

# La normatividad de lo mental y el rol de la segunda persona

## Tras las huellas de Donald Davidson

Autor:

Pedace, Karina S.

Tutor:

Pérez, Diana I.

2010

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

Posgrado

Tesis  
14-5-6

FACULTAD de FILOSOFIA y LETRAS  
Nº 859848  
31 MAR 2010 DE  
AGT. ENTRADAS

Tesis de Doctorado

**La normatividad de lo mental y el rol de la segunda persona**

**Tras las huellas de Donald Davidson**

Karina S. Pedace

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

Directora: Dra. Diana I. Pérez

A la memoria de Haydeé Bellagamba,  
querida *segunda* persona.

## ÍNDICE

Prefacio.....	1
---------------	---

### **Sección I: La filosofía de Donald Davidson, modelo para armar**

#### Capítulo 1: La normatividad como “cemento” del sistema davidsoniano

1.1. Terapia y Sistema.....	3
1.2. Las piezas del rompecabezas.....	9
1.3. La normatividad como “cemento” del sistema: el proyecto antropológico davidsoniano.....	17
1.4. En las huellas de Donald Davidson (o acerca de “Things that might have been”).....	21

#### Capítulo 2: La teoría de la interpretación radical..... 26

2.1. El diálogo de Davidson con Tarski: el desafío del tránsito desde la verdad al significado.....	29
2.2. El diálogo de Davidson con Quine: desde la traducción a la interpretación.....	35
2.3. ¿Comprender la comprensión?.....	42

#### Capítulo 3: El concepto de creencia y sus rasgos..... 48

3.1. Holismo.....	55
3.2. Caridad y Normatividad	
3.2.1. La caridad como antesala de la normatividad.....	67
3.2.2. Normatividad.....	77
3.3. El externismo de Davidson: la triangulación.....	86

### **Sección II: La filosofía de la mente de Donald Davidson**

#### Capítulo 4: El problema mente-cuerpo

4.1. El anomalismo de lo mental.....	95
4.2. La analogía entre las actitudes proposicionales y las actitudes numéricas y la <i>indeterminación</i> de la interpretación.....	103

4.3. La analogía entre las actitudes proposicionales y las actitudes numéricas y la normatividad de lo mental.....	116
--	-----

Capítulo 5: El problema de la relación entre el pensamiento y el lenguaje

5.1. Una “geografía” del problema.....	126
5.2. La realidad del lenguaje.....	135
5.3. La tesis de la conexión conceptual entre el pensamiento y el lenguaje.....	144

Capítulo 6: El problema de las perspectivas de primera y tercera persona de la

atribución mental. Subjetivo-Intersubjetivo-Objetivo.....159

6.1. La primera persona en el marco de la teoría de la interpretación radical: la autoridad de primera persona. Una “geografía del problema”.....	161
6.2. La autoridad de primera persona y el externismo davidsoniano.....	167
6.3. Primera y tercera persona en el marco de la atribución mental: el debate entre la teoría-teoría y la teoría de la simulación. Subjetivo-intersubjetivo-objetivo.....	181

**Sección III: En busca de nuevos horizontes**

Capítulo 7: Davidson y la dicotomía descriptivo/normativo: ¿el “cuarto dogma del empirismo”?

7.1. Tipos de debate y dicotomías. De la “discusión” a la “controversia”.....	191
7.2. La dicotomía descriptivo/normativo: su historia y su “colapso”.....	193
7.3. La dicotomía descriptivo/normativo en la filosofía de la mente de Davidson: ¿el “cuarto dogma del empirismo”?.....	204

Capítulo 8: La normatividad de lo mental y el rol de la segunda persona.....226

8.1. La perspectiva de segunda persona: repensando la atribución mental. El regreso de las emociones.....	227
8.2. Ser humano, ser persona, tener una mente.....	238
8.3. La relevancia moral de la perspectiva de segunda persona.....	260
8.4. La normatividad de lo mental y el rol de la segunda persona: conclusiones.....	268

Bibliografía.....	275
-------------------	-----

## PREFACIO

Este trabajo tiene una larga historia. En el año 2001 defendí mi tesis de licenciatura, uno de cuyos ejes fue la singular respuesta de Donald Davidson al problema mente-cuerpo: el *monismo anómalo*. Por entonces, me concentré en el examen de la normatividad de lo mental como aquel rasgo que, a su juicio, provee la clave de la peculiaridad de la psicología y hace que los conceptos mentales sean irreducibles a los físicos. Sin embargo, creo que sólo un tiempo después empecé a advertir la profundidad de su planteo. Recién cuando pude apreciar la sistematicidad de su filosofía apareció la “figura” que, en mi opinión, surge de su fascinante “rompecabezas” conceptual: su particular mirada acerca de qué nos hace ser quienes somos, su contribución antropológica. Las páginas que siguen pueden leerse como el intento de ir tras estas “huellas” que Davidson dejó en mí, para abrir un horizonte conceptual que, sin embargo, nos permita ir más allá de su propuesta.

Como no podía ser de otra manera, ha habido una multiplicidad de colegas y amigos que intervinieron en la gestación de las ideas que presento aquí. Eduardo Rabossi me dirigió en la primera etapa de la beca doctoral UBACyT con la que comencé a elaborar algunos de los temas de los que me ocupo. En esta ocasión, quiero honrar su memoria expresando mi gratitud por haberme permitido entrar en contacto, por primera vez, con la reflexión davidsoniana y por haberme iniciado en la investigación grupal. Asimismo, quiero agradecer a todos los miembros de los diversos proyectos que él dirigió junto a Aníbal Duarte, y que tuve el honor de integrar, por haberme ayudado en mis incipientes pasos en la filosofía de la mente.

A su vez, las reuniones del Grupo de Investigación en Cognición, Lenguaje y Conciencia, dirigido por Diana Pérez y Liza Skidelsky, me dieron la oportunidad de profundizar muchos de los aspectos de la filosofía de Davidson que discuto aquí. En este sentido, quiero agradecer el aporte y el estímulo que recibí de todos sus integrantes: Tomás Balmaceda Huarte, Sergio Barberis, Lucas Bucci, Federico Burdman, Mariela Destéfano, Sabrina Haimovici, Maximiliano Herrera, Ana Hulton, Carlos Limarino, Patricia Marechal, Andrea Melamed, Sheila Nesis, María del Carmen Perot y Christian Riopa. Por su parte, los encuentros que mantuvo el grupo con investigadores de la Universidad Nacional de Córdoba nutrieron

muchas de las reflexiones que aquí pongo en juego. Entre ellos, agradezco especialmente a Mariela Aguilera, Patricia Brunsteins, Laura Danón, Daniel Kalpokas, Carolina Scotto y Ana Testa.

Asimismo, las discusiones que mantuve al interior del proyecto de investigación acerca de “La experiencia de intersubjetividad y la perspectiva de segunda persona” fueron de suma importancia para madurar muchas de las ideas que presento aquí. En esta oportunidad, quiero reconocer las contribuciones que hicieron a mi línea de trabajo todos sus participantes: Federico Burdman, Silvia Español, Diego Lawler, Isabel Martínez, Diana Pérez y Favio Shifres. En ese contexto, también pude beneficiarme del intercambio de ideas con Antoni Gomila Benejam; vaya también a él mi gratitud.

Sin dudas, las reuniones que a lo largo de los dos últimos años mantuvimos con el grupo del programa de Reconocimiento Institucional de la Universidad de Buenos Aires: “La normatividad de lo mental: racionalidad e intencionalidad”, que tuve el placer de dirigir junto a Ana Hulton, resultaron absolutamente centrales para que pudiera clarificar y profundizar el “corazón” de la propuesta que presento aquí. Quiero expresarles mi inmenso agradecimiento a sus integrantes: Sergio Barberis, Federico Burdman, Valentina Contino y Abel Wajnerman.

Una mención especial merece Diana Pérez, directora de la tesis que presento a continuación. Las incontables discusiones que hemos mantenido, así como su lucidez, generosidad y apoyo constantes, resultaron decisivos para que este trabajo, finalmente, cobrara vida. Para ella, maestra y amiga, van mi admiración y total gratitud.

En un terreno estrictamente personal, quiero agradecer el apoyo que en esta empresa de largo aliento recibí de mis amigas: Eleonora Cresto, Ana Hulton y Anita Kalikies. Finalmente, quiero expresar mi profundo agradecimiento a mis padres, Teresa y Horacio, a mis hermanas, Romina y Marianela y, especialmente a mi hijo, Valentín, y a Sergio, mi compañero de ruta, sin cuyo amor, paciencia y compromiso, nada de lo que soy habría sido.

## **SECCIÓN I**

**La filosofía de Donald Davidson, modelo para armar**

# Capítulo 1: La normatividad como “cemento” del sistema davidsoniano

El método supone una sociedad de dos.  
(Davidson, 1984c/2004:35)<sup>1</sup>

## 1.1. Terapia y Sistema

Los “rompecabezas” estuvieron entre mis juegos favoritos en la infancia. Como es sabido, se trata de juegos que consisten en componer una determinada figura, combinando cierto número de piezas, en cada una de las cuales hay una parte de la figura. Desde aquella resonancia infantil, creo que el proyecto filosófico de Donald Davidson puede concebirse así: como un fascinante “rompecabezas” conceptual. En efecto, cada una de sus partes está íntimamente conectada con las restantes, componiendo una mirada original que reformula, y a la vez enlaza, debates que han sido centrales a una vasta porción de la filosofía. Prácticamente, no ha habido área de la misma en la que no haya incursionado, haciéndolo además de manera novedosa y profunda, tal como lo atestiguan sus contribuciones en filosofía de la mente y del lenguaje, epistemología, filosofía de la acción y antropología filosófica.

En una entrevista concedida a Giovanna Borradori (1994), el propio Davidson sugiere la caracterización “lúdica” que estoy haciendo de su proyecto filosófico: allí confiesa que aunque mucha gente pudo advertir sólo muy posteriormente que sus distintas propuestas se iban integrando entre sí, algunas de las ideas centrales ya las tenía en mente desde el principio, esperando su turno para ubicarse, como si fueran las piezas de un rompecabezas<sup>2</sup>.

Al hilvanar un *sistema* filosófico, Davidson hace una apuesta que no resulta muy popular entre los filósofos contemporáneos. El propio David Lewis escribe al respecto: “me hubiera gustado ser un filósofo no sistemático, que ofreciera propuestas independientes a una variedad de

---

<sup>1</sup> En todos los casos, ofreceré mis propias traducciones de los textos de Davidson, excepto en Davidson (1992).

<sup>2</sup> Para una consideración ya no meramente “lúdica” sino “artística” de la obra de Davidson, puede consultarse el artículo de Morris, R. “The art of Donald Davidson”, quien encuentra en ella “un edificio, a la vez, mínimo y complejo”. Al decir de Morris, quien ve su contribución como una pieza de arte conceptual: “comprendemos a Davidson como un filósofo, pero lo sentimos como un artista” (en Hahn 1999: 132).

tópicos” (1983: ix). Esta expresión parece resultar muy elocuente acerca del poco favor del que goza actualmente la sistematicidad<sup>3</sup>. Entre otros cargos, se le suele imputar que pone en jaque muchas de nuestras intuiciones filosóficas. Así, por ejemplo, en el caso de Davidson, que la oración “Ferdinando se casó con Cristina” deba entenderse como una cuantificación sobre eventos, que la psicología no pueda ser una ciencia precisa, que nadie pueda ser considerado como teniendo creencias en absoluto a menos que se pueda considerar que tiene creencias en su mayoría racionales y verdaderas, que para conocer el significado de una oración deba conocerse el lenguaje entero al que pertenece, éstas y otras tesis, pueden parecerle inaceptables a mucha gente.

Por otra parte, ser un filósofo sistemático no es fácil. Así, Evnine (1991) subraya que el sistema davidsoniano está sujeto a la irrupción de tensiones generadas internamente. A su juicio, esto se debe a que Davidson está comprometido simultáneamente con dos proyectos muy diferentes. Por un lado, encontramos un proyecto causal, explicativo, que reposa en su trabajo sobre los *eventos* y la *causación* y que busca explorar tópicos de la filosofía de la mente tales como la explicación y la producción de la acción, y la relación entre la mente y el cuerpo, en términos de estos conceptos fundamentales. Por otro lado, hay un proyecto interpretativo, hermenéutico, que reposa en su posición acerca de la importancia de la *racionalidad* para lo mental, e intenta describir nuestra interacción con agentes racionales como una interacción única, a partir de la aplicación del principio normativo de caridad a nuestra vida lingüística y mental.<sup>4</sup>

Evnine concede, sin embargo, que la obra de Davidson tiene virtudes innegables, que no consisten en la preservación de intuiciones filosóficas aisladas, sino en ofrecer una visión que ha provisto una dirección a la filosofía. No sólo en el sentido de que mucha gente trabaja hoy sobre problemas que estableció su agenda, sino porque ubicó ante nosotros un objetivo sustantivo: “explicar la importancia de la racionalidad y el rol que tiene en la lógica, la ciencia, el pensamiento y la vida” (Evnine 1991:174).

En “Can Analytic Philosophy Be Systematic and Ought it to Be?”, Michael Dummett (1978) vuelve sobre las huellas de Wittgenstein y Austin para ilustrar el rechazo al sistema y

---

<sup>3</sup> Sin embargo, hay honrosas excepciones. Ejemplo paradigmático de ello es la filosofía de Robert B. Brandom, tal como lo afirma, por ejemplo, en (Brandom 2002:1).

<sup>4</sup> El corazón de la lectura crítica que hace Evnine del proyecto davidsonino consiste en imputarle dos teorías distintas del contenido. Así, a su juicio, las tensiones que lo habitan dependen de que el proyecto causal se compromete con una “teoría realista del contenido”, mientras que el proyecto hermenéutico adhiere a una “teoría idealista del contenido” (1991: 176).

exhibir sus reparos acerca de la posición de que no es posible ninguna teoría sistemática del significado. A su juicio, el trabajo en filosofía puede ser “sistemático” en dos sentidos (1978: 455): por un lado, en tanto se ofrezca una teoría *articulada* (en la línea de B. Spinoza o I. Kant) y, por otro, en tanto se ajuste a *métodos* acordados de investigación y sus resultados sean aceptados o rechazados en función de dichos criterios, esto es, en tanto se ciña a criterios compartidos de investigación y testeo. El aspecto sistemático de la filosofía de Davidson que deseo destacar aquí es aquel que se vincula, en especial, con el primero de los sentidos que deslinda Dummett: se trata de un proyecto finamente articulado, desde una propuesta unificada acerca del pensamiento, el significado y la acción. El costo de dicha unificación, sin embargo, no ha estado exento de fuertes críticas.

En el mapa que nos ofrece Rorty (1997), el movimiento intelectual llamado “filosofía analítica” puede concebirse como oscilando entre dos concepciones de la filosofía: por un lado, como *terapia* y, por otro, como *constructora de sistemas*. Así, su tarea ha pendulado entre sacarse de encima “pseudo-problemas” o proponer explicaciones sistemáticas. ¿Dónde ubicar la contribución de Donald Davidson? Su originalidad consiste, en mi opinión, en que se extiende entre ambas vertientes, esto es: algunos aspectos de su contribución son terapéuticos y otros resultan propios del análisis conceptual sistemático. Sin embargo, a juicio de Rorty esta doble filiación genera una tensión inaceptable que, en su opinión, debería resolverse a favor del polo “terapéutico”.

En efecto, Rorty sostiene que varias doctrinas por las cuales Davidson se hizo famoso en los '80 son más afines a las del último Wittgenstein, que a las perspectivas de pensadores sistemáticos como Carnap o Tarski; por ejemplo, su idea de que la mayoría de nuestras creencias deben ser verdaderas (según veremos, un aspecto del “principio de caridad”) y que no hay distinción entre saber un lenguaje y saber cómo lidiar con el mundo [*know our way around in the world*]. Ahora bien, a juicio de Rorty, podemos retener los argumentos davidsonianos para estas tesis aún si no tenemos ningún interés en lo que Dummett llama una “teoría sistemática del significado” (esto es, el tipo de teoría que el último Wittgenstein consideró implausible e innecesaria y de la que, a su vez, resulta paradigmática una teoría tarskiana de la verdad para un lenguaje natural, tal como la que abraza Davidson). Precisamente, uno de los principales reparos que ofrece a propósito de la obra de Davidson es la propensión de éste a conectar todos sus argumentos con el proyecto interpretacionista. Pero, ¿podemos quedarnos con los argumentos y

prescindir de la teoría de la interpretación radical? ¿Podemos comprender la relevancia de introducir el principio de caridad sin ella? Yo creo, en contra de Rorty, que los argumentos aludidos reposan de un modo tal en la empresa interpretacionista que si nos deshacemos de ella, nos deshacemos del corazón de la propuesta davidsoniana. A mi juicio, la complejidad y riqueza (composicional, sistemática y productiva) que, según veremos, resulta distintiva de la conducta lingüística quiebra la analogía que sugiere Rorty entre hablar un lenguaje y andar en bicicleta, al punto de que lo que se requiere es una especificación también peculiar; esto es, una especificación que nos proporcione la estructura - recursiva- de esta práctica característicamente humana. Tal es, precisamente, la enorme contribución que Davidson encuentra en una teoría tarskiana de la verdad como vehículo para una teoría del significado: ilumina la manera en que el significado resulta aprendible (e infinitas emisiones, en consecuencia, interpretables) para seres finitos como nosotros<sup>5</sup>.

A mi juicio, en cambio, los aspectos terapéuticos y sistemáticos de la filosofía de Davidson deben ser vistos como complementarios y en ello, precisamente, radica gran parte de su atractivo: a la vez que nos muestra cómo se desmoronan muchos de los problemas

---

<sup>5</sup>En el capítulo próximo volveré, especialmente, sobre este punto. Por ahora, simplemente me interesa señalar que en su respuesta a Rorty (en Hahn 1999) el propio Davidson insiste en que los modos tarskiano y wittgensteiniano pueden ser reconciliados. Allí enfatiza la importancia de un trasfondo lingüístico institucionalizado contra el cual se comprende la conducta verbal desviada. Los miembros de la comunidad de habla comparten una constelación de hábitos de habla y tienen expectativas correspondientes acerca de lo que otros en la comunidad quieren significar mediante lo que dicen. Asimismo en Davidson (1986/2005:) apunta en esta dirección de complementariedad al subrayar que no es suficiente contar con una teoría formal para comprender al hablante. Por su parte, Glock (2003) sostiene que “en un nivel táctico”, hay interesantes similitudes entre Wittgenstein y Davidson: por ejemplo entre la discusión de Wittgenstein acerca de seguir una regla y la discusión en torno a la interpretación radical, o en sus reacciones acerca de la distinción analítico-sintético. También encuentra paralelos a “nivel estratégico”. En modos diferentes, ambos son parte del así llamado “giro lingüístico” de la filosofía analítica: le adscriben un rol filosófico central al lenguaje, aunque por razones diferentes. Ambos poseen concepciones del lenguaje y de la mente que se apartan del platonismo y el cartesianismo. Finalmente, ambos hay importantes temas pragmatistas, especialmente, el énfasis en la importancia filosófica de la acción humana. Sin embargo, es justo señalar que, a juicio de Glock, las diferencias son igualmente relevantes, e insiste en que, frente a la confianza de Davidson respecto del poder de la lógica formal, el último Wittgenstein fue escéptico acerca del valor filosófico del análisis lógico.

tradicionales que heredamos de la modernidad propone tesis novedosas profundamente interconectadas.

Según interpreta nuestro autor la situación de la filosofía en el presente, estamos asistiendo a un cambio fundamental de enormes consecuencias. En la raíz misma de este cambio se halla, según Davidson, la crítica a que está siendo sometida la concepción tradicional de las relaciones entre la subjetividad y el mundo objetivo. Desde Descartes hasta nuestros días, esta concepción se ha basado en la postulación de entidades mediadoras entre ambos términos de la relación: las ideas de Descartes y Locke, las impresiones e ideas de Hume, las intuiciones y conceptos de Kant, los datos sensoriales del positivismo lógico. Más allá de las diferencias que, sin duda, pueden trazarse entre estos expedientes, todos ellos coinciden en su función de mediación entre el sujeto y el mundo objetivo; todos ellos presuponen, asimismo, la posibilidad de distinguir, en el marco del pensamiento y conocimiento humanos, entre los conceptos y un contenido no contaminado por ellos. Uno de los aspectos fundamentales del pensamiento davidsoniano está constituido por la crítica a dicha separación entre los conceptos y un material neutro, no conceptualizado (cfr. especialmente Davidson 1988/2001). Esta crítica conlleva un ataque frontal a la concepción epistemológica que arranca en la modernidad y a la concepción de la mente asociada a ella, así como un cambio de rumbo decisivo en la reflexión filosófica. Con este cambio de rumbo no sólo se pone en cuestión la inteligibilidad de la idea de un dato sensorial absolutamente libre de conceptualización, sino también la comprensión tradicional de los conceptos mismos y de su función, y el hecho de que esta función haya sido definida bajo el supuesto de la separación entre ellos y un elemento no conceptual; definición que es ahora sometida a revisión crítica. La idea de que los conceptos son formas o estructuras de organización de un material conceptualmente neutro pierde contenido desde el momento en que no hay tal material en espera de organización. Para Davidson el dualismo en cuestión, esto es, el dualismo esquema-contenido es “él mismo un dogma del empirismo, el tercer dogma. El tercero y tal vez el último<sup>6</sup>, pues si lo abandonamos no resulta claro que quede ya algo distintivo que merezca el nombre de empirismo” (cfr. Davidson 1974/1984:189).

En Davidson (1988/2001) señala que una de las consecuencias de abandonar las dicotomías esquema-contenido y subjetivo-objetivo no es que se evaporen los problemas

---

<sup>6</sup> Retomaré este punto en la sección III, donde indagaré la persistencia en Davidson de un posible “cuarto dogma del empirismo”.

epistemológicos, sino que los problemas que parecían salientes cambian. Así, las preguntas dejan de ser de epistemología tradicional para pasar a ser preguntas acerca de la naturaleza de la racionalidad. Se trata de preguntas, nos dice, que como las preguntas epistemológicas que reemplazan, no tienen una respuesta final, pero a diferencia de las preguntas que sustituyen, a éstas vale la pena tratar de responderlas.

El marco novedoso en el que han de inscribirse dichas respuestas va a ser su proyecto unificado del pensamiento, el significado y la acción. A la vera del mismo, según veremos, vamos a asistir a una filosofía de la mente post-ontológica en la que estados de la mente como creencias, deseos, intenciones, no son entidades, sino modificaciones de una persona; describimos estas modificaciones relacionando la persona con una oración, pero no se trata de que haya objetos abstractos que estén “ante la mente” de la persona a la que se le atribuye el pensamiento (Davidson 1989/2001). Los supuestos de la interpretación obligan a concebir el contenido de nuestras creencias básicas como un evento u objeto público, y no como una entidad intermedia entre el sujeto y el mundo. Los contenidos de nuestras creencias básicas son parte del mundo público e intersubjetivo, y no objetos entre éste y nosotros. La concepción representativa de la mente como un espectáculo privado y accesible sólo al “ojo interior” de cada sujeto se ve asimismo, por las mismas razones, desacreditada por el análisis de los supuestos necesarios de la interpretación y la comunicación intersubjetiva, y con ella la legitimidad de los problemas epistemológicos a que ha dado lugar, como el escepticismo, el conocimiento de otras mentes o la existencia del mundo externo. Davidson concibe, así, su propia obra como parte del movimiento que conduce a una transformación profunda de los supuestos en que se ha movido la reflexión filosófica desde los tiempos de Descartes.

Las importantes consecuencias que se desprenden de la filosofía davidsoniana dependen, pues, crucialmente de su *peculiar concepción de la mente humana* desde la perspectiva de la interpretación del lenguaje y de la conducta, es decir, desde la perspectiva del proceso por el que un sujeto trata de hallar sentido en la conducta y las emisiones lingüísticas de otro. De ello resulta la concepción de la mente bajo el principio constitutivo de la racionalidad y bajo el supuesto de la veracidad de las creencias. Como señala acertadamente Moya (1992), la mente es un producto de la interpretación y de la comunicación intersubjetiva; es, en suma, aquello que atribuimos a los demás para hacernos inteligible su conducta, lingüística y no lingüística. Como veremos, esta concepción de la mente sitúa su estudio más allá del alcance de un modelo

explicativo inspirado en las ciencias naturales. Las ciencias que se ocupan de la acción intencional humana han de proceder de modo holista e interpretativo, ajustando sus resultados a la evidencia en desarrollo bajo la guía del carácter globalmente coherente de la vida mental y la conducta de los agentes. Este modo de proceder las separa –tal como veremos en el capítulo 4– de la búsqueda de leyes y de la explicación nomológica que caracteriza las ciencias de la naturaleza.

