto tenga un determinado conocimiento K sobre las posibilidades de la fuente. Así, una señal cualquiera tiene un determinado contenido informacional de que s es F si la probabilidad de que sea esto el caso es igual a la unidad *caeteris paribus* k , donde k es el conocimiento que tiene el receptor acerca de la fuente de información. La información siempre es relativa al conocimiento que posee el receptor acerca de la fuente de emisión. Llegados a este punto surge un importante problema. Si el conocimiento procede o se justifica en base a los flujos informacionales ¿no supone una petición de principio el hacer referencia a nociones epistémicas para justificar el propio conocimiento? De una manera más concreta, para poder adscribir un predicado a un determinado particular es necesaria la posesión previa de ese predicado o concepto pero ¿cómo se posee?. Este tipo de cuestiones serán abordadas de una manera más adecuada a través del concepto de función en su Dretske (1988), aunque para entonces ya habrá adoptado un punto de vista más teleológico.

Aunque objetiva, la información no es unívoca; el contenido informacional de una señal no es único, ya que la propia señal puede llevar incorporados otros contenidos, aunque en cualquier caso depende de una correlación nomológica. La información puede estar incorporada en virtud de leyes naturales o de leyes lógicas. En el primer caso diremos que la información está nómicamente incorporada,

¹⁵⁴ Dretske no es partidario de una teoría de la percepción de la creencia pura, es decir, una teoría que considera que todos los elementos presentes en la percepción son de carácter conceptual, sino que separa los conceptos de *sentience* y *sapience*.

mientras que en el segundo lo está de modo analítico. Una misma señal conlleva una variedad indefinida de contenidos informacionales, de hecho, aunque todas esas informaciones puedan guardar algún tipo de relación entre sí, es obvio que son diferentes informaciones. Formalmente podemos expresar esta idea diciendo que una señal r porta la información de que s es F pero también de que s es G o de que t es G , dado que existe alguna relación entre ser F y ser G . Por ejemplo, mi percepción de que el objeto que está delante de mí es una naranja incorpora la información de ser una naranja, ser de color naranja y ser redonda o ser comestible, etc. Que se escoja un determinado contenido y no otro dependerá del conocimiento K que tenga el sujeto, así como de sus motivaciones o necesidades en ese momento.

1.1.3. Contenido informacional y actitudes proposicionales

Hasta el momento he desarrollado la idea de un contenido informacional. Con este contenido informacional tenemos ya una forma de intencionalidad muy básica y natural. Sin embargo, el contenido presente en nuestros pensamientos es algo mucho más complejo. Dretske (1986: 179) señala que las creencias son estructuras que tienen un contenido semántico¹⁵⁵. El contenido semántico de una estructura se define como la porción de información que está en formato digital, es decir, conceptualizada. La información en formato digital exhibe un alto grado de intencionalidad respecto de la información que pueda estar contenida también en la señal.

Las creencias se generan a través del modo de codificación de la información por el sistema. Así, por ejemplo una señal lleva la información de que a es un cuadrado (Fa) de forma digital y no análoga. La señal también conlleva otras informaciones que están presentes de forma analítica como que a es un paralelogramo, o que a tiene cuatro lados. Todo sistema manipula señales en la medida en que está inmerso en flujos informacionales, cuando las señales manipuladas por el sistema tienen un formato digital decimos que el sistema alberga creencias, el contenido de estas creencias es

¹⁵⁵ Cf. también Dretske (1981:196).

precisamente la porción de la información de la señal en formato digital. Así el contenido semántico define el contenido doxástico. Las estructuras semánticas o creencias poseen propiedades como su sensibilidad a determinados fragmentos de información y no otros. Es decir, lo que cae dentro del concepto es aquello que puede ser conceptualizado y lo que cae fuera no, así hay un determinado rango de estímulos que serán conceptualizados y otros no, aunque estén presentes en la misma fuente informacional. Por ejemplo, una señal que al ser procesada por el sistema es conceptualizada bajo el concepto CUADRADO tal que *Fa* forma una representación en formato digital en el sistema cuyo contenido es *Fa*. Llamemos a esta representación *R*. *R* no es sensible por ejemplo al color del cuadrado, cuadrados rojos, azules o amarillos producirán *Fa* con independencia del color, el color es causalmente irrelevante respecto de *R*. También *R* es insensible a la información incorporada en el objeto en tanto que cuadrado, por ejemplo su propiedad de ser paralelogramo. Así las propiedades que causan *R* no son las que, por ejemplo, causan *S*. Estas propiedades hacen que el contenido de aquello que causa *Fa* sea "*Fa*" y no "*Ga*". Dicho de otra manera, así se asegura la selectividad del contenido y nos permite diferenciar el simple acto de ver de la percepción¹⁵⁶. Así, el proceso que va desde las señales a las creencias puede esquematizarse del siguiente modo:

Señal → Representación Análoga → Representación Digital
→ Contenidos Semánticos

Una señal informacional no es una estructura simple, sino que puede ser vista como un conjunto de círculos concéntricos. Cada uno de los círculos en un fragmento de información incorporada a la señal.

¹⁵⁶ Cf. Dretske (2000: 97 y ss.). En Dretske (1981: 181-182) se señala que: "*The process just described, the process of generating different semantic structures from one and the same internal analog representation, is meant to correspond to the way the perception of a red square (internal primary representation of the red square) can yield different beliefs about the object (different semantic structures) depending on the background, experience, training, and attention of the subject. The kind of semantic structure evoked by the incoming signal determines how the system interprets what it perceives(...).*"

Por lo tanto, el contenido semántico de una estructura no es único, ya que una señal puede conllevar más información incorporada de manera nómica o analítica. Pero el contenido semántico sólo tiene que ver con la capa -el último círculo- más externa de la señal que está completamente digitalizada. Esta matización nos aporta un criterio de relevancia en la información. Por ejemplo, cuando un sujeto está leyendo un libro recibe mucha información; no sólo recibe algo discreto, sino que el flujo informacional conlleva otros elementos como sonidos ambientales, matices lumínicos, la textura del papel, otras conversaciones que se cruzan, etc. Pero lo que el sistema informacional acepta como contenido semántico no es eso, sino lo que tiene un formato digital¹⁵⁷, es decir, el contenido de la lectura y no los otros elementos. Esta distinción exhibe la diferencia existente entre lo conceptual y lo no conceptual en la experiencia perceptiva. Obviamente que sea lo conceptual dependerá también de los recursos atencionales o del conocimiento previo del sujeto. Así si el sujeto quiere leer, el contenido conceptual es la lectura. Pero si es Sherlock Holmes o el Padre Brown puede ser cualquier otra cosa si el libro en cuestión ha pertenecido a Moriarty o Flambeau, es decir, algo no es intrínsecamente conceptual o no conceptual, sino conceptualizable o no conceptualizable. La generación de nuestras creencias depende del trabajo de nuestros procesos cognitivos sobre los datos de los sentidos¹⁵⁸. No todos los sistemas son capaces de albergar contenido

¹⁵⁷ La diferencia entre representaciones análogas y digitales puede ser ilustrada mediante la siguiente analogía: una representación análoga es equivalente a una imagen, mientras que una representación digital se asemeja más a un enunciado. De otra manera, lo análogo tendría que ver con lo no conceptual y lo digital con lo conceptual, con aquello que puede ser parte de un enunciado. Dretske (1981) define estos conceptos de la siguiente manera: "*I will say that a signal (structure, event, state) carries the information that s is F in digital form if and only if the signal carries no additional information about s, no information that is not already nested in s's being F. If the signal does carries additional information about s, information that is not nested in s's being F, then I shall say that the signal carries this information in analog form.*" Dretske (1981:137).

¹⁵⁸ "*What endows some systems with the capacity to occupy states which have, as their semantic content, facts about some distant source is the plasticity of the system for extracting information about a source from a variety of physically different signals.*" Dretske (1981:187).

conceptual (o semántico), éste sólo es posible para aquellos sistemas que exhiben un alto grado de plasticidad.

Lo dicho anteriormente no explicaría por qué podemos albergar creencias falsas, ya que la información que lleva una señal no puede ser falsa por definición. Si el significado de una estructura -creencia- depende de su origen informacional, entonces puede suceder que el tipo de estructura sea verdadero y los ejemplares no respondan a las mismas condiciones de verdad. Dretske (1981: 193-194) señala que ocurre lo mismo con el aprendizaje de conceptos: uno aprende que un concepto designa a un individuo por ostensión, por ejemplo "esto es un gato" pero al aplicar a otros individuos el mismo concepto puede fracasar. El concepto tipo obviamente es verdadero, pero ¿lo son los ejemplares subsiguientes? No, son ejemplificaciones erróneas. El error así depende de la competencia del sujeto¹⁵⁹. Bien la creencia bien el concepto pueden ser activados por información a la que son sensibles pero que no se corresponde con la clase de referencia del concepto o de la creencia. Obsérvese que la distinción podría analizarse igualmente en términos de conocimiento directo y conocimiento por descripción.¹⁶⁰ No hay conocimiento directo erróneo del mismo modo que no hay información errónea, pero sí puede haber conocimiento descriptivo- proposiciones en las que aparecen involucrados conceptos y particulares- erróneo. Como mostraré más adelante esta explicación del error se ve enriquecida en 1988 con la aportación del concepto de función biológica a la caracterización de las actitudes¹⁶¹.

¹⁵⁹ Cf. Los comentarios de Fodor (1987: 103) respecto de la explicación del error según Dretske (1981). Según Fodor la explicación de Dretske centrada en los procesos de aprendizaje sería insuficiente puesto que no hay una manera clara de determinar cuándo un sujeto está aprendiendo un concepto y cuando ya lo ha aprendido. Además, durante el periodo de aprendizaje podemos encontrarnos con casos de disyunción, es decir, en virtud de la robustez de las representaciones podemos encontrarnos con el caso de que no todas las instanciaciones de A representen el objeto A, sino que sean causadas por los objetos A o B.

¹⁶⁰ Cf. Russell (1912).

¹⁶¹ Hay que señalar que en Dretske (1981) encontramos de algún modo las ideas seminales que desarrollará en su texto de 1986. Así se hace depender el contenido de una creencia de su etiología y es en base a ésta a partir de la que se explica el error. "*A certain type of structure acquires its content, the sort of content we associate with belief, by its informational origins. An internal structure develops*

1.2. La caracterización teleológica del contenido en Dretske

Uno de los problemas principales a los que se enfrenta la teoría anteriormente expuesta es que es incapaz de dar razón de la representación errónea ya que las representaciones de GATO son causadas por los gatos. Si bien la teoría planteada a partir de 1988 no es una nueva teoría sí cabe decir que Dretske reelabora de algún modo los planteamientos iniciales de 1981 explotando los recursos disponibles en lo que Fodor (1987) ha denominado la teoría causal cruda (*The Crude Causal Theory*). Así, Dretske se aproxima a planteamientos más acordes con la teleosemántica en su *Explaining Behaviour* (1988), donde reformula su teoría del contenido mental de 1981 basándose de alguna manera en el concepto de función biológica para evitar el problema de la disyunción. En mi opinión es una cuestión sólo de matices acerca de cómo se fija el contenido y qué dependencia tienen los distintos estados mentales respecto de las funciones biológicas, el aprendizaje o las convenciones. El rasgo realmente novedoso es la mención a las funciones biológicas.

La teoría defendida en *Explaining Behaviour* evidencia un realismo intencional neutral ante las tesis de la hipótesis del lenguaje del pensamiento¹⁶². De hecho, este posicionamiento está justificado desde la defensa explícita del externismo que hace Dretske y que le impediría aceptar puntos de vista individualistas como la hipótesis del lenguaje del pensamiento.

1.2.1. Tipos de sistemas representacionales

Dretske (1986) parte de una concepción representacional de la mente y justifica la existencia de representaciones a partir de la

(during learning) as a system's way of completely digitalizing information about, say, the F-ness of things. This way of encoding information as the semantic content of an output-determining structure) makes the information so encoded relevant to explaining the system's behaviour." Dretske (1981:201).

¹⁶² Cf. Sterelny (1990:40)

necesidad de explicar la conducta. Así, la explicación de la conducta de un organismo depende de una atribución de estados mentales que son causa de esa conducta.

Las representaciones se articulan en sistemas representacionales cuyo propósito es indicar o referir a estados de cosas en el mundo. Los sistemas representacionales se clasifican en tres grupos. Así tenemos Sistema de tipo I, o sistemas convencionales cuyos elementos representacionales son símbolos. Los sistemas de tipo II emplean alternativamente signos y símbolos, es decir, están a medio camino entre la convencionalidad que sería el rasgo determinante de nuestra conducta representacional lingüística y el significado natural. Finalmente los sistemas de tipo III son sistemas naturales de representación. Por tanto la clasificación va desde el significado no natural al significado natural, siguiendo la terminología griceana. El contenido de los distintos tipos de sistemas representacionales se define a partir de la funcionalidad biológica de los mismos que, según Dretske (1986: 77) consiste en indicar. Sin embargo, indicar no es lo mismo que representar. El término "representación" se asemeja al significado no natural de Grice (1957) y, tal como estipula el propio Grice, pueden no satisfacerse las condiciones para que un hablante no logre significar no naturalmente que p , en tal caso estamos ante una situación de representación errónea (*misrepresentation*). No es posible, por el contrario, una relación de indicación errónea, la indicación es siempre verdadera.

Los sistemas convencionales de representaciones de tipo I

En estos sistemas representacionales, las representaciones no poseen poderes representacionales intrínsecos; no son signos naturales, sino signos no naturales o símbolos. Su intencionalidad no es originaria, sino derivada de las convenciones que regulan su uso. Ejemplos de estos símbolos son las letras del abecedario o las marcas viales. Nuestra conducta lingüística y nuestros contenidos abstractos dependen de este tipo de sistemas. Su funcionalidad consiste en indicar pero ésta deriva de nuestros propósitos comunicativos.

Los sistemas convencionales de representaciones de tipo II

Estos sistemas ya no manipulan símbolos, sino signos naturales. Estos sistemas no representan lo que sus signos componentes indican aisladamente, sino aquello para lo que tienen una función. Son ejemplos de estos sistemas los instrumentos de medida como termómetros, barómetros, velocímetros, indicadores de presión, etc. La relación de indicación es en parte natural y en parte convencional. Por ejemplo, un termómetro mide la temperatura mediante una correlación nomológica entre la altura de la columna de mercurio y la temperatura ambiental en una escala puramente convencional. Un termómetro de mercurio sólo puede medir la temperatura, no la felicidad o la presión atmosférica. El tamaño de la columna es un signo natural como lo es el color amarillento de la piel en los casos de ictericia; los grados Celsius o Kelvin son materia de convención. La funcionalidad que atribuimos a los sistemas de tipo II es en parte sobrevenida a la relación de indicación y en parte a nuestros propósitos comunicativos; así en primer lugar tenemos el poder indicador intrínseco de los distintos elementos y luego el propósito que adscribimos al sistema. Son sistemas en los que se mezcla lo artificial con lo natural. De este modo adscribimos funciones representacionales a signos naturales. Aunque estos sistemas no están exentos de error. Pueden fracasar si no se cumplen las condiciones *caeteris paribus* bajo las que son diseñados o pueden estar mal calibrados.

Los sistemas naturales de representación. Sistemas de tipo III

La semanticidad de estos sistemas es completamente intrínseca a los signos que lo componen pues sus signos poseen sus propias funciones que derivan del desarrollo de los indicadores y del uso que hace de los mismos el sistema. En estos sistemas nos encontramos con una fuente originaria de intencionalidad¹⁶³. Este tipo de sistema es por ejemplo, heredado en los animales y forma parte de sus repertorio

¹⁶³ La intencionalidad de los sistemas de Tipo II y Tipo I dependería de alguna manera de esta intencionalidad originaria y sería construida a través del desarrollo ontogenético del individuo y de los distintos procesos de aprendizaje.

de conductas estereotipadas. Ejemplos de estos sistemas los encontramos en las reacciones aversivas de mamíferos como los felinos a través del mecanismo piloerector, la conducta de apareamiento de las aves o las señales de feromonas captadas a través del órgano vomeronasal presente en roedores. Las representaciones que producen estos sistemas forman parte de la herencia biológica del organismo del animal. Del mismo modo a estos sistemas les atribuimos funciones biológicas en el sentido de que está muy claro qué es lo que deben hacer. Si bien, en este último respecto, Dretske (1986:63-64) es muy cauto y no se pronuncia acerca de cómo deba interpretarse el concepto de función biológica y admite que existan funciones derivadas de la evolución de la especie y otras derivadas del desarrollo del animal. La no falibilidad de este tipo de sistemas se podría explicar como una causa del éxito adaptativo de las diferentes especies que los poseen.

1.2.2. La explicación del error

El error no es posible en la relación de indicación. El error tiene que ver únicamente con las representaciones que se construyen sobre los indicadores o signos naturales. Para que una determinada representación sea errónea debe pertenecer a sistemas de tipo I o de tipo II. Un caso de representación errónea en sistemas de tipo I lo encontraríamos cuando no se satisfacen por ejemplo las condiciones para que una emisión lingüística posea un significado no natural o por una falta de pericia en la manipulación y uso de los símbolos pertenecientes al sistema de tipo I. En el caso de los sistemas de tipo II podemos encontrar con situaciones erróneas referidas al componente simbólico convencional pero no a la relación de indicación. Por ejemplo, puede suceder que el termómetro o cualquier otro sistema análogo esté mal calibrado o que simplemente funcione de modo erróneo por un error de diseño. En cambio no nos encontramos con signos naturales que indiquen falsamente, no hay humo sin fuego como no hay infección sin fiebre o sarampión sin manchas, salvo casos anómalos que no prosperan evolutivamente. Pero estos sistemas pueden fracasar si se localizan fuera de lo que es su medio habitual o normal. Por ejemplo, una brújula sobre un imán no señala el Norte, la

rana disparará igualmente su lengua ante moscas que ante pseudomoscas en una situación experimental -una situación por definición no natural- igualmente podemos inducir fiebre alta a un paciente sin necesidad de infectarlo, por ejemplo, administrándole anfetaminas. En cualquiera de estos casos está presente una anomalía que hace que el mecanismo productor de representaciones no funcione adecuadamente¹⁶⁴.

1.2.3. El papel de los procesos de aprendizaje

Gran parte del repertorio conductual de humanos¹⁶⁵ y de algunos otros animales es aprendido de un modo u otro, bien sea simple condicionamiento reflejo o asociativo o ya involucre condicionamiento operante o aprendizaje vicario. Aunque cuando se habla de aprendizaje se entiende siempre condicionamiento operante. En muchos casos la conducta es estereotipada y depende de la existencia de mecanismos cuya función es indicar ciertos elementos en el medio. Así, por ejemplo, parece que algunas aves reconocen siluetas como peligrosas y éstas elicitan su conducta de alerta y huida. En tales casos el mecanismo indicador indica lo que indica porque ésta es su función. Pero no nos encontramos con una dotación para el reconocimiento de los

¹⁶⁴ Dretske (1986: 69) señala que los casos de representación errónea dependen de dos factores: la condición bajo la que se representa el sistema el mundo y el modo en que se lo representa. Más concretamente: "*Misrepresentation depends on two things: the condition of the world being represented and the way that world is represented. The latter, as we have seen, is determined, not by what a system indicates about the world, but by what it has the function of indicating about the world. And as long as there remains this indeterminacy of function, there is no clear sense in which representation occurs.*"

¹⁶⁵ Dretske señala inequívocamente el papel que desempeña el aprendizaje en la conducta humana. Explicar el contenido de las actitudes que desencadenan la conducta implica explicar el papel de los procesos de aprendizaje en su génesis: "*The reason learning is so central to intelligent behaviour, to the behaviour of people, is that learning is the process in which internal indicators (...) are harnessed to output and thus become relevant -as representations, as reasons- to the explanation of the behaviour of which they are part. It is in the learning process that information-carrying elements get a job to do because of the information they carry and hence acquire, by means of their content, a role in the explanation of behaviour.*" Dretske (1986:104).

elementos específicos que constituyen una caja de Skinner. El animal únicamente posee un buen funcionamiento de sus sentidos que le permiten discriminar y, posteriormente por ensayo y error o por algún otro procedimiento de solución de problemas, que, por ejemplo, quedaría explicado por la teoría de la percepción ecológica o por los modelos de solución de problemas, el animal asocia una respuesta motora con la obtención de ciertas consecuencias. Así la conducta es posible porque en primer lugar se detectan ciertas condiciones presentes en el medio. En algunos casos la detección de las condiciones que elicitan una conducta son innatas, en otros, aprendidas. En estos últimos se aprovecha lo innato para construir otro tipo de conductas más complejas mediante los procesos de aprendizaje. Los procesos de aprendizaje introducen una nueva normatividad en las representaciones que controlan la conducta al definir el papel causal de ciertos mecanismos de identificación y asociarlos con respuestas conductuales beneficiosas para el organismo¹⁶⁶.

Una de las limitaciones del estudio de Dretske acerca de los procesos de aprendizaje es que sólo explora el papel del condicionamiento operante. Sin embargo, los procesos de aprendizaje vicario también desempeñan un papel muy importante en la producción de conductas complejas.

1.2.4. Cómo se fija el contenido

La asignación de funciones a los distintos tipos de sistemas representacionales considerados debería ser suficiente para la fijación del contenido. Así, una representación tiene el contenido que tiene en virtud de que representa aquello que debe representar. Si la representación es un símbolo su contenido viene determinado por el propósito que convencionalmente los miembros de la comunidad han

¹⁶⁶ "An animal cannot learn to behave in the prescribed way -it cannot learn to coordinate its output (*M*) with condition *F* -unless an internal indicator of *F* is made into a cause of, a switch, for *M*. This is why learning of this sort must recruit indicators of *F* as causes of *M*." Dretske (1986:101) *F* es un estado de cosas en el mundo y *M* una conducta.

estipulado que posea. La fijación del contenido de sistemas de tipo I como el lenguaje hablado depende de convenciones. En los sistemas de tipo II y tipo III el contenido se fija en base a flujos informacionales nomológicamente establecidos. En los sistemas de tipo II las funciones que permiten fijar el contenido son muy similares a aquellas que adscribimos a los artefactos y dependen de las capacidades intrínsecas a los mismos dentro de un plan de diseño¹⁶⁷. Aunque particularmente interesante es el caso de la fijación del contenido de los sistemas de tipo III, aquí la relación de indicación es, como ya he apuntado, puramente natural. En el caso de los organismos depende de los propósitos naturales diseñados por el proceso evolutivo, es decir, depende por entero de la funcionalidad biológica de la que están dotadas las representaciones que producen. Ejemplos de estos sistemas pueden ser los mecanismos exteroceptores, o los mecanismos de orientación de algunos animales. Los sistemas de representación de tipo III sólo pueden arrojar un contenido indeterminado; pues el contenido determinado sólo puede surgir en sistemas cognitivos complejos en los que están presentes procesos de aprendizaje.

Una limitación de la estrategia dretskeana es que sólo podemos atribuir a un organismo la referencia del contenido y no el sentido, es decir, podemos identificar el tópico y no el comentario, empleando su propia terminología. Si bien es éste un problema común en todas las teorías causales.

2. Semánticas basadas en el consumidor *down up*

Millikan y Neander son dos autoras que ya han aparecido en el capítulo dedicado a las funciones como proponentes de sendas teorías histórico-etiológicas. Sin embargo, como ya adelanté, su contribución

¹⁶⁷ Respecto de Wright, Dretske señala lo siguiente: "*The functions we assign to instruments are similarly restricted to what the instruments can do, or if Wright (1973) is correct, what (in the case of artifacts) we think they can do. This makes the functions of type II restricted in a way that those of Type I systems are not restricted. It is this fact, together with the fact that once a device has given such a function it performs without any help from us, that makes such systems only partly conventional.*" Dretske (1988: 60-61).

principal y la razón de que elaboren las citadas teorías responde a un interés semántico. La teoría de las funciones que ambas autoras desarrollan es el punto de partida de su semántica que, propiamente, puede ser definida como teleosemántica, ya que las consideraciones a propósito de la naturaleza de las representaciones mentales están en conexión directa con el valor biológico que éstas exhiben y que se codifica en el concepto de función. Es decir, un enfoque puramente teleológico del contenido de las representaciones mentales. Ambas teorías intentan perfeccionar los planteamientos de la semántica causal-informacional y responder a los principales problemas que éstas conllevaban, como pueda ser el problema de la indeterminación del contenido. Así, la teleosemántica puede ser vista como la continuación natural y mejorada de los planteamientos iniciales de la semántica informacional en la que, por otra parte, se apoya. El término *down up* viene motivado porque estas autoras parten de las funciones biológicas de los sistemas representacionales para dar cuenta del contenido de deseos y creencias. Es decir, deseos y creencias sobrevienen a las funciones de los sistemas que comercian con estas entidades.

2.1. La teleosemántica de Millikan

En el capítulo II hice una escueta presentación de los principales trabajos de Millikan y señalé que su dedicación al estudio de la teoría de las funciones estaba subordinado a intereses más ambiciosos pertenecientes de lleno a la Filosofía de la mente o del lenguaje. En efecto, ninguna de sus obras es propiamente una obra de Filosofía de la Biología, antes bien, la temática que se aborda en todas ellas tiene que ver con las dos disciplinas referidas anteriormente, en las que demuestra un amplio y profundo conocimiento. En esta sección incido sobre el tratamiento que esta autora hace del contenido mental.

Los posicionamientos de Millikan a propósito del tema del contenido mental aparecen ya elaborados en su *Language, Thought and Other Biological Categories* (LTOBC) de 1984, su primera obra y punto de referencia de toda su carrera posterior. LTOBC es un curioso texto en el que la primera parte del mismo está dedicada a una teoría de las funciones biológicas y la segunda y la tercera a temas de semiótica y filosofía del lenguaje. LTOBC es la causa de que se relacione el

concepto de función biológica con la filosofía de la mente y del lenguaje. En este sentido es una contribución original ya que nadie antes había relacionado de esta manera tan íntima y de modo tan detallado cómo se podían imbricar ambos conceptos. LTOBC en lo que no toca a las funciones biológicas detalla las nociones de sentido, referencia, verdad, intensión, índices, descripciones, cuantificadores y conectivas lógicas, o la teoría de la identidad. Es decir, nos encontramos con un elenco de gran parte de los principales tópicos a los que se ha dedicado la Filosofía del lenguaje en buena parte del siglo XX. La presencia del tema del contenido no resulta extraña si se tiene en cuenta que algunas de las aportaciones de Millikan a la Filosofía del lenguaje pueden ser vistas como una naturalización del programa de Grice. Sin embargo, lo que la diferencia de éste último no es sólo la naturalización, sino el externismo extremo que defiende. En estos dos aspectos coincide plenamente con Dretske.

Su *White Queen Psychology and Other Essays for Alice* (1993) recoge una serie de artículos aparecidos tras la publicación de LTOBC. Es ésta una exposición que no sustituye a LTOBC, tal como advierte la autora, pero que sí la facilita enormemente y permite, además, una aproximación más adecuada a sus ideas donde vamos a encontrar de manera precisa una teoría del contenido mental expuesta de modo autónomo y al margen de consideraciones estrictamente pertenecientes a la Filosofía del lenguaje. En este texto, además, las ideas de Millikan aparecen clarificadas a la luz de las respuestas y comentarios surgidos tras casi diez años del debate filosófico iniciado con la publicación de LTOBC.

On Clear and Confused Ideas. An Essay about Substance Concepts, aparecido en 2000, presenta las aportaciones de Millikan respecto al estudio del concepto de substancia. En este texto, de ecos aristotélicos y en línea con su peculiar interpretación del realismo, Millikan entiende que entre los conceptos más importantes que podemos albergar, está el de substancia porque nos permite individualizar particulares.

Durante el curso 2002 Millikan impartió las conferencias Jean Nicod, fruto de éstas es su *Varieties of Meaning*, aparecido en el año 2004. En *Varieties of Meaning* encontramos una discusión a fondo sobre el concepto de signo natural y signo convencional, tópico que ya había

sido anticipado en LTOBC y que es una constante en toda la obra de esta autora. Recientemente Millikan ha vuelto sobre el tema del lenguaje -tema que nunca ha abandonado- y en 2006 ha publicado su *Language: A Biological Model*, obra en la que recoge diversos artículos aparecidos entre 1993 y 2006.