Dada la sistematicidad que anima el proyecto de Davidson, un primer objetivo que persigo en el presente capítulo es presentar, a muy grandes trazos, la relación entre las tesis que sostiene en diversas áreas de la filosofía. En segundo lugar, voy a ofrecer las razones en virtud de las cuales creo que aquello que opera como “cemento” de su sistema filosófico, aquello que lo cohesiona y le confiere un singular “espesor”, esto es, la “figura” que surge del rompecabezas cuando las piezas se acomodan adecuadamente, es su proyecto antropológico, su particular mirada acerca de qué nos hace ser quienes somos. En este sentido, mi propuesta será considerar su teoría de la interpretación como el hilo conductor del sistema y como el marco en el que voy a insertar mi indagación, que estará circunscripta a una esfera de problemas a la que Davidson ha sido especialmente sensible: la filosofía de la mente. Finalmente, explicitaré mis razones para abrazar el marco conceptual davidsoniano en el tratamiento de los tres problemas de los que voy a ocuparme en el presente trabajo, a saber: el problema mente-cuerpo, el problema de la relación entre el lenguaje y el pensamiento y el problema de las perspectivas de primera y tercera persona en la atribución mental.

## **1.2. Las piezas del rompecabezas**

Lepore y Ludwig (2006) consideran que a pesar de que la mayor parte de la producción de Davidson se encuentra dispersa en la forma de artículos aparecidos en revistas y volúmenes colectivos, es posible encontrar en ellos un programa filosófico unificado. Dicho programa, a su juicio, puede concebirse como organizado en torno a dos proyectos conectados: el primero, concerniente a la comprensión de la naturaleza de la agencia humana y el segundo, relativo a la naturaleza y función del lenguaje y su relación con el pensamiento y el mundo<sup>7</sup>. Veamos en qué consiste cada uno de ellos, e intentemos precisar sus relaciones.

---

<sup>7</sup> En una línea consonante lo interpretan, *inter alia*, Caorsi y Silva Filho (2008) quienes consideran la elaborada e influyente reflexión de Davidson como articulada, por un lado, sobre algunos de los principales conceptos de la filosofía de la acción (como razones, causas, eventos y acciones) y, por otro, sobre conceptos indispensables para la

A mediados de los años sesenta, Davidson comenzó a desarrollar de manera paralela dos proyectos que en ese momento no se veían conectados entre sí. En la filosofía de la mente y del lenguaje, se trataba de una teoría del significado para las lenguas naturales basada en el modelo tarskiano de las oraciones-T. El interés de Davidson era ofrecer una teoría del significado que pudiese explicar qué necesitamos saber para poder interpretar las oraciones de un hablante en una lengua dada y cómo llegamos a saberlo. El otro proyecto pertenecía a la teoría de la acción y, en este caso, se trataba de ver si las razones por las cuales realizamos una acción pueden explicar causalmente esa acción. Empecemos, entonces, por este último (que, en rigor, fue el que apareció primero en orden de publicación).

Como es sabido, Davidson adquirió notoriedad filosófica recién en 1963 con la publicación de “Acciones, razones y causas”. A mi juicio, en dicho ensayo ya se encuentra prefigurada la mayor parte de las tesis que luego irá desplegando y puliendo delicadamente, como en “filigrana”, para la conformación de su sistema filosófico<sup>8</sup>. Así, el corazón del argumento que ofrece, presenta los dos grandes pilares de su filosofía: el principio de *causalidad* y el principio de *racionalidad*. A partir de ellos, Davidson va a ir tejiendo la trama de su sistema. En efecto, allí no sólo presenta su posición en torno a la explicación de la acción intencional sino que, al hacerlo, también anticipa su singular respuesta al problema mente-cuerpo, que ha de estar inscrita en la teoría de la interpretación (en la que, según veremos, convergen su filosofía de la mente y del lenguaje), de la que han de desprenderse, a su vez, sus tesis epistemológicas, y hasta su posición acerca de la objetividad los valores. En este espíritu, sus dos proyectos- ambos sumamente técnicos y elaborados- no sólo terminan integrándose en una teoría unificada del pensamiento, el significado y la acción, sino que muestran insospechadas consecuencias en múltiples regiones del pensamiento. Así, por ejemplo, es sobre la base de la teoría de la interpretación, que Davidson ofrece argumentos contra el escepticismo global, el relativismo y la inconmensurabilidad. Asimismo, objeta tanto el realismo metafísico

---

filosofía del lenguaje (como significado y verdad). A su juicio, esta reflexión proveyó una base para sus tesis en filosofía de la mente, metafísica y epistemología. Hay, en su opinión un “lazo que une esas dos áreas, otorgando unidad a su pensamiento, razón por la cual, la filosofía después de Davidson no puede ignorar el hecho de que la creencia, la racionalidad y el lenguaje están íntimamente vinculados” (2008:7).

<sup>8</sup> En cierta ocasión leí, a propósito de la obra del poeta Roberto Juarroz, que podríamos pensar que escribió un solo poema en toda su vida, y que el resto de su obra es un perfeccionamiento, una serie de variaciones de aquella contribución inicial. Al releer este ensayo temprano de Davidson, me resulta casi imposible no extrapolar esa idea: creo que en él podemos encontrar en ciernes los elementos centrales de su sistema conceptual, que despliega sutil y finamente en sus ensayos posteriores. Esta es la razón por la que voy a detenerme, con cierta morosidad, en su presentación.

como el antirrealismo y llega incluso a sostener que los valores morales pueden ser objetivos. Por otra parte, el análisis davidsoniano de la irracionalidad es particularmente importante para una comprensión filosófica del psicoanálisis<sup>9</sup> dado que, como es sabido, en ese terreno Freud dejó muchos cabos sueltos que no siempre han sido atados por las generaciones de estudiosos del psicoanálisis.<sup>10</sup>

Veamos, entonces, cuál es la pregunta con la que nos interpela Davidson al comienzo de “Acciones, razones y causas”: “¿qué relación hay entre una razón y una acción cuando una razón explica una acción ofreciendo la razón del agente para hacer lo que hizo?” (1963/1980:3). La posición que se desea defender allí es que las explicaciones de una acción mediante razones (explicaciones a las que propone llamar *racionalizaciones*) son una forma de explicación causal, siendo las razones causas de la acción. De este modo, pretende rehabilitar una tesis “antigua y de sentido común”<sup>11</sup>, aunque reelaborada adecuadamente, frente a la ortodoxia imperante en las discusiones acerca de la explicación de la acción.<sup>12</sup> En efecto, Davidson concibió este artículo como una reacción contra la doctrina ampliamente aceptada en esos días en el mundo filosófico anglosajón según la cual las razones no son causas de la acción.<sup>13</sup>

Las razones a las que apelamos para explicar por qué alguien hizo lo que hizo, justifican la acción en cuestión en tanto nos conducen a ver algo que el agente vio (o pensó que vio) en su acción y, de este modo, la tornan inteligible para nosotros.<sup>14</sup>

---

<sup>9</sup> Al respecto, puede verse Cavell, M. (1996).

<sup>10</sup> Cabe señalar que hasta la teoría literaria se ha visto afectada por el pensamiento de Davidson. La manera cómo él entiende la metáfora (como un fenómeno pragmático, y no semántico), modificó notablemente la concepción de la metáfora aristotélica tradicional.

<sup>11</sup> La resonancia “antigua” de la tesis se remonta al intento de Aristóteles –que, no obstante, tendrá que ser debidamente reformulado– de resolver la “misteriosa” conexión entre razones y acciones mediante la introducción del deseo como un factor causal.

<sup>12</sup> Al defender que no es necesario abandonar la idea antigua de que las racionalizaciones son una forma de explicación causal, Davidson dice tomar distancia *inter alia* de: Gilbert Ryle (1949), G.E. M. Anscombe (1957), Stuart Hampshire (1959), Hart y Honoré (1959), William Dray (1957), Anthony Kenny (1963) y A. I. Melden (1961).

<sup>13</sup> La concepción no causal de las razones contó con una importante adhesión, en especial, entre diversos filósofos wittgensteinianos quienes, según Davidson, encontraron en *Los cuadernos azul y marrón* una de las fuentes principales para su defensa – y que aunque publicados recién en 1958, ya circulaban ampliamente desde mediados de los ’30-.

<sup>14</sup> Davidson sostiene que dar la razón por la cual el agente hizo lo que hizo consiste en ofrecer lo que él denomina la “razón primaria” por la que el agente realizó la acción, esto es, el par constituido por una *actitud favorable* del agente hacia determinada clase de acciones (bajo esta noción incluye, paradigmáticamente, *desear*) y la *creencia* de que su acción pertenece a esa clase. El apelar a una razón primaria es, precisamente, lo que distingue a las racionalizaciones de otro tipo de explicaciones (Cf. Davidson, 1963/1980: 7 y 8). Más adelante, Davidson modifica esta concepción enraizada en la doctrina aristotélica del silogismo práctico ante la dificultad que plantean las situaciones de conflicto entre deseos y/o creencias, en las que el agente tiene razones tanto a favor como en contra

Davidson sostiene que algunos filósofos al advertir que las explicaciones causales no revelan este elemento de justificación que aportan las razones, han concluido que el concepto de causa que se aplica en otras partes no puede aplicarse a la relación entre razones y acciones y que el patrón de justificación proporciona *per se*, en el caso de las razones, la explicación requerida.

En defensa del carácter no causal de las razones, sus partidarios habían esgrimido una importante cantidad de argumentos, entre los cuales se destacó el así llamado “argumento de la conexión lógica” entre razones y acción. Su estructura general puede reconstruirse del modo siguiente: (i) La causa debe ser lógicamente distinta del efecto<sup>15</sup>; (ii) una razón para una acción no es lógicamente distinta de la acción; (iii) se concluye, entonces, que las razones no son causas de las acciones.<sup>16</sup>

Una de las variedades específicas que adquirió este argumento se encuentra en A. I. Melden (1961: caps. 8-10): si se ofrece como razón para una acción el deseo de llevarla a cabo, no se está ofreciendo una causa de la acción, dado que el concepto mismo del deseo (por ejemplo, mi deseo de *encender la luz*) contiene el concepto de la acción que explica (mi *encender la luz*) y se supone que los conceptos de la causa y el efecto son recíprocamente independientes. En el caso del deseo de encender la luz no podemos entender la naturaleza de ese deseo sin incluir el concepto de la acción deseada y, en consecuencia, el deseo en cuestión no puede ser la causa de esa acción. En suma: mi deseo de encender la luz explica que encienda la luz, pero esa explicación no es causal.

Como señala Carlos Moya (1992:18), en su respuesta al argumento de la conexión lógica, la estrategia de Davidson consiste, básicamente, en aceptar buena parte de las premisas de sus adversarios y en negar, sin embargo, que de ellas se desprenda la conclusión que ellos pretenden. La movida básica de Davidson consiste, entonces, en concederle a sus interlocutores que las racionalizaciones difieren de las explicaciones en las ciencias naturales, en el sentido de

---

de un determinado curso de acción (si se tratase de una consecuencia lógica de las razones del agente, esto supondría que en tales situaciones el agente infiere conclusiones contradictorias y lleva a cabo acciones incompatibles, resultado que llevó a Davidson a revisar su concepto original de acción intencional). Para un pormenorizado tratamiento de su filosofía de la acción puede verse Moya, C. (1990).

<sup>15</sup> A esta primera premisa subyace la concepción humeana de la relación causal según la cual los términos de dicha relación son eventos distintos e independientes, no habiendo entre ellos otro vínculo de unión más que la regularidad y la constancia con que se nos presentan juntos en la experiencia.

<sup>16</sup> Davidson (1963/1980: 13) afirma que pueden encontrarse distintas versiones de este argumento, por ejemplo, en Kenny (1963), Hampshire (1959), Winch (1958), Peters (1958) y Melden (1961). Y señala, asimismo, que en una de sus versiones, el argumento estuvo inspirado en el tratamiento que Ryle (1949) hizo de los motivos.

que las razones *justifican* las acciones al explicarlas, pero su punto es que de ello no se sigue que la explicación no sea *también – y necesariamente- causal*. De esta suerte, sus objeciones van a concentrarse en la tesis según la cual la justificación es un tipo de explicación tal que no necesita recurrir a la noción ordinaria de causa. A juicio de Davidson, si pensamos que la justificación sólo abarca que el agente tiene ciertas creencias y actitudes a la luz de las cuales la acción es razonable, estamos dejando afuera algo esencial, puesto que alguien puede tener una razón para una acción y, sin embargo, realizar la acción en cuestión sin que esta razón sea la razón por la que la hizo (Davidson 1963/1980:9). En consecuencia, sostiene que debemos trazar una distinción entre las meras “razones *para*” y las “razones *por*” las cuales una acción fue hecha, distinción que desde su perspectiva puede hacerse inteligible sólo en términos causales: la razón *por* la cual una acción fue hecha es aquella que la *causó*; una razón que favorece una acción - esto es, una “razón *para*”-puede fallar en causarla o en contribuir causalmente a ella, en cuyo caso no tiene ningún rol en la explicación. Veámoslo mediante un ejemplo. Ud. tenía una razón para cruzar la calle porque quería comprar un ramo de jazmines en la florería de enfrente. Pero Ud. cruzó efectivamente la calle porque vio que una anciana trastabilló y cayó sobre la vereda. Su deseo de ayudar a la anciana, no su deseo de jazmines, causó y, por ello, explica, su cruce de la calle. Ambos deseos son razones para la acción, pero sólo uno de ellos la explica. Según Davidson, el rol justificatorio de una razón depende, pues, de su rol explicativo.<sup>17</sup>

En suma, según la tesis davidsoniana, una racionalización exhibe los dos aspectos siguientes. Por un lado, justifica o racionaliza la acción al dar la razón por la cual el agente hizo lo que hizo; la idea es que a la luz de una razón una acción “se revela como algo coherente con ciertos rasgos del agente [...] y se muestra al agente en su papel de Animal Racional” (Davidson 1963/1980:8). Por otro lado, dice por qué la acción tuvo lugar y en virtud de que ese “porque” tiene carácter causal, la racionalización es una forma de explicación causal ordinaria (Davidson 1963/1980:11-12).

---

<sup>17</sup> Pero no a la inversa, cf. Davidson (1963/1980: 8). Según esta perspectiva, pareciera que -en principio- el papel explicativo de las causas es primario y básico, en tanto las razones derivan su rol justificatorio a partir de la eficacia causal que desempeñan. No obstante, está lejos de ser claro exactamente por qué Davidson cree que el rol justificatorio de las razones es derivado de, o dependiente de su rol explicativo, ni tampoco está claro que quiere decir “dependencia” en este contexto. Asimismo, al sostener que “la causa y el efecto forman el tipo de patrón que explica el efecto, en un sentido de “explicar” que comprendemos tan bien como cualquier otro”, pareciera considerar como no problemático ni controvertido tomar a la causación como relación explicativa, punto que, en mi opinión, amerita aclaraciones. No obstante, en virtud del propósito que persigo en este capítulo, a saber, iluminar algunas de las conexiones que le otorgan carácter sistemático a su contribución, no voy a detenerme en estas dificultades. Simplemente deseo enfatizar la centralidad que tienen en su propuesta las categorías de *racionalidad* y *causalidad*.

De este modo, su original comprensión de la naturaleza de la agencia humana combina dos elementos que se habían considerado incompatibles: la idea de que las explicaciones de la acción en términos de razones poseen una función de justificación, junto con la idea de que son una forma de explicación causal.

En entrevista concedida a Ernie Lepore (2004), Davidson dice que “Acciones, razones y causas” fue, simplemente, el resultado de darse cuenta de que nadie tenía un buen argumento en contra de las teorías causales de la acción. Y afirma que aquello que vinculó a lo que originalmente eran dos proyectos que no estaban claramente relacionados (uno en filosofía de la acción y el otro en filosofía del lenguaje y de la mente) fue advertir la relevancia que tiene para la filosofía de la acción el análisis de las *actitudes proposicionales*.

Según veremos, para Davidson, las creencias y los deseos son estados que atribuimos a los demás en el proceso de interpretación de su conducta, y en este proceso, el contenido de esos estados ha de ser tal que la conducta aparezca como racional en relación con ellos; de otro modo no hemos conseguido entenderla como acción intencional. Por ello es esencial la forma en que describimos las creencias, los deseos y las acciones. Una acción puede estar justificada a la luz de ciertos deseos y creencias cuando se describe de cierta manera y no cuando se describe de otra. El énfasis en la concepción de la mente y de la conducta intencional se sitúa, pues, en la descripción que una persona hace de otra con el fin de entenderla. Formulado en otros términos: la concepción davidsoniana de la mente está dominada por la perspectiva de tercera persona, por la perspectiva del proceso por el cual cada uno de nosotros trata de entender a los demás y ellos a nosotros. La mente, podríamos decir, es lo que atribuimos a los demás para hacernos inteligible su conducta, lingüística y no lingüística. El problema, pues, no afecta sólo a la comprensión de la acción, sino también del lenguaje.

A diferencia de las perspectivas introspectiva y platónica, la filosofía davidsoniana del lenguaje no pretende encontrar algo –representaciones mentales o entidades objetivas ideales– que haga significativa el habla. La pregunta davidsoniana no es ¿qué es el significado? ni ¿qué hace significativa la emisión de ciertos sonidos? Sino más bien la siguiente: dado que los seres humanos somos animales que hablamos ¿cómo podemos entender lo que decimos? El problema del significado se convierte en el problema de la interpretación y de la comunicación entre los hablantes.

La investigación davidsoniana de la comunicación y de la interpretación lingüística es, según veremos, heredera del análisis quineano de la traducción radical, que en Davidson pasa a denominarse *interpretación radical*. No se trata de un mero cambio terminológico: podemos saber que una oración traduce a otra sin saber qué significa ninguna de las dos; en cambio, la interpretación de una oración ha de proporcionarnos su significado. El intérprete radical pretende construir una teoría del significado de las emisiones aparentemente lingüísticas de un sujeto cuyo lenguaje le es completamente desconocido. Como veremos, situar el punto de partida del análisis de la interpretación en esta situación extrema es un artificio metodológico destinado a poner de manifiesto los aspectos implicados en la comunicación normal entre los seres humanos. La ventaja que reviste es que nos permite evitar que pasen inadvertidos presupuestos importantes de la comunicación, cosa que puede suceder fácilmente si analizamos la comunicación en el caso de sujetos que comparten un lenguaje y una cultura.

De acuerdo con esto, la tarea del intérprete radical consiste en elaborar una teoría de la verdad acerca de las emisiones que pretende interpretar, es decir, cuyo significado pretende conocer. Esta teoría debe dar como resultado teoremas que expresen, para cada oración que se interpreta, las condiciones en que esa oración es verdadera. Formalmente, los teoremas en cuestión son enunciados bicondicionales que son verdaderos en el caso de que ambas partes del bicondicional lo sean. Asimismo, según veremos, la teoría ha de estar ceñida por ciertas restricciones, entre las cuales será central una de las dos categorías centrales del sistema davidsoniano, que vimos aparecer seminalmente en “Acciones, razones y causas”: el principio de *racionalidad*.<sup>18</sup>

Por su parte, el otro pilar que pone en juego en el ensayo de 1963, el principio de *causalidad*, nos permite anticipar algunos aspectos de su pretendido naturalismo y efectuar ciertas aclaraciones preliminares. Davidson concede a sus contendientes que entre las descripciones de las razones y la acción puede haber una conexión conceptual o “lógica”, pero su punto es que esto no impide que las razones puedan ser causas de la acción. La clave de su respuesta reside, a mi juicio, en una herramienta teórica que atraviesa su sistema y a la que echa mano recurrentemente para distintos fines: la distinción entre el plano extensional y el intensional. De este modo, es posible, retener a la acción intencional humana dentro de la

---

<sup>18</sup> Dado que en el próximo capítulo voy a ocuparme, especialmente, de la *teoría de la interpretación radical* como marco para la tríada de problemas de los que voy a ocuparme en el presente trabajo, postergo su tratamiento detallado hasta entonces.

dimensión causal –extensional de la naturaleza física y, a la vez, concebirla como una conclusión justificada a partir de determinadas razones del agente, dadas ciertas descripciones de su conducta; tal es la dimensión intensional del planteo davidsoniano. Como se trata tanto de razones como de causas, ciertos eventos se convierten, pues, en acciones intencionales y libres.<sup>19</sup> En otros términos: la acción intencional es un proceso causal como cualquier otro, aunque la describimos de un modo tal que la singularizamos frente a otros procesos causales, confiriéndole el aspecto racional que le es característico.

Allí donde Davidson desconfía de la concepción naturalizada de la filosofía de Quine y hacia el carácter omniabarcante de la ciencia, no presenta tendencia alguna hacia la eliminación del discurso mental. La adopción de una tesis naturalista (ya no metafilosófica, sino metafísica) no conlleva dicha eliminación a menos que sea acompañada por un fuerte compromiso científicista, el rechazo de la existencia de, por ejemplo, estados y eventos mentales y la reducción del discurso mental al discurso físico. Ello no se sigue si se admite, como Davidson, que los eventos pueden tener, además de la descripción física, otras descripciones verdaderas no menos legítimas desde el punto de vista epistemológico. En este sentido, cabe subrayar que el naturalismo davidsoniano no va acompañado, como en Quine, de una naturalización de la reflexión filosófica e, incluso, según veremos, resulta compatible con la defensa de tesis tradicionalmente vinculadas al humanismo, como la autonomía de las ciencias sociales y humanas frente a las ciencias naturales y de los eventos mentales frente a las leyes físicas. Asimismo, cabe señalar que al rechazar el dualismo de esquema-contenido, y con él, el empirismo, Davidson se concede libertad para la reflexión puramente conceptual. En efecto, Davidson emplea sin reservas el modo de investigación *a priori*, provocando con ello la reacción de los representantes del empirismo y de la filosofía naturalizada. Así, por ejemplo, su desconfianza en las perspectivas de la psicología como ciencia natural obedece, como veremos, a razones vinculadas con la pura reflexión conceptual. Tal como advierte Moya, no se trata de que Davidson menosprecie en algún sentido los métodos o resultados de la investigación científica, sino que es consciente de que, cuando se llega a determinadas cuestiones, este tipo de investigación no puede servir de ayuda. La obra de Davidson no representa, pues, únicamente una ruptura con el empirismo, sino también una marcada recuperación de la autonomía de la

---

<sup>19</sup> En el capítulo 4, a instancias de su peculiar tratamiento del problema mente-cuerpo, voy a volver sobre esta interpretación de su filosofía en clave de la herencia kantiana que él mismo hace expresa, esto es, en términos del problema de cómo conciliar la necesidad natural con la libertad.

reflexión filosófica frente al discurso científico. Se trata de una obra en la que, según hemos visto sucintamente, hay una íntima conexión entre las ideas en filosofía de la acción, filosofía de la mente, filosofía del lenguaje y epistemología, a un punto tal que es difícil aislar una tesis de otra, un argumento de los otros. Dada esta fina articulación, es tentador intentar buscar el “cemento” del sistema. Tal es el objetivo que persigo a continuación.

### 1.3. La normatividad como “cemento” del sistema: el proyecto antropológico davidsoniano

Al inscribir a Davidson en el “panteón” de aquellos filósofos que buscan la sistematicidad, hice mención de David Lewis quien, según vimos participa –aunque, a su pesar- de la misma titánica empresa. A menudo, se sostiene que en Lewis aquello que parece operar como principio unificador del sistema es la hipótesis del realismo modal. En efecto, se afirma que a partir de ella se da, desde la dimensión metafísica, una respuesta orgánica a las preguntas por el contenido lingüístico y mental, a las nociones de necesidad y contingencia, causación, probabilidad, ley física y persistencia a través del cambio. Ahora bien ¿hay algún principio unificador en el sistema davidsoniano que desempeñe un papel análogo al de la tesis realista modal en Lewis? En otros términos: ¿hay una tesis o conjunto de tesis que permita ofrecer una respuesta orgánica a una variedad de cuestiones tales como el problema mente-cuerpo, la relación entre acciones, razones y causas, el contenido del pensamiento y el lenguaje, el abandono del escenario epistemológico tradicional y su impacto para la cuestión escéptica y para la autoridad de la primera persona, la resignificación de las relaciones entre las dimensiones subjetiva-intersubjetiva-objetiva? Creo que la respuesta es positiva: el hilo conductor del sistema davidsoniano es la *teoría de la interpretación radical*. A la luz de ella, se articulan las posiciones de Davidson en filosofía de la mente, filosofía del lenguaje y epistemología, y se traza el puente hacia la filosofía de la acción (a partir de complementarla con la teoría de la decisión, para armar su teoría unificada<sup>20</sup>) y emerge la figura del “rompecabezas”: su peculiar *concepción antropológica*. Y, como veremos a continuación, no estoy sola en esta opinión.

---

<sup>20</sup> Volveré sobre este punto en el capítulo 3.

En efecto, E. Lepore y K. Ludwig (2006:19) afirman que la contribución de Davidson tiene tal interconexión que, para apreciarla, debemos leer sus ensayos como una obra única, considerando cada uno a la luz de los otros. Así, a su juicio, finalmente veremos emerger una *concepción sistemática del hombre como un animal lingüístico racional*.

Por su parte, también en opinión de S. Eynine (1991) debemos buscar el principio unificador del sistema davidsoniano en el proyecto antropológico que surge de la teoría de la interpretación radical. Él nos recuerda que los principales intereses filosóficos de Davidson son el lenguaje, lo mental y la acción, esto es, todos los ingredientes de una antropología filosófica. De este modo, el trabajo técnico en lógica, en semántica formal, su indagación metafísica sobre la naturaleza de los eventos y la causación, aunque son importantes e interesantes por derecho propio para Davidson, son primariamente el aparato necesario para la consecución de su proyecto antropológico. Según veremos en detalle, nuestro autor cree que la mente, la acción y el lenguaje están todos sujetos a un conjunto de principios: los principios normativos de la racionalidad. Dichos principios están relacionados con los estados mentales, las acciones y las emisiones lingüísticas que gobiernan de modo tal que, según argumenta Davidson, no se puede decir de alguien que tiene estados mentales, lleva a cabo acciones intencionales o es usuario de un lenguaje, a menos que sea racional.

Kirk Ludwig (2003) también considera que la unidad de la obra de Davidson reposa en el rol central de su propuesta antropológica, esto es, en el papel que juega la reflexión acerca de cómo somos capaces de interpretar a los demás, acerca de las actitudes proposicionales, y nuestra posición epistémica con respecto a nuestras propias mentes, las mentes de los otros y el mundo que nos rodea.

Asimismo, a juicio de H. J. Glock (2003)<sup>21</sup>, la antropología filosófica tiene un lugar preponderante dentro de la contribución davidsoniana. En efecto, sostiene que lo que conecta la

---

<sup>21</sup> Cabe señalar el interés que reviste este libro a partir de una de las provocadoras tesis que defiende su autor, a saber: las posiciones de Quine y Davidson se captan mejor vistas *en conjunción*. La mayor parte de los ensayos existentes que han explorado el contrapunto entre ambos autores se concentra, a juicio de Glock, en tópicos específicos, más que en los rasgos generales de sus sistemas filosóficos. Asimismo, hay varios libros dedicados a presentar e iluminar las conexiones entre diversos aspectos de la obra de Quine o Davidson (v.gr. Gibson 1982; Hookway 1988 y Orenstein 2002 ofrecen introducciones a Quine; en tanto que lo propio hace Eynine 1991 con Davidson). No obstante, señala Glock, no hay un libro sobre *ambos*. Adicionalmente, el grueso de la bibliografía secundaria no desafía ciertos supuestos básicos de sus obras, supuestos que devinieron ortodoxias en la filosofía analítica post-positivista. Algunas de las objeciones que han aparecido en los artículos críticos son convincentes, dice, pero a menudo resultan escolásticas y relativas a detalles, más que a una mirada panorámica sobre sus contribuciones. Frente a este estado de la cuestión, el libro de Glock se pretende, entonces, como el primer texto

posición de Davidson con la de su maestro, Willard van O. Quine<sup>22</sup>, es una concepción singular del ser humano y de la acción humana, en el sentido de que ambos sostienen que el pensamiento y la racionalidad están encarnados en la conducta lingüística. Por esta razón, su antropología filosófica gira en torno a una concepción del lenguaje. Más aún, en su opinión, esta concepción del lenguaje combina la semántica formal desarrollada por Tarski y Carnap, con la idea pragmatista de que el lenguaje es, esencialmente, una forma de conducta humana. Ambos consideran al lenguaje como un proceso de comunicación e interacción, pero insisten en que puede ser elucidado con la ayuda del cálculo lógico abstracto.

Ahora bien, frente a esta suerte de consenso a propósito del rol central que posee la teoría de la interpretación radical como clave del proyecto antropológico davidsoniano, me interesa señalar dónde reside el núcleo de mi planteo. La intuición que vertebra mi indagación de la filosofía de Davidson es que aquello que opera como “cemento” de su sistema, aquello que lo cohesionan, que le confiere un singular “espesor” y que permite apreciar en qué sentido somos “animales racionales”, es la *normatividad*. A mi juicio, el aspecto más controvertido y original de la obra de Davidson es la idea de que las criaturas que han de ser vistas como teniendo creencias, hablando un lenguaje y llevando a cabo acciones intencionales, deben ser racionales. Creo, pues, que a la luz de este peculiar rasgo que es la *normatividad*, es posible advertir que la “figura” que surge del rompecabezas cuando las piezas se acomodan adecuadamente, es su proyecto antropológico, su particular mirada acerca de qué nos hace ser quienes somos.

Sin embargo, en mi opinión, la *normatividad* no ha recibido un tratamiento adecuado al interior de la obra de Davidson. Según él mismo concede, lo que ha:

---

dedicado a *ambos*, concebido como un intento de elucidar y ponderar sus contribuciones en filosofía del lenguaje, filosofía de la mente y metafísica. Lo anima un espíritu crítico: a su juicio, los argumentos de Quine y Davidson a menudo no son convincentes y sus conclusiones están, en parte, erradas. Como Frege, Russell y Wittgenstein, Quine y Davidson son grandes filósofos analíticos, no porque hayan ofrecido demostraciones incuestionables, sino porque han socavado supuestos profundamente establecidos de un modo original e iluminador.

<sup>22</sup> En muchos lugares Davidson presenta su pensamiento como mera continuación y elaboración de las ideas de Quine. Mucho de su vocabulario, su aparente confianza en la lógica y la semántica formal, enfatizan aspectos que pueden filiarse en él. No obstante, en el capítulo próximo, consideraré varios aspectos en los que Davidson se aparta claramente de Quine, y en los que se revela como un real innovador filosófico, portavoz de un pensamiento sumamente original.

[d]icho acerca de las normas que gobiernan nuestras teorías de la atribución intensional es, en los puntos cruciales, *crudo, vago o demasiado rígido*. La manera de mejorar nuestra comprensión de tal comprensión es mejorar nuestra captación de los estándares de racionalidad implícitos en toda interpretación del pensamiento y la acción (1980b/2004:166; mis itálicas).

Frente a la “crudeza”, la “vaguedad” y la “rigidez” con las que el propio Davidson asume haber tratado a la *normatividad* se impone, entonces, la tarea de articular de un modo preciso este rasgo que porta la clave para captar la peculiaridad de un *nosotros humano*. En este sentido, creo que es importante destacar que a pesar de la profusa cantidad de trabajos publicados recientemente sobre este tópico, no hay un acuerdo razonable entre los autores acerca de cómo entender la *normatividad*. Es por ello que, uno de los objetivos rectores del presente trabajo, es el de contribuir a esclarecer en qué sentido atribuir estados intencionales es atribuir estados normativos y en qué medida responder a dicha pregunta involucra responder a la cuestión acerca de qué nos hace ser quienes somos.

#### 1.4. En las huellas de Donald Davidson (o acerca de “*Things that might have been*”)<sup>23</sup>

##### Things that might have been

Pienso en las cosas que pudieron ser y no fueron.  
El tratado de mitología sajona que Beda no escribió.  
La obra inconcebible que a Dante le fue dada acaso  
entrever, ya corregido el último verso de la Comedia.  
La historia sin la tarde de la Cruz y la tarde de la cicuta.  
La historia sin el rostro de Helena.  
El hombre sin los ojos, que nos ha deparado la luna.  
En las tres jornadas de Gettysburg la victoria del Sur.  
El amor que no compartimos.  
El dilatado imperio que los Vikings no quisieron fundar.  
El orbe sin la rueda o sin la rosa.  
El juicio de John Donne sobre Shakespeare.  
El otro cuerno del Unicornio.  
El ave fabulosa de Irlanda, que está en dos lugares a un tiempo.  
El hijo que no tuve.

Jorge Luis Borges (*Historia de la noche*)

Davidson ha sido, a mi juicio, uno de los filósofos más influyentes y originales del siglo XX. Su gravitación sobre las derivas de la filosofía analítica resulta innegable. Según vimos, con “Actions, reasons and causes” revolucionó el debate en filosofía de la acción; a partir de “Truth and Meaning” y “Radical Interpretation” contribuyó a redefinir la discusión acerca del enfoque adecuado para la semántica del lenguaje natural. Con “The Myth of the Subjective” desafió el escenario epistemológico cartesiano. Y a partir de “Mental Events” reconfiguró la discusión en filosofía de la mente.

Actualmente, disponemos de presentaciones generales de su pensamiento que - frente a lecturas fragmentarias- han hecho justicia al carácter sistemático y comprehensivo de su obra, tales como las de Lepore y Ludwig (2005), Joseph (2004), Caorsi (2001) y Evnine (1991). Asimismo, contamos con indagaciones de su filosofía del lenguaje como, por ejemplo, las ofrecidas por Moretti (2008), Ramberg (1989), Lepore (1986), Malpas (1992) y Hernández Iglesias (1990), sólo por mencionar algunas de ellas. También hay exploraciones de su metafísica en Vermazen y Hintikka (1987) y Lepore, E. & McLaughlin, B. (1985), su filosofía de la acción como, *inter alia*, la de Moya (1990), así como de su epistemología (al respecto, pueden

---

<sup>23</sup> Le agradezco a Diana Pérez esta bella y perturbadora referencia poética, que revela un hecho profundo de nuestras vidas: la infinita contingencia de lo que hacemos, y lo que no.

consultarse por ejemplo, Preyer, Siebelt y Ulfig (1994), Zeglen (1999) y las excelentes compilaciones de Hahn, L.E. (1999) y Kotatko, P., Pagin, P. & Segal, G. (2001).

Tal como sostuve más arriba, las importantes consecuencias que se desprenden de la filosofía davidsoniana dependen crucialmente, en mi opinión, de su peculiar concepción de la *mente*. Con ella, pretendió socavar la imagen que está a la base de los dualismos modernos, esto es, un concepto de la mente con estados y objetos privados. Ahora bien, hasta donde sé, aún no disponemos de una obra que se haya ocupado de exhibir las conexiones de su *filosofía de la mente*, en el sentido en que pretendo hacerlo aquí.<sup>24</sup>

En efecto, en las páginas que siguen, voy a intentar iluminar las relaciones conceptuales que hay al interior de su filosofía de la mente y a examinar sus consecuencias. Para ello, el recorrido que propongo es el siguiente. En el capítulo 2 presentaré la teoría de la interpretación radical, que será mi “hilo de Ariadna” para mostrar la figura que, a mi juicio, surge del “rompecabezas” conceptual de Davidson cuando las piezas están adecuadamente ubicadas. A los efectos de exponer dicha teoría, me concentraré, por un lado, en su inscripción en el programa semántico que hace del concepto de *condiciones veritativas* de una oración la clave para desarrollar las ideas de *significado y comprensión* y, por otro, en las *restricciones* surgidas a la luz de la consideración de la situación del *intérprete radical*. De la conjunción de estos dos aspectos, emerge, a mi juicio, la originalidad de la contribución davidsoniana, cuyos requisitos voy a presentar, en primer lugar, para abocarme, luego, al objetivo y la naturaleza de la propuesta resultante.

En el capítulo 3 me voy a concentrar en los tres rasgos centrales de la noción davidsoniana de *creencia*, a saber: *holismo, normatividad y triangulación*. En cada caso, distinguiré variedades y propondré una articulación allí donde en la posición de Davidson no quedan caracterizados de modo preciso y claro. Esta presentación ha de resultar un rodeo imprescindible a los efectos de emprender la tarea de la siguiente sección, que consistirá en analizar el impacto de dichos rasgos en las respuestas davidsonianas a tres problemas tradicionales de la filosofía de la mente: el *problema mente-cuerpo*, el *problema de la relación*

---

<sup>24</sup> Desde luego, soy plenamente consciente de los excelentes aportes críticos que en esta área han hecho al respecto, por ejemplo, Rorty (1999), McDowell, J. (1998), Ramberg, B. (2000) y Kim (1989), sólo por citar algunos de los autores que se sintieron interpelados por su “monismo anómalo”. Pero, hasta donde llega mi conocimiento, ninguno de ellos se ocupó de exhibir las conexiones y explorar las consecuencias de su tratamiento de los tres problemas de filosofía de la mente de los que voy a ocuparme aquí, ni tampoco de elucidar la *normatividad de lo mental* bajo la clave que voy a proponer aquí.

*entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la atribución mental.*

En el capítulo 4 presentaré la peculiar respuesta de Davidson al problema mente-cuerpo: el *monismo anómalo*. En primer lugar explicitaré la tesis central de esta posición, que propondré descomponer en dos subtesis: la teoría de la identidad de casos, y la de la irreducibilidad de los conceptos mentales a los físicos. Luego de reconstruir la prueba davidsoniana de la primera de las tesis mencionadas, me detendré en el argumento a favor del anomalismo de lo mental a los efectos de examinar posibles candidatos para dar cuenta de la peculiaridad de lo mental. Para ello, a continuación, consideraré una analogía a la que Davidson apela en varios artículos: la analogía entre las actitudes proposicionales y las actitudes numéricas. Mostraré que el rasgo que provee la clave de la singularidad de la psicología está dado por la *normatividad de lo mental*, de modo que resultará crucial clarificar en qué consiste.

En el capítulo 5 me ocuparé de la posición de Davidson en torno al problema de la relación entre el pensamiento y el lenguaje. En primer lugar, ofreceré una “geografía” de la discusión. Previo a considerar la *tesis de la interdependencia entre pensamiento y lenguaje* que defiende Davidson, examinaré cuál es la concepción del lenguaje que quiere atacar, consideraré su argumento en contra de dicha posición y trataré de clarificar su perspectiva. A continuación, examinaré el rol que desempeñan en su argumentación a favor de la conexión conceptual entre el pensamiento y el lenguaje los tres rasgos del concepto de creencia que presenté en la primera sección de este trabajo, a saber: el *holismo*, la *triangulación* y la *normatividad de lo mental*.

En el capítulo 6 retomaré un punto señalado en la primera parte del trabajo: Davidson pretende que la asunción de la perspectiva de *tercera persona*, como metodológicamente básica, deje lugar para el reconocimiento de la autoridad de la *primera persona*. Pasaré, entonces, a ocuparme del tratamiento davidsoniano de la dimensión subjetiva desde el marco provisto por la teoría de la interpretación radical y a indagar de qué modo puede compatibilizarse con su externismo. A partir de estas reflexiones, a continuación, voy a evaluar cuál sería la posición de Davidson en un debate en el que no intervino expresamente, pero en el que clásicamente se lo interpreta de manera “teoricista”, a saber: el debate entre los defensores de la teoría de la simulación y los defensores de la teoría-teoría, en torno al problema de las perspectivas de primera y tercera persona de la atribución mental. Intentaré argumentar en contra de la interpretación “teoricista” estándar de Davidson, en una dirección que señalará allí donde él, a mi

juicio, no terminó de explorar, a saber: el rol de la *intersubjetividad* en su concepción social e interactiva de la mente. Según veremos, frente a las perspectivas de primera y tercera persona, en mi planteo cobrará una relevancia central la perspectiva de *segunda persona* de la atribución mental tal como se la introdujo recientemente en la filosofía de la mente.

De mi examen crítico del tratamiento davidsoniano de los tres problemas considerados, surgirá un supuesto en el que, a mi juicio, reposa su filosofía de la mente, y que la compromete con consecuencias indeseables: la dicotomía descriptivo/normativo. En sección final del trabajo presentaré, entonces, mi propuesta. Ésta ha de consistir en ofrecer un *correctivo* y un *complemento* al fascinante entramado conceptual de Davidson, de modo que podamos repensar su controvertido abordaje de los problemas de los que nos hemos ocupado: esto es, el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la atribución mental. El correctivo consistirá en despojarnos de la dicotomía descriptivo/normativo (que presentaré como el “cuarto dogma del empirismo”) y el complemento residirá en la adopción de la perspectiva de segunda persona de la atribución mental.

En el capítulo 7 argumentaré que si despojamos a la filosofía davidsoniana de la mente de este “cuarto dogma del empirismo”, no sólo habremos de tornar más “davidsoniano” a Davidson, sino que abriremos un horizonte conceptual para proponer una elucidación de aquel rasgo central de su planteo, que él mismo asume haber dejado “vago”: la *normatividad de lo mental*. A tal efecto, ofreceré una historia de la dicotomía descriptivo/normativo y de su “colapso”, y alegaré razones, tanto externas como internas al planteo de Davidson, para desembarazarnos de ella.

Finalmente, mi invitación en el capítulo 8 será a que clarifiquemos la *normatividad de lo mental* a partir de una perspectiva que el propio Davidson “bautizó” pero que, en mi opinión, no exploró en toda su hondura: la perspectiva de *segunda persona* de la atribución mental. A tal efecto, en primer lugar, ofreceré una caracterización de dicha perspectiva tal como se la introdujo recientemente en filosofía de la mente. Y, a continuación, entablaré un diálogo con las posiciones de D. Dennett y P. F. Strawson para mostrar de qué modo, una vez que nos despojamos de la dicotomía descriptivo/normativo, podemos repensar los tres problemas de los que nos hemos ocupado (el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la

atribución mental) y trazar continuidades, allí donde para Davidson se abren hiatos y la discusión se detiene *a priori*.

Pienso, pues, en el libro que Davidson no escribió. Como en la literatura de James Joyce, que tanta fascinación ejerció sobre él<sup>25</sup>, me propongo leerlo como un acto de complicidad: pasar de ser espectadora a ser colaboradora del escritor. Humildemente, me uno, entonces, a Davidson en la invención de aquel libro que podría haber sido, pero no fue.

---

<sup>25</sup> Cfr. Davidson (1989b/2005) y (1993/2005).

## Capítulo 2: La teoría de la interpretación radical

Toda comprensión del habla de otro involucra interpretación radical.  
(Davidson, 1973/1984:125)

La teoría de la interpretación radical va a ser mi “hilo de Ariadna” para mostrar la figura que, a mi juicio, surge del “rompecabezas” conceptual de Donald Davidson, cuando las piezas están adecuadamente ubicadas. Esto es, voy a insertar su singular contribución antropológica en el marco de la teoría de la interpretación radical, en la medida en que constituye, en mi opinión, su explicitación de aquello que posibilita una práctica peculiarmente *humana*, la comunicación lingüística, aquella práctica en la que, característicamente, tratamos a alguien como a uno de *nosotros*<sup>26</sup>.

En “Theories of Meaning and Learnable Languages” (1965), Davidson propone lo que considera como un rasgo necesario de un lenguaje aprendible: debe ser posible dar una *explicación constructiva* del significado de sus oraciones. Siguiendo a Lepore y Ludwig (2006), llamaremos *requisito de composicionalidad* a esta condición que Davidson impone para una teoría del significado para lenguajes naturales. Conforme a él, una teoría del significado tiene que dar cuenta de cómo el significado de las oraciones depende del significado de las palabras. La idea que está detrás de este requisito es la siguiente. Dado que el número de enunciados pertenecientes a un lenguaje natural es potencialmente infinito, ningún sujeto como nosotros, esto es, con capacidades finitas, sería capaz de aprender el significado de esos enunciados de modo directo. Podríamos, no obstante, aprender a construir el significado de cada uno de esos enunciados en base a los significados de sus partes componentes. Pero para que esto resulte posible, el significado de los enunciados debe ser construible a partir del significado de las palabras que los componen. En otros términos: dado que somos seres finitos y que toma una cantidad de tiempo adquirir cada primitivo semántico, los lenguajes naturales –que tienen una

---

<sup>26</sup> En consecuencia, el propósito que persigo aquí no es el de un examen crítico de la teoría de la interpretación, sino su presentación en tanto telón de fondo en el que inscribo mi propuesta.

infinidad de oraciones no sinónimas- nos resultarían inaprendibles, a menos que sean comprendidos sobre la base de un número finito de primitivos y de reglas de composición.

A continuación, caractericemos, entonces, una teoría composicional del significado<sup>27</sup>:

[SC] Una teoría composicional del significado para un lenguaje L es una teoría formal que permite a cualquiera que comprenda el lenguaje en el que se formula la teoría comprender las expresiones primitivas de L y las expresiones complejas de L sobre la base de comprender las primitivas.

Una teoría composicional del significado debe comenzar, pues, con axiomas para las expresiones primitivas del lenguaje y generar una especificación del significado para cualquier expresión compleja en él. A su vez, los axiomas han de estar divididos en básicos y recursivos. Los recursivos nos indican cómo descomponer las expresiones complejas de nuestro lenguaje objeto en expresiones más simples –que aún pueden ser complejas- hasta arribar a los axiomas primitivos (esto es, básicos).

En “Truth and Meaning” (1967) Davidson propone emplear una teoría de la verdad para lograr el objetivo que persigue una teoría composicional del significado. Su original sugerencia consiste en que una teoría de la verdad del tipo que Tarski mostró cómo formular para lenguajes formales podría ser empleada para producir el apareamiento entre las oraciones del lenguaje en estudio y las oraciones del metalenguaje y que el conocimiento apropiado acerca de la teoría de la verdad nos pondría en posición de comprender cualquier oración del lenguaje en estudio sobre la base de comprender sus primitivos semánticos y las reglas para su combinación. La clave para advertir cómo es esto posible reside en la Convención T, punto en el que nos detendremos en la próxima sección.

Al mismo tiempo que Davidson hace esta sugerencia, propone que la teoría del significado debe ser verificable sin el conocimiento de las actitudes proposicionales detalladas del hablante. El objetivo de esta condición es prevenir la introducción subrepticia en los fundamentos de la teoría, de conceptos excesivamente emparentados con el concepto de significado. Tal como señala Caorsi (2001), de lo que se trata es de dar una fundamentación de la teoría de manera no circular. A tal efecto, Davidson apela al *punto de vista del intérprete radical*. Esta sugerencia está inspirada por el enfoque de Quine para el estudio filosófico del

---

<sup>27</sup> Sigo la propuesta de Lepore y Ludwig (2005), cap. 2.

lenguaje: el proyecto de la traducción radical (cf. Lepore y Ludwig 2005, cap.11). Tal como sostienen Lepore y Ludwig (2006:14) el punto que los liga es la idea de que en virtud de que el lenguaje es por su naturaleza un instrumento para la comunicación, el significado debe ser accesible desde el punto de vista *público*. No obstante, cabe anticipar que entre los dos proyectos hay diferencias relevantes en las que nos detendremos oportunamente.

En suma, a mi juicio, es posible discernir –al menos- dos aspectos en la teoría de la interpretación radical. Por un lado, (i) encontramos el intento de construir una teoría composicional del significado para un lenguaje natural, que ha de tener la *forma* de una *teoría tarskiana de la verdad* para demostrar cómo podemos comprender expresiones semánticamente complejas sobre la base de que comprendemos las semánticamente primitivas y sus combinaciones. Por otro lado, (ii) encontramos el rol que tiene la adopción de la *perspectiva del intérprete* respecto al hablante; un intérprete cuya evidencia disponible excluye toda información acerca del significado de las expresiones lingüísticas del hablante y todo conocimiento acerca de los contenidos de sus creencias, deseos y demás actitudes proposicionales.