Puede pensarse que el tema del contenido es merecedor de un tratamiento autónomo respecto de otros tópicos filosóficos que caerían de lleno en la Filosofía del lenguaje. Se podría añadir que, además, el estudio del contenido mental compete propiamente a la Filosofía de la mente y es compartido, como ya he mostrado, por disciplinas más especializadas como la Epistemología o la Teoría de la acción. Hacer estos distingos no es propugnar ningún tipo de bizantinismo, sino que tiene que ver con una exigencia de claridad filosófica. Así, aunque separable, en cuanto a su estudio, de la Filosofía del lenguaje, lo cierto es que las páginas que esta autora dedica a su caracterización son muchas menos que las que dedica a la discusión de los tópicos propios de la Filosofía del lenguaje y, sin embargo, es conocida más por lo primero que por lo segundo. No en vano la semántica de Millikan es de base intencional. Si se tiene en cuenta el resto de autores que abordo en este estudio, se puede pensar que es casi una anomalía y, por ello, se justifique el que radique la teleosemántica en las discusiones propias de la Filosofía de la mente y no en su prima hermana, la Filosofía del lenguaje.

Para Millikan el contenido de un estado intencional se determina a partir de la función biológica que puede ser adscrita a los diversos mecanismos con los que cuenta un sujeto que, en cierto modo, consumen representaciones. Así, la cuestión acerca de cómo se fija un determinado contenido no puede ser resuelta aludiendo a qué tipo de items causan ese mismo contenido, sino que tiene que ser dirimida en la pregunta acerca de qué mecanismos dispone el organismo para vérselas con la información relevante en su medio. Sólo así, según Millikan, es posible encarar el problema del error o de la disyunción y caracterizar adecuadamente el contenido mental. Estos problemas se evitan desde el momento en que asumimos que, en tanto que entidades biológicas, estos mecanismos tienen funciones biológicas, más concretamente, propio-funciones. Así la función propia adscrita a un determinado mecanismo consumidor de representaciones nos

permite fijar el contenido de las propias representaciones. Estas funciones son entendidas conforme a la teoría que ya expuse en el capítulo II, es decir, desde un punto de vista histórico-etiológico que incide en los procesos de *selección para*.

Las siguientes secciones exponen de una manera más detallada las ideas esbozadas anteriormente. En particular trato de responder a las siguientes cuestiones: ¿en qué se diferencia la semántica causal de la teleosemántica de Millikan y cuáles son los argumentos que la autora presenta frente a la primera? ¿Cómo entiende Millikan el concepto de representación? Y ¿Cuál es la relación entre el concepto de función propia desarrollado en el capítulo II y el problema del contenido mental?

2.1.1. De las semánticas causales a las semánticas de los mecanismos consumidores de representaciones

Las semánticas causales como la de Dretske explican la fijación del contenido apelando a un conjunto de circunstancias cuya concurrencia causa necesaria y suficientemente el contenido de la representación. Así los objetos "gato" producen las representaciones GATO y no PERRO. La atribución de un contenido erróneo depende de una confusión en las causas productoras de la representación, es decir, la representación GATO es errónea cuando no es producida por los objetos "gato". El enfoque desarrollado por Millikan (1984; 1993; 2004), a diferencia de aquellos presentados en la semántica causal o del indicador, está centrado en los sistemas consumidores de representaciones antes que en los productores de los mismos. Las circunstancias que producen representaciones sólo tienen sentido si existen mecanismos biológicos que aseguran su identificación normalmente¹⁶⁸, es decir, mecanismos que aseguren una identificación correcta de la clase de referencia definida por el concepto en cuestión.

¹⁶⁸ En la terminología de Millikan, "normalmente" significa conforme a una explicación normal (*Normal explanation*) es decir, una explicación que tiene en cuenta las circunstancias que históricamente han concurrido en la producción de los ejemplares de ítems pertenecientes a un determinado linaje y que por lo tanto considera un estándar de normalidad del ítem al definir un patrón de producción o ejecución. Cf. Millikan (1993: 55).

En este sentido, la propuesta de Millikan es una respuesta a los problemas que, como el problema del error, afectaban a la teoría causal o del indicador que presenté. Las teorías causales señalan que en la naturaleza existen diversos tipos de sistemas representacionales que detectan estímulos y los representan, estos sistemas tienen la función -en el sentido de función biológica- de detectar esos estímulos y no otros, por eso el contenido es normativo y representa lo causado por una determinada clase de referencia y no otra. El concepto de función biológica no es ajeno en absoluto a las semánticas causales, sólo que éste es empleado casi a un nivel intuitivo¹⁶⁹. El concepto de función biológica juega, como es de suponer, un papel mucho más importante para Millikan, pues para esta autora el contenido no se fija únicamente atendiendo al régimen de las causas que lo producen, sino que se centra en los mecanismos consumidores de representaciones, que son los que determinan qué cuenta como una representación y los que, en consecuencia, fijan el contenido. Los mecanismos consumidores de representaciones son producto de la selección natural y tienen funciones propias, la de detectar un determinado contenido y no otro. La normatividad del contenido no depende únicamente del régimen de las causas, sino de la propio-funcionalidad del mecanismo que consume la representación¹⁷⁰. Así, algo representa erróneamente (o es “*misrepresenting*”) en la medida en que no cumple su función (es “*malfunctioning*”).

El hecho de centrarse en los mecanismos consumidores de representaciones y no sólo en los productores debería llamar la atención, pues es, junto con el recurso constante a la teoría de las funciones histórico-etiológica, el rasgo más característico de la teleosemántica de Millikan. La razón hay que buscarla en el concepto

¹⁶⁹ Dretske (1988:65), por ejemplo, nos dice que hay funciones biológicas y que éstas juegan un papel relevante en la determinación del contenido, pero no nos dice qué es una función biológica, si bien parece asumir que éstas tienen algo que ver con la teoría histórico-etiológica, pero finalmente no se pronuncia. Por su parte, Sterelny (1990:123) insiste también en la falta de compromiso de Dretske con una determinada teoría de las funciones. A mi juicio no se trata de un descuido de Dretske, sino de una virtud que habría que explotar en favor de una buena teoría de las funciones tal como sugiero en el capítulo III.

¹⁷⁰ Esta postura no es admitida por autores como Fodor (1990: 66 y ss).

de signo y representación y en su vinculación con la intencionalidad. No todo signo está dirigido explícitamente hacia lo mental, de hecho algo es signo para alguien, aunque en la naturaleza se den fenómenos que naturalmente conllevan información. Los signos naturales como las manchas del sarampión tienen una existencia carente de intenciones comunicativas, son signos sólo para aquellos que son capaces de interpretarlos, pero no pretenden decir nada. El dirigirse o no dirigirse hacia lo mental no es una cuestión de nomicidad natural, sino de la capacidad de un organismo para ser sensible a una información que, en principio, es biológicamente relevante y esa información puede ser muy local. Por eso la simple enumeración de las condiciones necesarias y suficientes que deben concurrir para que una determinada clase de referencia produzca una representación es insuficiente para fijar el contenido puesto que el contenido es algo de naturaleza mental y tiene sentido únicamente para una mente. En este sentido, parece hacerse necesaria la alusión a los mecanismos consumidores de representaciones y no sólo a aquellos que las producen. Nos es dado tener experiencia de un rango limitado de objetos y, en principio, cabría afirmar que tendríamos sólo experiencia de aquellos contenidos que tienen un valor biológico. La capacidad para poder tener esos contenidos y no otros es el resultado de nuestro desarrollo filogenético. Así, puede argüirse, según la propia Millikan (2004), que la información tal como la describe Dretske está descontextualizada es, como ella afirma, libre de contexto. En su opinión ocurre lo contrario, la información es local¹⁷¹. Los signos naturales no son absolutos, los signos naturales son recurrentes sólo en su contexto. Tampoco dependen exclusivamente de leyes naturales sino de correlaciones estadísticas circunscritas a un tiempo y lugar concretos. La no arbitrariedad o recurrencia de los signos y su carácter relativo o sensible se explican porque, en general, muchas condiciones relevantes productoras de signos son persistentes y replicantes. Lo que importa, para que el signo cuente como signo de algo, es una

¹⁷¹Según Millikan (2004:35) un problema de la semántica libre de contexto defendida por Dretske es que tiene serias dificultades para la explicación de las representaciones que versan sobre individuos, centrándose sólo en representaciones predicativas del tipo *a es F*.

correlación entre todos los ejemplares del mismo tipo que se afianza a través de una relación de reproducción de la que sobreviene la propositividad del signo.

2.1.2. Signos, intencionalidad y representaciones

La preocupación por las taxonomías semióticas es una constante en la obra de Millikan. A fin de cuentas las representaciones y los diferentes términos lingüísticos son tipos de signos que, de un modo u otro, transportan información relevante para el organismo. Por otra parte, es obvio que la referencia a los signos es un tópico inevitable tanto en filosofía del lenguaje como en filosofía de la mente¹⁷². En las

¹⁷² Así lo entienden diversos autores dentro de la tradición analítica. Por ejemplo, encontramos referencias al concepto de signo en el *Tractatus* de Wittgenstein (Cf. Propositiones 3.11 y 3.12). de manera más concreta la discusión sobre el papel del signo la encontramos en el célebre *Meaning* de Grice (1957). Como ya se habrá notado, el programa de Grice es ciertamente una de las grandes influencias para la teleosemántica desde el momento en que se recurre a los contenidos mentales para dar cuenta del significado de los términos lingüísticos. *Meaning* no sólo desarrolla un paradigma acerca de lo que es un buen análisis filosófico, sino que también rescata una distinción que durante largo tiempo había permanecido en el olvido. En *Meaning* encontramos un análisis del término significado bastante novedoso para el ambiente filosófico de su época dominado bien por el positivismo lógico bien por la figura estelar del segundo Wittgenstein, que apela directamente al contenido mental para dar cuenta del significado de nuestras preferencias. Así, *significar* tiene mucho que ver con *querer decir*. Por otra parte, también hay en Grice una rehabilitación del concepto de signo como material primario sobre el que centrar la reflexión semántica; el significado pues, se atribuye a los signos. Así, Grice constata que hay dos grandes tipos de signos; signos naturales y signos no naturales. La diferencia entre ambos tipos de signos tiene que ver con su arbitrariedad o no arbitrariedad. Los signos no naturales guardan una relación arbitraria entre lo significado y el signo, de hecho ésta es la teoría sobre la que se apoyan las reflexiones de la lingüística clásica. Pero no todo signo guarda esa misma relación con lo significado, los signos naturales no guardan una relación arbitraria con lo significado, pues la conexión entre ambos elementos se establece al margen de cualquier convencionalidad. La noción de signo natural que emplea Grice es la misma que emplea en la semiología médica, un signo, en este sentido es una manifestación de un proceso natural bien normal bien patológico, a diferencia de un síntoma, que es una manifestación subjetiva del estado psicofísico del paciente. Los signos naturales no constituyen únicamente el objeto de estudio de la semiología médica, sino también de otras disciplinas científicas como la

distintas obras de Millikan hay un intento constante por clarificar la noción de representación que, a su juicio, aparece ya muy viciada. La vía de análisis que explota Millikan no es nueva, sino que consiste en recurrir a una terminología menos empleada en la tradición analítica y en la Modernidad como es la *se signo*. El concepto de representación tiene claras connotaciones mentalistas.

Millikan (1993: 103-121), desglosa el concepto ordinario de representación en otros conceptos a su juicio más precisos como suposiciones tácitas (*tacit suppositions*), iconos intencionales¹⁷³ (*intentional icons*), representaciones internas (*inner representations*) y frases mentales (*mental sentences*). La idea que persigue es caracterizar adecuadamente los fenómenos que habitualmente englobamos bajo el concepto más general de "representación".

Las suposiciones tácitas son un tipo de representaciones muy básicas que deben suponerse para que otros mecanismos del organismo funcionen correctamente. Este tipo de representaciones vienen implantadas en la estructura de nuestro organismo, y aunque pueden ser falsas dada una situación anómala -una situación de laboratorio, por ejemplo- no suelen fracasar en el cumplimiento de sus propósitos. Por ejemplo, los ritmos circadianos son mecanismos que responden ante determinadas condiciones del medio, así podemos decir que estos ritmos representan la duración del día y son necesarios para el desarrollo de otros repertorios conductuales del organismo. Lo mismo ocurre con las disposiciones de color en ciertos animales que elicitán determinadas conductas.

meteorología o la geología. No obstante, el programa de Grice dista mucho de interesarse por el significado natural, sino que centra sus intereses en el estudio del significado no natural, puesto que emitir un signo no natural es realizar una acción guiada por un tipo especial de intenciones, las intenciones comunicativas.

¹⁷³ En LTOBC la categoría semiótica principal es la de "icono intencional", si bien analiza otros tipos de signos como signo natural, signo improvisado o signo producido por un instrumento. El concepto de icono intencional bien puede valer a grandes rasgos por el de representación, sin embargo, un icono intencional no es un pensamiento si por tal entendemos un juicio o una estructura de sujeto y predicado.

Millikan define los iconos intencionales¹⁷⁴ como estructuras que mantienen una relación de isomorfía con un rasgo cualquiera presente en el medio del organismo. Un icono intencional, a diferencia de las suposiciones tácitas, no guarda ninguna correspondencia con los rasgos orgánicos adaptativos, sino que es adquirido como una respuesta adaptativa del organismo frente al medio. Un icono intencional es una estructura que se empareja normativamente con un estado de cosas del mundo y la representa. Los iconos intencionales son mecanismos que se supone que deben ser usados o interpretados por otros mecanismos cooperantes. Un icono intencional no es un enunciado ni una representación, aunque tanto enunciados como representaciones en sentido estricto se construyen sobre iconos intencionales. El ejemplo paradigmático al que Millikan recurre es la danza de las abejas. Otros ejemplos de iconos intencionales son las imágenes retinianas, o los maullidos de los gatos. En todos estos casos se da una regla de correspondencia entre lo producido por un determinado dispositivo orgánico y un rasgo o situación presente en el medio. Por lo tanto, el requisito que debe ser satisfecho para que algo cuente como icono intencional es que se dé una relación de correspondencia (*mapping rule*) entre el elemento representado y la representación que denominamos icono intencional, en este hecho radica su valor semántico. Por otra parte, los iconos intencionales no tratan con información natural, no son signos naturales¹⁷⁵ y, por lo tanto, pueden ser falsos, no corresponderse con la realidad tal como

¹⁷⁴ Es importante notar que la precisión del término "icono intencional" varía respecto de LTOBC (1984) a *White Queen ...* (1993). En LTOBC apenas se distingue entre un icono intencional y un enunciado, en *White Queen* la diferencia es mucho más clara. Un enunciado es un icono intencional, pero hay iconos intencionales que no son enunciados: "*But not all of the most striking features of sentences are exhibited by other intentional icons. One of the most striking features of sentences is exhibited only by sentences. That is the subject-predicate structure.*" Millikan (1984:96). Es decir, en LTOBC apenas se señala que existe una diferencia entre ambos términos, cuando en Millikan (1993) se incide claramente en esa diferencia.

¹⁷⁵ "*Intentional icons do not, as such or in general, carry "natural information" Nor do they "covary" with or "track" what they icon. Their definition makes no reference to how likely or unlikely they are actually to correspond to their designated environmental features, nor to how likely these features are to get mapped by them.*" Millikan (1993:107).

supondría su propia regla de correspondencia. En tal caso la falsedad radicaría en un malfuncionamiento del mecanismo productor del icono intencional en cuestión y no en ningún sesgo de razonamiento o inferencia incorrecta, de hecho, en el caso de los iconos intencionales, el ámbito inferencial está ausente. También es importante señalar que un icono intencional no es un signo natural; los iconos intencionales no conllevan información natural en el sentido en que no se da una covariación nómica tal como ocurre con los casos de las manchas del sarampión o el humo que indica la presencia de fuego - todos ellos signos naturales-. Pero tampoco dependen de la existencia de convenciones, sino de mecanismos seleccionados biológicamente para la producción de estos iconos. En este sentido sí hay una norma de ejecución que depende de la historia selectiva y que justifica la funcionalidad del mecanismo productor del icono intencional; no hay convencionalidad, pero sí una normatividad que regimenta las ejecuciones correctas de las incorrectas. En este sentido la normatividad de este tipo de fenómeno intencional no sobreviene únicamente a la presencia de un contexto social o al hecho de que haya reglas acatadas tácitamente por todos los miembros de una comunidad, sino que está presente en la propia naturaleza del organismo productor y consumidor de iconos intencionales.

Sin embargo, un icono intencional no es una representación en el sentido que emplea Millikan este concepto. Millikan denomina propiamente "representaciones" a los elementos que participan en los procesos de transformación de la información y no a los simples iconos intencionales en los que los procesos inferenciales están ausentes como es el caso de cualquier presentación objetiva ante los sentidos. Las representaciones comparten todos los rasgos atribuidos a los iconos intencionales pero, además, tienen la función de tomar parte en los procesos de inferencia mediata. En una inferencia mediata ocurre una transformación de la información a través de un término medio. La característica distintiva de las representaciones es la implicación de estas en los procesos inferenciales.

Millikan (1993: 89-90) establece dos condiciones que deben ser satisfechas para poder afirmar que un determinado ítem cuenta como una representación para un sistema. En primer lugar, la correspondencia entre lo representado y la representación es la regla, y

casi nunca la excepción, del funcionamiento habitual del mecanismo consumidor de representaciones. En segundo lugar, se aplica el principio fregeano de contexto para toda representación perteneciente al sistema representacional: las representaciones tienen significado en un contexto.

La diferencia entre iconos intencionales y representaciones quizás tenga más sentido planteada desde la dicotomía entre lo conceptual y lo no conceptual, dicotomía que abordé en el capítulo IV. Si bien, Millikan (1993: 103-121) no dice nada al respecto. Un icono intencional como pueda ser una imagen retiniana es un elemento intencional, pues apunta a algo que no es él mismo y puede ser tenido erróneamente - dos características definitorias de la intencionalidad-. Sin embargo, en tanto que imagen retiniana, no es algo conceptual, pues no tiene todavía el formato necesario para poder formar parte de una estructura proposicional y, por lo tanto, para tener un carácter inferencial. Para ello debería ser un concepto perceptivo y no una imagen. Lo no conceptual está, pues, al margen de la inferencialidad. No ocurre lo mismo con las representaciones tal como Millikan las define, pues éstas, desde el momento en que toman parte de los procesos de inferencia, son elementos conceptuales. Pero del hecho de albergar representaciones y en consecuencia poder realizar inferencias no se sigue que los organismos involucrados sean racionales. Según Millikan hay dos razones a favor de la afirmación anterior: en primer lugar porque hay inferencias que parten de premisas no conceptuales como las suposiciones tácitas. En tales casos no estamos ante una genuina inferencia, sino ante una reconstrucción lógica del proceso que da lugar a "conclusiones verdaderas".¹⁷⁶ En segundo lugar, puede ocurrir que el organismo sea incapaz de reconocer que alberga representaciones inconsistentes. La racionalidad va ligada al hecho de reconocer las inconsistencias. Así, para Millikan, no toda criatura que alberga representaciones es una criatura racional.

¹⁷⁶ "(I)t is likely that many inferences produce true conclusions from true premises by heavily relying on tacitly presupposed information about the organism's normal environment (...). This information would have to be made explicit in the premises of a correct logical reconstruction of inference." Millikan (1993:115).

Respecto de las representaciones también cabe preguntarse acerca de qué tipos de relaciones pueden mantener los consumidores con los distintos tipos de signos que emplean. Millikan (2004) señala tres tipos de relaciones representacionales. Un organismo puede albergar representaciones intencionales descriptivas, directivas o lo que Millikan denomina "*Pushmi-Pullyu*", representaciones directivas-responsivas. En el caso de las representaciones intencionales descriptivas lo que determina su contenido es la correspondencia con lo descrito; estas representaciones describen lo que es el caso. Las representaciones intencionales directivas tienen como fin producir un estado de cosas necesario para el consumidor. Es decir, este tipo de representaciones tiene una función directiva que consiste en dirigir las acciones necesarias para la satisfacción de los deseos del organismo. Por último, las representaciones *Pushmi-Pullyu* son a la vez descriptivas y directivas. Además, estas representaciones son las más frecuentes, pues son las que encontramos en las interacciones comunicativas entre los organismos.

Las representaciones mentales nos brindan la posibilidad de establecer inferencias, sin embargo, aunque la capacidad de establecer inferencias va ligada igualmente a la posesión de algún tipo de estructura enunciativa, Millikan no está comprometida con la hipótesis LOT. Tal como he mostrado, Millikan no admite que el pensamiento sólo sea algún tipo de evento lingüístico a modo de enunciados en el lenguaje del pensamiento, aunque sí considera de algún modo la existencia de frases mentales (*mental sentences*) al margen de los mecanismos computacionales que señala Fodor para su lenguaje del pensamiento. Una de las razones principales para descartar la hipótesis LOT es la dificultad de identificar los constituyentes proposicionales que deben combinarse para generar enunciados en mentalés¹⁷⁷. En su lugar propone analizar las proposiciones en términos de la estructura sujeto-predicado¹⁷⁸. Esta estructura también

¹⁷⁷ "Sentences are usually thought of as composed of parts that can be strung together to contribute to the meanings of the whole and that can be recombined to yield different meanings. But how literally should we interpret "parts" and "strung together"?" Millikan (1993:117).

¹⁷⁸ "I do suggest that we insist on subject-predicate structure. There must be possible transformations of the icon to correspond to other subjects having the same predicate,

permite transformaciones y da cuenta igualmente de la sistematicidad y generatividad de los enunciados mentales. Así, es posible adscribir diversos predicados a un mismo sujeto (Fa, Ga, etc.) o adscribir diferentes sujetos a un mismo predicado (Fa, Fb, Fc, etc.). Con la estructura sujeto predicado se puede dar cuenta de buena parte de la vida mental de los organismos racionales, sin embargo como creo haber mostrado para Millikan hay vida mental más allá de las actitudes proposicionales; representaciones, iconos intencionales y suposiciones tácitas son muestras de una intencionalidad gradual que hunde sus raíces en los organismos más sencillos.

2. 1.3. ¿Cómo fijar el contenido? Propiofuncionalidad y contenido mental

Hasta el momento he presentado una serie de rasgos a propósito de la teoría de Millikan como pueda ser su dependencia de la teoría causal o su clasificación semiótica. En esta sección evaluaré la importancia del concepto de función biológica dentro de la teoría del contenido de Millikan. Una caracterización muy somera de la propuesta de Millikan podría ser la siguiente: Millikan afirma que las funciones propias de los mecanismos consumidores de representaciones fijan el contenido de las mismas. Pero, ¿en qué sentido las funciones propias determinan o fijan el contenido de las entidades intencionales consideradas en la sección anterior? Es erróneo pensar que lo que fija el contenido es únicamente la función propia de una representación, de hecho para Millikan no es así. Lo que fija el contenido de las creencias de los organismos es la función propia que poseen los mecanismos consumidores de representaciones; estos sistemas sólo pueden ejecutar su función en presencia de la representación adecuada. Así la teleosemántica fundamentalmente está enfrascada en dar respuesta al problema de la normatividad y es, antes que nada, una teoría de la normatividad, pues la representacionalidad se hace depender de teorías más básicas como la teoría causal-informacional. La

and there must be transformations to correspond the same subject having other predicates.” Millikan (1993: 117-118).

teleosemántica tal como Millikan la entiende no se ocupa de explicar qué es una representación o cómo un organismo llega a tener representaciones, sino que es una investigación acerca de la determinación de su contenido, así la cuestión importante es ¿por qué la representación R tiene el contenido C y no el contenido C'? Para contestar a esta pregunta tenemos que contemplar dos elementos: (1) Las representaciones mentales pertenecen a un sistema que las produce y las consume, es decir, hay una economía de las representaciones, están sistematizadas. (2) Los organismos consumidores de representaciones no viven aislados, sino que se relacionan con un medio y están adaptados a él. Decir esto último implica admitir que los sistemas que consumen representaciones tienen funciones en el sentido acuñado por la propia Millikan que mostré en el capítulo II. Así lo que sucede con las funciones es que determinan o regimentan el uso que un sistema en tanto que miembro de una comunidad o nicho ecológico hará de las diversas representaciones con las que comercia. El contenido se determina en relación a los consumidores y no a la fuente informacional, a saber, el contenido se determina por el uso regimentado que hacen los consumidores de la fuente informacional y resulta fijado por el valor biológico que supone poseer esos contenidos y no otros. Por ejemplo, tener el contenido de que las bayas blancas del muérdago son venenosas contribuye a la proliferación del organismo, mientras que carecer de ese contenido implica una disminución de la probabilidad de llegar a la edad adulta y reproducirse. Del mismo modo, el contenido de la representación de la comida de la rana es COMIDA y no BOLITAS NEGRAS PEQUEÑAS, ya que sólo la comida contribuye a la eficacia inclusiva del organismo.

La presencia de condiciones normales en la determinación del contenido sobre la que tanto insiste Millikan es importante porque estas condiciones delimitan los rasgos ambientales con los que la representación se corresponde y que coinciden con presentaciones pasadas de los mismos eventos ante el organismo que los representa. Así, la referencia a condiciones normales es un garante de la normatividad del contenido de las representaciones.

El recurso a las condiciones normales también nos permite evitar caracterizaciones individualistas del contenido - Millikan es una

partidaria acérrima del externismo semántico-. El contenido depende de los mecanismos consumidores, pero se admite que puede haber variación, por eso es importante apelar a la igualdad de las condiciones normales bajo las que se ejecutan las funciones, pues las representaciones con contenido idéntico pueden ser diversas.

2.2. Neander: *Low Church* frente a *High Church*

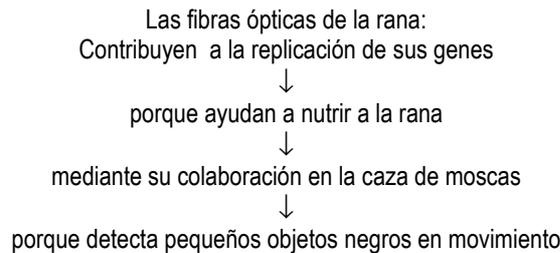
Los trabajos de Neander han ido encaminados a redefinir la teoría histórico-etiológica y a explotarla sistemáticamente en favor de una teoría naturalista del contenido mental fundamentada en la misma. A diferencia de Millikan o Papineau, no ha desarrollado sus investigaciones de un modo tan exhaustivo y, así, en su lugar, sus aportaciones consisten en artículos o contribuciones a obras colectivas en las que o bien se enfocan cuestiones muy concretas sobre teleosemántica o bien se hace una exposición exhaustiva de las diferentes teorías. Respecto de esto último hay que señalar que Neander (i.e. 2004; 2008) ha sido la autora que más esfuerzos ha invertido en divulgar el enfoque teleológico del contenido mental.

2.2.1. La teleología *Low Church*

Uno de los rasgos más sobresalientes de sus diversas aportaciones es el énfasis que pone en hacer de la teleosemántica una teoría semántica apta desde el punto de vista de la ciencia cognitiva. En *Malfunctioning and Misrepresenting* (1995) distingue lo que ella denomina un enfoque de *Low Church* frente a uno de *High Church*¹⁷⁹. Ambos enfoques son *down up* en el sentido de que parten de algo más básico como es la funcionalidad biológica para explicar la normatividad del contenido mental. Lo que diferencia a un enfoque *High Church* de una perspectiva *Low Church* es una cuestión acerca de qué nivel de

¹⁷⁹ El término *Low Church* o "baja iglesia" tal como lo emplea Neander es sinónimo de latitudinario o que concede amplitud. El término viene de las controversias jurídico teológicas de las comunidades anglicanas a propósito de la aproximación a la Iglesia Católica Apostólica y Romana (*High Church*) o, por el contrario, su necesario distanciamiento y asunción de posturas claramente reformadas.

descripción o análisis funcional debe ser tenido en cuenta como relevante a la hora de identificar el contenido de una representación. Consideremos el tan traído ejemplo de la retina de la rana cazadora de moscas. Podemos reconstruir el fenómeno de la siguiente manera: Las fibras ópticas de la rana contribuyen a la eficacia de la rana (a su probabilidad de tener descendencia) mediante su contribución a la alimentación, pues éstas contribuyen a que la rana pueda cazar moscas ya que detecta pequeños puntos negros en movimiento. Se puede describir lo anterior mediante el siguiente diagrama¹⁸⁰:



Bien pensamos que el contenido de la representación producida por la imagen retiniana de la rana es MOSCAS, bien PEQUEÑOS OBJETOS NEGROS EN MOVIMIENTO. Podemos considerar como relevante el nivel más básico (detección de puntos negros en movimiento) bien el nivel superior, cuyo valor ecológico es indubitable (comida = supervivencia = incremento de la probabilidad de ser eficaz). En el primer caso estamos ante una estrategia *High Church*, la estrategia habitual seguida por Millikan, en el segundo, la estrategia es *Low Church*.

Los argumentos a favor de la estrategia *High Church* parten de la distinción que establece Sober entre *selección de* y *selección para*. En efecto, se arguye que para los casos de indeterminación como el que ilustra el ejemplo anterior hay *selección para* detectar comida y no para cualquier objeto pequeño volador, sólo lo primero cuenta como contenido de la representación. En una situación experimental hipotética en la que la conducta predatoria (variable dependiente) fuera elicitada tanto por moscas (variable independiente₁) como por

¹⁸⁰ Adaptado de Neander (1995: 125).

pseudomoscas (variable independiente₂) nos encontraríamos ante un caso de representación errónea. Pues la rana no tendría los recursos cognitivos necesarios para distinguir las moscas de las pseudomoscas simplemente porque en su medio natural no hay pseudomoscas. Obsérvese que la posibilidad de error no desaparece, pero sí se minimiza; en condiciones normales las ranas cazan moscas y no otros objetos.

Sin embargo, el enfoque *Low Church* afirma que el contenido se fija en el nivel más básico de los anteriormente considerados. Así la disyunción entre mosca y pseudomosca es aparente; la rana únicamente se representa pequeños objetos redondos volantes. Neander advierte que por nivel básico debemos entender aquel que ya no puede ser descompuesto en un análisis funcional tipo Cummins, es decir, en el caso de la rana es el mecanismo de detección y elicitación de la conducta predatoria y no la representación o MOSCA o MOSQUITO o MOSCÓN. Según Neander esta postura es más aceptable desde el punto de vista de la ciencia cognitiva ya que no implica asumir más que lo que el investigador tiene presente sin necesidad de extravagantes historias selectivas o perspectivas ecológicas ajenas a los intereses más directos. La propia Neander (2006) insiste sobre la importancia de desarrollar la línea *Low Church* en la idea de que nos aproxime al trabajo de los científicos cognitivos¹⁸¹. Tal línea tiene la virtud de que nos permite localizar el contenido de una manera precisa dentro del sistema. Si bien esto no constituye propiamente una teoría del contenido, sino sólo una táctica metodológica. Por ejemplo, en el caso de la rana podríamos decir que el contenido se localiza a partir de la zona T5(2)s del *tectum* óptico¹⁸². Esto nos permite hablar de lo que se representa la rana o contenido de la representación de la rana en términos de reconocimiento de la presa o reconocimiento de estímulos insectiformes, o de estímulos con forma de presa. Así, al poder atribuir a las células de la zona T5(2)s del

¹⁸¹ "(I)f the contents of mental representations are to play a role in explaining cognitive capacities, they must cohere with the relevant information processing. At the risk of laboring the point, let me stress that **this is not offered as a theory of mental content.**" Neander (2006:180) (La negrita es mía).

¹⁸² Cf. Neander (2006: 175 y ss).

téctum óptico la función de ser las neuronas encargadas del reconocimiento de las presas podemos localizar el contenido al menos de una manera suficientemente atractiva para la ciencia cognitiva. Finalmente, otra de las ventajas de un enfoque *Low Church* es que no requiere considerar actitudes proposicionales, sino sólo estados con contenido como es el caso de la rana. La rana, a diferencia de los seres humanos no realiza inferencias sobre el contenido de sus perceptos que puedan sesgar sus actuaciones.

2.2.2. La determinación del contenido

Para Neander las representaciones a las que he aludido determinan su contenido basándose en las funciones de respuesta ante estímulos de las que están dotados nuestros sistemas representacionales. En el transcurso de la evolución del organismo estos sistemas representacionales fueron seleccionados para responder ante determinados estímulos relevantes dentro de la economía del organismo. Esta caracterización es, en lo esencial, coherente con las propuestas de Millikan, pero Neander (1996) señala dos puntos que la diferencian respecto del enfoque High Church: (1) el objeto de estudio es más próximo, lo que debe ser priorizado es la descripción inmediata de las funciones de los mecanismos de respuesta ante estímulos. (2) únicamente podemos acometer (1) describiendo las propiedades causalmente eficaces en el proceso, es decir, debemos atenernos a lo que de hecho hacen esos sistemas y no proceder desde perspectivas más amplias o que resaltan en primer lugar el valor ecológico del mecanismo. Puede pensarse que la estrategia de Neander no sólo tiene un interés conciliador con la ciencia cognitiva, sino que también es posible una lectura epistemológica, pues el acceso a la historia selectiva es más difícil que el conocimiento o análisis del funcionamiento del sistema en cuestión. Si bien es ésta una línea que no parece explotar. Suponer, como hace Millikan que la rana debe saber algo acerca de lo que cuenta como comida de rana o que los ratones de campo tienen un conocimiento acerca qué clase de organismos constituyen un peligro, es demasiado suponer. Para el caso de la rana una caracterización del contenido de las representaciones adecuada desde el punto de vista de la ciencia

cognitiva o de la neuroetología apelará antes a pequeños objetos que se mueven que al concepto MOSCA.

2.2.3. La explicación del error

Neander cree que no es necesario recurrir a la noción de malfuncionamiento del sistema cognitivo para dar cuenta del error. No hay error en comer moscas o abejas si ambas cuentan como alimento, el sistema perceptivo de la rana no es tan exigente ni parece permitir tales discriminaciones. Así plantear la disyunción entre MOSCAS o PSEUDOMOSCAS es una situación artificial en exceso¹⁸³. Una virtud del enfoque de Neander es que insiste en la necesidad de distinguir entre las capacidades más básicas de representación de las más sofisticadas, del mismo modo que hay también tipos de errores con un grado de sofisticación mayor o menor. En el caso de organismos con capacidades representacionales complejas podemos encontrar explicaciones alternativas al malfuncionamiento en los casos de disyuntivitis como puedan ser la presencia de diversos sesgos (expectativas, reducción de la disonancia cognitiva, conocimientos previos, etc.). En definitiva, podemos tener contenidos erróneos sin que necesariamente se dé un funcionamiento erróneo de nuestras capacidades.

2.2.4. Comentarios finales

Al final de su Neander (2006) desarrolla un elenco de posibles críticas contra su caracterización informal y *low church* del contenido. Una de las más interesantes es la que vierte contra los argumentos de Millikan (2000: Apéndice b). Millikan encabeza lo que es la facción *high church* y, por lo tanto, una caracterización como la de Neander no puede sino ser rechazada. Para Millikan no es suficiente con que

¹⁸³ "In the case of the frog that feeds itself by catching flies by detecting "small dark moving things", for example: so on the account I favour the frog is not in error when it snaps at bee-bees. Nonetheless, it can make mistakes by malfunctioning." Neander (1996:122).

exista una capacidad por ejemplo para reconocer aquellos rasgos que cuentan como pertenecientes a la comida de las ranas o que para un ratón podrían suponer un peligro. Millikan supone que las ranas deben tener un conocimiento acerca de lo que es comida de rana o lo que realmente representa un peligro para un ratón como pueda ser un halcón, elementos que no aparecen en la propuesta de Neander. Para Neander las suposiciones de Millikan son erróneas, lo importante para la rana es comer y para el ratón guarecerse de sus enemigos y no la representación de la comida o de los depredadores¹⁸⁴.

3. Papineau: el enfoque *top down*

David Papineau es un filósofo de amplio espectro que, al igual que Dretske, ha centrado su actividad investigadora en diversos ámbitos, entre los que cabe destacar la filosofía de la ciencia. En las últimas décadas ha presentado trabajos muy sobresalientes sobre filosofía de la mente que desarrollan tópicos actuales como el estudio filosófico de la conciencia¹⁸⁵. Papineau (1987; 1993) ha desarrollado su teoría teleológica del contenido a la par que Millikan, aunque las diferencias, como se verá, son notables. La caracterización del contenido mental en términos teleológicos aparece en Papineau en el marco de una investigación mucho más amplia a propósito del naturalismo filosófico que se recoge en su libro de 1993 *Philosophical Naturalism*. Esta obra puede ser considerada como la exposición más significativa de su teoría sobre el contenido mental, si bien, las reflexiones en la misma línea están presentes desde 1987, cuando publicó su *Representation and Explanation*. Desde entonces Papineau se ha mantenido fiel a una caracterización teleológica del contenido mental no composicional¹⁸⁶

¹⁸⁴ Cf. Neander (2006: 190).

¹⁸⁵ Cf. Papineau (2002).

¹⁸⁶ La no composicionalidad de Papineau hace referencia a que no trata de aislar o identificar conceptos mentales que puedan combinarse formando una especie de lenguaje del pensamiento. En su lugar trata tanto creencias como deseos como totalidades.

que parte de la consideración de la capacidad de desear de los organismos para explicar la economía mental de los mismos.

3.1. La alternativa *top down*

Tanto el enfoque de Millikan como de Neander podían ser calificados de enfoques *down-up*, ya que partían del concepto de función biológica para explicar el contenido de las diferentes representaciones mentales. Por el contrario, el de Papineau es *top-down*; parte de la concepción de sentido común o *folk* de nuestra vida mental, es decir, de la estructura deseos-creencias. Según Papineau la teoría teleológica del contenido sólo puede funcionar si igualamos los efectos biológicamente ventajosos derivados de la posesión de un determinado rasgo con aquellos efectos que satisfacen los deseos que pueda albergar un organismo. Así, en la determinación del contenido lo importante no es una panfuncionalidad. Por la mera atribución de funciones no determinamos contenidos, sino que sólo los podemos determinar centrando el estudio en aquellos organismos que son poseedores de creencias y deseos. La semántica de Papineau no es sólo *top down*, sino que, además, no es composicional -a diferencia de la semántica de Millikan o Neander-, pues parte de la consideración de las creencias y deseos como totalidades y no admite un estudio que dé razón de las mismas descomponiéndolas en unidades más simples. Aunque coincide con Millikan en que el estudio sobre el contenido debe hacerse sobre el uso que hace el organismo de las representaciones mentales y no sobre su producción. Es decir, la producción de representaciones tiene un sentido dentro de la economía cognitiva del organismo. Ese sentido consiste en guiar las acciones que satisfacen los deseos del organismo. Pero frente a Millikan, Papineau se centra en los resultados exitosos de la satisfacción de los deseos que pueda tener un organismo y no en el éxito biológico entendido al margen de aquellos. Es decir, la funcionalidad por sí sola sin estar circunscrita a los deseos que eliciten la conducta es insuficiente por sí sola y arroja indeterminaciones.

Frente a la estrategia seguida por Neander, Papineau (1997) considera que es una solución parcial y muy próxima a los planteamientos de la teoría causal-informacional. Si bien considera que Neander acierta en el estudio acerca de qué se representa un organismo como la rana, olvida la cuestión realmente importante desde el punto de vista de la teleosemántica; averiguar qué se supone que debe detectar el sistema perceptivo de la rana. Si la teleosemántica debe centrarse en los mecanismos consumidores de representaciones tal como señala Millikan, lo importante no son los efectos que produce el sistema (i.e. la representación de las bolitas negra voladoras), sino identificar el resultado que constituye la función propia del sistema. No obstante, tanto la perspectiva *Low Church* de Neander como la *High Church* de Millikan arrojan indeterminaciones y son sensibles a los contraejemplos de la rana que ya se han considerado anteriormente.

La estrategia de Papineau centrada en el esquema de la psicología popular también tiene algunas limitaciones; admitir una psicología en términos de creencias y deseos implica que el proyecto teleosemántico sólo es viable para organismos que posean dicha psicología. Esto, evidentemente es una merma del alcance que pretenden tanto Millikan como Neander para sus teorías, pero tiene la virtud de que logra evitar casos de indeterminación como el ejemplificado en la sección anterior a propósito del contenido de la rana. Dicho de otra manera, Papineau (1997) considera que sólo a los organismos dotados de creencias y deseos podemos adscribir un contenido determinado; la fijación del contenido requiere una estructura previa de creencias y deseos.

Una consecuencia de lo anterior es la defensa del realismo intencional; apelar a la psicología popular implica una defensa de la existencia de las actitudes. En su opinión esto se justifica porque sólo si tienen existencia pueden ser resultado del proceso evolutivo. La selección natural no puede actuar sobre entidades inexistentes, así, la estrategia de la intencionalidad derivada que propone Dennett (1987) o las perspectivas antirealistas son, en general, inadecuadas. De este modo son conciliables psicología popular y naturalismo filosófico.

3.2. La determinación del contenido

Como ya he mencionado Papineau parte del esquema de creencias y deseos ya que concibe las representaciones mentales como una clase de adaptaciones. La solución para dar cuenta de la semántica de nuestros estados mentales es una teoría teleológica del contenido en la que el contenido de las creencias sobreviene al contenido de los deseos. Es decir, albergamos creencias porque somos capaces de desear en función de nuestras necesidades biológicas. El propósito principal de las creencias consiste en colaborar con los deseos en la causación de la conducta orientada a su satisfacción; creencias y deseos son los elementos del mismo mecanismo adaptativo, a saber, la mente. Sin embargo, ambas entidades - creencias y deseos- no están al mismo nivel, sino que se establece un orden de prelación en el que el contenido de los deseos ocupa el primer lugar. Mientras que las creencias tienen condiciones de verdad, los deseos tienen condiciones de satisfacción. La condición de satisfacción de un deseo viene determinada por el estado de cosas que lo satisface. Los deseos desempeñan la función de combinarse con las creencias para generar conductas que los satisfagan. Así mi deseo de chocolate se combina con mi creencia de que hay chocolate en la nevera y desencadena mi conducta consumatoria de ingerir chocolate. Los deseos dependen de los propósitos biológicos que los producen. Es decir, no se desean imposibles como dormir en la luna o comer nubes sino que lo deseado está en función de las necesidades básicas del organismo. El contenido de las creencias depende así del contenido de los deseos, pues fijamos el contenido de las primeras en base al de estos últimos o, lo que es lo mismo, en base a sus condiciones de satisfacción que. De este modo se relacionan funcionalidad biológica y contenido mental; el contenido de los deseos es favorecido por la selección natural. Un organismo desea lo que desea porque desear unas cosas y no otras incrementa la eficacia biológica o resulta más adaptativo. Así, el propósito de los deseos debe entenderse en términos histórico-etiológicos. Si albergamos el deseo de comer cosas dulces esto es porque en el pasado hubo una selección para las cosas dulces porque proveían al organismo de una mayor eficacia biológica. Por lo tanto podemos decir que ha habido una selección para desear un

determinado rango de cosas y no otras. Papineau señala que hay tantas historias evolutivas como deseos; mi deseo de comer pizza es distinto de mi deseo de comer pasta, los mecanismos selectivos han favorecido determinados deseos o la capacidad para formar determinados deseos¹⁸⁷. Por otra parte, las creencias están copresentes con los deseos; albergamos creencias en tanto que somos capaces de representarnos nuestros deseos y estas funcionan como intermediarios en la satisfacción de los primeros. La selección también actúa sobre la capacidad de formar las creencias que covarían con los deseos y que representan sus objetivos; las creencias guían las acciones necesarias para la satisfacción de los deseos. Así, podemos decir que la creencia de que p es el estado seleccionado para covariar con p y el deseo de que p es el estado seleccionado para provocar p ¹⁸⁸. Tenemos las creencias que tenemos en la medida en que son buenas para satisfacer nuestros deseos. Esto incluye cierta retroalimentación sobre el medio, uno se representa estados de cosas presentes en el medio y actúa teniendo en cuenta las dificultades o ventajas que pueden suponer para la satisfacción de los deseos. La habilidad de crear creencias o de representarse estados de cosas es doblemente directiva y se configura en torno a nuestras capacidades racionales. Tanto los deseos como las creencias son estados mentales producto de la selección natural. La verdad de nuestras creencias garantiza la satisfacción de los deseos.

La exposición previa no implica ningún compromiso con el innatismo. Es decir, un organismo no nace con un repertorio fijo de conductas que satisfacen sus deseos, sino que también aprende. La capacidad de aprendizaje se construye sobre los mismos mecanismos; uno aprende para satisfacer sus deseos o necesidades. Sin embargo, el propósito de las creencias resultantes del aprendizaje sí es resultado de la selección natural en la medida en que la capacidad de aprendizaje lo es.

¹⁸⁷ "(S)ome past selection mechanism has favoured that desire -or, more precisely, the ability to form that type of desire- in virtue of that desire producing that effect." Papineau (1993:59).

¹⁸⁸ Cf. Papineau (1993: 94).

El enfoque de Papineau centrado en el uso de las representaciones puede resultar un tanto extraño si consideramos qué cuenta como condición de verdad de una creencia. Obviamente no es lo representado; aquí no se aplica el concepto clásico de verdad, sino que tiene que ver con la satisfacción de los deseos. La verdad es una propiedad ineliminable de las creencias (en contra de los posicionamientos deflacionarios y minimalistas). La condición de verdad de una creencia es aquella condición que garantiza que las acciones basadas en esa misma creencia resultarán satisfactorias desde la perspectiva de los deseos que motivan las acciones. Para Papineau (1993:71) las acciones basadas en creencias verdaderas satisfacen los deseos a los que las primeras apuntan. La verdad de nuestras creencias tiene dentro de la economía mental el propósito biológico de satisfacer los deseos subyacentes a la misma. Las condiciones de verdad de una creencia nos permiten determinar el contenido de esa misma creencia; cuando las identificamos sabemos cuál es el contenido de la creencia.

Por otra parte, si bien este criterio podría resultar excesivamente pragmático y perspectivista, Papineau insiste en que la verdad de las creencias debe entenderse así, pese a las dificultades que pudieran plantearse. Papineau (1993: 72 y ss.) señala, entre otras, las siguientes posibles objeciones a su formulación de la condición de verdad. En primer lugar considera la posible inadecuación de la misma a las creencias no instrumentales. En su opinión, su formulación cubre también las creencias no instrumentales, ya que la verdad de las mismas se mantiene porque podrían resultar en un momento dado prácticas; así las acciones basadas en creencias verdaderas tienden a ser exitosas. Es una cuestión de validez inferencial. Las creencias no directamente instrumentales tienen que ser válidas en tanto que pueden ser premisas de inferencias prácticas cuyos resultados sí son verdaderos en virtud de la condición de verdad estipulada. Si no fuera así, la condición de verdad sería inaplicable. En cuanto a la cuestión acerca de si las creencias falsas son capaces de satisfacer deseos, Papineau no ofrece una buena solución a este problema reafirmando en su formulación pragmática de las condiciones de verdad. Así, sostiene que por condición de verdad de una creencia debemos entender la condición que garantiza para todos los ejemplares de los

tipos relevantes que las acciones que son consecuencia de esa creencia resultarán satisfactorias. Por último aborda los casos de incertidumbre. En estos casos en los que el agente toma decisiones partiendo de una información incompleta no podemos garantizar que se cumpla la condición de verdad anteriormente estipulada; ésta sólo es aplicable a creencias que parten de situaciones certeras.

La verdad desde un punto de vista naturalizado no es otra cosa; las creencias de un organismo son verdaderas cuando satisfacen sus deseos. Desde la perspectiva de la Biología la verdad de las creencias es eso, aunque posteriormente pueda quedar un lugar para consideraciones más teoréticas y alejadas de los intereses inmediatos de los organismos. Un organismo debe pensar que las creencias que alberga son verdaderas si bajo esas creencias recae la responsabilidad de guiar las acciones que deben mantenerlo vivo. A mi juicio Papineau confunde la verdad con la justificación; el organismo debe estar justificado, la cuestión de la verdad o no importa o se subordina a la justificación. Por otra parte, la justificación desde este punto de vista dependería del buen funcionamiento de los sistemas sensoriales con los que el organismo está dotado.

También se debe distinguir entre propósitos normales y propósitos especiales de las creencias. Los propósitos normales son los que están orientados a la satisfacción de los deseos; así tengo el deseo de comer chocolate y me represento la tableta de chocolate en la nevera, del mismo modo que ante una bajada de glucosa sentimos la necesidad de comer cosas dulces y nos representamos esas cosas dulces y no, por ejemplo, un plato de huevos fritos. La teoría teleológica está centrada principalmente en este tipo de representaciones motivadas por deseos normales. Sin embargo, también es posible que, por ejemplo, ante un incendio creamos que vamos a morir; la condición de satisfacción de esta creencia no es el deseo de morir, sino todo lo contrario. La función de representarnos el peligro inminente es la de mantener el organismo a salvo. La creencia de ser herido o resultar quemado no surge en aras a satisfacer el conjunto de hechos que constituyen su condición de verdad, sino a su evitación. En tal caso identificamos un propósito especial en la creencia de que vamos a resultar heridos. Los propósitos especiales de las creencias tienen como finalidad asegurarse de que los individuos que albergan deseos incorrectos no los llevarán a cabo,

por ejemplo una temeridad o un arrojo excesivos. Así albergamos ciertas creencias falsas en virtud de que la selección natural diseña los sistemas biológicos para un último fin como la transmisión del material hereditario. De hecho no decidimos en base a la transmisión de nuestro material hereditario, sino en base a la satisfacción de nuestros deseos. Un animal no come porque vaya a tener descendencia, sino porque tiene hambre. La conducta sexual humana es un caso paradigmático de cómo deciden los agentes racionales con el fin de satisfacer sus deseos sexuales al margen de su escaso interés reproductivo.

4. Comentarios finales: Una comparación de las teorías expuestas en este capítulo

Las diversas teorías teleosemánticas que he examinado comparten una serie de rasgos en común:

- (1) Todas ellas parten del realismo intencional, es decir, admiten que nuestra vida mental puede caracterizarse adecuadamente en términos de creencias y deseos (quizás a excepción de Neander).
- (2) Todas ellas o bien rechazan la hipótesis del lenguaje del pensamiento -Millikan- o bien son indiferentes a ellas -Dretske-.
- (3) Todas ellas -salvo Dretske, que no se pronuncia al respecto- creen que el concepto de función biológica tal como lo define la perspectiva histórico-etiológica es el concepto adecuado para garantizar la fijación del contenido mental.
- (4) Todas parten de puntos de vista externistas, aunque con matices.

También resultan virtuosas en algunos aspectos y débiles en otros. El punto específico en el que cada una de ellas incide es frecuentemente un aspecto descuidado por las otras, a la vez que cada una de ellas se enfrenta a una serie de dificultades concretas. A continuación ofrezco una breve síntesis de los aspectos más relevantes de cada una de ellas

Respecto de Dretske es importante señalar que no define qué debe ser entendido por función biológica, sólo dice que existen funciones biológicas en tanto que adscribimos fines a los mecanismos que producen la conducta. Esta consideración dice mucho acerca de cómo entiende Dretske la fijación del contenido. Así, no puede decirse que el concepto de función tal como lo entienden los partidarios de la teleosemántica fije el contenido de las distintas representaciones. Las funciones es lo que adscribimos a los distintos sistemas y fijan en la medida en que detectan información en uno u otro formato. Que detecten la información es algo que se explica bien evolutivamente, ya sea a través del desarrollo ontogenético o de los diversos procesos de aprendizaje. Así, la dependencia del concepto de función biológica es sólo en unos casos inmediata.

En segundo lugar, Dretske estipula diversos grados de intencionalidad y diversos sistemas representacionales que manejan la información de una manera específica en cada situación. Esta clasificación de los diferentes tipos de sistemas representacionales (Dretske, 1988) es también muy clarificadora, pues permite una reflexión muy precisa sobre la intencionalidad y la semántica de estos tipos de representaciones y posibilita evaluar las capacidades y límites de los sistemas que los poseen. Este tipo de taxonomías no las encontramos en los otros autores que he examinado.

Respecto de Millikan el rasgo más sobresaliente es la importancia que concede al uso que hacen los organismos de la información y no a la información en sí misma. Esta precisión conlleva la modificación parcial de algunos puntos de vista de la semántica informacional y la inclusión del concepto de función biológica de manera decisiva. Sin embargo, no creo que logre hacer frente a la objeción principal, a saber, el problema de la indeterminación. De hecho el uso que hace Millikan del concepto de función biológica es muy complejo; pues parece que finalmente cualquier representación tiene de algún modo una función propia directa o derivada. Además, debe observarse que la atribución de funciones es a ítems biológicos y que su propósito o propio-funcionalidad depende del proceso de *selección para* tal como apunta la teoría histórico-etiológica. En este sentido lo importante es la existencia de relaciones de familia entre los distintos ítems que se propagan de modo que los individuos seleccionados tienen una mayor

descendencia y son más eficientes que aquellos no seleccionados. Este proceso, sin embargo, aunque podría ser adecuado para ítems claramente biológicos es extrapolado a otras entidades como los memes. Millikan intercambia habitualmente la adscripción de funciones a ítems biológicos y a ítems biológicamente reproducidos como a ítems de naturaleza cultural y explica su persistencia y normatividad a través de la misma teoría. Es decir, a aquello de lo que podemos decir que no tiene un origen directamente biológico, por ejemplo, una determinada preferencia lingüística adscribimos una función bajo los mismos presupuestos que lo haríamos si se tratara de un ítem biológico. En tales casos la propio-funcionalidad es, según la propia Millikan, derivada. Pero a mi modo de ver esto es una respuesta *ad hoc*. Tal como señalan Houkes y Veermas (2003) hay razones para dudar de que la adscripción funcional en entidades culturales y artificiales¹⁸⁹ y, en consecuencia, la normatividad dependa de los mismos parámetros empleados en biología. Igualmente podemos dudar de que los procesos de evolución biológica sean extrapolables a las creaciones culturales como la ciencia¹⁹⁰ o la tecnología. Así una crítica interesante a Millikan vendría de la mano de aquellos que han argumentado en contra de las explicaciones evolucionistas de los fenómenos culturales. En mi opinión la transición de la funcionalidad biológica a la funcionalidad y normatividad de las creaciones culturales es un paso ilegítimo. Resulta más plausible una explicación en los términos de Dretske y basándose en la diversidad de sistemas representacionales que el continuo naturaleza cultura con el que Millikan está comprometida. Aunque, como argumentaré en el siguiente capítulo, la principal amenaza a la que la teoría de Millikan está expuesta es la acusación de indeterminación y esa ha sido la vía tradicional de criticar sus posicionamientos.

La aportación de Neander es valiosa por cuanto trata de aproximarse a la ciencia cognitiva. Para algunos como Papineau (1997) aunque precisamente ahí radica también su principal debilidad ya que se aproximaría a planteamientos más cercanos a la teoría informacional causal que a la teleosemántica. Aunque en lo esencial

¹⁸⁹ Cf. Jaume (2005).

¹⁹⁰ Luján (2007) a propósito del cambio científico.

sigue los planteamientos de Millikan sí es digno de destacar que ofrece una buena estrategia metodológica para la determinación del contenido desde la perspectiva de la ciencia cognitiva. Por otra parte, una deficiencia clara en los trabajos de Neander es que no se centra en cuestiones ciertamente polémicas como si es necesario postular una estructura proposicional de las representaciones mentales. Todo ello hace del enfoque *Low Church* una perspectiva extremadamente básica.