Ahora bien, en discusión con Rorty, Davidson (1998) afirma que, lejos de considerar que estas dos cuestiones deriven en modos antitéticos de filosofar (para Rorty (1997), tarskiano y wittgensteiniano, respectivamente), constituyen dos aspectos *interdependientes* de la misma empresa interpretativa: la adopción de la perspectiva del intérprete *requiere* –entre otras cosas- un marco provisto por las estructuras de la semántica formal.

A los efectos de exponer la teoría de la interpretación, me concentraré, pues, por un lado, en su inscripción en el programa semántico que hace del concepto de *condiciones veritativas* de una oración la clave para desarrollar las ideas de *significado* y *comprensión* y, por otro, en las *restricciones* surgidas a la luz de la consideración de la situación del *intérprete radical*. De la conjunción de estos dos aspectos, emerge, en mi opinión, la originalidad de la contribución davidsoniana, cuyos requisitos expondré, en primer lugar, para abocarme, luego, al objetivo y la naturaleza de la propuesta resultante.

## 2.1. El diálogo de Davidson con Tarski: el desafío del tránsito desde la verdad al significado

En la Introducción a *Inquiries into Truth and Interpretation* (1984), Davidson nos interpela con la siguiente pregunta: ¿qué es para las palabras significar lo que significan? A su juicio, tendríamos una respuesta a esta cuestión si supiéramos cómo construir una teoría que satisficiera dos requisitos. En este apartado me abocaré al primero de ellos, en tanto que haré lo propio con el segundo en la siguiente sección.

El **primer requisito** establece que la teoría en cuestión debe: proporcionar una interpretación de todas las emisiones reales y potenciales, de un hablante o grupo de hablantes. Vayamos, pues, por partes.

Tal como vimos, desde la perspectiva de Davidson, una teoría del significado que ignore el requisito de composicionalidad estaría soslayando algo central al concepto mismo de lenguaje –al menos, en tanto aprendible por parte de seres finitos-. Así, al formular esta condición, Davidson acepta sin ambages el principio de composicionalidad de Frege<sup>28</sup>, conforme al cual el significado de una expresión compleja es una función del significado de sus partes componentes. No obstante, para dar cuenta de la aprendibilidad del lenguaje no basta con este principio. Como hemos advertido más arriba, en tanto somos seres finitos, nos resulta imposible aprender de modo directo el significado de un conjunto infinito de expresiones. Se hace, pues, necesario establecer que tanto el número de elementos semánticos primitivos, como el número de reglas semánticas sean finitos.

Lo que Davidson (1967) se va a preguntar es si alguno de los candidatos fregeanos resulta plausible para ser considerado como el significado de las expresiones de un lenguaje. Por un lado, si se considera el significado como *Bedeutung* (denotación), la aplicación del principio de composicionalidad conduce al siguiente problema: si identificamos el significado de una proposición con un valor de verdad, es inevitable considerar las oraciones que tengan igual valor de verdad como sinónimas. Tendríamos, pues, la consecuencia contraintuitiva de que tener que aceptar como sinónimas, por ejemplo, “Barack Obama es el presidente de los Estados Unidos de América” y “Cristina Fernández de Kirchner es la presidenta de Argentina”

---

<sup>28</sup> Sigo en este punto el señalamiento de Caorsi (2001: 8). No obstante, como es sabido, Davidson encontrará insatisfactorias las consecuencias que se extraen de la aplicación de este principio en la teoría de Frege.

(formuladas en 2009). Por otro lado, el principio de composicionalidad resulta igualmente problemático cuando se caracteriza el significado en términos del *Sinn* (sentido) fregeano. En este caso las consecuencias no son ya, a juicio de Davidson, contraintuitivas, sino triviales. Así, por ejemplo, a propósito de “Teeteto vuela” la perspectiva fregeana sostendría que dado el significado de “Teeteto” como argumento, el significado de “vuela” proporciona el significado de “Teeteto vuela” como valor. En su opinión, la vacuidad de la respuesta resulta obvia.

Frente a este estado de la cuestión, Davidson ensaya la hipótesis de que tal vez la dificultad se haya seguido de identificar el significado con alguna entidad particular, lo que ameritaría explorar una tercera estrategia que no pase por la consideración de cuál sea el estatus del significado. Supongamos, entonces, que el significado de las palabras es lo que establece un diccionario: no tendríamos el significado de las oraciones, pero sí el de sus partes componentes. Supongamos, adicionalmente, que poseemos una sintaxis recursiva para nuestro lenguaje L. ¿No podríamos en este caso conjeturar que si añadimos a esta sintaxis un diccionario que nos dé el significado de los elementos atómicos, tendríamos la teoría semántica buscada? No, en opinión de Davidson: “el conocimiento de las características estructurales que hacen significativa a la oración, más el conocimiento de los significados de las partes últimas no darán como resultado el conocimiento de lo que significa una oración” (1967/1984:21). Las oraciones de creencia son reveladoras de esta dificultad: su sintaxis, nos dice Davidson, es sencilla, pero el agregado de un diccionario no modifica el problema semántico consistente en que ni siquiera podemos dar cuenta de las condiciones de verdad de tales oraciones partiendo de nuestro conocimiento del significado de las palabras que contienen. En consecuencia, una sintaxis recursiva sumada a un diccionario no es necesariamente semántica recursiva.

Cerrados estos caminos, Davidson explicita su compromiso con una visión holista del significado, lo que lo llevará, tal como señala Caorsi (2001: 14-15), a conferirle otro papel al principio de composicionalidad:

Si el significado de las oraciones depende de su estructura y entendemos el significado de cada ítem de la estructura sólo como una abstracción de la totalidad de las oraciones en las que éste opera, entonces podemos dar el significado de cualquier oración (o palabra) solamente dando el significado de todas las oraciones y palabras del lenguaje. Frege dijo que una palabra tiene significado en el contexto de una oración; siguiendo esta idea debería haber agregado que sólo en el contexto del lenguaje una oración (y por lo tanto una palabra) tiene significado (1967/1984:22).

*Prima facie*, pareciera que el modo de poner en práctica el principio de composicionalidad sería partir del significado de las palabras para arribar al significado de las oraciones. Tal es la estrategia que Davidson denomina “teoría del bloque constructivo”. No obstante, una vez que se abraza una perspectiva de cuño holista tal como la que hemos consignado mediante la cita de arriba, parece que no tenemos modo de acceder al significado de las palabras, si no accedemos primero al significado de las oraciones. El punto de partida, en consecuencia, ha de ser otro. Dado que, como es obvio, no podemos partir de la totalidad de las oraciones del lenguaje en virtud de que éstas constituyen un conjunto infinito, hemos de hacerlo desde las oraciones particulares. Los rasgos semánticos de las palabras se van a abstraer de los rasgos semánticos de las oraciones. Ahora bien, las oraciones son aquello a lo que tenemos acceso en su contexto de uso: los hablantes de una lengua regularmente no emplean palabras aisladas, sino oraciones. Aquello que va a permitirnos el acceso al significado serán las emisiones de oraciones por un hablante en un tiempo dado. Lo que diferencia a esta estrategia de una teoría del bloque constructivo es que las palabras y su referencia han de tener el estatus de constructos hipotéticos, cuya finalidad es explicar el significado de los enunciados, pero la confrontación de la teoría con las situaciones de uso sólo podrá efectuarse mediante las preferencias de oraciones. Esta distinción tiene su correlato en la manera en que Davidson deslinda entre la explicación *dentro* de la teoría y la explicación *de* la teoría. En el primer caso, los rasgos semánticos de una oración se especifican vía la estructura postulada y conceptos semánticos como el de referencia; en el segundo caso, de lo que se trata es de interpretar la teoría como un todo, en cuyo caso partiremos de los rasgos semánticos de las oraciones que han de ser conectados con los fines y actividades humanas.

¿Cuál tiene que ser, pues, la forma canónica de los enunciados de la teoría que han de proveer una interpretación de las oraciones del lenguaje? Si nos ceñimos al requisito de que una teoría del significado debe dar el significado de toda oración *s* del lenguaje, una propuesta posible sería (1967/1984:22):

(1) *s* significa *m*

donde “*s*” es el nombre en el metalenguaje de la oración del lenguaje objeto que se pretende interpretar y “*m*” el nombre, también en el metalenguaje, del significado de dicha oración.

Ahora bien, como ya adelantamos, la apelación a los significados como entidades conduce o bien a resultados contraintuitivos o bien a pseudo explicaciones. De esta suerte, esta vía no parece promisorio<sup>29</sup>. Pareciera, pues, que una opción plausible sería desembarazar a los teoremas de la teoría semántica de cualquier tipo de expresión que denote significados. Una manera de efectuarlo sería reemplazar a “m” por una oración en uso en el metalenguaje:

(2) *s* significa que *p*.

En este caso, nos topamos con otra dificultad: al tratar con la lógica del aparentemente no extensional “significa que” nos encontramos con problemas tan duros como los que la teoría pretende resolver (1967/1984:22).

¿En qué consiste, entonces, la solución que propone Davidson? A su juicio, el objetivo de una teoría del significado consiste en asignar a cada oración a interpretar una oración que nos dé su significado. La clave del problema residirá, pues, tal como advierte Caorsi (2001:18) en el emparejamiento de las oraciones adecuadas. En esta línea, Davidson propone la eliminación de la sospechosa expresión “significa que” en vistas de que dado su carácter no extensional, introduce a “p” en un contexto no extensional. De acuerdo con ello, la propuesta de Davidson estriba en *tratar extensionalmente la posición ocupada por “p”*:

La teoría habrá cumplido su objetivo si provee para cada oración *s* del lenguaje estudiado, una oración correspondiente (para reemplazar a “p”) que, de una manera por aclarar “dé el significado” de *s*. Un candidato obvio para esa oración es precisamente la misma *s*, si el lenguaje objeto está contenido en el metalenguaje. Si no es ése el caso, una traducción de *s* en el metalenguaje. Como osado paso final intentemos tratar extensionalmente la posición ocupada por “p”: para implementar esto, bórrese el oscuro “significa que”, provéase la oración que reemplaza a “p” con una conectiva oracional adecuada y suminístrese la descripción que reemplaza a “s” con su propio predicado. El resultado plausible es:

(T) *s* es T si y sólo si *p* (1967/1984:23).

Efectuadas estas transformaciones, “lo que le pedimos a una teoría del significado para un lenguaje L es que sin apelar a ninguna noción semántica imponga las restricciones suficientes al predicado “es T” como para que de ella se deriven todas las oraciones obtenidas del esquema T al reemplazar “s” por una descripción estructural de una oración de L y “p” por esa misma oración” (1967/1984:23).

---

<sup>29</sup>En García Suárez y Valdés Villanueva (1982) se enfatiza que, no obstante, no son los viejos prejuicios nominalistas respecto de las entidades abstractas, ni las quejas de Quine respecto de la oscuridad de sus condiciones de identidad, lo que lleva a Davidson a rechazar los significados, sino que- tal como sostiene en su (1967: 20-1)- “carecen de una utilidad demostrada”.

Conforme a ello, una teoría del significado va a ser una teoría que permita definir un predicado “T” que satisfaga las exigencias establecidas. Lo que la teoría haría sería, pues, fijar la extensión del predicado “T” (cuál sea el predicado no interesa si cumple con las condiciones establecidas y posee, en consecuencia, la misma extensión que T).

Habiendo llegado a esta instancia, resulta claro que el esquema que se ha establecido para la teoría semántica coincide exactamente con el esquema descrito en la Convención T de Tarski<sup>30</sup>. Esto es, el predicado “es verdadero” de Tarski posee exactamente la misma extensión que T, si adoptamos el supuesto de que la extensión de esos predicados está limitada a las oraciones de *L*:

Dos predicados cualesquiera que satisfagan esta condición *tienen la misma extensión*, luego si el metalenguaje es lo suficientemente rico, nada impide poner lo que yo llamo una teoría del significado bajo la forma de una definición explícita de un predicado “es T”. Pero (...) es claro que, las oraciones a las que se aplica el predicado “es T” serán justamente las oraciones verdaderas de *L*, puesto que la condición que hemos requerido para las teorías del significado satisfactorias es en esencia la convención T de Tarski, que pone a prueba la adecuación de una definición formal de verdad (1967/1984:23).

De este modo, Davidson arriba a la conclusión que constituye el núcleo de su programa: una teoría del significado para un lenguaje *L* muestra cómo los significados de las oraciones dependen de los significados de las palabras (y satisface, así, el requisito de composicionalidad) si contiene una definición (recursiva) de verdadero-en-*L*.

---

<sup>30</sup> Lepore y Ludwig (2006) resumen el punto de la manera siguiente. La Convención T requiere que la teoría de la verdad sea formalmente correcta y que tenga entre sus teoremas todas las oraciones de la forma (T) en las cuales “*s*” es reemplazada por una descripción de una oración del lenguaje en estudio, el lenguaje-objeto, construida a partir de sus partes significativas y “*p*” es reemplazada por una traducción de *s* en el metalenguaje, y “es T” está por el predicado de verdad del lenguaje-objeto.

(T) *s* es T sii *p*

La idea es que si sabemos que una teoría de la verdad satisface la Convención T, y que es consistente, sabemos que “es T” tiene la extensión del predicado veritativo para el lenguaje objeto. Si sabemos que la teoría satisface la Convención T y podemos seleccionar los teoremas adecuados, podemos reemplazar “es T sii” con “significa que”, para arribar al enunciado explícito del significado de *s*:

(M) *s* significa que *p*.

Esto es así porque la condición para que (M) sea verdadero es que “*p*” traduzca a “*s*”.

Recapitulemos, entonces, el derrotero davidsoniano<sup>31</sup>. Tenemos una teoría del lenguaje L que nos permite atribuir una propiedad semántica –“es T” – a cada oración en virtud de la manera en que esa oración se compone, mediante un número finito de aplicaciones de un número finito de modos de composición, a partir de un stock finito de elementos léxicos. A los efectos de que dicha teoría cuente como una teoría satisfactoria del significado se requiere que, sin el auxilio de otras nociones semánticas, imponga restricciones al predicado “es T” de manera que entrañe para cada oración de L, un teorema de la forma: (T)  $s$  es T sii  $p$ , donde  $s$  y  $p$  se reemplazan de la manera descrita. Una teoría así constituiría una definición recursiva del predicado “es T”. Davidson advierte entonces que el predicado “es T” es coextensivo con el predicado “es verdadero”. En consecuencia, las oraciones a las que se aplica ese predicado son, precisamente, las oraciones verdaderas de L, dado que la condición impuesta es la Convención T de Tarski, que rige la adecuación material de una definición semántica de verdad<sup>32</sup>.

En suma, en “Truth and Meaning” Davidson ofrece su defensa del argumento de que una teoría de la verdad, modificada para su aplicación a un lenguaje natural, puede usarse como teoría de la interpretación. Lo que allí propone, pues, de manera muy novedosa, es *invertir la dirección de la explicación tarskiana*: esto es, suponiendo la traducción, Tarski estaba en condiciones de definir la verdad; la idea de Davidson consiste en tomar a la verdad como básica y extraer una explicación de la interpretación. Su esperanza es que si tomamos las restricciones formales y empíricas de la teoría como un todo, las oraciones T individuales servirán para producir interpretaciones.

En efecto, para adaptarse a la exigencia de proveer una interpretación de las oraciones de un lenguaje L, las teorías T deben ser *empíricas*, lo que nos remite a las condiciones de aplicación de las mismas y, por tanto, al segundo requisito que les impone Davidson.

---

<sup>31</sup> Sigo las líneas directrices de la reconstrucción de Caorsi (2001, cap.2).

<sup>32</sup> Cf. García Suárez y Valdés Villanueva (1982).

## 2.2. El diálogo de Davidson con Quine: desde la traducción a la interpretación

En la Introducción de *Inquiries into Truth and Interpretation* (1984) Davidson plantea un **segundo requisito** que ha de cumplir una teoría que se proponga operar como teoría del significado: debe ser verificable sin tener conocimiento de las actitudes proposicionales detalladas del hablante.

Al plantearnos la cuestión de su verificación empírica pareciera que lo que tenemos que hacer es, entonces, apelar a los métodos estándar de confirmación de teorías empíricas. Estos métodos, como es sabido, imponen la comparación de algunas de las consecuencias de la teoría con el mundo. Ahora bien, las consecuencias contrastables de la teoría que estamos considerando son las oraciones  $T$  que constituyen teoremas de la teoría y que, según hemos visto, han de tener la forma:  $(T) s \text{ es } T \leftrightarrow p$ . La Convención  $T$  establecía que las relaciones entre  $s$  y  $p$  podían ser de dos tipos: (i)  $p$  es la misma oración denotada por  $s$ , en caso de que el lenguaje objeto esté incluido en el metalenguaje o (ii)  $p$  es una traducción de  $s$ , en caso de que el metalenguaje no incluya al lenguaje objeto. Dicha condición le aseguraba a Tarski que el *significado* de  $s$  y  $p$  sería el mismo. En vistas de ello, quedan excluidas de la Convención  $T$  oraciones tales como (S) “La nieve es blanca” es verdadera sii la hierba es verde. Dada la dirección de la explicación tarskiana que mencionamos más arriba, Tarski podía imponer estas restricciones a  $p$  en virtud de que su objetivo era definir la verdad y a tal efecto podía recurrir legítimamente a la noción de “igual significado”. Ahora bien, dada la dirección de la búsqueda davidsoniana, tal expediente no resulta lícito.

Una oración  $T$  establece que una determinada oración del lenguaje objeto es verdadera si y sólo si se cumplen ciertas condiciones de verdad. De este modo, lo que hace la oración  $T$  es establecer esas condiciones de verdad. Ahora bien, ¿cómo se determinan las condiciones de verdad de un enunciado emitido por un hablante? Como hemos anticipado, de lo que se trata es de dar una fundamentación de la teoría de manera no circular, esto es, de evitar la introducción solapada del concepto de significado y nociones afines en los cimientos de la teoría. A tal efecto, Davidson apela al *punto de vista del intérprete radical*, sugerencia inspirada por el enfoque de Quine para el estudio filosófico del lenguaje: el proyecto de la traducción radical. Así, lo que Davidson denomina “interpretación radical” es la interpretación de las emisiones de un hablante

de un lenguaje totalmente desconocido para nosotros. No obstante, hay diferencias relevantes entre ambos proyectos, en las que resulta conveniente detenernos (cf. Lepore y Ludwig 2006: 14-15).

En primer lugar, Quine concibió su reflexión en torno a la *traducción radical* como una manera de mostrar cómo construir un concepto de significado científicamente respetable, que es caracterizado en términos de disposiciones a la conducta verbal en respuesta a patrones de estímulos físicos. En contraste, Davidson concibe el proyecto de la *interpretación radical* como un modo de iluminar la noción ordinaria de significado (“la tarea de una teoría del significado como la concibo, no es cambiar, mejorar o reformar un lenguaje, sino describirlo y comprenderlo”; 1967/1984: 29), y la clave para la interpretación del hablante ha de residir, no en los estímulos proximales, sino en los eventos distales del entorno que causan que el hablante sostenga como verdadera una oración.

A su vez, mientras que el objetivo de un traductor radical es producir un manual de traducción, el objetivo de un intérprete radical es producir una teoría de la interpretación. Conforme a Davidson “...un manual de traducción sólo es un método para pasar de oraciones de un lenguaje a oraciones de otro, y de él nada podemos inferir acerca de las relaciones entre palabras y objetos” (1977/1984:221). Una teoría de la traducción supone: el lenguaje traducido (lenguaje objeto, LO), el lenguaje al que se traduce (lenguaje sujeto, LS) y el lenguaje en el que se traduce (metalenguaje) y establece una relación de mapeo de las oraciones de LO sobre las de LS. De esta suerte, mientras que una teoría de la traducción nos provee de una relación entre palabras de LO y LS, una teoría de la interpretación ha de establecer una relación entre palabras de LO y el mundo. Según hemos visto, desde la propuesta de Davidson dicha teoría va a consistir en una teoría de la verdad para el lenguaje del hablante, cuya verificación empírica requiere la adopción de la perspectiva del intérprete radical, punto de vista en cuya consideración pasamos a detenernos.

Lo que Davidson denomina “interpretación radical” es la interpretación de las emisiones de un hablante de un lenguaje totalmente desconocido para nosotros. El intérprete radical supone, entonces, que está frente a una actividad lingüística que le es imposible, por mera observación, entender. En tal caso –el conceptualmente interesante- su único elemento de juicio serán grupos de emisiones, *prima facie* oracionales, que sean *sostenidas como verdaderas* [*hold true*] por los presuntos hablantes. Inicialmente, entonces, el intérprete debe descubrir la actitud

del hablante hacia sus propias emisiones, esto es, si sostiene como verdadera una oración –o no– en circunstancias particulares. De este modo, si no asumimos al menos provisionalmente que al emitir un sonido el sujeto está profiriendo una oración y que al hacerlo pretende enunciar algo que sostiene como verdadero en las circunstancias en las que realiza dicha preferencia, no podemos siquiera intentar una tarea de interpretación.

Ahora bien, tal como señalan Amoretti y Vassallo (2008:11), sostener como verdadera una oración es ya una actitud semántica<sup>33</sup>, pero de acuerdo con Davidson se trata de una actitud que puede preceder a la interpretación. Por ejemplo, el intérprete puede saber que un sujeto sostiene como verdadera la oración “Sta piovendo” sin haber reconocido qué verdad específica entraña. Así, aún cuando el intérprete sabe que el hablante sostiene una oración como verdadera, todavía no es capaz de determinar lo que el hablante cree ni lo que significa su emisión efectiva. Imaginemos que durante una tormenta el hablante profiere: “Sta piovendo”. A los efectos de interpretar esta emisión y averiguar que significa “Está lloviendo”, el intérprete debe ser capaz de adscribirle al hablante la creencia de que está lloviendo; sin embargo, para atribuirle dicha creencia, el intérprete debe conocer el significado de sus palabras. En otros términos: supongamos que el intérprete sabe que el hablante sostiene una cierta oración como verdadera; si supiera el significado de la oración, podría establecer lo que el hablante cree y, conversamente, si supiera lo que el hablante cree, podría determinar el significado de la oración proferida. De esta suerte, la interdependencia de creencia y significado es inextricable. En consecuencia, resulta crucial encontrar una manera de romper el círculo que los liga. Según Davidson, podemos lograrlo definiendo una teoría general de la interpretación que pueda ofrecernos “simultáneamente una teoría de la creencia y una teoría del significado” (1974b/1984: 144) y que nos suministre “un método para mantener constante un factor mientras estudiamos el otro” (1975/1984: 167): la teoría en cuestión es, precisamente, la de la interpretación radical, en tanto que el método aludido es suministrado por un principio que Davidson toma de Quine: el *principio de caridad*.

---

<sup>33</sup> Creo que es en virtud de este carácter que Moretti (2008) califica a este método –consistente en recurrir a oraciones que los hablantes sostienen como verdaderas (y de las que cabe conjeturar que son de bajo nivel teórico)– como “penoso”, aunque concede que, pese a todo, es el expediente que posibilita la conexión entre significatividad y creencias. Y la idea es que si dicha conexión se corta, se plantea una dificultad relativa a la base empírica de la teoría: en efecto, si no existe la posibilidad de evaluarla según la razonabilidad del sistema de actitudes proposicionales que adjudica a los hablantes, su carácter empírico podría pasar a depender de un “mal mayor”, a saber: que deba contarse con un conjunto previo de intuiciones semánticas acerca de implicaciones y sinonimias.

En suma, la *evidencia disponible* consiste en que los hablantes del lenguaje a ser interpretado *sostienen como verdaderas* diversas oraciones en ciertos tiempos y bajo circunstancias especificadas. El problema ha de ser, pues, ¿cómo puede emplearse esta evidencia para respaldar una teoría del significado?

Tomemos un ejemplo provisto por el propio Davidson:

(T) “Es regnet” es verdadero en alemán cuando es proferido por  $x$  en el tiempo  $t$  si y sólo si está lloviendo cerca de  $x$  en  $t$ .

¿Qué podría contar como evidencia a favor de esta oración-T? Según hemos visto, podemos suponer que si un hablante de alemán emite “Es regnet” en un momento determinado, sostiene como verdadera esa oración en esas circunstancias. Aún no sabemos lo que significa la oración. No obstante, si consideramos que los aspectos relevantes de la situación de emisión poseen carácter objetivo, esto es, que son los mismos para nosotros que para el hablante alemán, vamos a estar en condiciones de dar un paso más en la interpretación de su emisión. Hasta aquí tenemos, entonces, que al emitir “Es regnet” el hablante alemán profiere una oración que sostiene como verdadera y que las condiciones determinantes de la preferencia nos resultan accesibles a ambos. En este caso, pareciera que el siguiente enunciado puede proveer evidencia a favor de (T):

(E) Kurt pertenece a la comunidad de habla alemana y Kurt sostiene como verdadera “Es regnet” el sábado al mediodía y está lloviendo cerca de Kurt el sábado al mediodía.

Si vamos a suponer que (T) provee una interpretación del significado de “Es regnet” en alemán tenemos que suponer que debe ser verdadero para todo hablante alemán ( $x$ ) y todo tiempo ( $t$ ). Así, pues, tenemos que interpretar que la evidencia a favor de (T) tiene que ser tal que respalde la generalización de (E). Esto es:

(GE) ( $x$ ) ( $t$ ) Si  $x$  pertenece a la comunidad de habla alemana,  $x$  sostiene que es verdadero “Es regnet” en  $t$  si y sólo si está lloviendo cerca de  $x$  en  $t$ .

Ahora bien, hemos anticipado que no podemos acceder al significado de un hablante sin conocer sus creencias y, conversamente, no podemos acceder a sus creencias sin conocer el significado de sus palabras. ¿En qué sentido, pues, la apelación al *principio de caridad* permite

sustraernos a este círculo? En otras palabras, ¿cómo es que dicho principio resolvería el problema de la interdependencia de creencia y significado tomando a la creencia como constante en tanto se resuelve el significado?

Este principio que se debe a Neil Wilson ya había sido empleado por Quine en *Word and Object* para la situación de traducción radical. Como veremos, Davidson, a su vez, lo toma de Quine, pero hará un empleo mucho más amplio del mismo<sup>34</sup>.

Recordemos, en primera instancia, qué se intenta comprender<sup>35</sup>: lo expresado lingüísticamente por alguien y, con eso, comprender algo acerca del sujeto que se ha expresado.

La descripción de más bajo nivel del fenómeno de la clase que tendrá que estudiarse será algo como “Esa mujer (sea H) emitió los sonidos tales y cuales”. La primera hipótesis que ha de formular el intérprete (I) es que H emitió una expresión lingüística y se van a privilegiar los casos de emisiones favorables a la hipótesis más específica: H emitió la *oración* S.

En virtud de que el intérprete, en caso de que expresara su comprensión de S, lo haría en su propio lenguaje (cuya sintaxis le es conocida), continuará mediante una presunción acerca de la forma de S elaborada a partir de sus conocimientos gramaticales. Todas las hipótesis sintácticas que I haga acerca de las expresiones de H determinarán una teoría sintáctica sobre el lenguaje de H (sea £) cuya validez empírica va a depender de que la teoría semántica (sobre £) que la utilice resulte empíricamente justificable. Hasta aquí, tenemos que H profirió la oración S, ahora bien ¿qué tipo de acto de habla realizó H? y ¿qué contenido tuvo ese acto? El intérprete empleará alguna presunción sobre fuerzas ilocucionarias para conjeturar, por ejemplo, que H *afirmó* D, siendo D una oración declarativa de £ que, supuestamente, exhibe el contenido de S. Resta interpretar D, darle significado, a los efectos de esclarecer ese contenido.

Según hemos visto en el apartado anterior, Donald Davidson propuso que la forma adecuada de dar significado a D tiene su clave en la generación de un teorema de la forma “D es verdadera si y sólo si p”, un tipo de teorema que Tarski nos ha enseñado de qué modo generar. Llegado a este punto, sobre la base de que H afirmó D y D es verdadera si y sólo si p, el

---

<sup>34</sup> Quine hace un uso del principio de un modo restringido a la determinación de las conectivas lógicas oracionales, en tanto que Davidson lo aplica, para usar sus propios términos, del principio al final. Asimismo, Amoretti y Vassallo (2008) señalan un punto que, en mi opinión, es digno de análisis: sostienen que mientras que en Quine el principio posee sólo estatus pragmático, en Davidson dicho estatus es constitutivo, puesto que para este último, no es una opción para la interpretación: sin él la interpretación no sería posible. Volveré sobre esta cuestión en el capítulo próximo, en el que voy a detenerme a considerar los rasgos de la noción de creencia.

<sup>35</sup> Sigo en este punto la sugerencia de Moretti (2008, cap.1).

intérprete puede interpretar la preferencia de S del modo siguiente: H *afirmó* que p y H *cree* que p.<sup>36</sup>

Ahora bien, estas últimas conclusiones dependen, por un lado, de que el análisis ofrecido para S, satisfaga una restricción *holista*: esto es, que pueda integrarse en una teoría que provea análisis para *un número suficientemente amplio de oraciones de £*<sup>37</sup>. En este punto parece tallar una restricción central de la empresa interpretativa: *el principio de racionalidad* entendido, fundamentalmente, en términos de consistencia lógica entre las actitudes proposicionales atribuidas a los hablantes. Tal como consideraremos en más detalle en el capítulo próximo, la identidad del contenido proposicional de cada actitud depende de sus relaciones (especialmente, lógicas) con los restantes contenidos. Según veremos, el examen de la actividad lingüística proporciona, a juicio de Davidson, un criterio muy importante para la determinación de las características de nuestra racionalidad. Un paso fundamental de ese examen consiste en dilucidar la estructura de nuestro lenguaje. Un factor esencial para esto, y el más ligado con la idea de racionalidad, es la aclaración de los principios inferenciales involucrados. Y el núcleo mínimo de tales principios, también el más relacionado con la idea de lenguaje, está dado por las reglas inferenciales deductivas, reglas que caracterizan una idea de coherencia entendida como ideal regulativo para la aceptabilidad y constituibilidad de conjuntos de creencias. De lo que se trata, entonces, es que la atribución de actitudes proposicionales a los hablantes de £ alcance la magnitud adecuada como para constituir una base empírica que confirme la teoría y, así, obtengamos la interpretación de D, interpretación que, como vemos, no quedará en absoluto confinada al bicondicional tarskiano correspondiente.

---

<sup>36</sup> A la vera de la teoría de la interpretación radical como un expediente que nos provee simultáneamente de una teoría del significado y una teoría de la creencia, la aserción y la creencia parecen ser dos caras de la misma moneda.

<sup>37</sup> Como advierte Moretti (2008) este holismo no significa que la teoría afirme la desmesurada tesis de que no se es competente en el manejo de una expresión si no se llegó a conocer el significado de las infinitas oraciones de las que puede formar parte (al cabo, el significado de todas las posibles expresiones del lenguaje); como el propio Davidson admite, hay relaciones que resultan más relevantes que otras a la hora de determinar un cierto contenido proposicional. Asimismo, Moretti señala que admitir que la teoría propuesta debe ser compatible con una explicación sensata de la aprendibilidad del lenguaje y de la comunicación exitosa no está en pugna con una teoría holista que no necesita atribuir al hablante competente un conocimiento completo del significado de ninguna expresión, y que tampoco implica sostener que todas las relaciones que una expresión tiene con todas las demás contribuyan con igual importancia a la determinación de la función semántica. Volveré sobre esta restricción holista en el capítulo próximo.

En resumen: hemos supuesto que estamos frente a una actividad presumiblemente lingüística que nos es imposible, por mera observación, entender. En tal caso nuestro único elemento de juicio serán grupos de emisiones, *prima facie* oracionales, que sean *sostenidas como verdaderas* por los presuntos hablantes. Y a fin de poder atribuir significado a esas emisiones, deberemos atribuir simultáneamente muchas creencias a los hablantes (requisito holista). Ahora bien, según la interdependencia ya aludida, atribuir creencias, interpretar emisiones, y caracterizar como lingüística cierta actividad, no son tareas que puedan separarse; por tanto, sólo es posible comenzar la construcción de la teoría interpretativa, suponiendo que se comparte con los presuntos hablantes un vasto conjunto de creencias.

Aquí es donde llegamos al empleo que Davidson hace del principio que toma de Quine. Para interpretar algo como lenguaje, es inexorable proceder conforme al *principio de caridad*: esto es, (i) considerar como *verdaderas* un *vasto número*<sup>38</sup> de oraciones que los hablantes del lenguaje que se está interpretando sostienen como verdaderas y (ii) asignar *condiciones veritativas comunes* a oraciones del lenguaje desconocido y del propio lenguaje. Si suponemos un repertorio de creencias básicas compartidas, podemos conjeturar que respecto de determinadas situaciones las creencias del hablante y las nuestras van a ser las mismas y, de este modo, nuestras creencias han de constituir el “puente” hacia la interpretación de las emisiones del hablante extranjero. En consecuencia, el intérprete nunca puede estar en posición de creer que los hablantes que interpreta tienen creencias radicalmente diferentes de las suyas.

Sólo entonces, esto es, cuando una caracterización tarskiana de las condiciones veritativas - del tipo que hemos expuesto en el apartado anterior - es puesta bajo las restricciones que se derivan de considerar la situación del intérprete radical, permite adscribir un conjunto razonable de actitudes proposicionales a los hablantes de un lenguaje particular y puede considerarse justificada la sustitución de “es verdadera si y sólo si” por “significa que” en los teoremas de la forma “D es verdadera si y sólo si p”. Llegada a este punto, una teoría de esa forma da todos los significados oracionales. Con lo cual, quedaría resuelto el problema que dio origen a la pregunta por la naturaleza del significado oracional.

---

<sup>38</sup> Para una defensa de una relación inseparable entre la caridad y el holismo davidsonianos puede verse Malpas, J. (2007). Asimismo, para una crítica respecto de la pretensión de derivar el holismo a partir del principio de caridad, véase Fodor y Lepore (1992). No obstante, en el capítulo próximo volveré sobre las relaciones que guardan entre sí estas restricciones de la empresa interpretativa.

Vemos, ahora, más detenidamente cuál es el objetivo y la naturaleza de la contribución davidsoniana.

### 2.3. ¿Comprender la comprensión?

Entre los objetivos que pretenden justificar el empleo de la “semántica de los lógicos” para estudiar el lenguaje natural, Moretti (2008) identifica los siguientes.

Por un lado (1), discriminar las categorías básicas de la más clara y simple teoría (o teorías) acerca del mundo que en nuestro lenguaje se ofrece (n) en un momento dado.

Por otro (2), contribuir a explicar aspectos centrales de nuestro comportamiento lingüístico o de nuestro lenguaje, en particular la idea de significación o la de comprensión lingüística. A su vez, al interior de esta empresa distingue los dos propósitos siguientes: (2.1) predecir o explicar la conducta verbal y (2.2) explicar la comprensión del hablante, pudiendo con esto referirse al modo como le es posible a un intérprete cualquiera caracterizar la comprensión que le atribuye al hablante o bien a las características de la comprensión que el hablante posee o que se atribuye a sí mismo. Ambos objetivos, sostiene, difieren de (2.3) especificar formalmente las relaciones de significación que vinculan entre sí a los símbolos que intervienen en la comunicación verbal. Asimismo, Moretti señala -en una escueta e inquietante nota- que “hay un proyecto conexo que, para muchos, suena más estimulante: “comprender –como distinto de explicar y predecir- la comprensión lingüística (huir del externalista Hempel hacia el hondo Dilthey)”.<sup>39</sup>

Cuando el objetivo es del tipo (2), Moretti sostiene que se obtiene una justificación para aplicar las técnicas lógicas al lenguaje natural a partir de la idea de que parte principal del análisis del lenguaje y la comprensión lingüística reside en la determinación de las relaciones entre significados lingüísticos (la estructura semántica). Ahora bien, el aspecto de acceso más directo del funcionamiento del lenguaje y la comunicación está dado por -como vimos que sostiene Davidson- las emisiones de oraciones y las relaciones que guardan entre sí. Y las

---

<sup>39</sup> Moretti remite a H. Schneider (1979).

relaciones más importantes entre oraciones son las relaciones inferenciales (su estructura lógica), que pueden verse como restricciones a las condiciones veritativas en conjuntos de oraciones. Por tanto, desde esta perspectiva, resulta adecuado intentar satisfacer el propósito 2 mediante la aclaración de la estructura lógica del lenguaje<sup>40</sup>.

Otra consideración conduce cerca de la semántica formal cuando el fenómeno que se intenta aclarar es el de la comprensión lingüística, pero sin prejuzgar sobre el papel que en esta tarea cabe a la noción de significado y demás conceptos emparentados; esto es, cuando se evitan los giros *intensionales* en tanto su esquema de comprensión es, como sostiene Davidson, parte del problema general que debe resolverse. Desde un punto de vista filosófico, nos recuerda Moretti, lo que entonces interesa no es interpretar un lenguaje particular; esto es, lo que filosóficamente importa es discutir *cuál sea la forma general de una teoría de la comprensión lingüística* para cada lenguaje particular.

Desde esta búsqueda, la tarea de una teoría semántica será caracterizar a través de un conjunto de especificaciones que sea finitamente determinable y empíricamente revisable, lo siguiente<sup>41</sup>: (a) los significados de las oraciones, (b) los significados de las expresiones no oracionales que componen oraciones, (c) las relaciones entre (a) y (b), (d) las relaciones entre los significados de las oraciones, con la implicación como paradigma y (e) las relaciones entre los significados suboracionales.

Con estos *desiderata*, el modelo de dicha teoría se ha de buscar en la semántica tarskiana que, según hemos visto, consiste básicamente en la caracterización de un predicado veritativo adecuado al lenguaje en estudio. Recordemos que las construcciones formales tarskianas fueron usadas, canónicamente, para dotar de significado a sistemas puramente sintácticos, a los efectos de asemejarlos apropiadamente –para fines lógicos- a los lenguajes naturales. De esta “semilla”, nos dice Moretti, seguramente creció la semántica davidsoniana para los lenguajes naturales: esto es, al poner entre paréntesis la intelección de un lenguaje cualquiera, tal como tiene lugar al adoptarse la perspectiva del intérprete radical, “buena parte de lo que queda es un sistema de conjeturas sintácticas, y los procedimientos debidos a Tarski son una interesante promesa de restitución de lo que falta” (Moretti 1996: 56).

---

<sup>40</sup> Moretti advierte que en este punto da la impresión de que la teoría davidsoniana produce un deslizamiento insuficientemente justificado desde el –supuestamente exitoso- análisis lógico del lenguaje motivado por preocupaciones onto gnoseológicas, esto es, cuando el objetivo es de tipo (1), hacia su estudio mediante procedimientos similares (la “semántica de los lógicos”) pero ahora con fines lingüísticos.

<sup>41</sup> Cf. Moretti (2008, cap.1).

Tal es, en mi opinión, el marco en el que hay que inscribir la teoría de la interpretación radical de Donald Davidson, esto es, como una propuesta semántica del tipo de *condiciones veritativas*, que ha de tener a la *teoría tarskiana* de la verdad como *forma o estructura básica*<sup>42</sup> a la que, no obstante, se ha de complementar con elementos ajenos a ella: las peculiares *restricciones* que impone la adopción de la perspectiva del intérprete radical para su aplicabilidad empírica. En consecuencia, cabe enfatizar, que la propuesta de Davidson no radica en identificar, lisa y llanamente, el significado de una oración con sus condiciones de verdad: éstas son un aspecto central de aquél, pero están lejos de agotarlo. Me interesa señalar, entonces, que las singulares restricciones que se derivan de la consideración de la adopción de la perspectiva del intérprete radical se encuentran en el corazón mismo de la contribución davidsoniana<sup>43</sup> y constituyen el aspecto que, en mi opinión, hace que en lugar de concebirla meramente como un intento de consecución del objetivo 2.2 (explicar la comprensión del hablante) - que es aquel en el que estaríamos inclinados a inscribirla *prima facie*<sup>44</sup>-, tal vez amerite que la ubiquemos en el seno del proyecto conexo que menciona Moretti, esto es: el de *comprender* - como diferente de explicar y predecir- la comprensión lingüística<sup>45</sup>.

---

<sup>42</sup> Sólo, pues, como producto secundario se ofrece, para un lenguaje particular, un modelo al que tendría que ser empíricamente equivalente cualquier teoría sobre ese lenguaje particular.

<sup>43</sup> En consonancia con esta indicación, Quintanilla (1995) concibe el proyecto de Davidson a la vera de las dos tradiciones identificadas por Strawson (1983), a saber, los teóricos de la semántica formal y los teóricos de la intención comunicativa. Desde esta perspectiva, por un lado, Davidson transita el camino abierto por Frege, Tarski, Carnap y Quine, en tanto abreva en la aplicación de la semántica formal a los lenguajes naturales y utiliza como instrumento principal el análisis de las condiciones de verdad con la finalidad de proporcionar el significado de la oración. Por otro lado, adhiere a aquella tradición centrada en las relaciones entre hablantes e intérpretes y los actos de habla que ellos pueden realizar, tradición que dirige su atención a las nociones de comunicación e intención y que fue desarrollada inicialmente por el segundo Wittgenstein, Austin y Grice.

<sup>44</sup> Soy consciente de que hay un sentido claro en el que se pretende que la teoría de la interpretación *explique* la comprensión lingüística. En efecto, como es sabido, el propio Davidson nos dice que desea tomar en serio la idea de “teoría” en “teoría del significado” y tal es la promesa que encuentra en la teoría tarskiana de la verdad. Más aún, nos advierte que en la medida en que las teorías de la interpretación son teorías empíricas de la verdad, sus axiomas y teoremas han de considerarse *leyes naturales*, punto que discute en “Reply to Foster”. No obstante, a mi juicio, esta empresa- relativa a lenguajes particulares- se da sólo dentro del marco que resulta relevante discutir desde el punto de vista filosófico, que no es ya la cuestión de interpretar un lenguaje particular, sino determinar cuál sea *la forma general de una teoría de la comprensión lingüística* para cada lenguaje particular. Esta última empresa es la que habilita, en mi opinión, la posibilidad de cifrar el objetivo de su contribución semántica en la *comprensión* - como distinta de la explicación y la predicción- de la comprensión. En un espíritu afín, Devitt sostiene que la perspectiva interpretacionista se acerca a la tradición europea de la *Verstehen*, (cfr. Devitt 1991, Cap.10).

<sup>45</sup> Como veremos, entre estas restricciones ha de estar el principio constitutivo de la racionalidad que introduce la *normatividad* de lo mental, rasgo que, a juicio de Davidson, segrega radicalmente los conceptos mentales de los de las ciencias naturales y los torna definicional y nomológicamente irreducibles a éstos (volveré sobre esta cuestión en el cap 4). De este modo, el proceso de interpretación que estaría involucrado toda vez que nos comprendemos (y que, según explicitaremos, en su desarrollo completo supone conceptos de actitud proposicional, cuya posesión es característicamente humana, propia de “animales racionales”) se revela como elusivo a todo intento de

Ahora bien, ¿de qué estrategia se vale Davidson en su intento de hacer comprensible la comprensión lingüística<sup>46</sup>, nuestro peculiar “modo de ser en el habla”<sup>47</sup>?

Andrew Carpenter (2002) ubica la estrategia davidsoniana en la tradición trascendental kantiana, en el siguiente sentido.

La estrategia trascendental kantiana puede esquematizarse así:

1. Se identifica un fenómeno sobre cuya existencia estamos de acuerdo;
2. Se investigan las condiciones necesarias para la posibilidad de dicho fenómeno (“análisis trascendental”) ;
3. Se examinan las implicaciones filosóficas del análisis resultante.

En el más célebre argumento trascendental kantiano, la deducción trascendental de las categorías, Kant identifica a la experiencia como el fenómeno a explorar (paso 1). Luego argumenta que las condiciones necesarias para la posibilidad de la experiencia incluyen la aplicabilidad de cada una de las doce categorías (paso 2). Finalmente, entre las muchas implicaciones de dicho análisis, se encuentran las respuestas kantianas al escepticismo cartesiano y humeano (paso 3).

A juicio de Carpenter, Davidson presenta un argumento trascendental de esta forma acerca de la posibilidad de la comunicación lingüística. Bajo esta sugerencia, creo que es plausible concebir el proyecto involucrado en la teoría de la interpretación radical del modo siguiente. Se adopta como punto de partida el fenómeno de la comunicación, esto es, se asume

---

explicación/predicción de corte “hempeliano”, dado que no podrá ser capturado en cuanto tal bajo las *leyes strictas* que gobiernan el mundo natural. Esta es la senda por la que veo plausible, en principio, pensarlo a Davidson como “abismándose” en las “profundidades” diltheyanas. Cuando discutamos su original respuesta al problema mente-cuerpo, el monismo anómalo, nos detendremos en la manera expresa en que el propio Davidson relaciona el anomalismo de lo mental con la polémica entre explicacionistas y comprensivistas y, asimismo, aludiremos a las conexiones que él mismo traza con el proyecto de Gadamer. No obstante, soy plenamente consciente de que esta filiación que aquí estoy meramente sugiriendo, requiere - entre otras cosas- introducir precisiones a propósito de la noción misma de “comprender”, que tal como la concibo, ha de estar íntimamente relacionada con el carácter *trascendental* del proyecto davidsoniano. En un espíritu consonante con esta sugerencia véanse -*inter alia*- Malpas, J. (2002) y Quintanilla, P. “La hermenéutica en diálogo con otras tradiciones”, disponible en <http://www.pucp.edu.pe/jornada/hermeneutica/docs/Quintanilla1.pdf>.

<sup>46</sup> Hablo, indistintamente, de “comprensión” o “comunicación” lingüística para capturar la idea de que entendemos las oraciones que intercambiamos (en términos de Moretti: la idea de que hay “fluidez dialógica”).

<sup>47</sup> Cf. Moretti, A. (2008. 57). No pretendo sugerir que Davidson considere que la tarea de echar luz sobre la comprensión lingüística se agote en la dimensión filosófica; de hecho, en “Belief and the Basis of Meaning” (1974b/1984) hace explícito que la teoría de la interpretación ha de ser un trabajo común del filósofo, el lingüista y el psicólogo (entendiendo que en tanto teoría diseñada para un lenguaje particular). Simplemente, intento precisar en qué consistiría la contribución estrictamente filosófica en la *comprensión* de la comprensión lingüística.

que entendemos las oraciones que intercambiamos (paso 1); a partir de explicitar cómo es posible que tengamos o adquiramos esa capacidad se muestra que somos capaces de comprender o dar significado a un plexo ilimitado de emisiones y de este modo se legitima conceptualmente la presuposición involucrada en el paso 1 (paso 2)<sup>48</sup>; se trazan implicaciones filosóficas a partir del análisis trascendental resultante (paso3).

En otros términos: se pretende validar conceptualmente la presuposición inicial mostrando que podemos construir teorías tarskianas interpretativas para nuestros lenguajes; a partir de allí se han de extraer consecuencias filosóficas diversas, entre las cuales me interesan, centralmente, las que impactan en el tratamiento davidsoniano de algunos problemas cruciales en filosofía de la mente, que consideraré en la sección II.

Tal como anticipé, mi interés no es hacer una exposición crítica del proyecto interpretativo davidsoniano (ardua y apasionante empresa)<sup>49</sup>, puesto que me alejaría del objetivo que persigo en este trabajo. Esto es, en lo que sigue, voy a concederle (“caridad” mediante) la adecuación de su análisis trascendental del fenómeno de la comunicación lingüística, para detenerme en la consideración de los principios constitutivos que en él se revelan (especialmente, el principio constitutivo de la racionalidad) a los efectos de examinar el impacto que han tenido en el tratamiento que hizo Davidson de algunos problemas cruciales de filosofía de la mente.

Para formularlo en los términos del esquema trascendental tripartito que presenta Carpenter: no me voy a ocupar del tránsito de 1 a 2, sino de algunos aspectos específicos involucrados en el pasaje de 2 a 3, a saber, de la extracción de implicaciones relativas al

---

<sup>48</sup> En consecuencia, no se pretende descubrir la estructura empírica que determina nuestra comprensión lingüística, ni proponer un procedimiento práctico para lograrla; la empresa filosófica parece resultar previa a estos objetivos, en la medida en que presuponen que existe tal comprensión y que la presuposición es legítima.