En cuanto a Papineau puede pensarse que resulta muy acertado al señalar la importancia del deseo o de la facultad de desear en la economía cognitiva de los diferentes organismos. Pero si el deseo es resultado de un proceso de selección lo es sólo en un sentido muy básico, pues es obvio que deseamos de una manera desordenada y en muchos casos nociva. Por otra parte quizás sea excesivo decir que hay diversas historias selectivas para los distintos deseos, en todo caso ha habido una selección para la capacidad de desear ciertas cosas en virtud de necesidades básicas del organismo como hidratos de carbono o proteínas. Esta última precisión me parece más adecuada y evita equívocos indeseables en la interpretación de la teoría de Papineau. Así, uno no desea chocolate porque sí, sino porque el sabor dulce del chocolate está correlacionado con el aporte de hidratos de carbono. Lo que uno desea es dulce y la posesión de ese deseo ha sido favorecida por la selección natural. Sin embargo, como inmediatamente se verá, de la capacidad general de desear dulce no se sigue que uno desea chocolate o caramelos. La indeterminación persiste y no es resuelta por la apelación al mecanismo selectivo. Es más, si a alguien que desea chocolate le ofrecen croissants de crema sus deseos quedan colmados. En cualquier caso la teoría de Papineau sólo es capaz de fijar el contenido de las creencias de modo muy general. El deseo tiene una función biológica máxima para cubrir unos mínimos como el mantenimiento de la homeostasis o la perpetuación de la especie. Eso explica por qué deseamos pero esa explicación es insuficiente para la explicación de nuestra economía cognitiva que si bien se construye sobre el deseo es mucho más sofisticada. Papineau (1993:64) es consciente de estos fenómenos. La satisfacción de nuestros deseos no siempre va aparejada de éxito biológico. La facultad de desear es un mecanismo bastante torpe que no resulta eficiente, aunque sí eficaz. Por eso Papineau circunscribe su teoría teleológica del contenido a los

propósitos normales de las creencias. Así apelar a la facultad de desear no es más que una explicación de la naturaleza de los distintos organismos dotados de una estructura de creencias y deseos. En mi opinión Papineau debería responder a los siguientes tres interrogantes: (1) ¿Qué grado de definición es necesario para poder decir que un organismo alberga “deseos definidos” y poder atribuirle una psicología de creencias y deseos? (2) ¿Qué especies tienen claramente una psicología de creencias y deseos? (3) ¿Puede haber deseos definidos que sean resultado de un proceso de selección natural? Si es así ¿exhiben el mismo grado de finura que el contenido?

Respecto de Millikan y Papineau podría pensarse que defienden de algún modo cierto innatismo. Aunque no es así, uno estaría tentado a admitir que ambos autores son fieles defensores de alguna corriente innatista minusvalorando los aspectos de aprendizaje en la instauración de nuevas creencias y repertorios conductuales. En el caso de Millikan la novedad se explica recurriendo al concepto de función propia derivada cuyo ejemplo paradigmático es el mecanismo pigmentador de la piel del camaleón.

Respecto de las teorías causales-informacionales hay que señalar que las teorías teleosemánticas siguen enfrentándose de un modo u otro a los mismos problemas como el error, el problema del contenido distal o el problema de las representaciones vacías (inexistencia intencional). Si bien es cierto que las teorías teleosemánticas tratan de responder a la amenaza de la indeterminación, su respuesta, como mostraré en el siguiente capítulo, no es satisfactoria, al igual que las soluciones que pudieran argumentarse en favor de los otros dos problemas.

En conclusión, las diferentes teorías teleosemánticas que aquí se han expuesto presentan una imagen del contenido mental bastante uniforme, centrada principalmente en la determinación del contenido por medio de la fijación de la referencia desde parámetros claramente naturalistas y que alguna plausibilidad puede tener. Las limitaciones y deficiencias de este proyecto serán examinadas con mayor detenimiento en el siguiente capítulo.

VI. Las limitaciones del proyecto teleosemántico

Contenidos

Introducción: Los problemas de las teorías teleológicas del contenido

1. El concepto de función biológica es inadecuado
2. El problema de la indeterminación
3. La independencia de la polémica entre internismo y externismo y el caso del hombre del pantano
4. La cuestión del contenido abstracto
 - 4.1. Naturaleza del contenido abstracto
 - 4.2. Contenido abstracto y aprendizaje
5. Utilidad biológica y contenido mental.
6. La cuestión del contenido no conceptual
7. La teleosemántica no explica la vida mental al completo

Comentarios finales

Introducción: Los problemas de las teorías teleológicas del contenido

Las diversas teorías teleosemánticas que he examinado comparten una serie de rasgos en común como son los siguientes: (1) Todas ellas parten del realismo intencional, es decir, admiten que nuestra vida mental puede caracterizarse adecuadamente en términos de creencias y deseos. (2) Ninguna caracteriza el pensamiento como un episodio exclusivamente lingüístico; todas ellas o bien rechazan la hipótesis del lenguaje del pensamiento -Millikan- o bien son indiferentes a ellas -Dretske-. (3) Todas ellas -salvo Dretske, que no se pronuncia al respecto- creen en que el concepto de función biológica tal como lo define la perspectiva histórico-etiológica es el concepto adecuado para garantizar la fijación del contenido mental. (4) Todas parten de puntos de vista externistas.

Estos supuestos de base tienen claras ventajas respecto de otras teorías rivales como:

(1) A pesar de ser teorías que parten del realismo de nuestras actitudes, la confrontación entre imagen científica y psicología popular no es tan acusada. El objeto de estudio no es la caracterización de nuestras actitudes, sino la fijación del contenido de las mismas bajo parámetros científicos. La teleosemántica integra la psicología popular en un marco aparentemente científico. Por lo tanto la cuestión de la fijación del contenido puede hacerse al margen de la discusión en torno al estatuto de la psicología popular.

(2) Siendo teorías naturalistas explican aparentemente bien el problema de la disyunción, puesto que hacen depender la normatividad del contenido de una noción biológica -una noción "científica"-. En principio éste es el principal argumento con el que cuentan las diversas teorías teleosemánticas. Pero, como más adelante mostraré, está viciado desde el principio; no se trata de defender la normatividad -algo que en el caso del contenido no se cuestiona- sino de la legitimidad de su defensa desde la perspectiva teleosemántica.

(3) Al no depender de ninguna de las variedades de la hipótesis LOT no tienen, en principio, por qué enfrentarse al problema de tener

que explicar el contenido en organismos alingüísticos, de hecho, eso no representa un problema para la teleosemántica. En este sentido la teleosemántica puede ir de la mano de la semántica del éxito (*success semantics*); si la segunda supone un procedimiento metodológico, la primera estipula las condiciones bajo las que acaecen las propiedades semánticas de los contenidos.

(4) También son neutrales respecto de la polémica funcionalista entre computacionalismo clásico y conexionismo, ya que exploran otra vía de caracterización de la semántica.

Sin embargo, las diferentes teorías teleosemánticas se enfrentan a problemas muy diversos, he aquí los que considero primordiales: (1) No emplean un concepto adecuado de función biológica. (2) El problema de la indeterminación persiste; el recurso a la funcionalidad biológica no logra superar el escollo de la indeterminación. (3) La cuestión externista es independiente. (4) No explican la fijación del contenido sofisticado o abstracto. (5) No explican toda la riqueza de nuestros estados intencionales. (6) No abordan la cuestión del contenido no conceptual y (7) No explican nuestra vida mental al completo. En las páginas que siguen desarrollo más detenidamente cada uno de los anteriores siete tópicos.

1. La inadecuación del concepto de función biológica

Habitualmente las críticas a la teleosemántica emplean como argumento principal el que el recurso a la funcionalidad biológica no evite el problema de la disyunción (Fodor 1987; 1990) o, de otra manera, insisten en que la indeterminación persiste y que por lo tanto son teorías ineficaces en la determinación del contenido. Esto ha llevado a autores como Fodor a plantear alternativas muy próximas a las teorías causales-informacionales y a los planteamientos teleosemánticos¹⁹¹. Pero mi crítica es más básica; el concepto de

¹⁹¹ Respecto de Dretske (1981) Fodor (1987) aduce que la teoría causal no puede enfrentarse adecuadamente al problema del error ya que por definición los ejemplares representacionales de GATO son causados por la clase de referencia que constituyen los gatos y por tanto no es posible tener instancias falsas de GATO. Pero el caso es que los distintos organismos albergan en ocasiones representaciones erróneas. Por lo tanto hay que admitir que la representación GATO es causada por

función biológica que emplean es erróneo. No hay nada que fijar con ese concepto a riesgo de hacer ciencia ficción. En efecto, tal como ya señalé en los capítulos 2 y 3, las teorías teleosemánticas parten de un concepto erróneo. Las teorías histórico-etiológicas no son adecuadas desde el punto de vista de la práctica científica por lo que supone una anomalía presentarse como naturalista y fundamentar la teoría que se

gatos y por entidades distintas de los gatos como los linceos. Esto último no es sino una manera de referirse a la robustez de las representaciones mentales. Respecto de las teorías teleosemánticas Fodor señala que las condiciones óptimas bajo las que según los partidarios de la teleosemántica se fija el contenido no nos ofrecen garantías de arrojar siempre resultados verdaderos; la utilidad biológica no es lo mismo que el concepto semántico de verdad. *"Clearly we can't define 'optimal circumstances' in terms of the teleology of those mechanisms; not if we're to define truth conditions as the ones that are satisfied when a symbol is caused in optimal circumstances. (...) The teleology story perhaps strikes one as plausible in that it understands one normative notion -truth- in terms of another normative notion -optimality. But this appearance of fit is spurious; there is no guarantee that the kind of optimality that teleology reconstructs has much to do with the kind of optimality that the explication of 'truth' requires. When mechanisms of repression are working 'optimally' -when they're working 'as they 're supposed to' -what they deliver are likely to be falsehoods."* Fodor (1987: 105-106)

En su lugar Fodor ofrece una teoría causal más refinada que es conocida como la teoría de la dependencia asimétrica. En líneas generales la teoría de la dependencia asimétrica afirma que las representaciones erróneas dependen de los objetos que correctamente instancian las representaciones y no viceversa, es decir, porque los gatos causan la representación GATO es posible que los no-gatos (i.e. los linceos) causen también la representación GATO. De este modo la causación de GATO por los gatos depende de una ley causal. Es decir, la causación de GATO tiene un carácter nomológico; los gatos causan GATO, los linceos también pueden causar GATO. Pero esta segunda ley está en dependencia respecto de la primera; la conexión entre los gatos y GATO es más pertinaz o sólida que la que se da entre linceos y GATO.

La teoría fodoriana es la gran alternativa a las teorías causales como la de Stampe o Dretske. En su contra tiene que es una teoría difícilmente confirmable ya que la teoría requiere que los sujetos representantes tengan disposiciones sincrónicas (Fodor, 1987:109) para albergar las representaciones producidas por gatos y linceos. Pero, en cualquier caso, no es esta la línea que han seguido otros autores partidarios de la perspectiva teleológica que en lugar de centrarse en argumentos basados en contrafácticos han puesto énfasis en la naturalización a través del concepto de función biológica.

propone en un concepto que hace tiempo que ha sido desestimado por la comunidad científica. Si se quiere ser consecuente con el naturalismo filosófico que defienden deberían aplicar otro concepto de función biológica. La razón no es porque los científicos lo emplean, sino porque, como ya he mostrado, creo que hay más argumentos a favor de una perspectiva no histórico-etiológica que en contra. Incluso si se trae a colación el controvertido tema de la normatividad se puede explicar satisfactoriamente que ésta no radica únicamente en la historia selectiva del ítem.

Del mismo modo que se podía acusar a las diversas teorías histórico-etiológicas de seguir el programa adaptacionista, podría decirse lo mismo respecto de la teleosemántica (Fodor, Neander). En mi opinión es ésta una acusación legítima, sólo que irrelevante. Como ya mostré en el capítulo III el adaptacionismo es una posibilidad que muy bien puede intepretarse como heurístico y no como una amenaza hacia las teorías biológicas. Si la teleosemántica ha de ser una aplicación del programa naturalista a las cuestiones semánticas ¿por qué no va a proceder del mismo modo que proceden las distintas ciencias? Esta podría ser una buena respuesta ante la acusación de adaptacionismo. Pero creo que no va con el espíritu que algunos partidarios de la teleosemántica han propugnado. La concepción que tenga Millikan de la ciencia no parece contemplar el recurso a teorías que funcionen como heurísticos, máxime cuando la propia noción de función que emplea proviene de Wright, quien, como mostré en el II capítulo, trata de llevar el análisis de Hempel sobre la explicación funcional más allá de un mero recurso heurístico. El adaptacionismo realmente es una amenaza si uno parte de presupuestos excesivamente realistas sobre la naturaleza de las teorías biológicas.

2. El problema de la indeterminación

La acusación de indeterminación no es nueva. Uno de los primeros en hacerla fue Fodor (1987; 1990) quien se percató de que ni el enfoque causal-informacional ni el teleológico podían enfrentarse a casos en los que la indeterminación era manifiesta. De hecho, los intentos posteriores de los autores estudiados se han centrado en

hacer frente a los casos de indeterminación propuestos por Fodor. Así Dretske (1988), Millikan (1993), Neander (1996) o Papineau (1997).

Pese a estos intentos hay autores que como Perelmann (2002) insisten en que la indeterminación persiste y que por lo tanto las teorías teleológicas del contenido fracasan respecto de la cuestión de la determinación del contenido. Así, aún cuando decidiéramos no mantener la crítica anterior basada en la inadecuación del concepto de función biológica, podría decirse que el problema de la indeterminación persiste.

En cualquier caso pueden mantenerse dos posturas: La primera estaría caracterizada por aquellos que como Millikan o Neander creen que todo contenido debe estar determinado de alguna manera, bien a través de un recurso *High Church* (Millikan) bien desde una perspectiva más modesta o *Low Church* (Neander). La segunda es la postura defendida de diferente manera por Dretske y Papineau. Ambos autores imponen ciertas restricciones para que un organismo posea un contenido determinado. Tal como he mostrado, en el caso de Dretske la determinación del contenido está definida de diversos modos según la naturaleza del propio contenido, es decir, según consideremos sistemas representacionales de tipo I, II o III. Pero en modo alguno la noción de función biológica constituye por sí misma la condición de fijación del contenido, sino sólo una condición más en determinados sistemas representacionales. En el caso de Papineau (1997), la determinación del contenido sólo es posible si se posee una psicología caracterizable en términos de creencias y deseos y depende fundamentalmente de los deseos que alberga el organismo a modo de funciones biológicas. En el caso de Millikan es quizás más obvio por el énfasis que pone en explicar cualquier manifestación intencional en términos de funciones propias o funciones propias derivadas. Admitamos que la teoría de Millikan es adecuada y que, en efecto, el contenido de, por ejemplo, la representación de la danza de las abejas, se determina en base a la función biológica que ésta tiene. Consideremos otro ejemplo, como organismo biológico supuestamente bastante complejo, estoy dotado de mecanismos que me permiten distinguir con un cierto grado de fiabilidad aquellas cosas comestibles de aquellas que no lo son, es más, mi sentido del gusto permite distinguir entre el queso y el bizcocho y, aún más, puedo distinguir

diversos tipos de queso como un mahonés de un queso azul. Pero no parece que mis papilas gustativas tengan la función de distinguir el queso azul del mahonés. No hay algo más allá de una función general para distinguir cosas comestibles de cosas no comestibles. En cualquier caso lo que puede ser determinado a través de la funcionalidad biológica tal como la entiende Millikan es la categoría general de COSA COMESTIBLE y no la riqueza conceptual de las subcategorías. La función de mi sistema digestivo es nutrirme y este objetivo se cumple con independencia de qué coma concretamente. Aunque defendiéramos que nuestra posesión de conceptos depende de nuestras habilidades cognitivas como puedan ser el aprendizaje, la memoria o el lenguaje, fenómenos que a su vez son funciones de las diversas partes de nuestro psiquismo. Sin embargo, el concurso de estas capacidades cognitivas es común a la posesión de cualquier concepto y no lo determina en tanto que concepto. Parece que a lo más que puede llegar una caracterización del contenido en los términos anteriormente indicados es a determinar algunas categorías máximamente generales. No dudo que mis gatos tengan categorías, distinguen, comida de no comida, sitio adecuado para descansar de sitio no adecuado, etc pero no creo que distingan entre dormir en un colchón de látex o uno de muelles, ni parecen muy sensibles a otros matices. Si nos fijamos únicamente en la funcionalidad biológica de nuestros mecanismos consumidores de representaciones nuestra vida mental es muy monótona.

Además, ¿a qué estructuras les son adscritas las funciones? Es obvio que cuando nos referimos a partes del organismo adscribimos funciones a cada una de ellas pero lo que ocurre con nuestro sistema cognitivo no es lo mismo. No hay un órgano de percibir queso azul, el contenido no es la circulación sanguínea o la función renal, es algo más difuso y particularista. O bien admitimos que en general no hay funciones, es decir, que son una quimera filosófica para explicar la ejecución regular de ciertas actividades o bien que éstas supervienen a estructuras muy concretas y tienen una realidad muy delimitada. Esto último parece poco factible, nadie ha identificado las estructuras concretas de mi percepción del queso azul y, además, seguro que no difieren de las que empleo para percibir el roquefort o el cabrales. Estas distinciones parecen más bien constituidas sobre el aprendizaje

que, a su vez, está construido sobre mi dotación de recursos cognitivos como organismo. Decir que las funciones de los mecanismos cognitivos determinan el contenido de mis representaciones equivale a decir bien poco acerca de la determinación del contenido. Más bien lo que ocurre es que el organismo actúa como un todo organizado y difícilmente puede ser atomizado en partes y partes a las que adscribamos funciones. Sin embargo esta perspectiva corre el riesgo de conducirnos a un holismo malsano; a riesgo de no querer decir nada concreto nos quedamos en la más absoluta de las indeterminaciones. El caso del contenido mental agudiza todavía más esta situación. Todo esto puede condensarse en la siguiente pregunta ¿a qué superviene una función? En la medida en que no identificamos muy claramente la base sobreviviente no podemos responder la cuestión. De otra manera, Millikan afirma que existen unos mecanismos consumidores de representaciones y que el contenido se fija en la medida en que estos mecanismos necesitan un cierto tipo de representaciones para ejecutar su función propia. Ahora bien ¿qué realidad tienen estos mecanismos? Tienen algún substrato neuronal diferenciado o más bien responden a una explicación funcional? Si como parece la respuesta correcta es la última entonces nos encontramos con que la naturalización que propone Millikan es parcial, pues aunque recurre a un concepto "científico" no podemos identificar ninguna estructura funcional concreta. La existencia de los mecanismos consumidores de representaciones es una hipótesis poco parsimoniosa; hay las mismas razones para adscribir funciones biológicas a estos sistemas como para decir que su funcionalidad no responde a un planteamiento biologicista.

Tanto el enfoque de Neander como el de Papineau son menos sensibles a las críticas puesto que son también menos ambiciosas. La teoría de Neander es menos vulnerable, pues parte de presupuestos *Low Church*. En este sentido la respuesta al problema de Fodor es clara; en el nivel más básico la rana se representa como bolitas negras pequeñas voladoras y no comida. La estrategia de Cummins permite fijar el contenido si descomponemos el sistema visual en los elementos más básicos y, además, es susceptible de una lectura desde presupuestos seleccionistas o teleológicos. De este modo se hace frente al problema de la indeterminación pero al elevado precio de

minimizar la dimensión teleológica. Así no hay indeterminación aparentemente; la rana se respresenta claramente las bolitas negras y... resulta que en su medio son comida. Por tanto nuevamente nos encontramos con comida o bolitas negras (y la disyunción es inclusiva). Papineau (1997) advierte que aunque la estrategia de Neander pueda resultar parcialmente satisfactoria, se aleja de una perspectiva teleológica y se aproxima peligrosamente a los presupuestos de la teoría causal-informacional y, por lo tanto, se expone a las mismas críticas que aquella como acabo de mostrar.

En el caso de Dretske el recurso explicativo de la funcionalidad biológica sólo es aplicable a los sistemas de tipo II y de tipo III y en estos últimos hay claras indeterminaciones que Dretske no niega, pero que tampoco le preocupan. El uso del concepto de función por este autor es mucho más discreto y cauto que el que hacen Millikan, Neander o Papineau. No se puede hablar de un contenido determinado más que en los sistemas de tipo I y en esos la funcionalidad biológica realmente poco importa, ya que son otros factores los que contribuyen a fijar el contenido como puedan ser los procesos de aprendizaje bajo condiciones ciertamente ideales. Aunque permeable a las críticas de Fodor (1987; 1990), no parece que tal como formula Dretske la cuestión de la fijación del contenido se vea realmente afectado. Para teorizar es necesario proponer condiciones ideales. En este sentido no creo que la teoría de Dretske se vea mermada por las réplicas de Fodor.

Respecto de Papineau ya argumenté que el deseo tiene un carácter de por sí bastante indeterminado. Como señalé en el capítulo anterior Papineau debe responder a tres interrogantes: (1) ¿Qué grado de definición es necesario para poder decir que un organismo alberga "deseos definidos" y poder atribuirle una psicología de creencias y deseos? (2) ¿Qué especies tienen claramente una psicología de creencias y deseos? (3) ¿Puede haber deseos definidos que sean resultado de un proceso de selección natural? Si es así ¿exhiben el mismo grado de finura que el contenido?

Identificar el grado de definición de los deseos puede resultar una tarea ciertamente difícil. Aunque sí hay deseos muy primarios que son resultado de un proceso selectivo. Por ejemplo, la conducta alimenticia puede estar guiada por las necesidades del organismo, sin embargo lo

que se desea son determinadas sustancias y no chocolate Nestlé o bombones Lindt. Las conductas estereotipadas en ese sentido son bastante inespecíficas, aunque menos de lo que cabría esperar; los niños con déficit de calcio lamen las paredes encaladas y las palomas y demás aves picotean los muros de caliza aunque no tengan predilección por la catedral de Palma o por la de Salamanca, algo que para la satisfacción del deseo es absolutamente irrelevante. Lo mismo puede ocurrir ante la necesidad de ingesta de alimento, en condiciones normales a uno le va bien cualquier cosa. Pero aunque así planteado pueda parecer una parodia de la teoría de Papineau lo cierto es que no albergamos deseos con tal grado de definición que fijen de una manera unívoca el contenido de nuestras creencias. La relación que puedan guardar los deseos con las creencias son motivacionales y la motivación, salvo en el caso de las necesidades básicas del organismo o motivos primarios –sueño, ingesta de líquidos y alimentos, sexo-, son poco definidas. Suponer que los deseos desempeñan la función de combinarse con las creencias para generar conductas que los satisfagan puede verse como una petición de principio, ya que al final recurrimos a las propias creencias para fijar el contenido de otras creencias. Por ejemplo, tengo deseo de comer chocolate y sé que hay chocolate en la nevera, mi deseo de comer chocolate se combina con mi creencia de que puedo obtener chocolate y así fijo el contenido de la creencia que dirige mi acción consumatoria. Puede verse desde otra perspectiva: los deseos o las motivaciones pueden actuar en la economía mental en tanto que conceptualizados, pero en cuanto ya están conceptualizados ya tienen un contenido determinado – pertenecen al ámbito de las razones- y son por lo tanto una razón para actuar. Así nos encontramos con que una razón fija el contenido de otra razón. Justo lo que un naturalista no puede consentir si lo que quiere es naturalizar la intencionalidad. Papineau se ve envuelto en este tipo de polémicas aunque no quiera. Si suponemos una economía de creencias y deseos en la mente del organismo entonces la naturalización a la que podemos aspirar es muy débil; ese es el precio que hay que pagar para poder hacer frente al problema de la indeterminación.

3. La independencia de la polémica entre internismo y externismo y el caso del hombre del pantano

Ser externista en el caso de los teleofuncionalistas no aporta nada. El externismo que defienden es independiente de las caracterizaciones del contenido mental que realizan. Así, sostengo que uno podría ser partidario del internismo y defensor de alguna forma de teleosemántica. Ni los argumentos clásicos del tipo Putnam-Burge, ni las situaciones contrafácticas como el caso del hombre del pantano constituyen argumentos decisivos para que nos decantemos por una u otra teoría de la individuación del contenido mental. En esta sección presentaré en primer lugar la polémica del hombre del pantano y los posicionamientos externistas de los partidarios de la teleosemántica para finalmente mostrar que sus argumentos son erróneos.

Una de las características que se observa en la exposición de las diferentes teorías teleosemánticas es que todas ellas son partidarias de un externismo extremo. Autores como Dretske (1981, 1986), Millikan (1984, 1993) o Papineau (1993, 2001) defienden fuertes posicionamientos externistas y consideran que una buena caracterización del contenido mental debe ser externista o no lo caracteriza en absoluto. La idea general que subyace tras el externismo defendido por los partidarios de las distintas teorías teleológicas del contenido es que el contenido de nuestras creencias y deseos depende de la historia selectiva de los mecanismos que los producen o consumen. Así, en la medida en que han sido seleccionados por cumplir una determinada función puede decirse que el contenido no sobreviene al propio individuo, sino a la historia selectiva del individuo. Por otra parte, los argumentos internistas como he mostrado requieren del concurso de la conciencia por parte de los organismos que albergan contenidos y evidentemente el grado de conciencia que emplean estos argumentos (i.e. Farkas 2008) sólo lo encontramos en humanos adultos normales. El resto de organismos puede albergar estados fenoménicos pero no estados conscientes al nivel requerido por los argumentos alineados con tesis cartesianas. Así, una de las razones por las que el externismo es aceptado por los partidarios del enfoque teleológico es porque en absoluto dependen de

la conciencia para sus atribuciones intencionales y por lo tanto los argumentos cartesianos se neutralizan fácilmente.

El argumento clave en favor del externismo lo encontramos formulado en Davidson (1987). Así, que el contenido se determina relacionamente puede ser mostrado a través del estudio del caso *swamp*; el hombre del pantano, se dice, no puede albergar más que contenidos fenoménicos pues carece de la historia selectiva apropiada para albergar otro tipo de contenido. De otra manera, sin historia selectiva sólo puede haber en el mejor de los casos un contenido informacional o fenoménico carente de normatividad.

El análisis del caso *swamp* ha constituido hasta la fecha una de las principales preocupaciones de los partidarios de la teleosemántica. Millikan (1984: 93 y ss.), antes de que Davidson formulara su experimento mental, defendía ya un externismo extremo basado en la consideración de dobles en el que no había lugar para las consideraciones de tipo cartesiano. Respecto de un hipotético hombre del pantano subraya su carencia absoluta de ideas, aunque no de estados conscientes. El hombre del pantano no tiene contenidos intencionales porque estos son categorías biológicas adquiridas a través de la historia selectiva.

Papineau (1993: 91 y ss.) señala que el hombre del pantano no comparte nuestros estados representacionales al carecer de la misma historia selectiva que su original. Quizás el enfoque más detallado y sofisticado sobre esta cuestión lo encontremos en Neander (1996; 2008) ya que analiza tanto la importancia del caso *swamp* para el concepto de función como para la caracterización teleológica del contenido mental, pues sostiene que la dimensión histórica es imprescindible para atribuir normatividad tanto a funciones como a contenidos. A diferencia de los autores considerados anteriormente, Neander cree que el hombre del pantano sólo tendría contenido en sentido estrecho Neander propone que consideremos los casos *swamp* desde la perspectiva más general de las funciones que atribuiríamos a sus diferentes órganos y no sólo a la determinación del contenido. Para esta finalidad idea un caso análogo al hombre del pantano que denomina *Swampcow*. Una vaca del pantano es, al igual que su análogo humano, una réplica accidental exacta molécula por molécula de una vaca particular carente de historia selectiva alguna. Entre las

diferencias que a primera vista afloran es que Tonina, nuestra vaca *swamp* réplica de Tonona, no puede ser clasificada como perteneciente a la especie de las vacas según los criterios taxonómicos más habituales o la carencia de propio-funciones de cualquier rasgo orgánico. Su corazón bombea la sangre y sus riñones la filtran pero, como señala Neander, lo hacen por cortesía, ya que ambos ítems carecen del trasfondo histórico que determina su propio-funcionalidad según la teoría histórico-etiológica.

Vayamos ahora a las consideraciones acerca de la vida mental de un hombre del pantano. Parece difícil afirmar que carece de cualquier intencionalidad, pero si admitimos que de algún modo tiene intencionalidad, esta afirmación sería inconsistente con las teorías teleosemánticas. Hay diversas posibilidades que deben ser consideradas respecto de su posesión de contenidos mentales: (1) Una primera opción consiste en admitir que no tiene contenidos mentales en absoluto, pues carece de la historia necesaria para fijarlos. (2) También podemos pensar que su vida mental está limitada a la posesión de estados fenoménicos muy básicos y poco más, es decir, incluso carecería de actitudes proposicionales al no poder tener un yo. (3) Finalmente también es lícito pensar que tiene creencias y deseos análogos a los de su original en la medida en que está dotado de las mismas disposiciones causales sobrevenidas a su estructura física, tal como le ocurre a su doble real. Es decir, se mantienen los roles funcionales en un sentido sistémico pero no las teleofunciones. Esta última tesis, rechazada por Millikan y Papineau, es la defendida por Neander (1996; 2008) tanto para el caso de las atribuciones funcionales a órganos como a representaciones. Lo más sobresaliente de esta situación es su anomia funcional y de contenido. En el caso particular del contenido nuestro doble carecería de la dimensión normativa, pues al obviarse la dimensión histórica no estarían presentes las normas que establecen qué cuenta como una representación correcta y qué no. Podríamos postular que tiene algún tipo de contenido, aunque posiblemente no podríamos hacer ninguna adscripción correcta de contenido. Así, es obvio que no tiene contenido en sentido amplio (*Broad Content*), pero sí en sentido estrecho (*Narrow Content*), es decir, aquel contenido sobrevenido a sus propios estados psicológicos individualizados y accesibles a su introspección. Por lo

tanto puede afirmarse que un caso *swamp* tendrá todas aquellas propiedades que supervengan al propio individuo de manera intrínseca y sincrónica y carecerá de todas aquellas propiedades cuya base sobreviniente sobrepase esos límites. Ésta puede ser una buena respuesta, pero previamente habrá que darle un carácter más substantivo, es decir, se hace necesario definir qué estados sobrevienen a las propiedades físicas intrínsecas de un organismo y qué estados no. Un autor puede considerar que la conciencia fenoménica o los *qualia* son sobrevenidos a partir de las propiedades individuales - como la propia Neander-, otro autor puede ser partidario de una teoría de la conciencia de corte más externista (i.e. Dretske 1995). Hay que señalar que la respuesta anterior tiene sentido dentro del posicionamiento *Low Church* de Neander, puesto que para la ciencia cognitiva puede resultar más operativo tratar con representaciones localizadas en sistema concretos. En este sentido el contenido en sentido amplio no puede ser abordado por la ciencia cognitiva, aunque quizás deba ser supuesto para dar cuenta de las propiedades semánticas inherentes a las representaciones que un determinado sistema maneja. El argumento es muy similar al de Fodor (1987) y en cierto modo fue anticipado por Dretske (1981: 214) quien señalaba que si bien los significados no están en la cabeza, sí lo están los conceptos. Si bien Dretske no aceptaría una teoría del componente dual como sí hace Fodor.

Los argumentos anteriores a favor del externismo fueron respondidos de una manera contundente por Braddon-Mitchell y F. Jackson (1997) quienes piensan que la historia selectiva es insuficiente para la determinación del contenido. En su artículo de 1997 "*The teleological theory of Content*" señalan que la teleosemántica, como teoría que trata de caracterizar el contenido en términos de historia selectiva, es insatisfactoria. Si la historia selectiva es un rasgo esencial del contenido entonces debe haber un contenido determinado para cada historia selectiva. Pero puede suceder que los mecanismos que causan contenidos y que han sido desarrollados evolutivamente sean en realidad los mismos para muchas historias selectivas posibles. Así, señalan Braddon-Mitchel y Jackson, lo que importa no es la historia selectiva en tanto que historia selectiva, sino que el sistema que maneja las representaciones sea de una determinada clase con la

historia evolutiva correcta. Es decir, lo que se busca es la compatibilidad con un determinado entorno. La razón que subyace a este argumento es que la historia selectiva no es una propiedad esencial o constituyente, ya que no afecta a la identidad transmundana. Así, podemos afirmar que un objeto tiene una función sin necesidad de considerar su historia selectiva. Este argumento evidentemente supondría una merma del poder normativizador que atribuyen los partidarios de la teleosemántica a la teoría histórico-etiológica y atentaría igualmente contra esta teoría de las funciones biológicas. Una distinción importante que elaboran estos dos autores es la de contenido informacional -o contenido sobrevenido individualmente- y contenido selectivo -contenido individuado relacionamente tal como suponen los partidarios del externismo-. La teleosemántica se centraría únicamente en el contenido selectivo y entendería que el contenido informacional es reducible al primero. Sin embargo, como que es posible que diversas historias selectivas fijen el mismo contenido no tiene sentido preguntarse por la historia selectiva y carece del peso que a menudo le atribuyen los partidarios de las teorías teleológicas.

En 2001 Papineau publica "*The Status of Teleosemantics or How to Stop Worrying about Swampmann*" en donde trata de contestar a las alegaciones de Braddon-Mitchell y Jackson tratando de desactivar su argumento modal en favor de que otras historias selectivas no tienen por qué arrojar contenidos diferentes. Papineau piensa que el núcleo central de las teorías teleosemánticas es la afirmación de que los roles de las creencias y deseos se realizan por estados resultantes de procesos de selección en el mundo presente; si hubiera sido otro el escenario selectivo entonces los estados mentales también tendrían una identidad diferente.

Como ya sugerí en el capítulo IV, el externismo es la posición dominante en la filosofía de la mente actual, aunque quizás sea erróneo plantearlo como una tesis sin fisuras o que no admita ningún matiz. Sin embargo, si analizamos los distintos argumentos veremos que en la mayoría de los casos no hay razones para adoptar un punto de vista u otro sobre la individuación de las representaciones mentales. El argumento de Fodor (1987) respecto de la teleosemántica es neutral ya que la teleosemántica no precisa de la hipótesis LOT y, por lo tanto, no es necesario suponer unos átomos del pensamiento que deban ser

combinados para formar frases en mentalés. Los argumentos de Bach (1988) o Farkas (2008) sí podrían ser relevantes, pues en ambos casos se pone en cuestionamiento el externismo defendido por Burge o Putnam respecto de lo que suponen sus respectivas caracterizaciones cartesianas de la mente. Pero tiene en su contra que la teleosemántica no parte de presupuestos cartesianos. Así, la perspectiva de primera persona que Farkas (2008) reivindica como característica distintiva de lo mental es superflua desde un punto de vista teleosemántico. Lo mental no es lo que parte desde un punto de vista, sino lo que es intencional, aunque esta intencionalidad sea gradual. La idea de un punto de vista no parece tener mucho sentido desde la perspectiva de un animal carente de estados conscientes o capacidades introspectivas. Por supuesto que puede atribuirse un punto de vista, pero ése es irrelevante desde una perspectiva epistemológica; el animal conoce en tanto que está bien adaptado y ese conocimiento es de un tipo muy básico. Por otra parte, aunque el animal se comunica, no emplea signos lingüísticos, así las críticas vertidas desde la consideración de los lenguajes públicos (Bach, Crane) son igualmente irrelevantes para la teleosemántica. Como he mostrado, las teorías teleosemánticas tienen por objeto caracterizar la semántica a un nivel muy básico. Un organismo dotado de conciencia y lenguaje no es el tipo de objeto idóneo para estas semánticas naturalistas.

En mi opinión el argumento del que parten en líneas generales los partidarios de la teleosemántica en su defensa del externismo es defectuoso no solo porque proceda a partir de un concepto erróneo de función biológica que invalida la fijación del contenido, sino porque como señalan Braddon-Mitchel y Jackson (1997) la historia no es una propiedad esencial de ninguna entidad por lo tanto no es constitutiva ni de la funcionalidad ni del contenido. Podemos atribuir funcionalidad y contenido sin necesidad de considerar la historia selectiva, únicamente considerando cuál es la situación actual en la que se desenvuelve el organismo. El problema es que los partidarios de las teorías histórico-etiológicas confunden causación con constitutividad. Papineau (1993: 87 y ss.), por ejemplo argumenta su defensa del externismo a partir de la necesidad de la existencia de los objetos que constituyen la referencia de las creencias. Es decir, Papineau confunde causación con constitutividad. En mi opinión del hecho de que haya actitudes

cuyo contenido es causado por objetos externos al sujeto no se sigue que debemos caracterizar las actitudes en términos externistas, puesto que lo normal es que el contenido sea causado por objetos y que verse sobre estos. El internismo puede caracterizarse desde un punto de vista cartesiano, pero el cartesianismo no niega que haya ideas adventicias causadas por algo ajeno al sujeto, sólo niega su justificación epistémica.

Podemos hacer frente al argumento del hombre del pantano. De hecho no es un argumento tan extraño. La cuestión no es si el hombre del pantano tiene contenidos mentales, sino si los llegará a tener. En mi opinión el hombre del pantano y su original Donald Davidson, son dos entidades diferentes, pero nada impide que sus órganos y sus facultades mentales operen de la manera correcta. El hombre del pantano es capaz de aprender como lo fue en su día Donald Davidson, es capaz de todo lo que es capaz Donald Davidson y si su memoria no se guarda fuera de su cerebro es capaz de recordar todo aquello que podría recordar Donald Davidson en las mismas circunstancias.

Propongo considerar una situación análoga más próxima: preguntémonos ¿qué contenidos poseía la oveja Dolly? Exactamente los mismos que su original. Dolly reconoce la comida como comida, tiene miedo cuando el resto de ovejas también tiene miedo, por ejemplo, si se acercara un lobo, nada en Dolly es diferente salvo que Dolly es exactamente una copia de otra oveja que no ha sido obtenida por los medios habituales. ¿En qué se diferencia el caso del hombre del pantano? En nada. El experimento de Davidson ha sido, además, forzado. Davidson lo planteaba únicamente para señalar que las cuestiones relativas a la normatividad del contenido mental no supervenían al individuo, sino que las determinaba el contexto lingüístico y cultural en el que se encontraba. Pero Davidson en ningún momento quiso ir tan lejos. Así, incluso si trajéramos en una máquina del tiempo al mismo Platón y le habláramos griego clásico con nuestra desagradable pronunciación erasmiana nos llegaría a entender tras una serie de intentos.

4. La cuestión del contenido abstracto

La teleosemántica no explica la fijación del contenido más abstracto. Los ejemplos considerados son a menudo ejemplos que tienen relación con objetos relevantes para el medio del animal y cuyo acceso epistémico es perceptivo. Así, podemos considerar la fijación del contenido de las creencias acerca de qué cuenta como COMIDA y qué no, acerca de qué estímulos representan PELIGRO para el animal y cuáles COBIJO o qué estímulos pueden ser categorizados bajo el concepto de OBJETO y cuáles bajo el de PROPIEDAD. Los casos de indeterminación, a pesar de todo, pueden resultar persistentes tal como muestra el enfoque *Low Church* de Neander respecto de los posicionamientos de Millikan o las críticas de Perlmann (2002); algo es una bolita negra y voladora o es una mosca, o también es ambas cosas. En el mejor de los casos la teleosemántica puede ser una buena teoría para explicar la fijación del contenido perceptivo, pero una muy mala teoría para dar cuenta del contenido complejo o de los conceptos abstractos. Por ejemplo, es muy dudoso de que aquellas creencias o deseos que no tienen por objeto satisfacer las necesidades primarias puedan resultar fijadas por las funciones que adscribimos a nuestros sistemas cognitivos. Esto es incluso así aún cuando decidamos definir la funcionalidad bajo parámetros distintos a los estipulados por la familia histórico-etiológica. En cualquier caso, se defina como se defina la funcionalidad, daremos con algún tipo de adaptación y el conjunto de adaptaciones no puede ser tan amplio como para poder incluir en él creencias alejadas del ámbito perceptivo como puedan ser aquellas formadas con ideas abstractas como los conceptos jurídicos o los conceptos científicos. Ningún animal salvo el hombre necesita de estos tipos de conceptos, incluso el propio hombre ha vivido muchos siglos sin ellos. Ello no implica que no podamos rastrear en la naturaleza las bases de nuestros juicios morales o de nuestra racionalidad científica. Hay suficientes estudios que demuestran más bien lo contrario como por ejemplo los de De Waal (1996). Pero admitir que la posesión de creencias sofisticadas como "El agua hierve a una temperatura de 100°" o "La transducción del estímulo visual depende de la rodopsina" o "La incapacidad jurídica de un individuo debe decretarla un juez" dependen de nuestra naturaleza biológica no implica afirmar que se

reduzca a ella. La fijación del contenido de estas creencias, como la fijación del concepto de número primo no depende más de funciones biológicas que del hecho de que tengamos cuatro extremidades y no ocho. Cualquier extrapolación del concepto de función o incluso de función propia derivada tal como sugiere Millikan, es innecesaria y no aporta ninguna razón substantiva para naturalizar este tipo de conceptos, sino una jerga inútil y confudente.

Hay una diferencia entre una adaptación y aquellos factores que constituyen un genuino contenido: el contenido de las creencias no se agota en sus condiciones de verdad o la clase de estados de cosas a que refieren. Las adaptaciones carecen del grado de especificidad que atribuimos a las creencias. Tampoco tienen algo así como un sentido y una referencia ni se ven involucradas en contextos opacos.

Por otra parte, a menudo identificamos las adaptaciones con rasgos anatómicos o pautas fijas de conducta y no con contenidos, aún cuando podamos pensar que la conducta de un determinado animal está guiada por algún tipo de estado representacional. No se poseen contenidos del mismo modo que se poseen rasgos anatómicos. Ya en el capítulo III señalé el riesgo que corrían los partidarios del adaptacionismo de atomizar el organismo en conjuntos ridículos de adaptaciones. Qué sea un órgano y qué sea una parte es hasta cierto punto una cuestión que depende de la perspectiva del investigador; la anatomía no trata exclusivamente con clases naturales. Del mismo modo una extrapolación de estas categorías al ámbito de las actitudes puede estar sujeta a teorizaciones muy alejadas de la práctica científica actual.

Por otra parte, las distintas adaptaciones de los animales contribuyen en el mejor de los casos a una modificación parcial y limitada del medio; el animal no sólo asimila el medio, sino que se acomoda a él por medio de algún tipo de intervención. Pero las creaciones culturales han logrado modificar el medio de una manera substantiva. Una posible respuesta ensayada por algunos autores, principalmente Dretske (1988) consiste en señalar la importancia de la capacidad de aprendizaje en la fijación de nuestras actitudes. El aprendizaje obviamente es posible porque estamos dotados de mecanismos biológicos cuya función es producirlo, del mismo modo que el ballet es posible porque poseemos un organismo articulado y

sometido a un control motor suficientemente fino. Pero afirmar que poseemos mecanismos cuya función propia es producir o facilitar el aprendizaje o que las conductas aprendidas poseen funciones propias derivadas supone un abuso del concepto de función biológica en el problema de la determinación del contenido. Equivale a decir que pensamos o que albergamos ideas porque está en nuestra naturaleza humana pensar o albergar ideas. En suma, podemos aceptar difícilmente que el contenido de algunas representaciones está fijado a través de la función propia de los mecanismos que las producen o consumen. Pero extrapolar esta idea al contenido de las actitudes que involucran conceptos abstractos es una tarea implausible.

4.1. Naturaleza del contenido abstracto

Por contenido abstracto entenderé el contenido de aquellas actitudes proposicionales que involucran uno o más conceptos abstractos, conceptos cuya clase de referencia no responde a una entidad o propiedad perceptible a través de la acción directa y no mediada de los sentidos o a entidades y propiedades cuya existencia es ideal como sucede con los conceptos matemáticos, científicos, jurídicos o filosóficos; no así con algunas operaciones lógicas cuya existencia cabe suponer en la mente de organismos medianamente inteligentes. Los contenidos abstractos son poseídos de manera casi exclusiva por los miembros adultos y normales de la especie humana. Parece que salvo el concepto de substancia que está presente en los niños durante el periodo sensoriomotor y algunos animales, el resto de organismos carece por completo de otros conceptos abstractos.

De lo anterior se desprende que el contenido abstracto no es de naturaleza perceptiva ni tiene relación alguna con el medio natural. Tampoco tiene que ver con la simple ejecución, sino que su posesión implica de algún modo una intencionalidad de segundo orden tanto en la aplicación como en la adquisición de esos conceptos. Además, la referencia de los contenidos abstractos, si existe de algún modo, difícilmente puede tener algún impacto causal sobre nuestras

mentes¹⁹². Por tanto difícilmente puede ser explicado de un modo naturalista. Además, podría darse el caso de que la referencia es vacía, el número π no existe en el mundo perceptible; nadie puede identificar cosas que son ejemplares de ese número, del mismo modo que nadie cree que cuatro guijarros son el número cuatro. En este aspecto la teleosemántica no tendría que dar únicamente razón de cómo puede fijarse el contenido abstracto recurriendo al concepto de función biológica, sino que también tendría que explicar cómo podemos albergar contenidos de objetos no existentes y que, en consecuencia, son incapaces de ejercer poder causal alguno sobre nuestras mentes. Obsérvese, por lo tanto, que aquello que puede ser representado según una interpretación estándar de las teorías teleosemánticas, es aquello que está presente en el medio del organismo y que representa de algún modo un tipo de ventaja. Además, el contenido abstracto difícilmente puede ser atomista, los conceptos abstractos son conceptos complejos; no podemos localizar elementos simples ni en el mundo de las ideas ni en el mundo sensible que constituyan la clase de referencia adecuada para estos conceptos. Por ejemplo, uno no puede aplicar el concepto de número primo si no es capaz de poseer el concepto de número natural o el de divisibilidad. Tampoco se puede poseer el concepto de ciudadanía si no se posee el de persona jurídica.

Sobre su fijación ya he señalado que puede decirse que se construye sobre el aprendizaje como todas las conductas y su fijación no depende de ninguna relación causal -no hay clases de cosas que causan contenidos abstractos- sino de la pericia del usuario. En este aspecto es importante señalar que los procesos de aprendizaje que involucran la adquisición de conceptos abstractos son procesos de aprendizaje estructurados, no espontáneos. Tampoco es un aprendizaje individual, sino que es de carácter social y, a menudo, institucional, pues se aprenden generalmente en instituciones educativas. Además, las experiencias que posibilitan la adquisición de este tipo de contenidos son experiencias sociales en los que no hay flujos de información natural sino que, en cualquier caso, los significados que intercambian los hablantes son significados no naturales en el sentido acuñado por Grice (1951). El aprendizaje de

¹⁹² Cf. Neander (2006: 376).

conceptos abstractos también implica algún tipo de mecanismo de retroalimentación o de control consciente y emplea grandes recursos memorísticos, principalmente a nivel declarativo. Es un contenido además de índole claramente proposicional y lingüístiforme y no se resuelve únicamente en un repertorio estructurado de respuestas motrices condicionadas a estímulos presentes en el medio del animal. Por lo tanto, una teoría que no incida sobre los aspectos anteriores es insuficiente para explicar este tipo de conceptos.

No creo que ninguno de los autores considerados en la sección anterior estuviera en desacuerdo con las afirmaciones vertidas anteriormente. Si bien es cierto que los distintos autores expuestos y comentados en el capítulo V no suelen profundizar en exceso sobre este tipo de creencias. Millikan (1984: 48) declina abordar estas cuestiones. Sin embargo ¿cómo puede naturalizarse el contenido si dejamos de lado gran parte de lo que constituye el contenido de nuestras actitudes? Los humanos vivimos en un mundo cultural, no natural. Las creencias que albergamos tienen que ver con valores estéticos, científicos, morales y no sólo con elementos perceptivos de nuestro medio que, por supuesto, ya no es natural. Papineau (1993: cap. 6) señala las limitaciones de su teoría teleosemántica del contenido argumentando que ésta se circunscribe únicamente a las creencias sobre el mundo natural. La prudencia de Papineau no impide que no se plantee la cuestión de qué clase de naturalismo filosófico propone. Pues si para aquel tipo de contenido que constituye la mayor parte de nuestras actitudes no encontramos una explicación naturalista, ¿para qué nos sirve lo que podemos considerar casi una explicación de sentido común a la luz de nuestra imagen científica acerca de cómo actúan y qué piensan el resto de organismos? Además, con su teoría no tenemos un argumento inequívoco para la fijación de los distintos contenidos básicos como ya mostré en la sección 2 de este mismo capítulo.

4.2. Contenido abstracto y aprendizaje

Decir de la teleosemántica que tiene serias dificultades para enfrentarse a la explicación de este tipo de contenido puede verse como una crítica barata a la misma. Los propios autores considerados

en el capítulo anterior no han centrado sus esfuerzos en la explicación de contenidos sofisticados, sino que obviamente han empezado por lo más sencillo. En su defensa puede pensarse que si lo que persiguen es una naturalización de la intencionalidad la estrategia que debe ser seguida radica en explicar aquel tipo de contenido que poseemos en común con especies más próximas filogenéticamente o más alejadas. En cualquier caso es la estrategia científica habitual; si un investigador desea investigar el mecanismo del sueño es muy probable que inicie sus investigaciones con ratones de laboratorio. De este modo extrapolamos los resultados de la investigación en una especie a otras y justificamos el procedimiento idealmente porque asumimos que existe una relación filogenética entre las diversas especies. Desde el punto de vista de la teleosemántica se dirá que lo importante es que existe intencionalidad más allá del mundo humano. Y no les falta razón. Pero que haya intencionalidad y que podamos rastrearla hasta en especies ciertamente alejadas desde el punto de vista filogenético no implica que demos con una pauta máximamente general para fijar los contenidos en todo el continuo intencional que acabamos de identificar. Las limitaciones de este proyecto, como ya se ha podido observar, proceden incluso de los propios partidarios de los enfoques teleológicos (i.e. Papineau). Ciertamente el caso del contenido abstracto es un caso extremo pero en el mundo humano constituye antes la regla que la excepción.

Fácilmente se podría argumentar que la fijación de las creencias cuyos contenidos son abstractos dependen de procesos sofisticados en los que interviene la conciencia y el aprendizaje tal como apunté en la sección anterior. Esta línea de contraargumentación la emplea la propia Millikan, si bien hay que advertir que la insistencia de los partidarios de la teleosemántica sobre el papel del aprendizaje en la fijación del contenido es muy limitada (Millikan, 1984: 46-49 y 1993: 95, 132-133; Dretske, 1986: 95 y ss.; Papineau, 1993: 64, 171 y ss.)

En el caso de Millikan la contraargumentación está viciada desde el inicio. Aunque, por ejemplo, la propia Millikan insinúa las limitaciones de su proyecto en este terreno¹⁹³, pues hace depender la fijación de otras creencias más básicas resultantes de procesos de aprendizaje

¹⁹³ Cf. Millikan (1984:48).

instrumental - como por ejemplo la conducta de guía del perro lazarillo, o el conocimiento de mis gatos de que no se deben subir a la encimera de la cocina- de un uso cuestionable del concepto de función propia como es el de función propia derivada¹⁹⁴. Las funciones propias derivadas no son funciones propias en absoluto aunque Millikan diga lo contrario, pues no dependen expresamente de procesos de selección sino que se construyen sobre procesos de selección. Millikan aplica ilegítimamente la ley de transitividad. Para Millikan las conductas aprendidas tienen funciones propias derivadas de la necesidad de reducción de ciertas necesidades que sí tienen funciones propias. Es, por ejemplo, lo que sucede en el condicionamiento operante. La recompensa tiene que ver siempre con la satisfacción de una necesidad.

Dretske (1986) es el autor que más se expone sobre la importancia del aprendizaje y quizás el que ofrece una solución menos comprometida, pues la dependencia de la fijación del contenido respecto de las funciones es muy laxa y lo que cuenta es la transmisión de información. Pero en el caso de los contenidos sofisticados la transmisión de información deja de ser algo natural; no podemos hablar de información natural -no son signos naturales- sino que es algo puramente convencional, con lo que la naturalización del contenido está de más.

En cualquier caso apelar o no al concepto de función para explicar la fijación del contenido es insuficiente. Si las funciones son entendidas en un sentido histórico-etiológico no podemos fijar el contenido porque no hacemos justicia a lo que realmente es una función en Biología y,

¹⁹⁴ Neander (2006: 384) señala respecto del uso de la noción de función propia derivada en Millikan lo siguiente: "*Exactly how, or whether, this idea could determine contents for sophisticated representational states (e.g., beliefs in humans) remains obscure. But perhaps an analogy with literal mapping might help. Suppose we (intentionally) select a system for mapping terrain. This involves our selecting mapping rules. However, once we have done so, we can apply these rules to generate indefinitely many maps that represent many terrains. We do not need to select a fresh set of mapping rules for each new map. Along similar lines, the idea might be that natural selection can select certain mapping rules (or respects of isomorphism), which a cognitive system can then exploit to produce indefinitely many novel representations. Millikan's response to the problem of novel concepts (e.g., the concept of an electron, when the concept was first proposed) is along the same lines.*"

además, el rango de creencias que hipotéticamente podríamos fijar sería -salvando el problema de la indeterminación- excesivamente limitado. Nos encontramos por lo tanto con que según Millikan el contenido de creencias sofisticadas debe ser fijado recurriendo a funciones propias derivadas, que son unas entidades muy discutibles. Las funciones propias derivadas son antes una invención *ad hoc* que unas funciones propias genuinas. Su normatividad procede de la función propia de la que derivan y ese paso lo único que hace es diluirla aún más. Apelar como hace Millikan a los mecanismos de recompensa del aprendizaje instrumental no es sino apoyarse en nuestra naturaleza sensible máximamente general. Además, el tipo de explicaciones que aporta Millikan carecen de una virtud epistémica esencial; no son explicaciones parsimoniosas. Decir que el contenido es normativo y decir que su normatividad depende de una conjunción enorme de funciones propias derivadas es lo mismo que no decir nada o decir mucho a costa de no explicar.

En general, aquellas teorías que son menos comprometidas tienden a ser menos problemáticas; por ejemplo, aquellas que como Dretske (1986) o Papineau (1993) no se pronuncian sobre temas concretos o establecen distinciones minuciosas. Son menos falsables y... más filosóficas.

5. Utilidad biológica y contenido mental

Volviendo al tema de las relaciones entre adaptaciones y funciones, la teleosemántica fácilmente podría admitir que toda creencia en tanto que resulta fijada por un mecanismo funcional es una adaptación y, por lo tanto supone de algún modo algún beneficio o utilidad. Esta interpretación fácilmente podría ser acusada de seguir de manera ultraortodoxa el programa adaptacionista tal como apunté en el capítulo III. Pero no voy a explotar esta línea de argumentación, sino que creo que hay una manera más refinada de abordar la cuestión. Al afirmar la anterior tesis a propósito de cómo interpretar no sólo las creencias sino su fijación parece que afirmamos también que para la teleosemántica no hay sitio para el contenido inútil (Millikan, 2006). Todo contenido, cabría suponer, tiene un uso o, de otra manera, su posesión implica algún beneficio para el poseedor, ya que es fijado

bajo determinadas restricciones evolutivas de carácter seleccionista. La idea que algunos autores como Millikan defienden es que en cualquier caso podemos rastrear la senda de la funcionalidad biológica para todos los contenidos que podamos albergar. Éstos, finalmente, resultan fijados conforme al modelo que la propia Millikan propone. Así, se mire como se mire, uno al final topa con la biología y de algún modo su explicación conduce a la idea de que albergamos contenidos que de alguna manera son útiles, pues resultan fijados a través de las funciones propias de los mecanismos que los consumen. Sin embargo, creo que esta idea es errónea: hay ideas buenas y hay ideas nocivas. El contenido no se agota en la utilidad. Que los chinos piensen que tener hijas no es bueno es una idea nociva desde un punto de vista puramente ecológico, que algunos musulmanes consideren positiva la ablación del clítoris es igualmente pernicioso y moralmente censurable o que la gente crea en el tarot y confíe decisiones importantes a un procedimiento carente de las más mínimas garantías epistémicas. Estas creencias no sólo son complejas, sino que además son inútiles, no suponen ninguna adaptación al medio. Poseer creencias no es sinónimo de poseer recursos útiles. Podrá contraargumentarse que he apuntado directamente a las creencias religiosas y que incluso podría sostenerse que estas constituyen algún tipo de contenido abstracto. De hecho la creencia religiosa y el contenido abstracto comparten algunas características. Pero creo que en lo esencial el argumento es correcto, pese al recurso a este tipo de creencias. Millikan (2006) afirma respecto de esta importante cuestión que la cuestión del contenido inútil constituye un serio problema para la teleosemántica. Millikan trata de hacer frente a las objeciones que parten del argumento de la inutilidad de algunos contenidos incidiendo en el sentido amplio de la determinación biológica de estos. Pero esta estrategia es análoga a la que emplea para justificar la existencia de funciones propias derivadas. Es decir, la propositividad biológica de estos contenidos inútiles sería derivada de la genuina propositividad y, por lo tanto también estarían de alguna manera fijados por las funciones propias de los mecanismos que las consumen. Pero la referencia a las funciones propias no deja de ser excesivamente laxa. De este modo hablar de fijación del contenido a través de funciones propias derivadas no es más que una estrategia oscurantista. El problema de Millikan es querer derivar la

normatividad doxástica de la normatividad funcional. Pero como ya he mostrado aunque este paso fuera legítimo, Millikan no dispone de una buena explicación de la normatividad funcional. Su concepto de función propia es inexacto. No sólo arroja indeterminaciones, sino que desde el punto de vista científico es inaplicable. No hay ejecuciones funcionales que respondan a los requisitos que Millikan estipula.