<sup>49</sup> Entre algunas de las críticas más relevantes a la propuesta davidsoniana puede verse, *inter alia*, Lepore & Ludwig (2005). Allí, en la Parte II argumentan que Davidson no logró justificar *a priori*, como tendría que haber hecho, que la interpretación radical pueda funcionar para *cualquier* hablante. Asimismo, la Parte III - Metafísica y Epistemología- es mayormente crítica de las conclusiones que Davidson traza en dichas áreas a partir del proyecto de la interpretación radical. Entre otras cuestiones, objetan que el punto de vista del intérprete radical pueda explicar la *autoridad de primera persona*, y sostienen que Davidson tampoco ha logrado mostrar cómo es que la interpretación radical es una *condición necesaria del concepto de objetividad*. Su objeción central es que no se ha mostrado que el proyecto de interpretación radical sea posible, y por tanto, sus implicaciones epistemológicas y metafísicas carecerían de apoyo suficiente. Puede consultarse, a su vez, el diagnóstico de Moretti (2008), que parece consonante con el de Lepore y Ludwig (2005) en varios puntos. Para Moretti desde la propuesta davidsoniana, se corre el riesgo de conseguir entender sólo cómo conviene diseñar para cada uno de los miembros de un amplio conjunto de lenguajes naturales una teoría capaz de representar lo que llamaríamos intuitivamente sus nexos de significaciones, sin por eso alcanzar a dar cuenta del fenómeno de la comprensión.

problema mente-cuerpo, la relación entre pensamiento y lenguaje, y el problema de las perspectivas de primera y tercera persona de la atribución mental.

Resumamos, entonces, el derrotero que hemos transitado hasta aquí. El punto de partida es que comprendemos cierto lenguaje  $\mathcal{L}$ , y la idea es que podemos legitimar conceptualmente esta presuposición mostrando cómo es posible, para nosotros, construir empíricamente una interpretación de las oraciones de  $\mathcal{L}$ . La contribución davidsoniana intentó mostrarnos (i) que podemos llevar a cabo un proceso constructivo que se ajusta a las formas tarskianas (que, en especial, produce teoremas de la forma “D es  $\emptyset$  si y sólo si p”) y (ii) que podemos controlar su adecuación interpretativa a través de nuestros recursos cognitivos hasta llegar a una instancia en la que resulte razonable creer que nos suministra, a la vez, *un* predicado veritativo  $\mathcal{L}$  e interpretaciones para las oraciones de  $\mathcal{L}$ . Logrado dicho estadio hemos de tener motivos para creer que el predicado  $\emptyset$  es el predicado “verdadera” (y que, entonces, p da condiciones veritativas de D) y que el predicado relacional “es  $\emptyset$  si y sólo si” puede sustituirse por “significa que” (y que, entonces, p da el significado de D)<sup>50</sup>.

De este modo, parece ser que al cabo de la “teoría” de la interpretación radical, nos quedamos con lo que ya teníamos pero, ahora, no tendríamos motivo para creer que se pueda tener una teoría más explicativa; no hay necesidad de ella<sup>51</sup>. En este sentido, el mérito principal de esta vía, a juicio de Moretti (2008: 68), es el de ofrecer cierta legitimación teórica del concepto de lenguaje, al sostener que su contenido se agota en la determinación de la forma general de una teoría empíricamente accesible, para la comprensión de un lenguaje natural cualquiera, dado que cuando se tiene tal cosa *se ha mostrado cómo es posible que tengamos un lenguaje*, esto es, cómo es posible que no haya límite para la comprensión de oraciones nuevas. De manera que, en esencia, lo que tendríamos no es una “teoría” sino sólo una manera aparentemente teórica de mostrar el funcionamiento correcto del lenguaje.

Desde esta perspectiva, acuerdo con el diagnóstico de Moretti (2008): la teoría tarskiana (complementada, en mi opinión, con las restricciones que se siguen de la adopción de la perspectiva del intérprete) parece resultar un recurso heurístico para *comprender* como lenguaje cierta práctica. La teoría tarskiana aparece, entonces, no como “genuina” teoría sino como recurso para adquirir una habilidad (una habilidad que, dice Davidson (1998) en respuesta a

---

<sup>50</sup> He seguido en este punto la reconstrucción propuesta por Moretti (2008: 16-7).

<sup>51</sup> En el sentido de que, según afirma Davidson (1998) en su respuesta a McDowell, parece que aquí “hemos llegado a algo que le pone un fin a un cierto regreso”.

Rorty, sin embargo, sólo puede describirse apelando al marco provisto por una semántica formal). Formulado en otros términos: aparece como la constitución de un marco dentro del cual formular teorías, esto es, como dando la posibilidad de teorizar y, por ende, no como teoría<sup>52</sup>.

El carácter constitutivo de la semántica se muestra, entonces, en la comunicación lingüística; frente a esto, la reflexión que nos deja con lo que ya teníamos, pese a que no posee carácter informativo, lejos de ser inútil, puede ayudarnos a orientar la mirada hacia ese carácter constitutivo.

Tal es el sentido, en el que al comienzo del presente capítulo he sostenido que concibo *la teoría de la interpretación radical* como la *explicitación* davidsoniana de aquello que hacemos al llevar a cabo una práctica peculiarmente *nuestra, humana*, tal como es la comunicación lingüística. De este modo, encuentro pertinente tomar a la teoría de la interpretación radical como el “hilo de Ariadna” que nos permitirá ver el entramado conceptual de la original contribución antropológica de Davidson, la figura que, en mi opinión, surge de su “rompecabezas conceptual”.

---

<sup>52</sup> Este modo de entenderla hace ver que “la reflexión semántica conduce a formular trivialidades, a lo sumo explicitaciones, porque de ese modo muestra que es posible la comprensión de significados lingüísticos que creemos tener” Moretti (2008:70).

### Capítulo 3: El concepto de creencia y sus rasgos

El *holismo*, el *externismo* y el rasgo *normativo de lo mental* se sostienen o caen juntos (...) creo que los elementos normativos, holistas y externistas en los conceptos psicológicos no pueden ser eliminados sin cambiar radicalmente de tema.  
(Davidson, Nicod Lectures, V, mis itálicas).

Como vimos en el capítulo anterior, la interpretación radical requiere asumir ciertos supuestos no sólo respecto del significado de las expresiones, sino también respecto de la relación entre las creencias, las oraciones y el mundo. En especial, hemos visto cómo la interpretación del lenguaje de un hablante supone, dada la manera peculiar en que se vinculan significados y creencias, la asignación simultánea de condiciones de verdad a sus emisiones y de estados mentales al hablante.

Ahora bien, dado que la única evidencia con la que cuenta el intérprete radical son las conductas efectivas y posibles de los hablantes<sup>53</sup>, al emprender la tarea de interpretación radical

---

<sup>53</sup> Uno de los puntos que Fodor y Lepore (1992) cuestionan es el supuesto que consideran común a Quine y Davidson, según el cual nada *puede ser* un lenguaje a menos que la interpretación radical sea posible. A tal efecto, una de las cuestiones que atacan (pp. 72-6) es el carácter poco transparente de la noción de “intérprete” que Davidson tiene en mente: no se nos dice de modo preciso en qué consiste la “virginidad” del investigador. Los lingüistas de campo reales explotan las intuiciones de sus informantes, e invariablemente se acercan al análisis de la lengua extranjera con un *background de supuestos teóricos muy poderosos*. A juicio de ellos, es enteramente implausible asumir que estos supuestos de *background* no jueguen ningún rol esencial en la selección de las teorías-T por parte del lingüista. Su punto no es sólo que funcionan en el “contexto de descubrimiento” sino que también funcionan en el “contexto de justificación” como parte del *corpus* de restricciones empíricas a las que tiene que adecuarse la teoría-T preferida. De este modo, para Fodor y Lepore, serán las observaciones disponibles de la conducta del informante *más los supuestos de background* aquello que conducirá a la elección de una teoría -T para la lengua extranjera. A partir de esto, nos dicen, *no se sigue* que la interpretación radical sea posible en los términos pretendidos por Davidson: esto es, que la elección de la teoría-T pueda hacerse sobre la base de *mera* evidencia observacional (junto con las restricciones generales que pueda haber sobre la inferencia teórica como tal). Frente a esta objeción, Evnine (1991:133) divide las aguas entre el proyecto de Quine y el de Davidson: “mientras que Quine piensa que el asentimiento es puramente conductual, Davidson admite que la evidencia es irreduciblemente intencional”. Una indagación más profunda acerca de la licitud de este punto de partida davidsoniano excedería, no obstante, los límites de nuestro trabajo, de modo que esta observación vale como mero señalamiento; tal como expresé en el capítulo anterior, mi intención no es hacer una consideración crítica de la teoría davidsoniana de la interpretación radical en sí misma, sino de sus implicaciones para la filosofía de la mente.

es necesario construir una teoría de la acción en forma simultánea con la construcción de una teoría del significado, esto es, el intérprete deberá partir no sólo de ciertos supuestos pertenecientes a una teoría del significado, sino también de supuestos pertenecientes a una teoría de la acción intencional. De este modo, la teoría de la interpretación podría considerarse como una teoría combinada del significado y la acción, esto es, una teoría que apunte tanto a asignar interpretaciones a las oraciones del hablante como contenidos a sus acciones. La evidencia básica para lo primero, como hemos visto, está constituida por las oraciones sostenidas como verdaderas [*hold true*], en tanto que para lo segundo, está dada por las preferencias del hablante exhibidas en sus acciones. En este sentido, en distintos trabajos<sup>54</sup> Davidson hace explícito su objetivo de ofrecer lo que denomina una “teoría unificada” en la que “el significado, la creencia y el deseo serán tratados como elementos completamente coordinados en una comprensión de la acción”<sup>55</sup>.

Hasta aquí, hemos supuesto que la identificación de las actitudes del hablante de sostener oraciones como verdaderas sería suficiente para dar lugar a una interpretación de cada oración en su lenguaje. Dichas actitudes nos ofrecen la pista inicial sobre cómo interpretar oraciones ocasionales [*occasion sentences*]. Dado que cualquier término puede aparecer en una oración ocasional, es plausible suponer que las interpretaciones asignadas a las oraciones ocasionales nos proveerán evidencia para las interpretaciones de todos los términos del lenguaje en cuestión. No obstante, la interpretación correcta de los términos “más o menos remotos de lo que es inmediatamente observado” (Davidson 1980b/2004: 157) dependerá, especialmente, de cómo el hablante trata las oraciones ocasionales como apoyando las oraciones que los contienen y qué oraciones ocasionales trata como consecuencias de las oraciones que los contienen. Esto es, la interpretación de los términos cuyas condiciones de aplicación están guiadas por una teoría compleja en la que están insertos dependerá de trazar las conexiones de apoyo entre las oraciones

---

<sup>54</sup> Cf. vgr. Davidson (1980b/2004), (1984b/1984).

<sup>55</sup> Davidson (1980b/2004: 152). En el capítulo anterior, enfatice que la teoría de la interpretación radical no es una mera teoría de condiciones de verdad, sino que a los efectos de su aplicabilidad empírica hay que complementarla con elementos ajenos a ella, a saber: las *restricciones* que surgen de la adopción de la perspectiva del intérprete radical. En este sentido, sugerí que la contribución davidsoniana debía pensarse a la vera de sus diálogos con Tarski y Quine, respectivamente. Su teoría unificada nos invita, ahora, a precisar dichas restricciones, en las que Davidson pone en juego lo que nos hace peculiarmente *humanos*, mediante un diálogo al que hay que sumar los aportes de los teóricos de la decisión. En este sentido, Lepore y Ludwig (2005: 257) sostienen que la teoría unificada, lejos de reemplazar el procedimiento de la interpretación radical presentado inicialmente en (1973/1984), constituye un refinamiento del mismo.

acerca de lo inmediatamente observable y las oraciones más remotas (digamos, oraciones teóricas).

Dado que las relaciones de apoyo son típicamente más débiles que las de implicación, lo que se requiere es una medida del grado en el cual un hablante trata a una oración ocasional dada -contra el *background* del resto de su teoría- como apoyando a una oración teórica. Esto a menudo se reflejará en cómo incrementa o decrece su confianza en la verdad de una oración teórica con el correspondiente incremento o merma de su confianza en la verdad de las oraciones ocasionales. Una vez que se ha reconocido, pues, la relevancia del grado de creencia en este contexto, podemos advertir que será de central importancia también para el caso de interpretar oraciones ocasionales (dado que, con frecuencia, sostenemos como verdaderas oraciones acerca de eventos en nuestro entorno con bastante menos que una confianza perfecta). En consecuencia, debemos reparar no sólo en las actitudes de sostener oraciones como verdaderas, sino también en los grados de confianza que un hablante tiene en la verdad de sus oraciones y, similarmente, para las actitudes optativas hacia las oraciones, el grado en el cual el hablante desea o prefiere que la oración sea verdadera. Una teoría de la interpretación completamente articulada requerirá, entonces, el empleo de grados de confianza en la verdad de las oraciones (probabilidades subjetivas) y de deseabilidades [*desirabilities*] relativas asignadas a la verdad de las oraciones<sup>56</sup>.

Este punto introduce una dificultad que, hasta ahora, he soslayado: si la asignación correcta de interpretaciones requiere que identifiquemos probabilidades subjetivas y deseabilidades, ¿cómo podemos hacerlo sin suponer lo que la teoría de la interpretación debe proveernos? En otros términos, hemos visto que para resolver el problema de quebrar el círculo entre la creencia y el significado, Davidson elige una actitud que considera como vector de ambos: la actitud de sostener como verdadera una oración. El problema más general que ahora nos concierne puede tratarse en forma paralela: necesitamos identificar una actitud que pueda emplearse como evidencia simultánea para las probabilidades subjetivas, las deseabilidades y los significados. La solución propuesta por Davidson reside en la actitud de preferir una oración como verdadera [*prefer true*] en lugar de otra. Esta actitud es, pues, el análogo de la actitud de sostener una oración como verdadera para la teoría de la decisión, a partir de la versión de

---

<sup>56</sup> A los efectos de reconstruir la motivación de Davidson para abrazar el marco provisto por la teoría de la decisión, sigo los lineamientos principales de Lepore y Ludwig (2005).

Richard Jeffrey que Davidson adopta como marco<sup>57</sup>. Su singular teoría resultará, entonces, de la unión entre la teoría del significado- cuyos lineamientos centrales vimos en el capítulo anterior- y una versión de la teoría de la decisión. En este sentido, su teoría unificada puede concebirse a mi juicio, a partir de sus diálogos con Tarski y Quine, por un lado, y con Jeffrey, por otro<sup>58</sup>. En opinión de Davidson, la teoría del significado y la teoría bayesiana de la decisión están “evidentemente hechas una para la otra” (1980b/2004:158). No obstante, no basta con establecer esta dependencia mutua, dado que ninguna de las teorías puede ser desarrollada primero sobre la base de la otra. La teoría de la decisión debe ser liberada del supuesto de un conocimiento del significado determinado independientemente, en tanto que la teoría del significado requiere de una teoría del grado de creencia a los efectos de hacer un uso serio de las relaciones de apoyo evidencial. Por un lado, la teoría de la decisión normalmente es testeada pidiéndole a un sujeto que elija entre opciones descriptas por un experimentador, esto es, el sujeto elige entre oraciones y el experimentador supone que sabe lo que esas oraciones quieren decir para el sujeto. La teoría unificada abandona este supuesto al tomar los datos como elecciones entre oraciones ininterpretadas. Por otro lado, respecto de la teoría del significado, aún cuando la actitud de sostener como verdadera fuera base suficiente para ella, según hemos visto, persiste la dificultad de que esta actitud no posee una conexión obvia o directa con la acción; el problema no estriba tanto, nos dice Davidson, en que sólo puede ser observada la acción, sino en que nada puede contar como una razón para suponer que alguien sostiene una oración como verdadera sin suponer mucho, a su vez, acerca de sus intenciones, propósitos o valores. Estamos forzados, entonces, a traer el deseo o la preferencia, tanto como la creencia, al centro de la escena. En otras palabras: el sostener oraciones como verdaderas fue la base evidencial para la teoría de la creencia y el significado; al introducir la preferencia en la base evidencial, estamos en condiciones de incorporar el deseo en la teoría. Así, nos dice Evinine (1991:114), al emplear las elecciones de la gente entre pares de oraciones, Davidson ha esbozado un método para asignar creencias y deseos al hablante y significados a las oraciones entre las que elige.

---

<sup>57</sup> Cabe señalar que Jeffrey escribió la primera edición de *The Logic of Decision* en Stanford entre 1961 y 1964, y según consignan Lepore y Ludwig (2005: 252), por entonces estuvo en contacto con Davidson (de hecho, agradece su ayuda en el prefacio, conjuntamente con la colaboración prestada por Patrick Suppes).

<sup>58</sup> Davidson (1995/2004) remite a la versión de Jeffrey, R. (1990) y, asimismo, a la teoría de F.P.Ramsey (1931). Por razones de pertinencia y claridad expositiva, voy a omitir deliberadamente los detalles técnicos de estas propuestas para concentrarme, más adelante, en el tratamiento conceptual de uno de los rasgos propios de la noción de creencia que, a mi juicio, guarda una clara inspiración en dichos aportes: la normatividad en tanto rasgo de lo mental.

En suma, a juicio de Davidson, se requiere una teoría unificada que nos dé el grado de creencia, las utilidades en una escala y una interpretación del habla sin suponer ninguna de ellas.<sup>59</sup> Esta teoría no pretende describir cómo nos interpretamos efectivamente, sino que se ofrece, según hemos dicho, “como una especulación sobre qué hace interpretable al pensamiento y al lenguaje”, y si somos capaces de contar una historia acerca de cómo esto es posible en los términos de su propuesta unificada, “podemos concluir que las *restricciones* que la teoría ubica *sobre las actitudes* pueden articular algunos de sus *rasgos filosóficamente significativos*” (1995/2004: 128; mis itálicas). Precisamente, dispuesto a caracterizar dicha teoría en “Could there be a science of rationality?” Davidson hace explícitas las *tres restricciones fundamentales* que la teoría impone *sobre las actitudes proposicionales* y que, según veremos, puesto que no son meramente metodológicas sino que resultan restricciones *constitutivas*, se nos ofrecen como los rasgos propios de lo mental: *holismo, normatividad y externismo* (Cf. Davidson 1995/2004: 129-130).<sup>60</sup>

En este capítulo, voy a centrarme en el análisis de dichos rasgos a partir del concepto de la actitud proposicional por excelencia para Davidson: el concepto de *creencia*<sup>61</sup>. Sostengo que se trata del concepto de actitud proposicional por excelencia en la medida en que la *creencia* resulta básica para Davidson en el siguiente sentido:

---

<sup>59</sup> Dados mis propósitos, no voy a detenerme en la consideración crítica de la teoría unificada que propone Davidson. En Davidson (1995/2004: 131-4) pueden verse sus reconstrucciones de las objeciones principales que Chomsky y Fodor le formularon a dicha teoría y al método de la interpretación radical.

<sup>60</sup> Al tematizar cada uno de los rasgos en cuestión, volveré sobre el crucial carácter *constitutivo* de lo mental que, a mi juicio, revisten. No obstante, cabe adelantar la siguiente reflexión general en torno a la manera en que, en mi opinión, están entrelazadas la dimensión metodológica y la constitutiva en el planteo davidsoniano. Los conceptos de actitud proposicional resultan conceptos teóricos en virtud de que sólo cobran sentido en el seno de la teoría de la interpretación radical. Hemos visto que, dadas sus restricciones, dicha teoría supone la perspectiva del intérprete radical, esto es, la perspectiva pública, atributiva. Ahora bien, a mi juicio, la adopción davidsoniana de dicha perspectiva involucra un solapamiento entre las dimensiones *epistemológica* y *metafísica*, en el mismo sentido sugerido por Fodor y Lepore (1992: 59): la idea básica es que es *metafísicamente* constitutivo de los hechos acerca del contenido que deben ser accesibles para alguien en la situación *epistemológica de un intérprete radical*. Si esto es así, las restricciones que la teoría impone no conciernen simplemente a la atribución, sino también a la identidad misma de las actitudes proposicionales, es decir, no son meramente metodológicas, sino que resultan constitutivas de aquello acerca de lo cual versa la teoría, es decir, de las *actitudes proposicionales* (entre las cuales será central la *creencia*, por las razones que expondré a continuación). Tal es el espíritu que me anima para considerar las tres restricciones a las que alude Davidson (1995/2004) como rasgos propios del concepto de creencia.

<sup>61</sup> En el mismo espíritu, Fodor y Lepore (1992: 118-9) recogen esta reflexión davidsoniana bajo el rótulo de la “centralidad de la creencia”.

Cada creencia no solamente requiere de multitud de otras creencias para darle contenido e identidad, sino que toda otra actitud proposicional depende para su singularidad de una multitud similar de creencias [...] La creencia –de hecho, la creencia verdadera- juega un papel central entre las actitudes proposicionales. Así, pues, voy a hablar de todas las actitudes proposicionales como pensamientos (1982/2001:98).

Ahora bien, ¿por qué me interesa, especialmente, el concepto de *creencia*? En las páginas que siguen, voy a concentrarme en dicho concepto (o indistintamente, en el concepto de *pensamiento*) en tanto es una de las piezas “esenciales”, “insustituibles” del “rompecabezas” conceptual que nos propone Davidson<sup>62</sup>. Como hemos visto en los dos capítulos anteriores, dada su centralidad en el sistema davidsoniano el concepto de *creencia* es abordado desde múltiples vertientes disciplinares. Resulta crucial para su filosofía de la acción (en tanto *razón* que justifica y, a la vez, *causa* la acción), esencial para su propuesta epistemológica coherentista (según la cual sólo otra creencia puede ser razón para una creencia), y asimismo, resulta un dispositivo conceptual insustituible como base del significado, según su teoría de la interpretación radical. Ahora bien, mi interés específico en esta “pieza” reside en que creo que el concepto de *creencia* opera, asimismo, como una suerte de “esfinge” que contiene la clave para alumbrar aquello que, a juicio de Davidson, nos hace especiales, esto es, para que surja la “figura” del rompecabezas que, tal como ya he sugerido, es su contribución antropológica.

En lo que sigue, pues, voy a presentar cada uno de los tres rasgos del concepto de creencia: *holismo*, *normatividad* y *externismo*. En cada caso, distinguiré variedades y propondré una articulación allí donde en la posición de Davidson no quedan, en mi opinión, caracterizados de modo claro y preciso. Esta presentación resulta un rodeo imprescindible a los efectos de emprender la tarea de la próxima sección, esto es, analizar el impacto de dichos rasgos en las respuestas davidsonianas a tres problemas tradicionales de la filosofía de la mente: el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la atribución mental<sup>63</sup>.

---

<sup>62</sup> Desde la perspectiva de Davidson, nuestro pensamiento y nuestro lenguaje están contruidos por “algunas piezas conceptuales esenciales” y, por ello insustituibles, tales como: bien, verdad, conocimiento, objeto físico, causa, evento y creencia (cf. Davidson 1998b/2001).

<sup>63</sup> Cabe anticipar que soy plenamente consciente de que los problemas aludidos no son privativos de la filosofía de la mente. Por ejemplo, el problema mente-cuerpo ha sido considerado, canónicamente, parte de la vasta reflexión metafísica, en tanto que la relación entre el pensamiento y el lenguaje también ha sido, clásicamente, transitada desde la filosofía del lenguaje. La demarcación que trazo es, pues, a los efectos de acotar el “mapa” conceptual de

### 3.1. Holismo

“Holismo” se dice de distintas maneras en la teoría unificada propuesta por Donald Davidson. A mi juicio, es posible discernir, al menos, las dos variedades siguientes: por un lado, (I) un holismo *metodológico*, que constriñe la teoría<sup>64</sup> y, por otro, (II) un holismo *constitutivo*, que constriñe aquello acerca de lo cual versa la teoría<sup>65</sup>; esta variedad se subdivide, en mi opinión, en: (II.1) holismo semántico (bajo el “ropaje” del holismo del *contenido mental/lingüístico*) y (II. 2) holismo psicológico (en tanto holismo de las *actitudes proposicionales*). Vayamos, pues, por partes<sup>66</sup>.