6. La cuestión del contenido no conceptual

En el capítulo IV se trató la dicotomía entre lo conceptual y lo no conceptual como una característica relevante que constituía uno de los seis *desiderata* que en ese mismo capítulo se proponían para evaluar una teoría del contenido. Las teorías teleosemánticas nada pueden decir respecto del contenido no conceptual sino admitir su existencia. El contenido no conceptual no caería bajo su ámbito de análisis pues no resultaría algo que debiera, en principio, ser fijado y, por tanto, no requeriría del concepto de función biológica. Por otra parte, los autores considerados no abordan este tópico con la excepción de los estudios de Dretske (2000) acerca de la teoría de la percepción, aunque de manera independiente respecto de sus consideraciones teleológicas.

Sobre el contenido no conceptual pueden formularse las tres siguientes cuestiones: (1) ¿Acaso el contenido no conceptual no podría tener condiciones de fijación? (2) ¿No tendría una función biológica? Y (3) ¿Qué sucede respecto de los organismos que carecen de conceptos o hasta qué punto pueden relajarse los criterios de posesión de conceptos?

Respecto de lo primero parece que la respuesta es negativa: no hay condiciones de fijación para el contenido no conceptual. Su posesión es por definición extraconceptual y por lo tanto no hay nada que fijar. Está fuera de la ley si con esta afirmación se quiere decir que cae por entero fuera del ámbito de las reglas o del entendimiento. Su existencia es condición necesaria pero no suficiente para que el entendimiento aplique sus categorías; pone la materia, pero no la forma que permite tener un pensamiento. Por otra parte, cuando se cumplen las condiciones para la posesión de conceptos siempre queda un remanente de no conceptualidad. Si se quiere presentar así, es "la

imagen retiniana del observador" si hemos de hacer caso de la metáfora fregeana. Frente a la afirmación anterior puede decirse que algunas teorías teleosemánticas son incapaces de distinguir entre el ámbito de las intuiciones y el ámbito de los conceptos o de las razones. Hay organismos en lo que las condiciones de posesión de conceptos no resultan satisfechas y curiosamente es en estos casos donde el recurso a las funciones más explicativo resulta para explicar o determinar de alguna manera sus estados mentales. Frente a estos la explicación de la posesión de conceptos queda infradeterminada si aplicamos las soluciones teleosemánticas, justo lo que la teleosemántica desearía evitar. Por otra parte, lo no conceptual en humanos tiene un carácter residual; lo no conceptual no entra en el espacio de las razones, sino que pertenece por completo a las intuiciones. No tiene una dimensión cognitiva a diferencia de lo que ocurre en organismos más básicos.

Respecto de la segunda cuestión difícilmente puede atribuirse una función biológica a algo que como el contenido no conceptual es por definición indeterminado. Parece más un efecto o algo derivado que algo resultante de un proceso selectivo. De otra manera puede decirse que no ha habido ningún proceso de selección para el contenido no conceptual. Por otra parte la cuestión de su adaptabilidad permanece igualmente ignota. La posesión de conceptos claramente supone una adaptación al medio si por tal entendemos que la posesión de conceptos resulta más eficaz que la no posesión.

Finalmente, respecto de la tercera cuestión puede argumentarse que la funcionalidad puede fijar sólo muy laxamente -es decir, casi una pseudofijación, si se me permite el uso de este término- el contenido de un estado mental sin que ese estado sea conceptual, pues no se cumplen las condiciones estipuladas como la restricción de generalidad. Así el organismo detecta ciertos estímulos y no otros sin que haya conceptualización y ese proceso detector depende del buen funcionamiento de sus sistemas cognitivos y es resultado del proceso de selección natural. Pero aunque el contenido está determinado al no satisfacerse las condiciones para la posesión de conceptos no puede decirse que se fije ningún concepto; no hay posesión de conceptos. Esa determinación es muy vaga y no se da a nivel conceptual, pero es suficiente para la economía cognitiva del organismo. Así, que se

detecte algo como comida no implica que se posea el concepto de comida; la simple identificación de algo máximamente genérico no conlleva la posesión de habilidades conceptuales al menos en un sentido muy estricto pues la simple identificación no garantiza la posesión de estructuras proposicionales. En el mejor de los casos se posee un protoconcepto. Si bien esta afirmación debe enfrentarse a la acusación de que permite atribuir estos protoconceptos a las amebas y demás organismos carentes de un sistema nervioso. Por lo tanto es mejor restringir al máximo las condiciones para la posesión de conceptos y no propugnar atribuciones tan laxas. Así, en conclusión, las limitaciones de las distintas teorías teleosemánticas van siendo manifiestas a medida que ascendemos en la escala evolutiva y nos topamos con organismos más complejos.

7. La teleosemántica no explica la vida mental al completo

Y podría añadirse que tampoco lo pretende, pues sólo es una teoría acerca de la fijación del contenido. Sin embargo, el recurso constante al concepto de función parece no quedarse sólo en eso. En estas líneas exploro los presupuestos metafísicos que subyacen a la caracterización de lo mental de la que parte la teleosemántica entendidos como una metacrítica a la misma desde los dos posicionamientos básicos a partir de los que se puede caracterizar el ámbito de lo mental: la teoría aristotélica frente a la teoría cartesiana. La teoría teleosemántica necesita antes de un enfoque aristotélico que cartesiano. Si bien asumir este enfoque tiene un precio elevado.

Si algo queda claro en torno a las teorías naturalistas del contenido es que no ciñen la vida mental al ámbito de la conciencia. De otro modo, son aristotélicos, no cartesianos. Aunque la perspectiva aristotélica es más idónea para una naturalización, de hecho ya de por sí contempla la continuidad de niveles en la naturaleza, la mente humana adulta resulta mejor caracterizada desde la perspectiva cartesiana. Como apunté al inicio del capítulo IV, nuestro acceso cotidiano a lo mental es a través de la conciencia. La vivencia consciente de nuestros contenidos mentales es paradójicamente lo más familiar y es lo que la concepción cartesiana destaca frente a la caracterización aristotélica.

Para Aristóteles el alma (*psyché*) es el principio de la vida y se posee gradualmente. Tienen alma, si por tal concepto entendemos aquello que informa a una substancia orgánica¹⁹⁵ y que hace que sea lo que es, tanto las plantas¹⁹⁶ como los animales o el hombre de manera más específica. El concepto aristotélico de alma puede desligarse hasta cierto punto relevante de las ideas de inmortalidad o incorporeidad. Por otra parte, su división tripartita es coherente con una gradación de la naturaleza tal como sugiere la idea de *Scala Naturae*. De este modo se procede de lo común a todos los organismos hasta lo más específico. Esta idea de gradualidad tiene que ver con el principio *Natura non facit saltum* que de algún modo ha pervivido hasta nuestros días en algunas interpretaciones de la teoría de la evolución. El alma nutritiva es compartida por las plantas, los animales y el hombre. El concepto de alma nutritiva hace referencia a las funciones nutritivas que tienen los organismos vivos. El concepto de alma nutritiva recoge los requisitos más básicos para la vida, la autonutrición y el desarrollo de los organismos¹⁹⁷. Por su parte el alma sensitiva es gradual en su complejidad, es decir, si lo propio de la facultad sensitiva es precisamente la sensación (*aisthesis*) es obvio que no todos los animales poseen las mismas capacidades sensoriales pues no todos ellos tienen cinco sentidos. Aunque también pertenece a esta parte la facultad de desear (*epithymía*) lo que para la tradición escolástica eran los apetitos sensibles¹⁹⁸. Lo propio del viviente animal es el deseo. Si bien el deseo es distinto para los seres irracionales que para los racionales. En los primeros el deseo es algo que responde meramente a la sensación, a la evitación del dolor y a la consecución del placer. En los segundos la facultad desiderativa está abajo control racional y por lo tanto no se habla de mero deseo, sino de una acción racional fruto

¹⁹⁵ El alma es la entelequia de un cuerpo organizado u organismo tal como Aristóteles afirma en *De Anima II*, 1 "Queda expuesto, por tanto, de manera general qué es el alma, a saber, la entidad definitoria, esto es, la esencia de tal tipo de cuerpo."

¹⁹⁶ Cf. *De Anima II*, 2. "Y llamamos facultad nutritiva a aquella parte del alma de que participan incluso las plantas."

¹⁹⁷ Aristóteles en *De Anima II*, 1 define la vida del modo siguiente: "Solemos llamar vida a la autoalimentación, al crecimiento y al envejecimiento."

¹⁹⁸ Cf. *De Anima II*, 3. "Baste por ahora con decir que aquellos vivientes que poseen tacto también poseen deseo".

de la deliberación (*bouleesis*). Finalmente tenemos el alma racional, poseída de modo exclusivo por los seres humanos sobre la que ya hablé en el capítulo IV.

La imagen aristotélica de la mente es bastante completa, Aristóteles sí ofrece lo que podríamos definir como una imagen de nuestra vida mental al completo. Además, como puede verse nuestra psicología popular es heredera de la teorización aristotélica pues contempla tanto las creencias como los deseos y justifica la acción desde el conocido silogismo práctico¹⁹⁹. También tiene la virtud de que dentro del mismo ámbito natural destaca la posesión de capacidades racionales justificatorias de la acción y constituidas sobre elementos más simples como las propiedades destacadas respecto del alma sensitiva. Para Kenny (1989), que sigue la tradición aristotélico-tomista, lo mental está relacionado de una manera más específica con el intelecto y la voluntad²⁰⁰. En este sentido destaca dos importantes elementos que, sin embargo, también han sido incorporados de un modo peculiar y muy restringido por la perspectiva cartesiana. Para Descartes tanto la inteligibilidad que se supone al entendimiento como la voluntad cuya libre determinación constituye uno de los atributos de la *res cogitans* son únicamente posibles desde la conciencia reflexiva. Sin embargo, no hay que olvidar que esta exigencia se propone a partir de un argumento epistemológico que considera únicamente la justificación de nuestro conocimiento. Es decir, determinamos la esencia de lo mental o de la *res cogitans* porque sólo del hecho de poseer contenidos de conciencia podemos tener una garantía epistémica indudable. Así, fácilmente caemos en el riesgo de confundir una razón epistémica con una caracterización metafísica. Es obvio que este énfasis en la dimensión epistémica es un rasgo propio de la modernidad filosófica. El principal problema de la perspectiva cartesiana es el dualismo, lo que Ryle ha denominado el fantasma en la máquina. Sin embargo, creo que

¹⁹⁹ Cf. *Eth. Nic.* 1144a 31.

²⁰⁰ Según Kenny (1989 (2000: 49-50) "*Podemos desear disponer de un término para referirnos al racimo de nuestras capacidades sensoriales de modo que "mente", según mi uso de la palabra, se refiere al racimo de capacidades ujos miembros fundamentales son el intelecto y la voluntad. La palabra más adecuada parece ser "psique". Si la adoptamos, podemos decir que aunque en la tierra sólo los humanos tienen mente, tanto los humanos como otros animales tienen psique.*"

podemos aceptar un cartesianismo atemperado sin comprometernos con el dualismo, en la práctica es esto lo que muchos filósofos han hecho. Este cartesianismo atemperado parte de la siguiente tesis epistemológica y descarta los compromisos ontológicos del cartesianismo: lo que mejor conocemos, lo más familiar, son nuestros contenidos de conciencia, para nosotros la nota definitoria de lo mental es el ser conscientes. De este modo lo mental se define desde la conciencia²⁰¹. Obsérvese que ni en Aristóteles ni en Descartes aparece el concepto de intencionalidad, no se presenta la mente, al menos las facultades de la sensación y de la intelección en Aristóteles o la conciencia en Descartes, como poseyendo el rasgo característico de la intencionalidad. Fue curiosamente un nearistotélico como Brentano quien incidió en este punto como marca de lo mental. Sin embargo, el concepto de intencionalidad es clave para la definición de lo mental. Se suele señalar que el rasgo distintivo de lo mental es la intencionalidad no la conciencia.

La contribución positiva de la perspectiva aristotélica - no de la teoría propiamente aristotélica- es que no circunscribe la intencionalidad a la conciencia. Sino que abarca un ámbito más amplio y que claramente hace justicia a la explicación de las capacidades o habilidades cognitivas que poseen otros organismos. En este sentido la perspectiva aristotélica constituye el marco metafísico desde el que se debería proceder a una naturalización. Por el contrario, pese a las dificultades de la perspectiva cartesiana como puedan ser su dualismo o el mecanicismo al que relega los organismos no humanos, ésta tiene la virtud de que resalta lo que para los humanos constituye nuestra aproximación más directa a lo mental: la conciencia reflexiva. La perspectiva cartesiana así es fenomenológicamente más apta para una caracterización de la vida mental. Cuenta además con un potente argumento epistemológico: conocemos directamente -en el sentido russelliano del término- qué cuenta como genuinamente mental. Sin

²⁰¹ Los peligros de esta afirmación fueron advertidos por el propio Descartes, el solipsismo. Sin embargo es esta una amenaza constante, si atribuyéramos conciencia a otros seres nos encontraríamos con el mismo problema, es decir, ninguna perspectiva racional está al margen del problema del solipsismo, pero tampoco hay por qué prestarle demasiada atención

embargo, la perspectiva cartesiana no posibilita una ciencia de la mente; no se puede hacer psicología desde presupuestos cartesianos, sólo fenomenología. De este modo la mente no es naturalizable a riesgo de perder lo genuinamente mental. Al menos en un sentido estricto y no exento de dificultades metodológicas considerables para lograr sumir los objetivos de una naturalización sólida. La teleosemántica tiene la virtud de asumir tesis aristotélicas como pueda ser su caracterización amplia de lo mental, es decir, no tiene reparos en adscribir estados mentales a organismos no humanos y resalta la continuidad de niveles en la naturaleza -si bien las razones son antes darwinianas que aristotélicas-. Pero tiene la dificultad de dejar de lado la conciencia. Bien puede pensarse que ésta es algo residual en la vida mental, pero es obvio que al margen de su historia filogenética sí determina de un modo substantivo nuestra caracterización de lo mental.

Finalmente, del mismo modo que podemos fijar la función de un mecanismo consumidor de representaciones podemos también atribuir una función a otros estados mentales. Así podemos pensar que las emociones tienen una función y que la conciencia tiene una función. Sin embargo, aquí entran en conflicto algunas de las asunciones básicas de la teleosemántica como que la noción de función tenga que ver con la perspectiva histórico etiológica. Así, si bien podemos atribuir una funcionalidad a la conciencia es discutible pensar que ésta es una adaptación definida en términos históricos.

Como ya he señalado, Millikan rechaza abordar el tema de la conciencia. Neander no lo trata. Únicamente lo abordan Papineau (1993: cap. 4) y Dretske (1995: cap. 4). Se puede argumentar que hay razones evidentes para no centrarse en el tema de la conciencia. Las teorías naturalistas del contenido parten antes de una concepción aristotélica de la mente que de una concepción cartesiana. La conciencia no es sinónimo de lo mental, término mucho más amplio que abarca más fenómenos. Lo mental es fundamentalmente lo intencional y para dar cuenta de fenómenos básicos de intencionalidad no necesitamos suponer ningún proceso consciente en el organismo al que atribuimos contenidos intencionales. No es que el naturalismo esté reñido con el estudio de la conciencia, sino con la conceptualización cartesiana de la mente entendida como conciencia.

Comentarios finales

El enfoque teleosemántico es correcto en espíritu, pero no en la letra. Creo que hay elementos muy valiosos en la teleosemántica y también que hay una serie de modificaciones que vale la pena hacer y otras ideas que hay que desechar. Entre los elementos que hay que mantener está el realismo intencional. Sin embargo, hay que advertir que éste no es un rasgo exclusivo de la teleosemántica, pues son muchas las teorías del contenido que se mantienen fieles al realismo intencional (Fodor, *Success Semantics*, etc.). Hay razones suficientes para defender el realismo intencional, quizás una de las más importantes sea de orden metodológico. Así, las explicaciones a las que proceden psicólogos y etólogos necesariamente han de estar basadas en la atribución de estados representacionales tal como señala Bermúdez (2003). La psicología popular, pues, no es tan acientífica como sus detractores suponen.

La concepción gradual del ámbito de lo mental es también una virtud de los posicionamientos teleológicos. En este sentido cualquier reflexión acerca de lo mental no puede obviar ni la perspectiva ontogenética -psicogénesis individual- ni filogenética. La perspectiva científica debe estar de algún modo presente en la reflexión filosófica.

Sin embargo, también hay elementos que, aunque valiosos, habría que modificar:

(1) Las funciones no son prescindibles, pero tampoco tan imprescindibles como suponen los autores teleosemánticos. Coincido con estos autores en que es una buena estrategia para explicar la fijación de algunos contenidos muy básicos más allá de la simple causalidad -Dretske- o de estrategias asimétricas -Fodor-.

(2) No necesitamos el concepto de función biológica histórico-etiológico. Como ya he señalado no es un concepto científicamente adecuado. Podemos conseguir lo mismo con una noción más amplia y que hace justicia a lo que los científicos entienden por función sin que por ello se ponga en peligro la normatividad. Aunque no creo que la

normatividad de los ítems biológicos sea extrapolable a la de las creencias. Esta modificación substantiva nos va a permitir eliminar el problema (4).

(3) La teleosemántica no es una teoría que explique el contenido mental en su totalidad (en contra de Papineau) ni una teoría naturalista de la fijación del contenido (en contra de Millikan o Neander), sino una teoría mucho más modesta. La teleosemántica es una teoría naturalista que explica sólo la fijación de algunos contenidos. No es una teoría todo terreno, hay contenidos que no se fijan a través de las funciones. Entre ellos los contenidos abstractos. Este es un precio que hay que pagar si se quiere seguir hablando de conceptos sofisticados.

(4) No hay por qué ser externistas. Podemos señalar que el recurso a las funciones es un buen método para la fijación de según qué creencias o contenidos y, no obstante, respaldar el internismo. Esta cuestión no es en realidad una asunción con la que haya que comprometerse, sino que son dos cuestiones independientes, aun cuando todos los autores teleosemánticos sean empedernidos defensores del externismo.

Por lo tanto, y siguiendo las cuatro tesis anteriores, entiendo que el concepto de función, adecuadamente definido, es clave en la determinación de la normatividad de muchas de nuestras actitudes básicas o cuyo contenido perceptivo es clave, aunque no es un concepto pannormativizador. Hay otras fuentes para fijar tanto el contenido de un concepto como de una creencia. Por lo tanto se hace necesario distinguir aquello que puede ser fijando apelando al concepto de función y aquello que escapa a esa estrategia fijadora. En mi opinión, el concepto de función es apto para fijar sólo categorías muy generales como la de objeto o propiedad o macroconceptos como cosa comestible, cosa peligrosa. Igualmente respaldaría la tesis que hace depender las representaciones de los deseos. Es decir, aquellas representaciones para las que fijamos un contenido son las que responden a la satisfacción de deseos o necesidades básicas, a saber, las que tienen que ver con la satisfacción de nuestros motivos primarios como el hambre, la sed, el sueño o el sexo. Esto es lo que

hacen todos los animales... salvo el hombre. De este modo la cuestión del contenido abstracto es irrenunciable. No hay fijación funcional, sólo una determinación genérica que va diluyéndose progresivamente a medida que las actitudes consideradas involucran conceptos más abstractos y dependientes de la convención.

VII. Conclusiones

La tesis principal defendida en este trabajo es que el concepto de función biológica es insuficiente para la fijación del contenido a diferencia de lo que proponen los partidarios de la teleosemántica. Que esto es así se demuestra no sólo porque los partidarios del enfoque teleológico emplean un concepto de función discutible, sino porque, además, la normatividad que pretenden para éste difícilmente puede ser trasladada al ámbito intencional. Sin embargo, no es esta la única contribución, más bien es el resultado de un cúmulo de aportaciones más modestas a las que llego tras el análisis de los diferentes tópicos que componen esta investigación. Dividiré la exposición de estos resultados en dos apartados correspondientes a cada una de las dos partes de este trabajo.

La primera parte de este trabajo ha estado dedicada al análisis de la teleología y, de manera más concreta, al concepto de función biológica. El análisis de la teleología que presento en el primer capítulo tiene un doble objetivo: Por una parte pretendía analizar el contexto y señalar la problemática en la que resurge el pensamiento teleológico en la Filosofía contemporánea y, particularmente, en la Filosofía de la ciencia. Así, creo haber mostrado suficientemente que, lejos de ser una doctrina agotada, goza de cierta buena salud, particularmente cuando se amalgama con el concepto de función en Biología. Por otra parte, las teorías teleosemánticas son también teorías teleológicas del contenido, es decir, teorías que recurren al concepto de función para explicar cómo se fija el contenido de las actitudes que pueda albergar un organismo. Por lo tanto, una crítica a los posicionamientos teleosemánticos como la que emprendo a partir del capítulo IV debe

partir necesariamente de un análisis exhaustivo del concepto que se presenta como central para estos. Así, llego a los siguientes resultados positivos en el análisis previo del concepto de función que fundamenta mi crítica a la teleosemántica:

1. En primer lugar, en el capítulo I, tras un análisis de tipo histórico acerca de los diversos posicionamientos clásicos (Platón, Aristóteles, Tomás de Aquino, Ockham, Suárez, etc), planteo la cuestión acerca de si existe algo así como una teleología natural o simplemente debemos contentarnos con un procedimiento heurístico tal como manifiesta Kant. Ambas posturas son defendibles. La primera de ellas ha caído claramente en desgracia desde que Darwin hiciera públicos sus descubrimientos, la segunda ha sido defendida por Hempel y ha desembocado en la pléyade de teorías que estudian el concepto de función biológica como un intento de recuperar esa teleología defenestrada bajo presupuestos compatibles con nuestra imagen científica. Así, con el paso de las explicaciones funcionales al estudio del concepto de función biológica se legitima en nuestros días el discurso teleológico más allá de la estrategia heurística que proponían Kant y Hempel.

2. En segundo lugar, y tras analizar tanto el contexto en el que surgen las teorías contemporáneas a propósito del concepto de función biológica, como un elenco significativo de las mismas (capítulos I y II), afirmo que el concepto de función ha sido tratado de una manera errónea por los filósofos. Estos se han empeñado en acuñar un concepto inservible para la investigación científica o, al menos, para la práctica de la Biología. Así, ni la propuesta histórico-etiológica ni la teoría sistémica resultan enteramente satisfactorias y se ven abocadas al fracaso. Ninguna de ellas responde, además, claramente a la práctica científica habitual de los biólogos. Respecto de los intentos de instanciación (Griffiths, 1993; Walsh, 1996; Buller, 1998) puede decirse lo mismo y en lo tocante al intento unificador de Kitcher (1993) cabe verlo como una estrategia nominalista en la que cualquier atisbo de unidad es aparente, pues el concepto unificador de diseño se aplica de manera diferente tanto desde una perspectiva sistémica como a partir de una perspectiva histórico-etiológica. En su lugar, pese a que las diversas razones que han ofrecido los filósofos sean poderosas, coincido con Sober (1993) en que un análisis correcto del mismo se ha

de dirimir en el terreno del concepto de adaptación. El concepto de función, como se intuye en el capítulo II y se manifiesta en el III, es un concepto dependiente del de adaptación. Sólo las adaptaciones son conceptos genuinamente pertenecientes a la teoría sintética de la evolución y no las funciones. Qué sea una función dependerá hasta cierto punto de cuál sea el efecto de una adaptación. Esta idea ha permanecido de manera más o menos velada en la obra de algunos autores como Millikan, Neander o Godfrey-Smith, pero apenas ha sido desarrollada. Únicamente Sober (1984; 1993) ha señalado la importancia del concepto de adaptación para delimitar claramente el concepto de función biológica. Pero, tal como he expuesto en el capítulo III, hay razones sobradas para pensar que el análisis de Sober acerca del concepto de adaptación es erróneo y así lo muestro en la sección 1.3. del capítulo III. En su lugar me decanto por un concepto de adaptación no historicista como el que acuñan Reeve y Sherman (1993) siguiendo a Bock y Von Wahlert (1965). Así, una adaptación es la variante fenotípica más eficaz de entre todas las presentes en una determinada población al margen de la historia selectiva del rasgo en cuestión. De esta manera la relación entre las funciones y las adaptaciones es clara y no necesita de peligrosas – y a menudo ficticias- constricciones historicistas. Decir de un rasgo que es adaptativo es decir que es funcional y su funcionalidad consiste en la contribución que realiza éste, o conjunto de ellos, a la economía global del organismo portador. No hay funciones al margen de las adaptaciones tal como han sido definidas, o al menos no como la mayoría de los filósofos piensan. Una función es, en la mayoría de los casos, lo que hace una adaptación, ya se defina este último concepto en términos históricos o en términos de eficiencia. Por otra parte, las resonancias sistémicas de la definición anterior son obvias, aunque creo que, a diferencia de la postura de Cummins, hay razones para pensar que es una definición inmune a las críticas habituales, principalmente porque es claramente normativa como a continuación voy a argumentar.

3. En tercer lugar, se ha insistido sobremanera en la importancia de la normatividad para acuñar una buena teoría de las funciones basándose en la idea de que toda adscripción funcional es una adscripción normativa. Así, la normatividad se ha impuesto entre los

filósofos como el criterio máximo para evaluar si una teoría determinada era o no una buena teoría de las funciones. En lo esencial ésta es una buena estrategia, aunque también es importante señalar que esta estrategia está viciada por la Filosofía de la mente, pues deriva inmediatamente del uso que le han dado los cultivadores de esta disciplina a este concepto científico. La consecuencia directa de esta tendencia ha sido que el énfasis puesto en la normatividad ha favorecido la preeminencia de las teorías histórico-etiológicas sobre las demás, argumentando (i.e. Millikan 1984; 1993) que la única vía para salvaguardar la normatividad era el recurso a la historia de selección del ítem funcional. Aunque hay soluciones intermedias (i.e. los intentos de instanciación), no es necesario apelar ni a las teorías histórico-etiológicas ni a las referidas soluciones intermedias para mantener la normatividad. La normatividad es una propiedad sobrevenida al contexto y dependiente en gran medida de la perspectiva del observador. En este sentido, y tal como muestro en el capítulo III (2.1.1 y 2.1.2), las críticas que contra el concepto de normatividad han vertido algunos autores como Davies (2001) o Cummins (2002) no son aceptables. Por otra parte, el rechazo de la dimensión normativa de las funciones supone una gran pérdida porque es la dimensión más característica de muchos de los fenómenos que pretendemos estudiar a partir de este concepto. La moralidad, el derecho o el contenido tienen un carácter intrínsecamente normativo y un estudio naturalista de estos fenómenos debería preservarlo. Además, como se expone en la sección 2.2 del capítulo III, existe una confusión grave respecto de los conceptos normativos que no ocurre con un concepto igualmente normativo como es el de "ley científica", empleado en Filosofía de la ciencia. Si uno está legitimado para afirmar que no sólo hay simples correlaciones sino que existen leyes, también debería estar legitimado para declarar la existencia de fenómenos normativos en algún sentido en la naturaleza no intencional. La investigación desarrollada en la citada sección tiene por objeto legitimar el discurso normativo que tan extraño puede resultar para algunos.

Finalmente, esta primera parte concluye con una aportación substantiva a la teoría de las funciones que desarrollé anteriormente y que he bautizado con el nombre de "Funciones sistémicas normativizadas". Así, cuando observamos que las cosas han sucedido

de un determinado modo durante un lapso de tiempo más o menos largo podemos suponer una norma o pauta de acción que actúa en ese conjunto de fenómenos. En tal situación cualquier acaecimiento que se aparta de la norma es anormal. Es más, si el análisis de la situación lo permite, se pueden identificar una serie de constantes ambientales que, a modo de cláusulas *caeteris paribus*, constituyan un marco nomológico que expliquen el rol causal de la interacción entre el organismo y el medio. Así hay normatividad dentro del ámbito natural sin necesidad de recurrir a heurísticos, aunque tampoco al margen de la perspectiva del observador. La solución que ofrezco es kantiana, ya que trato la existencia de las normas naturales como un *como sí* referido al mundo orgánico. En él, sostengo, al igual que Kant, que en la naturaleza no hay más que procesos mecánicos; no está presente ningún tipo de propositividad ni cabe apelar a fuerzas vitales misteriosas. Las normas naturales que adscribimos a las funciones únicamente describen lo que de hecho ocurre dado un determinado régimen ambiental al que sobrevienen.

La primera parte podría verse como un estudio autónomo respecto de la segunda. En este sentido, si el concepto de función es importante es porque con él vemos que se revive en parte una discusión muy antigua como es la de la teleología. Es ésa una posibilidad y así puede entenderse si se quiere. De este modo los tres primeros capítulos podrían leerse autónomamente y verse como un estudio más o menos minucioso sobre una cuestión que no deja de ser una tormenta en un vaso de agua para muchos otros filósofos, pero hay que decir que la Filosofía contemporánea está actualmente plagada de este tipo de fenómenos. Sin embargo, si el concepto de función es importante es porque ha sido una noción explotada sistemáticamente para solucionar problemas relativos a la Filosofía de la mente. Es decir, el concepto de función entra en el edificio de la Filosofía por la puerta de la Filosofía de la mente y no por la ventana de la Filosofía de la Biología donde, respecto de los grandes tópicos de esta disciplina, no ocupa un lugar destacado.

La segunda parte de esta investigación examina la cuestión del contenido mental (cap. IV) y se centra en una exposición y análisis de las soluciones teleosemánticas (cap. V) para, finalmente, señalar sus límites y deficiencias haciendo uso de diversos argumentos (cap. VI).

En la Filosofía de la mente contemporánea encontramos diversos intentos naturalizadores de la intencionalidad. Entre ellos ocupan un lugar destacado las teorías teleológicas del contenido mental, que tratan de explicar los fenómenos más sobresalientes de nuestras actitudes como su carácter normativo apelando al concepto de función biológica para fijar el contenido. Sin embargo, como ya adelanté al inicio de estas páginas, el planteamiento está viciado de origen: las teorías teleosemánticas no son teorías aptas para caracterizar la fijación del contenido mental.

A mi juicio hay un criterio adecuado para evaluar una teoría del contenido mental que esbozo a través de seis *desiderata* en el capítulo IV. Así, una teoría del contenido mental para ser adecuada debería satisfacer los siguientes seis *desiderata*:

D.1. Una teoría del contenido debe dar razón del carácter selectivo del contenido

D.2. Debe ser capaz de individualizarlo adecuadamente, es decir, debe dictaminar si el contenido se individualiza en base a propiedades relacionales o si, por el contrario, responde a una caracterización internista.

D.3. Debe dar razón de su dimensión normativa enfrentándose eficazmente a puzzles como el de la disyunción.

D.4. El contenido involucra aspectos tanto conceptuales como no conceptuales o cualitativos. Una teoría del contenido debería explicar qué papel epistemológico desempeñan los aspectos cualitativos de la posesión de estados intencionales.

D.5. Debe explicar qué relación se establece entre contenido mental y contenido lingüístico

D.6. Debe dar cuenta de su dimensión práctica

En el capítulo IV, junto a la exposición de estos *desiderata*, analizo una serie de cuestiones acerca de la naturaleza del contenido mental. En lo que respecta a la relación entre pensamiento y lenguaje considero que la perspectiva que aborda el pensamiento como evento lingüístico (Sellars, 1956) tiene claras virtudes. Sin embargo, posicionamientos más extremos como el de Fodor (1975) son controvertidos, pues, como señala Bermúdez (2003), no hacen justicia

a las capacidades cognitivas de criaturas alingüísticas. Por otra parte una dificultad importante respecto de la hipótesis del lenguaje del pensamiento es que está a medio camino entre la Filosofía y la Ciencia cognitiva. Si bien puede ser visto esto como una virtud, creo que también puede inducir a cierta confusión al no clarificar suficientemente las diferencias existentes entre los niveles doxástico y subdoxástico.

En lo tocante al problema de la individuación la principal dificultad a la que se enfrentan los planteamientos externistas es a la de justificar adecuadamente la compatibilidad entre el acceso privilegiado y la causación mental de nuestras acciones con el externismo que defienden. Los diferentes argumentos estudiados (Bach, 1988; Fodor 1987; Crane, 1991; o Farkas, 2008) inciden en la misma dificultad, la imposibilidad de que el contenido sea causa de nuestras acciones y que a la vez se individualice relacionamente. Es éste último un argumento poderoso que cuenta además con los recursos adecuados para solventar las dificultades semánticas que pueda acarrear el análisis de la práctica lingüística tal como señala Crane (1991). Mi opinión al respecto es que es los conceptos siguen estando en la cabeza y con toda probabilidad también los significados. Si bien también señalo que la defensa de una perspectiva cartesiana, como hacen Bach y Farkas, incide excesivamente en la conciencia como rasgo de lo mental, presupuesto que no es fácil de asumir desde posturas naturalistas que pretendan atribuir actitudes a organismos distintos de los seres humanos adultos y normales.

Respecto de la normatividad, lo conceptual y lo no conceptual o el holismo, procedo analizando el significado de estos conceptos dentro de una teoría general del contenido. Así, el aspecto normativo es el principal y, como se verá en los dos siguientes capítulos, sobre el que inciden las diversas teorías teleosemánticas del contenido. Por lo que respecta a la dicotomía entre los aspectos conceptuales y no conceptuales apunto que si bien sólo estos últimos tienen un papel claro y determinado en lo que respecta a la teoría del contenido, también los primeros deben ser dignos de estudio, particularmente en aquellos organismos en los que la atribución de actitudes es discutible o problemática.

El capítulo V constituye una exposición de cuatro teorías teleosemánticas del contenido (Dretske, Millikan, Neander, Papineau).

En este capítulo señalo que la teleosemántica debe entenderse más que como una teoría general del contenido, como una teoría acerca de la fijación del contenido cuyo recurso principal es la noción de función biológica.

A mi juicio la teleosemántica está comprometida, además, con una serie de tesis substantivas acerca del estudio y de la naturaleza de los estados intencionales como son:

T.1 Naturalismo filosófico

T.2 Realismo intencional respecto de las actitudes.

T.3 Son partidarias de la teoría representacional de la mente pero no se comprometen con la hipótesis del lenguaje del pensamiento.

T.4. Atomismo; no son partidarios del holismo semántico.

T.5. Están comprometidas con el externismo.

Por otra parte, clasifico las teorías teleológicas del contenido que analizo en el capítulo V en las siguientes categorías:

(I) Teorías informacionales: semánticas basadas principalmente en el concepto de información como las que presentan Stampe o Dretske.

(II) Semánticas del consumidor como las de Millikan, Neander o Price.

(III) Teorías no combinatorias como la de Papineau.

(IV) Teorías combinatorias modestas como la de Sterelny.

El núcleo de las distintas teorías hay que verlo en el grupo de las semánticas centradas en el consumidor. Estas semánticas son las que hacen un uso más sistemático del concepto de función y las que, por decirlo de alguna manera, más arriesgan. Las críticas que realizo en el capítulo VI son coextensivas a todas las teorías, aunque se centran principalmente en el grupo anteriormente mencionado.

En mi opinión las teorías mencionadas anteriormente no satisfacen los *desiderata* aludidos en el capítulo IV. Particularmente identifico las siguientes deficiencias que expongo sistemáticamente en el capítulo VI. Así, las teorías teleosemánticas deben enfrentarse a los siguientes cargos:

(1) No emplean un concepto adecuado de función biológica.

- (2) Son susceptibles de casos de indeterminación
- (3) La cuestión externista es independiente.
- (4) No explica la fijación del contenido sofisticado o abstracto.
- (5) No aborda la cuestión del contenido inútil.
- (6) No aborda la cuestión del contenido no conceptual.
- (7) No explica nuestra vida mental al completo.

El primero de ellos es obvio a la luz de lo expuesto en los capítulos II y III. Y considero que es el cargo más importante que invalida en gran medida los intentos naturalizadores y normativizadores de la teleosemántica. Además, la crítica aquí presentada invalida de manera más eficaz los intentos teleosemánticos sin necesidad de recurrir a argumentos como los de Fodor (1987; 1990). Sin embargo, los seis restantes creo que también representan importantes dificultades para el cumplimiento del proyecto teleosemántico, ya que si se trata de ofrecer no sólo una explicación acerca de cómo se fija el contenido, sino también de la naturaleza del contenido, será preciso satisfacer los desiderata señalados. Brevemente, al margen de las discrepancias ya obvias sobre el concepto de función, mis conclusiones respecto de los cargos de los que acuso a la teleosemántica son las siguientes.

Las teorías teleosemánticas, pese a que se enfrentan contra los casos de indeterminación, no logran hacerles frente. Las indeterminaciones persisten, el concepto de función biológica no permite dar cuenta de la fijación del contenido.

La cuestión externista no supone ningún aditamento substancial a la doctrina teleosemántica. No hay razones que inviten a decantarse por un posicionamiento u otro en lo tocante a cómo se individúa el contenido. Así, teóricamente uno podría ser partidario de la teleosemántica siendo individualista.

La teleosemántica es una mala teoría para explicar la fijación y normatividad de aquellas creencias que involucran conceptos sofisticados como puedan ser los conceptos científicos y jurídicos. No creo que ningún concepto de función pueda ser de alguna utilidad en la fijación de dichos contenidos. Los conceptos abstractos no guardan relación inmediata con nuestra constitución biológica; la biología no tiene nada que aportar al respecto.

El quinto cargo tiene que ver con la ausencia de correlación entre albergar creencias y suponer que todas ellas son resultado de algún mecanismo biológico que las produce o las consume en virtud de su funcionalidad. Los humanos de modo específico albergamos creencias inútiles, creencias cuyo valor biológico es nulo. Si bien esto ocurre de manera casi exclusiva en humanos no debería ser pasado por alto, pues la caracterización del contenido se hace desde el ámbito humano y no animal. La teleosemántica no contempla ningún caso de contenido inútil y carece de los recursos necesarios para su explicación.

En cuanto a lo no conceptual hay que decir que es más importante por aquello de fenoménico que tiene que por el hecho de su mera consideración. Lo fenoménico no tiene por qué tener ninguna función biológica, podría entenderse como un epifenómeno, pero del mismo modo que sucede con el punto anterior, es un aspecto esencial de nuestras vidas. Por otra parte, dentro de la economía cognitiva de los distintos animales lo fenoménico también parece desempeñar algún papel epistemológicamente relevante. El último punto hace referencia a la conciencia que no es un tópico abordado desde la teleosemántica y que difícilmente puede tener una explicación evolutiva. Con el concepto de vida mental al completo hago referencia a estos aspectos que, como la conciencia, quedan descuidados por el enfoque teleológico. En conclusión, el diagnóstico precedente podría entenderse como una serie de consideraciones substantivas acerca del alcance y límites de la teleosemántica.

Finalmente este trabajo puede verse como una crítica a cierta manera de entender la actividad filosófica. De otra manera, es una crítica contra cierto naturalismo filosófico exacerbado. Los problemas filosóficos son filosóficos porque no son susceptibles de un tratamiento científico, si lo fueran, ya no serían problemas filosóficos, sino cuestiones abiertas a una práctica experimental sistemática. En eso consiste la miseria del saber filosófico, pero también su grandeza; la Filosofía, tal como señalaba Wittgenstein, lo deja todo tal como está. La Filosofía ni es una ciencia empírica ni aspira a serlo. Su ámbito es el de la reflexión puramente conceptual y es en ese mismo terreno en el que ha de dirimir sus cuitas. Lo anterior no significa que deba hacerse filosofía al margen del conocimiento científico, ni que la garantía epistémica que buscamos para nuestras teorías científicas deba ser

dejada de lado en la reflexión filosófica. No en vano el mejor conocimiento del que disponemos es el conocimiento científico. No creo que ningún filósofo serio haya procedido de espaldas a la actividad científica de su tiempo, más bien ha sucedido lo contrario. No es tarea que nos competa el suplantar a los buenos científicos. Así, nuestra mejor contribución es modesta y se circunscribe a la honrosa tarea de poner orden en los conceptos con los que pensamos la realidad y en afinar nuestros métodos justificatorios de las diversas teorías de las que disponemos. La investigación que he presentado en las páginas que constituyen este trabajo ha tenido por guía las ideas anteriormente expuestas acerca de la actividad filosófica.

Apéndice

La teoría sintética de la evolución

1. La teoría darwinista de la selección natural

Charles Darwin (1809-1882) es el fundador de la moderna teoría de la evolución por medio de la selección natural²⁰². El desarrollo de su

²⁰² La gran idea de Darwin no puede entenderse sin el trabajo previo de los naturalistas, geólogos y paleontólogos de los dos siglos anteriores a él. Un problema importante era explicar la cronología de la tierra. Por esa época el arzobispo Usher publicó su *Cronología Sagrada*, y dedujo que la tierra se creó en el 4004 a.C. Pero esta datación era incompatible con la evidencia del registro fósil que, por otra parte no sólo era una amenaza a la cronología bíblica, sino también al concepto fijista de especie. El registro fósil sí permite elaborar una cronología precisa apoyándose en la estratigrafía, que es la base de datación de la edad de la Tierra. Que haya fósiles de organismos inexistentes indica que las especies no han sido creadas de un modo perfecto y que, en consecuencia, algunas han desaparecido. Algo que para un creacionista es inaceptable.

También hay que destacar la importancia del surgimiento de la Anatomía Comparada debida a George Cuvier (1769-1832), quien formula su ley de correlación, según la cual todas las estructuras del cuerpo de un animal son dependientes entre sí, de modo que la modificación de una de las partes produce el cambio en el resto del organismo. En 1698 Edward Tyson diseccionó a un chimpancé y concluyó que se parecía en un 75% al hombre. Con esto nos alejamos de la idea de *scala naturae* y nos acercamos más a la idea arbórea del parentesco de las diferentes especies.

George Louis Leclerc, conde de Buffon (1707-1788), fue quizás uno de los primeros naturalistas en plantearse seriamente el fijismo. Era contemporáneo de Linneo y destacado opositor de sus ideas. En Buffon hallamos ideas protoevolucionistas. En 1776 publica su *Ensayo sobre la degeneración de los animales* en el que abandona las ideas fijistas. Considera que los distintos organismos pueden originarse unos a partir de otros por degeneración, así la evolución surge por degradación de los organismos a partir de un tipo ideal. La doctrina de Buffon es el transformismo, aunque nunca llegó a ser una teoría completa. También se interesó por el problema de la historia de la Tierra tratando de establecer etapas geológicas.

teoría responde a varios factores, como la atmósfera intelectual de la primera mitad del XIX y el hecho de que se embarcara en la expedición del Beagle de 1831 a 1836, durante la que acumularía una gran cantidad de observaciones sobre la variedad de especies encontradas a lo largo de su viaje. Lo que más le llamó la atención fue la adaptación de las distintas variedades de una misma especie a su medio. A su llegada a Gran Bretaña inició la reflexión y ordenación sistemática de todo el material que había acumulado durante los años de la expedición con vistas a la elaboración de una obra que abordara científicamente la cuestión del origen de las especies. En 1858, cuando su obra estaba ya prácticamente lista, tuvo conocimiento de que el naturalista norteamericano A. R. Wallace (1823-1913) había llegado a conclusiones casi idénticas de manera independiente, y en el mismo año presentaron ambos sendas comunicaciones en la Sociedad Lineana de Londres donde dieron a conocer sus ideas. En 1959 publicó *El origen de las especies por la selección natural*. Las influencias cruciales en el desarrollo de la teoría darwiniana fueron cuatro:

Louis Moreau de Maupertuis (1698-1759) fue otro defensor del transformismo. Sostenía que la Naturaleza contenía todas las variedades posibles para que surgieran rasgos nuevos en los organismos y que sólo el azar o la intervención humana las ponía en escena. En su *Sistema de la Naturaleza* (1751) intenta una reconciliación entre el creacionismo y la investigación científica.

La idea de un tronco común a partir del cual surgieran todas las especies no es debida sólo a Darwin. Su abuelo, el también naturalista Erasmus Darwin (1731-1802), defendió en su *Zoonomía* (1794-1796) la posibilidad de que todos los organismos tuvieran un origen común y que se hubieran diversificado al ir perfeccionándose progresivamente. Sin embargo la primera teoría evolutiva completa se la debemos a Lamarck (1774-1829). En 1809 publica *Filosofía zoológica*, en la que aparece su teoría evolutiva. Lamarck traza un esquema evolutivo que va de los organismos más simples, que aparecen por generación espontánea, a los más complejos. La naturaleza, considera, busca la perfección: en esto todavía mantiene las ideas aristotélicas. Las exigencias del medio provocan en los organismos la modificación o aparición de nuevos órganos como mecanismo adaptativo. Estos caracteres adquiridos en base a una supuesta ley del uso o del desuso son además heredables, es decir, Lamarck creía en la herencia de los caracteres adquiridos, idea que, por otra parte, la ciencia actual desmiente.

1. La lectura de la obra de Lyell (1797-1875) *Principios de Geología* (1830), el libro con más impacto de la época. Lyell postula la hipótesis del uniformitarismo: los procesos naturales que han conformado la historia de la tierra son los mismos en todas las épocas. Frente a los catastrofistas defendía el gradualismo en los cambios geológicos a través de millones de años. Fue una influencia decisiva en Darwin, quien lo leyó a bordo del Beagle. Darwin decide aplicar el mismo principio gradualista al mundo viviente.
2. La lectura de los ensayos de Malthus sobre la población le suscitaron la idea de la lucha por la supervivencia.
3. La experiencia de los criadores en el proceso de selección para la consecución de razas con rasgos diferenciados le condujo a plantear la idea de una selección natural al margen del proceso de selección artificial.
4. La existencia de afinidades entre los seres vivos procedente de los estudios de embriología y anatomía comparada le condujeron a la idea de un origen común de todos los seres vivos.

La teoría de Darwin queda bien reflejada en este párrafo de su Introducción a *El origen de las especies* (1859):

“Como de cada especie nacen muchos más individuos de los que pueden sobrevivir, y como, consiguientemente, hay que recurrir con frecuencia a la lucha por la existencia, se deduce que cualquier ser, si varía, aunque sea levemente, de algún modo provechoso para él, bajo las complejas y a veces variables condiciones de la vida, tendrá mejores probabilidades de sobrevivir, y de ser así seleccionado naturalmente. Según el vigoroso principio de la herencia, toda variedad seleccionada tenderá a propagar su forma nueva y modificada.”

El mérito de Darwin consiste en haber desarrollado una teoría absolutamente mecanicista que diera cuenta de la diversidad de las especies. El argumento de Darwin es el siguiente:

1. Existe una lucha constante por la supervivencia debido a que los animales tienen una tendencia a reproducirse más allá de las capacidades del medio, agotando fácilmente los recursos.

2. Existen variedades inter e intraespecíficas. Algunas de éstas se adaptan mejor al medio que otras.
3. Los organismos dotados de rasgos menos adaptativos no se reproducirán, favoreciendo la selección a los mejor adaptados.

A todo esto hay que añadir la existencia de un linaje común y la variabilidad del ambiente que hace que se seleccionen nuevos rasgos. La suma de todos estos factores promueve la divergencia de las especies.

II. La teoría sintética de la evolución

La teoría sintética de la evolución resulta de añadir una base genética a la teoría darwiniana y tuvo su origen en la década de los años veinte del pasado siglo con el nacimiento de la genética de poblaciones. Particularmente importante es la obra del biólogo de origen ruso, más tarde nacionalizado americano, Theodosius Dobzhansky (1900-1975), quien en 1937 publicó *La Genética y el Origen de la Especies*, donde presenta una síntesis de la teoría de la herencia con el concepto de selección natural. Hay que tener en cuenta que el propio Darwin no tenía una teoría científica para explicar la herencia. Dobzhansky tiene la virtud de conjugar los resultados matemáticos de los teóricos precedentes como Wright, Haldane o Fisher con los resultados experimentales y observacionales.

La teoría de la evolución posibilita explicaciones no finalistas en el ámbito de la vida. Es un modo distinto de responder a los diversos interrogantes sin necesidad de apelar a causas finales. En este sentido constituye una revolución comparable a la física newtoniana; por otro lado, la teoría de la evolución posibilita en las ciencias biológicas cierta unificación.

¿En qué consiste la evolución? O ¿Qué es evolución? Evolución, en principio, significa cambio. Ahora bien, este cambio se entiende en su aplicación a los seres vivos y, más concretamente, a un tipo de cambio que implica dejar descendientes con ciertas modificaciones y con cierto grado de diversificación. Hay que advertir que la teoría de la evolución,

tal como la entendemos hoy día no tiene absolutamente nada que ver con la idea de progreso.

Futuyama define la evolución como el cambio en las características de las poblaciones de organismos, o grupos de tales poblaciones, a través del curso de sucesivas generaciones.

“El mecanismo evolutivo está basado en tres puntos: (1) variabilidad heredable de ciertas características o rasgos, (2) variabilidad entre individuos de una misma población en cuanto al éxito reproductor, y (3) la correlación entre ambos tipos de variabilidad. Cuando existen ambos tipos de variabilidad, la correlación entre ellos determina el tipo de evolución que se está produciendo. Si la correlación entre la variabilidad heredable y la variabilidad en éxito reproductor es nula o muy baja, los cambios que se producirán de una generación a otra fluctuarán al azar, esto es lo que se conoce como “evolución neutral”. Si por el contrario, esa correlación es elevada, el cambio evolutivo no será al azar sino en la dirección en la que se incrementa el éxito reproductor, es decir, en el sentido en que se mejora la eficacia de la adaptación. Esta es la denominada “evolución adaptativa”. El motor de la evolución adaptativa es el mecanismo descrito por Darwin al que denominó “selección natural””

(Soler, 2003: 22).

2.1. Factores que desencadenan procesos evolutivos

El *qué* de la evolución no es cuestión de debate en el panorama científico actual, donde se admite una u otra versión de la misma, aunque sí se discute el *cómo*, es decir, los mecanismos que dan cuenta de cómo ha ocurrido la evolución.

La teoría de la evolución, tal como fue concebida por Darwin, se apoyaba en la selección natural. Es decir, el mecanismo que explica que los diferentes organismos cambien es la selección natural. Esto significa que las adaptaciones fenotípicas de los distintos organismos se deben al proceso de selección natural. Sin embargo, hoy día sabemos que la selección natural no es el único factor de cambio, aunque parece ser el más importante. Es decir, el cambio del acervo genético de una determinada población de individuos puede deberse a otros factores de cambio como:

1. *Mutaciones*. Cambios heredables del genotipo. Distinguimos dos tipos de mutaciones: (A) Mutaciones genéticas: cambio de una o más

bases en el proceso de replicación del ADN. (B) Mutaciones cromosómicas, como la delección o el sobrecruzamiento.

2. *Flujo genético*. Desplazamiento de alelos hacia dentro y fuera de una población. Puede colocar alelos nuevos en una población o cambiar las frecuencias génicas existentes. El efecto global en la población consiste en reducir las diferencias respecto de otras poblaciones, es decir, tiene un efecto homogeneizador (a diferencia de la selección natural).

3. *Deriva genética al azar*. Es un error de muestreo que consiste en el cambio de frecuencias génicas producido al azar. Tiene un papel destacado en la evolución de las poblaciones pequeñas, aunque su importancia relativa al compararla con la selección natural es polémica. Hay dos situaciones en las que se ha demostrado importante: (A) El efecto fundador y (B) cuellos de botella. Ambos son causas de deriva genética. En el primer caso nos encontramos con que una subpoblación que se separa de una población mayor puede que no sea representativa de la población de la que deriva. En el caso de los cuellos de botella nos encontramos que por causas ajenas a la población, ésta queda drásticamente reducida (i.e. una cacería masiva).

4. *Reproducción selectiva*. Es un caso de apareamiento no aleatorio (i.e. ánsares nivales que únicamente se aparean con ánsares nivales).

2.2. El concepto de selección natural

La selección natural es el principal mecanismo de la evolución de las especies. Puede ser un factor decisivo para promover la variabilidad de las poblaciones.

Selección natural y fenotipo. La selección natural opera, en principio, sobre el fenotipo. El fenotipo es, por lo general, el resultado de la expresión de muchos genes. Un mismo fenotipo puede ser debido a diversas combinaciones genotípicas; resulta extraño que un único alelo sea el responsable del éxito de un fenotipo, si bien se dan casos como en *Biston Betularia*. El fenotipo es dependiente del ambiente, es decir, no está únicamente determinado por el genotipo, sino por la relación del genotipo con el ambiente.

¿Cuándo ocurre selección natural? Siguiendo a Soler (2003: 129), podemos afirmar que en una población determinada ocurre selección natural si y sólo si se dan las siguientes condiciones:

- C1. Existencia de variación fenotípica en la población.
- C2. Existencia de variación en eficacia biológica como consecuencia de la relación directa entre los caracteres y la habilidad del portador para el emparejamiento y la fertilización, y en su fertilidad, fecundidad y/o supervivencia.
- C3. Existencia de caracteres heredables.

Tipos de selección natural. Empezaremos enumerando los casos de selección neutral: la eficacia del fenotipo es independiente de su frecuencia en la población.

- 1. Selección estabilizante: eliminación de los individuos con características extremas. Se da en todas las poblaciones de manera constante (p. ej. tamaño medio de las nidadas de estorninos).
- 2. Selección disruptiva: la selección opera para hacer aumentar las frecuencias de los fenotipos extremos de una población (p.ej. plantas que viven en suelos contaminados y en suelos no contaminados).
- 3. Selección direccional: aumento de las frecuencias de individuos con una característica alejada de la mediana (p.ej. resistencia a los insecticidas, *Biston betularia*)

Continuamos con los casos de selección no neutral:

- 1. Selección dependiente de la densidad: actúa disminuyendo la frecuencia de los fenotipos más abundante y aumentando la frecuencia de los menos comunes.
- 2. Selección sexual: acentuación de rasgos diferenciadores que tienen que ver con la obtención de pareja, no con el acto reproductor en sí. Hay que advertir que la selección sexual para Darwin era una fuerza separada de la selección natural. Sin embargo, hoy día, tal distinción parece no aceptarse. Principalmente desde el momento en que se

redefine el concepto de eficacia biológica en términos de números de hijos que sobreviven.

El resultado de la selección natural. Sólo la selección natural puede producir adaptaciones, el resto de mecanismos anteriormente expuestos inducen variabilidad en la población, pero no producen adaptaciones. Una adaptación puede entenderse como una situación de amoldamiento al ambiente; también puede referirse a un carácter determinado que permite al organismo adaptarse (p.ej. el oído); por último, también puede referirse al proceso evolutivo que da como resultado organismos mejor dotados para su ambiente. Una definición más precisa es que se trata de cualquier característica que posee un organismo que contribuye a su eficacia. Igualmente nos referimos a las adaptaciones no sólo como proceso, sino también como resultado: en este sentido una adaptación es el proceso por el que se desarrolla dicha característica.

2.3. La potencia de la selección natural: la eficacia biológica

La Teoría de la Evolución asume que las variaciones heredables se han de traducir en un incremento de la eficacia media del individuo para que se dé una adaptación, es decir, un incremento en viabilidad y fertilidad respecto de la generación anterior. Aquí ya tenemos un dilema, viabilidad y fertilidad pueden entrar en conflicto, cuando la selección actúa sobre la viabilidad las eficacias se interpretan como probabilidades (probabilidad de alcanzar la edad adulta y reproducirse), mientras que cuando la eficacia se interpreta en términos de fertilidad nos estamos refiriendo al número de crías (ejemplo de la cola del pavo real). El problema va más allá si tenemos en cuenta que la probabilidad no es una noción unívocamente definida. ¿Qué entendemos por eficacia? Hay que tener en cuenta que una de las acusaciones en contra de la Teoría de la Evolución era la de tautologocidad y precisamente se argumentaba que el concepto del que dependía tal acusación era el de eficacia. En definitiva, los problemas relativos al concepto de eficacia están en relación con su definición. La definición involucra dos tipos de problemas: (1) La definición de probabilidad. (2) La acusación de tautologocidad.

2.4. ¿Sobre qué nivel actúa la selección natural?