(I) Como señalan Lepore y Ludwig (2003:52), la sugerencia inicial de Davidson (1967) de que una teoría de la verdad que fuera extensionalmente adecuada (en el nivel de los teoremas-T) sería, *ipso facto* interpretativa, resultó incorrecta.<sup>67</sup> El objetivo de esa sugerencia fue identificar restricciones que pudiera satisfacer una teoría de la verdad, tal que no estuvieran

---

las discusiones en las que me interesa intervenir para ver cómo juega el concepto de *creencia* a la hora de arrojar luz sobre qué nos hace ser quienes somos.

<sup>64</sup> O “teorías” (asumiendo la indeterminación semántica). No obstante, en lo que sigue, para facilitar la exposición apelaré a la formulación en singular.

<sup>65</sup> Sigo en esta denominación la propuesta de Malpas (1992) quien distingue entre: holismo metodológico, una variedad de holismo que constriñe la construcción de la teoría, y holismo constitutivo, que constriñe aquello acerca de lo cual es la teoría, esto es, que “constriñe lo psicológico mismo”. De esta suerte, tanto los objetos mismos de la interpretación (creencias, deseos, etc.) así como la evidencia en la que está basada la interpretación están holísticamente restringidos. El holismo davidsoniano es así metodológico, en la medida en que importa una restricción sobre las teorías de la interpretación, y es constitutivo, en tanto el ámbito psicológico, que es el tema de la interpretación, está holísticamente constituido. Lo que le resulta evidente a través de una consideración de la interpretación es que los requisitos holistas que constriñen la metodología interpretativa son constitutivos de lo psicológico. En un espíritu consonante, Orlando (1999) ha sostenido que el holismo atributivo (metodológico) en conjunción con ciertos presupuestos de la interpretación -básicamente, la idea de que no hay contenido alguno al margen de la práctica interpretativa- implica el holismo semántico (constitutivo). Las dos tesis son, pues, menos claramente distintas de lo que en principio podría pensarse. Asimismo, también en Eynine (1991:121) puede encontrarse la sugerencia de que en la propuesta davidsoniana habría, al menos, dos tipos de holismo: el holismo que, con Malpas yo opto por llamar “constitutivo”, es a juicio de Eynine, un tipo de holismo “más profundo” que aquel que meramente restringe la teoría a los efectos de que resulte interpretativa.

<sup>66</sup> Antes de avanzar, me gustaría hacer la salvedad siguiente: estoy tratando de discernir entre las variedades de holismo de lo mental más o menos *explícitas* en el planteo de Davidson. No obstante, esto no será óbice para explorar si, además de estas variedades, hay otros tipos de holismo quizás *implícitos* o *meramente sugeridos* en el tratamiento davidsoniano de lo mental.

<sup>67</sup> Esta dificultad es aquella que Fodor y Lepore (1992: 62) rotulan como el “problema de la extensionalidad”. Una teoría de la verdad es extensionalmente adecuada si y sólo si todas las oraciones -T que implica son verdaderas. Ahora bien, la mentada adecuación extensional podría lograrse apareando cada oración de L con cualquier oración materialmente adecuada con independencia de lo que significa. El ejemplo clásico es una teoría de la verdad para el castellano que implique, *inter alia*, la oración- T W: “La nieve es blanca” es verdadera en castellano si y sólo si el pasto es verde. El punto es que se trataría de una teoría extensionalmente adecuada, pero que no captura los significados de las oraciones que describe y, en este sentido, no resulta interpretativa.

formuladas en términos semánticos, de manera de arrojar luz indirectamente sobre qué es lo que está involucrado en la comprensión de un hablante.

Ahora bien, si la adecuación extensional no es suficiente, surge entonces la pregunta acerca de qué restricciones adicionales debe cumplir una teoría de la verdad para un lenguaje natural a los efectos de ser interpretativa.<sup>68</sup> Según vimos en el capítulo anterior, Davidson vuelve sobre esta cuestión en “Radical Interpretation” (1973), regreso que es caracterizado en “Reply to Foster” (1976) como un intento de rigorizar la relación entre una teoría de la verdad y una teoría del significado. Al comienzo de “Radical Interpretation” Davidson nos interpela con las dos preguntas siguientes: (i) ¿qué es lo podríamos saber que nos permitiría interpretar a otros hablantes?<sup>69</sup> y (ii) ¿cómo podríamos llegar a saberlo? Esta segunda pregunta ha de ser respondida, según vimos, sobre la base de evidencia que incluye el conocimiento de la conducta efectiva y potencial de un hablante, pero que excluye el conocimiento del significado de cualquiera de sus términos y todo conocimiento detallado de sus actitudes proposicionales. Esto es lo que define la posición del intérprete radical y, de esta suerte, el proyecto de la interpretación radical. Tal como sostienen Lepore y Ludwig (2003), el objetivo de la descripción teórica del procedimiento de la interpretación radical es arrojar luz sobre los conceptos que usamos al interpretar a otros hablantes –entre los cuales, según veremos, será central el de *creencia*- relacionando la estructura teórica que los contiene con evidencia para su aplicación, que es descripta independientemente de su uso.

A propósito de la pregunta (i), hemos visto que la sugerencia de Davidson es que lo que podríamos saber que nos permitiría interpretar a un hablante es una teoría de la verdad para su lenguaje (propuesta que hemos presentado como surgida de su “diálogo” con Tarski), que satisfaga ciertas restricciones (propuesta que hemos presentado como surgida, al menos en parte, de su “diálogo” con Quine). Hemos señalado que según Davidson (1967) la sugerencia fue que el

---

<sup>68</sup> En efecto, también en (1974b/1984) Davidson sostiene que una teoría de la verdad producirá interpretaciones sólo si sus oraciones T enuncian condiciones de verdad en términos de lo que puede verse como “dar significado” de las oraciones del lenguaje objeto. Nuestro problema, dice, es encontrar en la teoría restricciones que sean lo suficientemente fuertes como para garantizar que ella pueda emplearse en la interpretación.

<sup>69</sup> Lepore y Ludwig (2003) nos advierten sobre el elemento contrafáctico de la pregunta y afirman - en consonancia con nuestra sugerencia en el capítulo anterior respecto de la contribución filosófica que aporta la teoría de la interpretación radical- que Davidson jamás afirmó que nuestra competencia en nuestro lenguaje esté constituida por conocimiento proposicional de una teoría composicional del significado explícita o una teoría de la verdad para el lenguaje. Sostienen que la teoría intenta, en cambio, “capturar” en su estructura, la estructura de una *habilidad práctica compleja*, la habilidad de hablar y comprender. (Cf. Ludwig ed. 2003: 62). Este punto será especialmente relevante cuando, en el cap. 6, cuestionemos la pertinencia de pensar la “teoría” de la interpretación radical desde una concepción excesivamente “teoricista”.

requisito apropiado para la teoría era su mera adecuación extensional, en el nivel de los teoremas-T. La sustitución de este requisito en “Radical Interpretation” nos interesa, especialmente, porque nos permite tematizar una primera variedad de holismo en la propuesta de Davidson: el *holismo metodológico*.

Leemos en Davidson:

Si supiéramos que una oración T satisface la Convención T de Tarski, sabríamos que es verdadera, y podríamos usarla para interpretar una oración porque sabríamos que la rama derecha del bicondicional traduce la oración a ser interpretada. Nuestro problema actual emana del hecho de que en la interpretación radical no podemos suponer que una oración T satisface el criterio de traducción. Hemos estado dejando de lado, sin embargo, que ya proporcionamos un criterio alternativo: este criterio consiste en que la *totalidad de las oraciones T* deberían ajustarse óptimamente (en el sentido descripto más atrás)<sup>70</sup> a la evidencia acerca de las oraciones consideradas verdaderas por los hablantes nativos. La presente idea es que lo que Tarski daba directamente por supuesto para cada oración T puede ser extraído indirectamente mediante *una restricción holística*. Si esta restricción es adecuada, cada oración T producirá en efecto una interpretación aceptable (1973/1984: 139; mis itálicas).<sup>71</sup>

Hemos visto que a los efectos de proveer una interpretación de las oraciones de un lenguaje L, las teorías T deben ser teorías empíricas. Hemos señalado, asimismo, al plantearnos el problema de su aceptabilidad, que pareciera que lo que debemos hacer es echar mano a los métodos estándar de confirmación de teorías empíricas. Como es sabido, dichos métodos imponen la comparación de las consecuencias de la teoría con el mundo. Para el caso que nos ocupa, las consecuencias contrastables de las teorías son las oraciones T que constituyen sus teoremas. En este sentido, lo que a mi juicio amerita, pues, que lo concibamos a Davidson comprometido con un *holismo metodológico* es que a los efectos de la aceptabilidad de la teoría lo que debe buscarse no son interpretaciones correctas de muestras<sup>72</sup>, sino la corrección material de la teoría como *totalidad*.

---

<sup>70</sup> Es decir, empleando los procedimientos del intérprete radical.

<sup>71</sup> Dados mis propósitos, no voy a detenerme a considerar si esta restricción resulta adecuada a los efectos de responder la pregunta (i) ¿qué es lo podríamos saber que nos permitiría interpretar a otros hablantes? En este sentido, puede verse *–inter alia–*. Lepore y Ludwig 2003: 53-54.

<sup>72</sup> Tal como advierte explícitamente en (1973/1984). En principio, esta reflexión podría tentarnos a pensar que podemos asimilar de manera directa esta variedad al holismo de la *confirmación*. No obstante, si nos remitimos al *locus* clásico de su formulación, a saber Quine (1951), pareciera que en su rechazo a la tesis reduccionista ínsita en el localismo de Carnap, la defensa del holismo de la confirmación que allí acomete Quine atañe a mucho más de lo que Davidson pone en juego (al menos, en el pasaje que citamos recién). Esto es, Quine (1951:41) está afirmando

La idea que parece estar detrás de este señalamiento es la siguiente. Ninguna teoría de la verdad adecuada para el castellano podría implicar W (“La nieve es blanca” si y sólo si el pasto es verde) y, a la vez, exhibir un rasgo que, según vimos en el capítulo anterior resulta clave en la propuesta davidsoniana de una teoría del significado: esto es, exhibir estructura composicional para todas las oraciones del castellano que contengan “nieve” o “blanca”; puesto que si una teoría que implica W respetase, a la vez, la estructura composicional, tendría que asignar condiciones de verdad incorrectas a otras oraciones en las cuales ocurrieran “nieve” y “blanca”.<sup>73</sup>

En este sentido, una primera moraleja<sup>74</sup> a extraer es que, después de todo, podemos tener una teoría extensional del significado en la medida en que en la derivación de las oraciones-T se respeten las relaciones estructurales entre las fórmulas; la otra moraleja es, precisamente, el holismo: “La nieve es blanca” tiene las condiciones de verdad que tiene dado que pertenece a un lenguaje que contiene “Esto es blanco” y “Esto es nieve” (e indefinidamente, muchas otras oraciones en las que ocurren “es blanca” y “es nieve”).

Como se recordará, Davidson distingue dos niveles explicativos: la explicación *dentro de* la teoría y la explicación *de* la teoría. Dentro de la teoría, las condiciones de verdad de las oraciones son especificadas, en virtud del principio de composicionalidad, haciendo referencia a la estructura interna y, por tanto, al concepto semántico de referencia, que involucra las relaciones entre los distintos tipos de palabras del lenguaje y los distintos tipos de entidades del mundo. Esto es, los axiomas a partir de los que se derivan las oraciones-T emplean las nociones semánticas de referencia y satisfacción: ellos incluyen axiomas tales como “‘nieve’ refiere a nieve” y “‘es blanco’ es satisfecho por todas y sólo las cosas blancas”. Pero cuando se trata de

---

que nuestros enunciados acerca del mundo externo han de afrontar el tribunal de la experiencia sensible no individualmente sino sólo como cuerpo completo (en tanto que Davidson sólo se está refiriendo a la totalidad de las consecuencias de una teoría específica, esto es, sus oraciones-T). Sin embargo, si acotamos la tesis a la formulación conocida como “Quine/Duhem” (cf. v.gr. Fodor y Lepore 1992), según la cual la confirmación es holista en la medida en que todo enunciado en una teoría determina (parcialmente) el nivel de confirmación de todo otro enunciado en la teoría, tal vez haya un sentido plausible en el que pueda alegarse la presencia de una variedad de holismo de la confirmación en la teoría unificada propuesta por Davidson (véase, por ejemplo, Evnine 1991: 15). No obstante, dado que, por un lado, el foco de mi indagación no es epistemológico y, por otro, dado que la variedad de holismo que me interesa es la *constitutiva* (puesto que es allí donde, a mi juicio, aparece la originalidad de Davidson a la hora de iluminar qué nos hace ser los “animales racionales” que somos), no voy a profundizar esta interesante línea de reflexión, que probablemente encuentre terreno fértil en el rechazo davidsoniano de la distinción analítico/sintético y en sus contribuciones más estrictamente epistemológicas (como, por ejemplo, 1983/2001).

<sup>73</sup> En particular, tendría que asignar condiciones de verdad incorrectas a oraciones en las que ocurren con demostrativos como “Esto es blanco” y “Esto es nieve”. En esta línea, tendría que afirmar, por ejemplo, “Esto es blanco” si y sólo si es verde.

<sup>74</sup> Sigo en este punto, los señalamientos de Fodor y Lepore (1992: 64-5).

establecer la aceptabilidad de la teoría, aquello que está vinculado con la conducta y el entorno es el concepto de verdad.

La razón de la centralidad conferida a la noción de verdad, en detrimento de la de referencia, se debe a la suscripción davidsoniana de la tesis de la inescrutabilidad de la referencia<sup>75</sup>, originada en Quine, de acuerdo con la cual la noción de referencia carece de contenido empírico objetivo y es, por tanto, relativa a un cierto esquema conceptual. De esta suerte, el concepto de referencia resulta ser un concepto intrateórico que tiene diferentes poderes deductivos para la asignación de condiciones veritativas; en otras palabras, tanto las relaciones entre nombres y objetos como las relaciones entre predicados y conjuntos de objetos sólo reciben contenido empírico a través de la asignación de condiciones veritativas a las oraciones en las que los nombres y los predicados figuran, es decir, en el nivel de los bicondicionales T.

Así, desde la perspectiva *holista metodológica* que Davidson pone en juego a la hora de construir una teoría de la verdad que sea interpretativa, el lenguaje nos conecta con el mundo no por medio de las palabras sino en el nivel de las emisiones oracionales, consideradas como el comportamiento lingüístico básico, y dado que una oración sólo tiene significado en el contexto del lenguaje al que pertenece, es la *totalidad de las oraciones T* la que debe ajustarse a la evidencia acerca de las oraciones sostenidas como verdaderas por el hablante. Según esta primera variedad de holismo, el intérprete debe presuponer, entonces, que los significados que adscribe están relacionados con otros significados, sin que esto implique *prima facie* que los significados en sí mismos son constitutivamente holistas. Este punto nos remite a la segunda variedad de holismo que encontramos en Davidson: (II) el *holismo constitutivo*, en la versión del holismo semántico (II.1).

Leemos en “Truth and Meaning”:

Si el significado de las oraciones depende de su estructura y entendemos el significado de cada ítem de la estructura sólo como una abstracción de la totalidad de las oraciones en las que éste opera, entonces podemos dar el significado de cualquier oración (o palabra) solamente dando el significado de todas las oraciones y palabras del lenguaje. Frege dijo que una palabra tiene significado en el contexto de una oración; siguiendo esta idea debería haber agregado que sólo en el contexto del lenguaje una oración (y por lo tanto una palabra) tiene significado (1967/1984:22).

---

<sup>75</sup> Sigo aquí los lineamientos principales de Orlando, E. (2001).

(II) En primer lugar, cabe una aclaración relativa a la denominación de esta variedad, que bajo el rótulo de “*holismo semántico*”, a mi juicio subsume tanto al holismo del *significado lingüístico* como al holismo del *contenido mental*, puesto que, en primera instancia, podría considerarse la posibilidad de distinguir el holismo entre aquel que concierne al *lenguaje* y aquel que concierne al *pensamiento*.

La razón por la que apelo a la denominación común de “holismo semántico” reside en la peculiar visión unificada del lenguaje y de la mente que posee Davidson y que plasma en su tesis de la *interdependencia entre lenguaje y pensamiento*, posición que defiende en varios trabajos<sup>76</sup> y que, según hemos visto, abraza a la hora de caracterizar la empresa de la interpretación radical, como involucrando la asignación simultánea de significado a las emisiones lingüísticas del hablante y de contenido a sus actitudes proposicionales. Así, leemos en Davidson:

[N]i el lenguaje ni el pensamiento pueden explicarse plenamente en términos del otro, y ninguno de ambos tiene prioridad conceptual. Los dos, por cierto, están relacionados en el sentido de que cada uno requiere del otro para su comprensión, pero la relación no es tan completa como para que alguno de ambos baste, incluso cuando se lo refuerza razonablemente, para explicar al otro (1975/1984: 156).

Precisamente, en “The problem of objectivity” afirma explícitamente que dicha tesis constituye *per se* una “forma de holismo”<sup>77</sup> a la vera de la cual carece de sentido discernir entre variedades relativas al lenguaje y al pensamiento. Su idea es que las distinciones cognitivas que somos capaces de expresar en el lenguaje deben ser distinciones que somos capaces de hacer en el pensamiento. Si un pensamiento implica lógicamente a otro, o provee un grado de apoyo

---

<sup>76</sup> Especialmente, en dos trabajos que abordaremos en detalle en el capítulo 5: (1975/1984) y (1982/2001). Obviamente, soy consciente de que a la vera de este señalamiento podría cuestionarse la adecuación de trazar una demarcación entre su “filosofía del lenguaje” y su “filosofía de la mente”. Tal delimitación es “prágmática” en el sentido de que simplemente apunta a señalar que, en lo que sigue, voy a acotar mi tratamiento del concepto de *creencia* a una tríada de problemas, tal como han sido tradicionalmente tematizados desde la filosofía de la mente, ámbito disciplinar cuya geografía conceptual queda, a mi juicio, fuertemente reconfigurada e inextricablemente unida a la de la filosofía del lenguaje, si uno está dispuesto a tomarse en serio la manera en que Davidson, según veremos, socava una “imagen incorrecta de la mente” y propone una nueva agenda de problemas.

<sup>77</sup> Cf. Davidson (1995b/2004: 11). ¿En qué sentido la tesis de interdependencia entre el lenguaje y el pensamiento constituye una forma de holismo? Davidson no lo aclara; no obstante creo que la idea central que pone en juego es su concepción *relacional* del *lenguaje y el pensamiento*, en el siguiente sentido: las atribuciones de estados intencionales y de actos de habla –con contenido proposicional– van de la mano, esto es, ninguna de ellas es inteligible con independencia de su relación con la otra. Como ya anticipé, volveré sobre esta cuestión en el cap. 5.

racional para otro, las mismas relaciones lógicas o de confirmación se mantienen entre las oraciones que expresan esos pensamientos. Y, a su vez, considera que una criatura no puede tener pensamientos a menos que tenga un lenguaje, esto es, a menos que sea intérprete del habla de otro. Estos hechos, a juicio de Davidson, son suficientes para hacer altamente plausible que cualesquiera sean las restricciones holistas que se mantienen para el pensamiento, se mantienen también para el lenguaje y viceversa. Davidson pretende, entonces, que lo que diga del holismo se sostenga indistintamente para el pensamiento y el lenguaje y, de este modo, me parece adecuado apelar al rótulo de “*holismo semántico*” para referirme, indistintamente, al del significado lingüístico y al del contenido mental.<sup>78</sup>

De acuerdo con una perspectiva estándar<sup>79</sup>, el *holismo semántico* es la tesis según la cual la identidad del contenido de una creencia (o el significado de una oración que lo expresa) está determinada por su lugar en la red de creencias u oraciones que comprende una totalidad (el lenguaje al que pertenece), esto es, depende de sus relaciones con muchas o todas las otras creencias (u oraciones). De este modo, el holismo semántico contrasta con el atomismo<sup>80</sup>, de acuerdo con el cual cada expresión simple puede tener significado independientemente de todas las otras expresiones, y con el molecularismo<sup>81</sup>, según el cual hay dependencias de significado, pero restringidas a partes más pequeñas, y a menudo, unidireccionales.

---

<sup>78</sup> Desde ya, cabe aclarar que el holismo del contenido/ del significado no agota el *holismo semántico*, que resulta una doctrina más abarcativa en tanto atañe no sólo a la propiedad de tener *contenido intencional*, sino también a las de: *ser un símbolo, ser un símbolo que pertenece al lenguaje L, tener un referente, ser semánticamente evaluable*, etc. Aquí me concentraré en el *holismo del contenido intencional* puesto que, junto con el *holismo de las actitudes proposicionales*, resultan las variedades de las que Davidson se vale explícitamente en los argumentos de los que voy a ocuparme.

<sup>79</sup> Véanse, por ejemplo, Block, N. (1998) y Pagin, P. (2006). En este último trabajo se consigna que las ideas holistas semánticas fueron introducidas en la filosofía analítica al comienzo de los '50, especialmente, con las contribuciones de Carl Gustav Hempel (1950) y Willard van Quine (1951). Asimismo, se menciona como otra fuente fundacional del holismo semántico el aporte de Wilfrid Sellars (1963) en torno a los juegos de lenguaje, que resultó inspirador para la -así llamada- semántica de rol conceptual o inferencial, presente en los trabajos de Harman (1974), Field (1977), Block (1986) y Brandom (1994).

<sup>80</sup> Según Fodor y Lepore (1992:7), la tradición atomista puede retrotraerse a los empiristas británicos, y encuentra su *locus* clásico en la obra del Círculo de Viena, así como en el aporte de Russell, B. (1921). La idea básica que se pone en juego es que las propiedades semánticas de un símbolo están determinadas sólo por sus relaciones con cosas en el mundo no-lingüístico; en este sentido, Fodor (1998) sostiene que la opción atomista es la única que permite superar los problemas que suscita el holismo.

<sup>81</sup> Un molecularista sostiene, que si hay creencias que compartimos, debe haber otras creencias que también compartimos, pero niega que *todas* nuestras otras creencias tengan que ser compartidas para que podamos compartir alguna de ellas. Un ejemplo paradigmático de esta posición es Dummett, M. (1978). Frente a sus pretensiones, Fodor y Lepore (1992) sostienen que todo intento de articular una posición molecularista conduce, o bien al holismo, o bien a la aceptación de la distinción analítico/sintético (esto es, o bien a la idea de que el contenido está dado por todas nuestras creencias o bien a la idea de que está dado por definiciones analíticas). Sin embargo, por ejemplo, Perry (1994) sostiene que hay una forma débil de molecularismo que elude estas críticas.

Según la definición oficial ofrecida por Fodor y Lepore' (1992) el *holismo semántico* es caracterizado como una doctrina acerca de las *condiciones metafísicamente necesarias* para que algo tenga *significado* o *contenido*. De este modo, comienzan su discusión insertando el holismo en este contexto metafísico. A tal efecto, distinguen entre propiedades *anatómicas* (*anatomic*) y propiedades *atomistas*. Una propiedad es anatómica sólo en caso de que si algo la tiene, al menos otra cosa la tiene (por ejemplo, la propiedad de ser hermana: si soy hermana, no puedo ser la única persona en el mundo que instancie dicha propiedad). Si una propiedad no es anatómica, es *atomista*: se trata de una propiedad que, en principio, podría ser instanciada por una única cosa (por ejemplo, las propiedades expresadas por los predicados: “ser lo último que comí...” “ser la única...descubierta”).

A su vez, entre las propiedades anatómicas están aquellas que son *muy* anatómicas u *holistas*: una propiedad holista es una propiedad tal que si algo la tiene, *muchas* otras cosas la tienen (por ejemplo la propiedad de *ser un número natural*: esto es, de acuerdo con el tratamiento estándar, los números naturales son definidos por referencia a la relación de sucesor: nada es un número natural a menos que haya un número natural que sea su sucesor, cuya existencia dependerá metafísicamente, a su vez, de que otro número sea su sucesor, etc. La idea es pues, que si hay números naturales debe haber una *infinidad* de ellos – a diferencia de la propiedad de ser hermana, que no requiere una infinidad de individuos que la instancien, sino al menos de dos).