El problema de las unidades de selección tiene que ver con la cuestión acerca de sobre qué nivel actúa la selección natural. Para la teoría darwiniana la selección opera sobre el individuo. Sin embargo, teóricamente la selección podría actuar sobre cualquier sistema en que se dieran las condiciones 1-3 anteriormente expuestas. Así la selección podría actuar a nivel de especie, de organismo o a nivel genético. Como se ve es una cuestión que fundamentalmente deberán determinar los biólogos. Sin embargo, hay una cuestión de fondo que sí afecta notoriamente a la Filosofía, particularmente desde la propuesta de la selección multinivel (la propuesta teórica que sostiene que la selección actúa a distintos niveles). Se admite que los seres vivos son sistemas organizados jerárquicamente. Pues bien, tanto la definición de sistema como la cuestión de la jerarquía requieren de un aparato conceptual muy fino, ya que si no caemos fácilmente en la afirmación de que la realidad es “compleja” con lo que apenas estamos aportando un grano de claridad, ya que lo que nos interesa es desentrañar dicha complejidad, no quedarnos extasiados con ella misma. La cuestión de la organización jerárquica es especialmente rica. Los organismos no son los únicos sistemas organizados jerárquicamente, también lo son los artefactos. El problema fundamental consiste en caracterizar ambas esferas de un modo preciso, particularmente en lo que se refiere a su funcionalidad (los artefactos tienen funciones del mismo modo que las adaptaciones resultan funcionales). En toda organización jerárquica asoma la cuestión acerca de cómo se produce dicha organización y qué relación se establece con los componentes del sistema, por tanto la cuestión esencial y que a día de hoy da mucho que hablar es la superveniencia. Sirva lo anterior como una serie de insinuaciones acerca de la importancia del tema. Ahora voy a exponer sucintamente cuales son las teorías al uso en este debate:

A) *Selección a nivel de grupo*. Esta perspectiva teórica afirma que lo que se selecciona es el grupo (especie, clado) no el individuo particular. Es particularmente importante por la repercusión que ha tenido en ciencias humanas. Uno de los temas que más interés ha

generado es la explicación de la conducta altruista de algunos organismos como los himenópteros.

El problema es el siguiente: si admitimos que una población o especie está formada mayoritariamente por individuos egoístas que únicamente tienen interés en maximizar su propio beneficio y tienden a reproducirse a la máxima tasa posible, acabarán al cabo de una serie de generaciones con los recursos disponibles y se extinguirán. Sin embargo parece que las especies son capaces de controlar la tasa de natalidad de los individuos evitando la destrucción de los recursos (Wynne-Edwards 1962). Actualmente, (Stearns 1986) se asume que la selección que pueda actuar a nivel de especie o de clado es muy débil, ya que la duración de su ciclo es enormemente larga en comparación con el de los genes o los individuos.

B) *Selección a nivel de gen*. Es la propuesta popularizada por Dawkins en su libro "El gen egoísta" que asume que la unidad de selección no es el individuo, sino el gen. Los organismos son los medios que emplean los genes para duplicarse y preservarse a lo largo del tiempo.

C) *Selección a nivel de individuo*. Es la perspectiva clásica defendida por Darwin y por los partidarios de la Teoría Sintética. Asume que la selección opera sobre el organismo.

Bibliografía

A) Obras generales de consulta (Diccionarios, Compendios e historias de la Filosofía)

N.B. Cuando se trata de una traducción indico la edición original entre corchetes.

FERRATER MORA, J. (1981). *Diccionario de Filosofía* (4 Vols.). Madrid: Ariel.

FRAILE, G. & URDÁNOZ, T. (1966). *Historia de la Filosofía Vol. III*. Madrid: Biblioteca de Autores Cristianos.

GILSON, E. (1965). *Historia de la Filosofía Medieval*. Madrid: Gredos. [1952].

GUTTENPLAN, S. (1994). *A Companion to the Philosophy of Mind*. Oxford: Blackwell.

GUTHRIE, W.K.C. (1991). *Historia de la Filosofía Griega* (6 Vols.). Madrid: Gredos. [1962].

MOLINER, M. (1998). *Diccionario de uso del español*. Madrid: Gredos.

REALE, G & ANTISERI, D. (1988). *Historia del pensamiento filosófico y científico* (3 Vols.). Barcelona: Herder. [1983].

B) Clásicos de la tradición filosófica y científica (hasta el s. XIX)

N.B. Entre paréntesis la abreviatura que empleo para citar sus textos.

AGUSTÍN DE HIPONA

Sobre el orden. Edición de V. Capanaga. Madrid: Biblioteca de Autores Cristianos. (1969).

ARISTÓTELES:

Acerca del Alma (De An.). Traducción de T. Calvo. Madrid: Gredos. (1978).

Ética a Nicómaco (Et. Nich.). Traducción de J. Pallí. Madrid: Gredos. (1985).

Física (Phys.). Traducción de G. R. De Echandía. Madrid: Gredos. (1995).

Metafísica (Met.). Traducción de T. Calvo. Madrid: Gredos. (1994).

Tratados de Lógica (Org.). Traducción de Miguel Candel. Madrid: Gredos. (2 Vols.) (1982/1988).

DARWIN:

El origen de las especies. Edición de J. Josa y traducción de A. De Zulueta. Madrid: Espasa (2001).

DESCARTES:

Discurso del Método. Edición y traducción de M. García Morente. Madrid: Espasa. (1997).

Meditaciones Metafísicas. Edición y traducción de M. García Morente. Madrid: Espasa: (1997).

KANT:

Crítica de la razón pura (KrV). Edición y traducción de Pedro Ribas. Madrid: Alfaguara. (1998).

Crítica del juicio (K.U.). Edición y traducción de M. García Morente.. Madrid: Espasa-Calpe. (1999).

LEIBNIZ:

Discurso de Metafísica. Edición y traducción de J. Marías. Madrid: Alianza. (1997).

PLATÓN:

Fedón. Traducción de C. García Gual. Madrid: Gredos. (1986).

Timeo. Traducción de M. López. Madrid: Gredos. (1992).

OCKHAM:

Antología filosòfica. Edición y traducción de J. Batalla. Barcelona: Edicions 62. (2003).

SUÁREZ, FRANCISCO:

Disputationes Metafísicas (Disp.). Edición y traducción de S. Rábade, S. Caballero y A. Puigcerver. Madrid: Gredos. (7 vols.). (1960/1966).

TOMÁS DE AQUINO:

Sobre los principios de la Naturaleza. Edición de A. Osuna y traducción de E. García Estébanez. Madrid: Biblioteca de Autores Cristiano. (2001).

C) Manuales y textos básicos de Biología y ciencias

- AYALA, F.J. & VALENTINE, J.W. (1983). *La Evolución en Acción. Teoría y procesos de la evolución orgánica*. Madrid: Alhambra. [1979].
- DÍAZ, J. A. S., T. (2000). *Zoología. Aproximación evolutiva a la diversidad y organización de los animales*. Madrid: Editorial Síntesis.
- DOBZHANSKY, AYALA, STEBBINS & VALENTINE. (1983). *Evolución*. Barcelona: Omega.
- FERNÁNDEZ-RAÑADA, A. (Ed.). (1993). *Física Básica*. Madrid: Alianza.
- FONTDEVILA, A. & MOYA, A. (2003). *Evolución. Origen, adaptación y divergencia de las especies*. Madrid: Editorial Síntesis.
- GRIFFITHS, MILLER, SUZUKI, LEWONTIN & GELBART. (2002). *Genética*. Madrid: McGraw Hill. Interamericana. [2000].
- HICKMAN,ROBERTS, LARSON, L'ANSON, EISENHOUR. (2006). *Principios Integrales de Zoología*. Madrid: Mc Graw Hill. [2006].
- SOLER, M. (Ed.). (2003). *Evolución. La base de la Biología*. Proyecto Sur de Ediciones.

D) Bibliografía general: artículos y libros

- ACERO, J. J. (1995). "Teorías del contenido mental". En F. Broncano (Ed.). *La mente Humana*. Madrid: Trotta. Pp.175-206.
- ALLEN, C., BEKOFF, M. & LAUDER, G. (Eds.). (1998). *Nature's Purposes*. Cambridge, Massachusets: M.I.T. Press.
- ARIEW, A., CUMMINS, R., & PERLMANN, M (Ed.). (2002). *Functions. New Essays in the Philosophy of Psychology and Biology*. New York: Oxford University Press.

- BACH, K. (1988). "Burge's New Thought Experiment: Back to the Drawing Room." *The Journal of Philosophy*, 85 (2), 88-97.
- BAKER, L. R. (1995). *Explaining Attitudes*. New York: Cambridge University Press.
- BEDAU, M. (1991). Can Biological Teleology be Naturalized? *The Journal of Philosophy*, 88 (1), 647-655.
- BEDAU, M. (1992). Where's the Good in Teleology? *Philosophy and Phenomenological Research*, LII(4), 781-906.
- BERMÚDEZ, J. L. (2003). *Thinking without Words*. New York: Oxford University Press.
- BERMÚDEZ, J.L. (2005). *Philosophy of Psychology*. New York. Routledge.
- BIGELOW & PARGETTER. (1987). "Functions." *Journal of Philosophy*, 84, 181-196.
- BOCK & VON WAHLERT (1965). "Adapatation and the form-function Complex." *Evolution*, 19, 269-299.
- BOCK, J. (1980). "The Definition and recognition of biological adaptation." *American Zoologist*, 20, 217-227.
- BOORSE, C. (1976). "Wright on Functions". *The Philosophical Review*, 85(1), 70-86.
- BRADDON-MITCHELL, D. & JACKSON, F. (1997). "The Teleological Theory of Content". *Australasian Journal of Philosophy*, 75(4), 474-489.
- BRADDON- MITCHELL, D. & JACKSON, F. (2002). "A Pyrrhic Victory for Teleonomy". *Australasian Journal of Philosophy*, 80(3), 372-377.
- BULLER, D.J. (1998). "Etiological Theories of Function: A Geographical Survey." *Biology and Philosophy*, 13, 505-527.
- BULLER, D.J. (Ed.). (1999). *Function, Selection and Design*. Albany: State University of New York Press.

- BURGE, T. (1979). "Individualism and the Mental". En Peter A. French, Th. Uehling y K. Howard, (Eds.). *Midwest Studies in Philosophy IV*. Minneapolis: University of Minnesota Press.
- BURGE, T. (1986). "Individualism and Psychology", *The Philosophical Review*, vol. 95, nº1, 3-45.
- CARRUTHERS, P. (1998). "Thinking in language?: evolution and a modularist possibility". En P. Carruthers y J. Boucher (Eds) *Language and Thought*. C.U.P.
- CARRUTHERS, P. (2009). "Language in Cognition". En E. Margolis, R. Samuels y S. Stich (Eds). *The Oxford Handbook of Philosophy of Cognitive Science*.
- CARTWRIGHT, N. (1999). *The Dappled World*. Cambridge: Cambridge University Press.
- CHALMERS, D. J. (2002). "The Components of Content". En D. J. Chalmers, (Ed.). *Philosophy of Mind. Classical and Contemporary Readings*. Oxford: Oxford University Press. Pp. 608-633.
- CLARK, A. (1994). "Language of Thought". En S. Guttenplan, (Ed.). *A Companion to the Philosophy of Mind*. Oxford: Basil Blackwell. Pp. 408-412.
- CLARK, A. & CHALMERS, D. (1998). "The extended Mind", *Analysis* 58:10-23.
- CRANE, T. (1991). "All Differences in the World", *The Philosophical Quarterly*, Vol.41, nº162, 1-25.
- CUMMINS, R. (1975). "Functional Analysis". *Journal of Philosophy*, 72, 741-765.
- CUMMINS, R. (2002). "Neoteleology" En Ariew, A., Cummins, R. y Perlman, M. (Eds.). *Functions, New Essays in the Philosophy of Psychology and Biology*. Oxford: Oxford University Press. Pp.157-172.

- DAVIDSON, D. (1987). "Knowing One's Own Mind". *Proceedings and Addresses of the American Philosophical Association*, 60, 441-58. También en D. Davidson (2001). *Subjective, Intersubjective, Objective*. Oxfors: Oxford University Press.
- DAVIES, P.S. (1994). "Troubles for Direct Proper Functions". *Noûs*, 28, 363-81.
- DAVIES, P.S. (2001). *Norms of Nature*. Cambridge, Massachusets: M.I.T. Press.
- DENNETT, D. (1987). *The intentional stance*. Cambridge, Massachusets: M.I.T. Press.
- DE WAAL, F. (1996). *Good Natured*. Cambridge, Massachusets: Harvard University Press
- DIÉGUEZ LUCENA, A. J. (1998). "Los compromisos del realismo científico". En P. Martínez Freyre, (Ed.). *Filosofía actual de la ciencia*, suplemento nº3 de "Contrastes" Málaga. Pp. 145-173.
- DIÉGUEZ LUCENA, A. (1998). *Realismo científico. Una introducción al debate actual en la filosofía de la ciencia*. Málaga: Servicio de Publicaciones e Intercambio Científico de la Universidad de Málaga.
- DIÉGUEZ LUCENA, A. (2005). *Filosofía de la Ciencia*. Madrid: Biblioteca Nueva y Universidad de Málaga.
- DIÉGUEZ LUCENA, A. (2005). "Realismo y antirrealismo en la filosofía de la biología". *Ludus Vitalis*, vol. 13, n.23, 49-71
- DÍEZ, J.A. y MOULINES, C.U. (1997) *Fundamentos de Filosofía de la Ciencia*, Barcelona: Ariel.
- DRETSKE, F. (1969). *Seeing and Knowing*. London: Routledge & Kegan Paul.
- DRETSKE, F. (1977). "Laws of Nature". *Philosophy of Science*, 44, 248-268.
- DRETSKE, F. (1981). *Knowledge and the Flow of Information*. CSLI Publications.

- DRETSKE, F. (1983). "Précis of Knowledge and the Flow of Information". *Behavioural and Brain Sciences* 6, n 1, 55-63.
- DRETSKE, F. (1986). "Misrepresentation". En Bogdan (ed.) *Belief*, Oxford: Oxford University Press.
- DRETSKE, F. (1988). *Explaining Behaviour. Reasons in a World of Causes*. Cambridge, Massachusets: M.I.T Press.
- DRETSKE, F. (1989). *Reasons and Causes*. *Philosophical Perspectives*, 3, 1-15.
- DRETSKE, F. (1994). "The explanatory role of information". *Philosophical Transactions*. Royal Society of London, 349, 59-70.
- DRETSKE, F. (1995). *Naturalizing the Mind*. Cambridge, Massachusets: M.I.T. Press.
- DRETSKE, F. (2000). *Perception, Knowledge and Belief. Selected Essays*. Cambridge: Cambridge University Press.
- EVANS, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
- FARKAS, K. (2008). *The Subject's Point of View*. Oxford: oxford University Press.
- FODOR, J. (1975). *The Language of Thought*. Nueva York: Thomas Y. Crowell.
- FODOR, J. (1987). *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge, Massachusets: M.I.T. Press.
- FODOR, J. (1990). *A Theory of Content and Other Essays*. Cambridge, Massachusets: M.I.T.
- FODOR, J. (1998). *Concepts*, Cambridge, Massachusets: M.I.T. Press.
- FODOR, J. & LEPORE, E. (1992). *Holism. A shopper's guide*. Oxford: Basil Blackwell.
- FREGE, G. (1879). "Sobre sentido y referencia", en García Suárez, *Frege. Ensayos sobre semántica y filosofía de la lógica*. Madrid: Tecnos.

- GARCÍA MADRUGA ET ALII (2002). *Psicología evolutiva II. Desarrollo cognitivo y lingüístico*. Vol 1. Madrid: UNED.
- GARCÍA SUÁREZ, A. (1997). *Modos de significar. Una introducción temática a la filosofía del lenguaje*. Madrid: Tecnos.
- GETTIER, E. (1963). "Is justified true belief knowledge?" *Analysis*, 23, 121-123.
- GODFREY-SMITH, P. (1993). "Functions: Consensus without unity". *Pacific Philosophical Quarterly*, 74, 196-208.
- GODFREY-SMITH, P. (1994). "A modern history theory of functions" *Noûs*, 28, 344-362.
- GOMILA, A. (1995). "Externalismo semántico y determinación del contenido: el enfoque teleológico de R. Millikan". *Análisis Filosófico*, XV, 1/2.
- GOODMAN, N. "The problem of Counterfactual Conditionals". *The Journal of Philosophy*, 44, 113-128.
- GOULD, S.J. y LEWONTIN (1979). "The sprandels of San Marco and the panglossian paradigm: A critique of the adaptationist programme". *Proceedings of the Royal Society of London*, B205, 581-598.
- GOULD, S. J. y VRBA, E. (1982). "Exaptation-a mising term in the science of form". *Paleobiology*, 8(1), 4-15.
- GREENBERG, M. (2005). "A New map of the Theories of Mental Content: Constitutive Accounts and Normative Theories". En Villanueva, E. y Sosa, E (Eds.). *Normativity*. Pp. 299-320.
- GRICE, P. (1957). "Meaning". *The Philosophical Review*, 66, 377-388.
- GRIFFITHS. (1993). "Functional Analysis and proper functions". *British Journal for the Philosophy of Science*, 44, 409-422.
- HACKING, I. (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.
- HARDCASTLE, V. (Ed.). (1999). *Where Biology Meets Psychology*. Cambridge, Massachusetts: M.I.T. Press.

- HARMAN, G. (1965). "The inference to the best explanation." *The Philosophical Review*, 74, 88-95.
- HECK, R. G. (2000). "Nonconceptual Content and the "Space of Reasons"." *The Philosophical Review*, 109 (4), 483 - 523.
- HEIL, J. (2004). *Philosophy of Mind. A contemporary introduction*. New York: Routledge.
- HEMPEL, K. (1996). *La explicación científica*. Barcelona: Paidós. [1965]
- HEMPEL, K. & OPPENHEIM, (1948) "Studies in the Logic of Explanation." *Philosophy of Science*, 15, 135-138.
- HOUKES & VEERMAS. (2003). "Ascribing functions to technical artifacts". *British Journal of Philosophy of Science*, 54 (2).
- JAUME, A.L. (2005). *Análisis del concepto de función en ciencia y tecnología*. Trabajo de Grado presentado en la Facultad de Filosofía de la Universidad de Salamanca.
- JAUME, A.L. (2008). "La cuestión del realismo en el análisis del concepto de función". *Teorema*, XXVII (I), 57-71.
- KENNY, A. (2000). *La metafísica de la mente*. Barcelona: Paidós. [1989].
- KIM, J. (2006). *Philosophy of Mind*. Westview.
- KITCHER, P. (1993). "Function and design." *Midwest studies in Philosophy*, 18, 379-397.
- KRISTELLER, P. O. (1982) *El pensamiento renacentista y sus fuentes*. México: Fondo de Cultura Económica. [1979].
- LAUDAN, L. (1981). "A refutation of convergent realism." *Philosophy of Science*, 48, 19-48.
- LEWENS (2007). "Adaptation." En Hull y Ruse (eds.). *The Cambridge Companion to the Philosophy of Biology*. Cambridge: Cambridge University Press. Pp. 1-21.

- LIZ, M. (1995). "La estructura de las representaciones mentales: una perspectiva integradora". *Análisis Filosófico*, XV(1/2), 135-166.
- LIZ, M. (1995). "Causalidad y contenido mental". En Broncano, F. (Ed.). *La mente humana*, Madrid: Trotta. Pp. 207-244.
- LIZ, M. (2001) *Perspectivas actuales en Filosofía de la Mente*. Tenerife, Gobierno de Canarias.
- LOEWER, B. (1982). "Knowledge and the Flow of Information". *Philosophy of Science*, 49(2), 297-300.
- LUJÁN, J.L. (2007). "Alcance y límites de los modelos evolucionistas de cambio científico." *Ludus Vitalis*, XV, 28, 3-20.
- MACDONALD, G. y PAPINEAU, D. (Eds.). (2006). *Teleosemantics*. Londres: Oxford University Press.
- MAYR, E. (1983). "How to Carry Out the Adaptationist Program?". *The American Naturalist*, 121(3), 324-334.
- MAYR, E. (1985). *The growth of biological thought*. New York: Harvard University Press.
- MAYR, E. (1988). *Towards a new Philosophy of Biology*. Cambridge, Massachusetts: Harvard University Press.
- MAYR, E. (2005). *Así es la Biología*. Barcelona: Debate. [1995].
- MAYR, E. (2006). *Por qué es única la biología. Consideraciones sobre la autonomía de una disciplina científica*. Buenos Aires: Katz Editores. [2004].
- MOYA, C. J. (2004). *Filosofía de la mente*. Valencia: Publicaciones de la Universidad de Valencia.
- MÜLLER, G.L. (1998). *Dogmática. Teoría y práctica de la teología*. Barcelona: Herder. [1995].
- NAGEL, E. (2006). *La estructura de la Ciencia*. Barcelona: Paidós. [1961]

- MC LAUGHLIN, B. (Ed.). (1991). *Dretske and his Critics*. Cambridge, Massachusetts: Basil Blackwell.
- MC LAUGHLIN, P. (2001). *What Functions Explain*. New York: Cambridge University Press.
- MILLIKAN, R.G. (1984). *Language, Thought and Other Biological Categories*. Cambridge, Massachusetts: M.I.T. Press
- MILLIKAN, R.G. (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge, Massachusetts: M.I.T. Press.
- MILLIKAN, R.G. (2000). *On Clear and Confused Ideas. An Essay about Substance Concepts*. Cambridge: Cambridge University Press.
- MILLIKAN, R.G. (2002). *Varieties of Meaning*. Cambridge, Massachusetts: M.I.T. Press.
- MILLIKAN, R.G. (2005). *Language: A Biological Model*. New York: Oxford University Press.
- NEANDER, K. (1995). "Misrepresenting & Malfunctioning". *Philosophical Studies* 79: 109-141.
- NEANDER, K. (1996). "Swampnan Meets Swampcow". *Mind & Language* Vol.11, nº1, pp.118-129.
- NEANDER, K. (2006). Naturalistic Theories of Reference, en Devitt, M. y Hanley, R. *The Blackwell Guide to the Philosophy of Language*, Blackwell Publishing, pp.374-391.
- NEANDER, K.(2006). "Content for Cognitive Science". En Macdonald, G. y Papineau, D. (Eds.). (2006). *Teleosemantics*. Londres: OUP. Pp. 167-194.
- NEANDER, K. (2008). "Teleological Theories of Mental Content: Can Darwin Solve the Problem of Intentionality". En M. Ruse (Ed.), *The Oxford Handbook of Philosophy of Biology*. Oxford: Oxford University Press.
- PAPINEAU, D. (1987). *Reality and Representation*. Oxford: Blackwell.

- PAPINEAU, D. (1993). *Philosophical Naturalism*. Oxford: Blackwell.
- PAPINEAU, D. (1994). "Content (2)". En S. Guttenplan (Ed.). *A Companion to the Philosophy of Mind*. Oxford: Basil Blackwell. Pp. 225-230.
- PAPINEAU, D. (1997). "Teleosemantics and indeterminacy". *Australasian Journal of Philosophy*, 76, 1-14.
- PAPINEAU, D. (2001). "The Status of Teleosemantics, or How to Stop Worrying about Swampman". *Australasian Journal of Philosophy*, 79(2), 279-289.
- PAPINEAU, D. (2002). *Thinking about Consciousness*. New York: Oxford University Press.
- PAPINEAU, D. (2005). "Précis of Thinking about Consciousness". *Philosophy and Phenomenological Research*, LXXI(1), 143-186.
- PEACOCKE, C. (1994). "Content (1)". En S. Guttenplan (Ed.). *A Companion to the Philosophy of Mind*. Oxford: Basil Blackwell. Pp.219-225.
- PRESTON, B. (1998). " Why is a Wing like Spoon? A Pluralist Theory of Function". *The Journal of Philosophy*, XCV, 215-254.
- PRICE, C. (2001). *Functions in Mind. A Theory of Intentional Content*. New York: Oxford University Press.
- PUJADAS TORRES, L. M. (2002). *La ascensión y la caída de la teoría funcionalista de la mente*. Palma de Mallorca: Universitat de les Illes Balears.
- PUTNAM, H. (1975). *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- PUTNAM, H. (1973). "Meaning and Reference." *The Journal of Philosophy* 70, 699-711.
- QUESADA, D. (1995). "La concepción teleológica de los estados mentales y de su contenido." En Broncano, F. (Ed.). *La mente humana*. Madrid: Trotta. Pp. 77-96.

- QUINE, W.v. O. (1951). "Two Dogmas of empiricism". *The Philosophical Review*, 60, 20-43.
- QUINTANILLA, M. A. (2005). *Tecnología un enfoque filosófico y otros ensayos de filosofía de la tecnología*. México: Fondo de Cultura Económica.
- REEVE & SHERMAN (1993). "Adaptation and the goals of evolutionary research." *Quarterly Review of Biology*, 68, 1-32.
- ROSENBERG, A. (1985). *The Structure of Biological Science*. Cambridge: Cambridge University Press.
- ROSENBERG, A. & KAPLAN, D.M. (2005). "How to Reconcile Physicalism and Antireductionism about Biology". *Philosophy of Science*, 72, 43-68.
- ROSSI, P. (1990). *Las arañas y las hormigas*. Barcelona: Crítica. [1986].
- ROWLANDS, M. (1997). "Teleological Semantics". *Mind*, 106 (422), 279-303.
- RUSE, M. (1990). *La filosofía de la biología*. Madrid: Alianza Universidad. [1973].
- RUSSELL, B. (1994). *Los Problemas de la Filosofía*. Barcelona: Labor. [1912].
- SELLARS, W. (1997). *Empiricism and the Philosophy of Mind*. Cambridge, Massachusetts: Harvard university Press. [1956].
- SHANNON, C.E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, Vol. 27, pp. 379-423.
- SOBER, E. (1984). *The Nature of Selection*. Chicago: University of Chicago Press.
- SOBER, E. (1996). *Filosofía de la biología*. Madrid: Alianza Ed. [1993].
- SOLER, M. (2003). (Ed.) *Evolución, la base de la biología*. Sevilla: Proyecto Sur de Ediciones.
- SOLER, J.J.. (2003). "Selección natural y Adaptación" En Soler, M. (Ed.). *Evolución, la base de la Biología*. Proyecto Sur de Ediciones. Pp. 127-157.

- STALNAKER, R. (1998). "What Might Nonconceptual Content Be?"
Philosophical Issues, 9, 339-352.
- STAMPE, D.(1977). "Toward a causal theory of linguistic representation".
Midwest Studies in Philosophy, vol. 9. Minneapolis: University of
Minnesota Press.
- STERELNY, K. (1990). *The Representational Theory of Mind. An Introduction*.
Oxford: Blackwells Publishers.
- TINBERGEN, N. (1963). "On the aims and methods of ethology". *Zeitschrift für
Tierpsychologie*, 27, 1-32.
- TUGENDHAT, E. & WOLF, U. (1997). *Propedéutica lógico-semántica*,
Barcelona: Anthropos.
- VAN FRAASEN, B. (1980). *The Scientific Image*. Oxford: Oxford University
Press.
- VON BERTALANFFY, L. (1976). *Teoría general de los sistemas*. Madrid: Fondo
de Cultura Económica. [1968].
- VON WRIGHT, G. H. (1971). *Explanation and Understanding*. Londres:
Routledge and Keagan Paul.
- VYGOTSKY, L. (1996). *Pensamiento y lenguaje*. Barcelona: Paidós. [1934].
- WALSH, D. M. (1996). "Fitness and Function". *The British Journal of Philosophy
of Science*, 47 (4), 553-574.
- WALSH, D. M. & ARIEW. A. (1996). "A Taxonomy of Functions". *Canadian
Journal of Philosophy*, 4, 493-514.
- WEST-EBERHARD, M.J. (1998). "Adaptation: Current Usages". En Hull, D.L. y
Ruse, M. *The Philosophy of Biology*. Oxford: Oxford University Press.
Pp.8-14.
- WIENER, N. (1961). *Cybernetics*. Massachusetts: M.I.T. Press.
- WILLIAMS, G.C. (1996). *Adaptation and Natural Selection*. Princenton, New
Jersey: Princenton University Press. [1966].

- WITTGENSTEIN, L. (1988). *Investigaciones Filosóficas*. Barcelona:Crítica. [1953].
- WRIGHT, L. (1973). "Functions". *The Philosophical Review*, 82, 139-168.
- WOODFIELD, A. (1976). *Teleology*. Cambridge: Cambridge University Press.

E) Recursos electrónicos o artículos en línea

- Brown, Curtis, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2008/entries/content-narrow/>>.
- Neander, Karen, "Teleological Theories of Mental Content", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/content-teleological/>>.