El punto que les interesa discutir en Fodor y Lepore (1992) es, específicamente, si las *propiedades semánticas genéricas* (esto es, propiedades cuya especificación puede considerarse como involucrando variables que ranguean sobre proposiciones, significados, contenidos) son anatómicas, de manera de explorar si hay una inferencia plausible desde su anatomismo a su holismo. En particular, van a ocuparse del *holismo del contenido*<sup>82</sup>, la tesis según la cual propiedades tales como *tener significado* (*o tener contenido intencional*) son holistas en el sentido de que ninguna expresión de un lenguaje puede tenerlas, a menos que otras expresiones (no sinónimas) también las tengan.

Ahora bien, dado que a juicio de Davidson (1995b) la verdad del holismo ha sido conspicuamente cuestionada por Fodor y Lepore (1992), se ve compelido a clarificar su

---

<sup>82</sup> Fodor y Lepore se concentran en esta variedad, puesto que sostienen que la mayoría de los argumentos que se han propuesto para el holismo semántico u holismo del significado son argumentos a favor del holismo *del contenido* (1992: 6).

compromiso con el holismo, sus variedades y las razones para abrazarlo<sup>83</sup>. Allí distingue entre las dos variedades que, en su opinión, completan su pintura del holismo de lo mental y que bautiza como: el holismo *intra-actitud* y el holismo *inter-actitud* (y que yo prefiero denominar como holismo semántico o del *contenido (lingüístico/mental)* y holismo *psicológico* o de las *actitudes proposicionales*, respectivamente)<sup>84</sup>.

Encontramos, en primer lugar, (II. 1) su *tesis holista del contenido* según la cual “la identidad de un pensamiento dado depende, en parte, de sus relaciones con otros pensamientos” (1995b/2004:11). Dado que esta variedad de holismo concierne a las relaciones entre varias creencias *dentro* de la categoría de creencia, Davidson la denomina “holismo *intra-actitud*”<sup>85</sup>(1995b/2004:13). La individuación de lo creído depende de otras cosas creídas, esto es, no podemos tener un pensamiento aislado porque cada pensamiento supone un contenido y dicho contenido está determinado por su relación con los contenidos de otras creencias atribuidas. Así, si las creencias son individuadas por sus relaciones con otras creencias, entonces algo no podría ser, digamos, la creencia de que las ballenas no son peces, a menos que esté relacionada con otras creencias acerca de cómo son las ballenas y los peces y, por supuesto, éstas a su vez tendrán ramificaciones adicionales. Esto es, sin un *background* suficiente de otras creencias, no podríamos tener la creencia de que las ballenas no son peces. En otras palabras: una creencia se identifica por su posición en un patrón de creencias; este patrón es el que determina de qué trata la creencia.

Ahora bien, lo que él tiene en mente, según veremos, “no es sólo la necesidad de que haya una *multitud* de pensamientos dentro de una misma actitud”, sino también “los *modos* en que estos pensamientos *deben estar relacionados entre sí, el tipo de estructura* que podemos esperar encontrar que constituye *la arquitectura de la creencia* o la arquitectura de la intención u otra actitud” (1995b/2004: 13-4). La pregunta más simple, dice Davidson, que podemos hacer

---

<sup>83</sup> Hasta donde sé, Davidson (1995b/2004) es el único artículo en el que Davidson desbroza explícitamente las variedades del holismo que suscribe.

<sup>84</sup> Las razones de la denominación que propongo tienen que ver con “espejar” el análisis estándar de los conceptos de actitud proposicional tal como aparece en los argumentos davidsonianos de los que me ocuparé en la segunda parte del presente trabajo, esto es, en términos de las nociones de adoptar una cierta *actitud* hacia (esto es, estar en una cierta relación con - tal como la de creer, desear, esperar, etc.) un cierto *contenido* (u objeto hacia el cual dicha actitud está dirigida). Cf. Russell (1921).

<sup>85</sup> Dada la ambigüedad de la noción de “pensamiento”, en el sentido de que alude tanto a la “actitud” como al “contenido”, cabe pues enfatizar que el holismo “intra-actitud” es relativo a la individuación del *contenido* (en tanto que la otra variedad que caracterizaré en breve, el holismo “inter-actitud”, concierne a las relaciones entre las distintas *actitudes*)

acerca del holismo, es si una criatura podría tener un único pensamiento, dado que si así fuera, sería plausible sostener que –incluso dado más de un pensamiento- cada pensamiento podría ser esencialmente independiente de otros pensamientos: podría no haber restricciones sobre las *combinaciones* de pensamientos que serían posibles. En ese caso, las relaciones entre los pensamientos podrían ser irrelevantes para el contenido del pensamiento. Frente a ello, veremos que las restricciones sobre sus combinaciones posibles resultan decisivas a la hora de arrojar luz sobre nuestro peculiar carácter de animales *racionales*<sup>86</sup>.

Según hemos adelantado, Davidson defiende una segunda variedad de holismo constitutivo: (II. 2) el holismo *inter-actitud* (que yo he preferido denominar, como holismo de las *actitudes proposicionales* u holismo *psicológico*<sup>87</sup>). En “The problem of objectivity” (1995b/2004:16) sostiene que el holismo “inter-actitud” es igualmente importante que el “intra-actitud” (u *holismo semántico* o *del contenido*, en mi denominación) y completa su pintura del holismo de lo mental. Esta segunda tesis concierne a las relaciones entre una categoría de pensamiento o juicio y otra (de allí, que sea “inter-actitud”): esto es, concierne, por ejemplo, a las relaciones entre los actos de creencia, deseo e intención.

Recordemos que, a juicio de Davidson, las diversas actitudes proposicionales requieren o dependen de la creencia para darle sustancia a su contenido. En efecto, Davidson sostiene que, por ejemplo la mayoría de nuestros deseos, depende de nuestras creencias. Así, nos dice, no

---

<sup>86</sup> La tesis holista del contenido ha sido objeto de varias críticas de las que me ocuparé al analizar su impacto en la filosofía de la mente davidsoniana. En esta instancia, simplemente me interesa advertir que es posible discernir, al menos, los siguientes argumentos en su contra: el argumento del aprendizaje del lenguaje, desarrollado por Dummett (1976 y 1991; una variante del mismo concierne a la posibilidad de comunicación, cf. Dummett 1973) y los argumentos presentados por Fodor y Lepore (1992) acerca de la inestabilidad o cambio total y en torno a la imposibilidad de contar con generalizaciones psicológicas. En efecto, Fodor y Lepore (1992:23) sostienen que las consecuencias drásticas en torno a, por ejemplo, la comunicación, el realismo científico y la explicación intencional, parecen depender de la tesis holista que relativiza las condiciones para el contenido a lenguajes o sistemas de creencia *enteros*, en el sentido de que Ud. no puede compartir ninguna de mis creencias a menos que comparta *todas* ellas. De este modo, creen que es esta versión fuerte, radical o global de la tesis holista la que vale la pena discutir. Por razones de orden expositivo y argumentativo, prefiero relegar, pues, para la sección siguiente mi crítica a la plausibilidad de leer en Davidson tal versión del holismo semántico, para favorecer una interpretación más moderada.

<sup>87</sup> La elección de esta adjetivación obedece a la peculiar concepción davidsoniana de lo mental, que abordaremos en la segunda parte de este trabajo y que, según veremos, importa una clara inflexión respecto de la tradición cartesiana. No obstante, cabe adelantar que dado que para Davidson el rasgo distintivo de lo mental será la intencionalidad (en el clásico sentido de Brentano, es decir, como direccionalidad hacia un objeto), las actitudes proposicionales – esto es, actitudes dirigidas hacia un contenido - serán el paradigma de lo mental o psicológico; de allí la calificación que propongo de esta variedad del holismo como “psicológico”. Una denominación alternativa – aunque en un espíritu consonante- a mi propuesta puede encontrarse en Lepore y Ludwig (2005: 212), quienes hablan de restricciones trans-modales, esto es, restricciones sobre las atribuciones de actitudes en diferentes modos psicológicos (creencia, deseo, intención, etc.)

querriamos hacer dinero a menos que creyéramos que eso nos pondría en posición de obtener cosas que necesitamos o valoramos; no deseáramos ir a la ópera a menos que, por ejemplo, creyéramos que lo disfrutaríamos. Y la lista continúa. Muchas actitudes, como estar complacida, enojada u orgullosa de que algo haya sucedido, dependen de la creencia verdadera de que ello haya sucedido. No podemos estar esperanzados de que ganaremos un premio, a menos que sepamos, o al menos creamos, que podemos ganarlo<sup>88</sup>.

Así, encontramos la tesis holista inter-actitud en Davidson cuando, por ejemplo, sostiene que:

Las actitudes proposicionales proporcionan un criterio interesante de racionalidad porque se dan solamente como un conjunto cohesionado. Es obvio que una rica pauta de creencias, deseos e intenciones es suficiente para la racionalidad; con todo, puede parecer demasiado riguroso hacer de esto una condición necesaria. Pero de hecho, la rigurosidad se basa en la propia naturaleza de las actitudes proposicionales, ya que tener una significa tener un amplio complemento. Una creencia requiere muchas creencias, y las creencias requieren de otras actitudes básicas tales como las intenciones y los deseos y, si estoy en lo cierto, que tengamos el don de lenguas. Esto no quiere decir que no haya casos dudosos. Sin embargo, el carácter intrínsecamente holista de las actitudes proposicionales hace que la distinción entre tenerlas y no tenerlas sea tajante (1982/2001: 96).

Lo que subyace, pues, a esta variedad de holismo parece ser la idea de que, en especial, las creencias y las preferencias del hablante resulten, en su totalidad, racionales en el sentido de ajustarse a los axiomas de la teoría de la decisión. Como hemos anticipado, el punto que aquí se pone en juego no es que sepamos como una cuestión de psicología empírica que los hablantes se ajustan a las normas de la agencia racional, sino que tales normas describen un patrón tal que si nos distanciáramos demasiado de él, no podríamos describir a alguien como agente, y consecuentemente, como creyente o hablante. En este sentido, la teoría de la agencia racional ubica restricciones *a priori* sobre los patrones de actitudes, probabilidades subjetivas, y órdenes de preferencia que asignamos a los hablantes y representa, de este modo, una restricción

---

<sup>88</sup> Todos los ejemplos aludidos se encuentran en Davidson (1995b/2004: 16). Es menos obvio, a su juicio, que no podría existir la creencia sin actitudes conativas, pero hay, no obstante, “un sentido claro que emerge a partir del estudio de la teoría de la decisión en torno a las probabilidades subjetivas, esto es, las creencias son finalmente extraídas de la preferencia o la elección, aunque –nos advierte– esto no es sostener con Hume que la creencia “es y debe sólo ser esclava del deseo” (1995b/2004: 16). Y finalmente, en virtud de su tesis de interdependencia entre el pensamiento y el lenguaje, a la que ya aludimos y que tematizaremos en el capítulo 5 más pormenorizadamente, sostiene que la creencia, el deseo y todas las otras actitudes proposicionales, dependen del lenguaje (1995b/2004:17).

poderosa sobre la interpretación. Estas restricciones son, entonces, holistas en el sentido en que deben ser impuestas sobre el sistema total de las actitudes del hablante.<sup>89</sup>

En suma, hay a mi juicio dos variedades de holismo de lo mental en el planteo de Davidson, que distinguen dos maneras diferentes en que una atribución de una actitud dada constriñe las atribuciones de otra. En primer lugar, hemos considerado el holismo del *contenido*, según el cual atribuir cualquier actitud requiere atribuir creencias y atribuir cualquier creencia requiere atribuir un rango adicional de otras creencias con contenidos relacionados. No obstante, de acuerdo con Davidson, sería impropio insistir en una lista particular o fija de creencias asociadas. En este sentido, no se trata de un holismo “rígido”, pues siempre puede haber explicaciones disponibles de por qué aparece una creencia y no otra, pero el punto es que sin tal patrón apropiado, no podemos sostener nuestra atribución. Esta variedad parece ser, pues, la posición que dota de “carnadura” a la tesis de la “centralidad de la *creencia*”.<sup>90</sup>

En segundo lugar, hemos considerado el holismo de las *actitudes proposicionales*, a la vera del cual encontramos restricciones sobre las atribuciones de actitudes en diferentes modos psicológicos, a saber, creencia, deseo, intención, etc. Así, atribuir un deseo a un agente requiere atribuir creencias, y viceversa; las actitudes vienen sólo en patrones de tipos entrelazados, cuyos contenidos están relacionados.

Tras este rodeo en torno al primero de los rasgos del concepto de creencia que hemos presentado, creo que es posible comenzar a avizorar que las conexiones entre el holismo y algunas preguntas profundas acerca de la comprensión que tenemos de nosotros mismos devienen íntimas y urgentes, como veremos en detalle al considerar el rol que desempeña dicho rasgo de lo mental en los argumentos de Davidson acerca del problema mente-cuerpo, acerca de la relación entre el pensamiento y el lenguaje y acerca de las perspectivas de primera y tercera persona de la atribución mental (sección II).

No obstante, antes de proseguir con nuestro itinerario y sumergirnos en la consideración del segundo de los rasgos del concepto de creencia, me parece conveniente detenernos en un punto que adelantamos en el capítulo anterior: la pretendida conexión entre el holismo y el **principio de caridad**, antesala de la *normatividad* de lo mental.

---

<sup>89</sup> Sin embargo, cabe anticipar que pese a que el sistema considerado como un todo debe ajustarse a dichas restricciones, las desviaciones locales son posibles.

<sup>90</sup> En la medida en que toda atribución de una actitud proposicional presupone adscripción de creencias, puesto que ninguna criatura puede tener ninguna actitud dirigida hacia un contenido a menos que tenga algunas creencias. Cf. Fodor y Lepore (1992: 18).

## 3.2. Caridad y Normatividad

### 3.2.1. La caridad como antesala de la normatividad.

En el capítulo anterior, mencionamos que el holismo y el principio de caridad (junto con la adopción de la actitud de sostener una oración como verdadera) son las distintas *restricciones* que Davidson le impone a la construcción de una teoría semántica para lenguajes naturales a los efectos de que la *estructura provista por una teoría tarskiana de la verdad* tenga aplicabilidad empírica como teoría de la interpretación, esto es, para poder relacionarla con un hablante o con una comunidad concreta de hablantes. Como he señalado, de la conjunción de ambos aspectos emerge, a mi juicio, la originalidad de su planteo a la hora de arrojar luz sobre las condiciones de la comprensión lingüística, aquella práctica en la que, característicamente, tratamos a alguien como a uno de *nosotros*.

Observamos, asimismo, el *rol* que desempeña el principio de caridad en el seno del procedimiento que adopta el intérprete radical, a saber: romper el círculo que liga el proceso de atribución de creencias con el proceso de asignación de significados a las emisiones del hablante. Davidson presenta este punto del modo siguiente:

Una fuente central de problemas es la forma en que creencias y significados conspiran contra la explicación de las emisiones. Un hablante que sostiene que una oración es verdadera [holds a sentence to be true] en una ocasión lo hace en parte por lo que él significa, o significaría, mediante una emisión de esa oración, y en parte por lo que él cree. Si sólo podemos proseguir en base a una emisión honesta, no podemos inferir la creencia sin conocer el significado, y no tenemos esperanza de inferir el significado sin la creencia (1974/1984: 142).

El desafío básico que debe afrontar la teoría davidsoniana de la interpretación es, pues, que no se puede asignar significado a las emisiones de un hablante sin saber lo que el hablante cree, en tanto que, a su vez, no se puede identificar creencias sin saber lo que significan las emisiones del hablante. De esto resulta, como ya hemos señalado, que debe ofrecerse una teoría de la creencia y el deseo y una teoría del significado al mismo tiempo. Davidson introduce, entonces, el principio de caridad para disipar el problema que suscita la interdependencia de la *creencia* y el *significado*: dado que la tarea de despejar ambas incógnitas es excesiva a la luz de

la magra evidencia provista por la actitud de sostener una oración como verdadera, sugiere que el intérprete mantenga “fijo un factor mientras se estudia el otro” (1975/1984). La caridad, precisamente, proporciona un método para mantener constante el factor *creencia* mientras se resuelve el factor *significado*; esto es, en virtud de dicho principio el intérprete presupone que la oración que el hablante sostiene como verdadera<sup>91</sup> generalmente expresa una creencia que es de hecho verdadera y que el hablante no tiene creencias distintas de las que él mismo tendría si estuviera en su lugar. Si suponemos que las creencias del hablante están de acuerdo con las nuestras y son de hecho verdaderas –al menos en los casos más simples y básicos- podemos emplear nuestras creencias acerca del mundo como guía hacia las creencias del hablante.

Ahora bien, el principio de caridad comprende tesis muy distintas y ha sido formulado de maneras muy diversas. En la contribución específica de Davidson lo encontramos caracterizado, recientemente, a partir de dos vertientes que denominó: el *principio de coherencia* y el *principio de correspondencia* (1991/2001).

El principio de *coherencia* lleva al intérprete a descubrir un cierto grado de consistencia lógica en el pensamiento del hablante, en tanto que el principio de correspondencia lleva al intérprete a considerar que el hablante está respondiendo a los mismos rasgos del mundo a los que el intérprete respondería en circunstancias similares. De este modo, “un principio dota al hablante de una *módica cantidad de lógica* y el otro lo dota en un cierto grado con lo que el intérprete considera que son *creencias verdaderas* acerca del mundo” (1991/2001: 211; mis itálicas). En otros términos, según el principio de coherencia la *caridad* es entendida como *racionalidad*: desde esta perspectiva, el intérprete está obligado a presuponer que el hablante es –más o menos- racional de acuerdo con sus propios *standards* de racionalidad. En consecuencia, el intérprete no puede evitar proyectar su propia lógica sobre el discurso, el pensamiento y la acción ajenos; la consistencia entre las creencias y entre éstas y la acción resulta un requisito esencial de la interpretación.

---

<sup>91</sup> Recordemos que la actitud de sostener una oración como verdadera, así como la actitud de preferir que una oración sea verdadera en lugar de otra, son ejemplos del tipo de actitudes que Davidson considera “no-individuativas”: esto es, aunque son psicológicas por su naturaleza, no distinguen los diferentes contenidos proposicionales que expresan emisiones distintas. De este modo, el supuesto de que podemos detectar tales actitudes no constituye, a su juicio, una petición de principio respecto de cómo dotamos de contenido a las actitudes, puesto que una relación como la de sostener o considerar/preferir verdadera, entre un hablante y una preferencia, es una relación extensional que podemos saber que se da sin saber lo que significa la oración en cuestión.

A su vez, según el principio de *correspondencia*, la *caridad* es asociada con la *verdad*: el intérprete está obligado a presuponer, asimismo, que la mayor parte de las creencias del hablante son verdaderas de acuerdo con sus propios *standards*<sup>92</sup>.

En suma, dada esta doble vertiente, el principio de caridad típicamente davidsoniano implica presuponer que: las creencias son *comunes* a hablantes e intérpretes y que la mayor parte de esas creencias compartidas son *verdaderas*. De esta suerte, si el supuesto hablante no es inteligible y mayoritariamente correcto, no es interpretable, esto es, no es realmente un hablante para el intérprete. Conforme a ello, las diferencias de tipo y grado entre los sistemas de creencias del hablante y el intérprete sólo pueden ser comprendidas a partir del acuerdo general. Del mismo modo, los errores sólo pueden ser comprendidos a partir de la verdad global<sup>93</sup>. En este sentido, afirma Davidson:

[C]ualquier creencia o conjunto de creencias particular acerca del mundo que nos rodea puede ser falsa, o falso. Lo que no puede suceder es que nuestra imagen general del mundo y nuestro lugar en el mismo esté equivocada, pues es esta panorámica la que informa el resto de nuestras creencias y las hace inteligibles, tanto si son verdaderas como si son falsas (1991/2001:214).

---

<sup>92</sup> Según algunos autores, esto es lo característico del aporte de Davidson. Véase, por ejemplo, Orlando, E. (1999). Sin embargo, dado el objetivo que persigo en el presente trabajo, esto es, iluminar la singular mirada sobre un *nosotros humano* que emerge de la pintura que Davidson traza de lo mental, considero que su aporte específico reside en la otra vertiente del principio de caridad, su concepción de la *racionalidad*, en la medida en que –según veremos– porta la clave para descifrar en qué consiste la singularidad de quienes somos “animales racionales”, a saber: la *normatividad* de lo mental. No obstante, soy consciente de que ha habido intentos de iluminar la normatividad a partir de la otra vertiente del principio de caridad. De hecho, McDowell (1998) liga de un modo estrecho *normatividad* y *verdad*. Allí defiende la idea de que el carácter desentremecedor de la verdad en la versión tarskiana que abraza Davidson hace que –contra la opinión de Rorty– la normatividad tenga lugar en la “superficie semántica”: esto es, a su juicio las relaciones de denotación y satisfacción están “normativamente cargadas” (en un sentido que, no obstante, McDowell deja sin especificar, aunque podríamos aventurar que tal vez esté poniendo en juego la manera en que dichas relaciones proporcionan *condiciones de corrección* para la aserción). No obstante, esta alusión a la normatividad que hace McDowell parece pretender arrojar luz sobre el modo en que la subjetividad está constitutivamente atravesada por la *objetividad*, en virtud de la conexión que guarda la intencionalidad con un orden extralingüístico. La otra vertiente de la caridad que es la que deseo explorar –la racionalidad– nos permitirá, en cambio, clarificar el tipo peculiar de normatividad que distingue los conceptos que empleamos para nuestra mutua comprensión a partir de la íntima conexión que establece entre subjetividad e *intersubjetividad*.

<sup>93</sup> A pesar de que en principio es posible presuponer que los miembros de una comunidad lingüística se comunican mutuamente sobre la base de creencias comunes pero erróneas, el punto de Davidson es que esto no puede constituir una regla general. En su opinión, aún cuando el concepto de creencia es el de un tipo de pensamiento que puede o no coincidir con lo que es el caso, los casos de falta de coincidencia sólo son comprensibles a la luz de la coincidencia general. Por consiguiente, como veremos en detalle en la segunda parte de este trabajo, tener el *concepto de creencia* implica ser consciente de la posesión de un conjunto mayoritariamente verdadero de creencias. Por esta razón, Davidson considera que la naturaleza misma de la creencia elimina la posibilidad del error masivo.

Volvamos, ahora, más pormenorizadamente sobre el rol que desempeña el principio de caridad en el procedimiento del intérprete radical. Las actitudes de sostener oraciones como verdaderas, que constituyen la evidencia inicial para el intérprete, estarán dirigidas a oraciones ocasionales (esto es, oraciones que el hablante a veces considera como verdaderas y a veces como falsas, en respuesta a eventos o condiciones de su entorno). Si consideramos las oraciones ocasionales, la pista principal que tenemos respecto de su interpretación consiste en las condiciones bajo las cuales los hablantes las sostienen como verdaderas. Dada nuestra identificación de dichas actitudes y dadas las condiciones bajo las cuales los hablantes sostienen como verdaderas oraciones ocasionales, podemos entonces establecer inductivamente leyes *ceteris paribus* de la forma:

(L) Para todos los hablantes  $S$ , tiempos  $t$ , *ceteris paribus*,  $S$  sostiene como verdadera  $s$  en  $t$  sii  $p$ <sup>94</sup>.

Llegado a esta instancia, el intérprete querrá pasar desde estas oraciones-L a algo más relevante para confirmar una teoría de la verdad, a saber, un enunciado que no sea acerca de las condiciones bajo las cuales los miembros de una comunidad lingüística *sostienen* oraciones *como verdaderas*, sino un enunciado acerca de las condiciones bajo las cuales esas oraciones *son verdaderas*. Dado que las oraciones-L no implican enunciados que establezcan tales condiciones, se deberá aumentar la teoría con un principio que nos permita zanjar este hiato, esto es, un principio que garantice la inferencia desde oraciones-L a oraciones que puedan ser empleadas para proyectar una teoría de la verdad interpretativa para todo el lenguaje:

(V) Para todos los hablantes  $S$ , tiempos  $t$ ,  $s$  es verdadera para  $S$  en  $t$  sii  $p$ .

Precisamente, el expediente propuesto por Davidson para zanjar el hiato es el *principio de caridad* que es introducido, pues, en virtud de que los patrones entre actitudes de sostener oraciones como verdaderas no son suficientes para fijar uno de los dos factores que determinan dichas actitudes (creencia y significado) a los efectos de resolver el otro. Sin embargo, cabe advertir que, así entendido recoge sólo una de las dos vertientes que, según vimos, componen el principio de marras. Es la introducción del *principio de correspondencia* la que parece estar

---

<sup>94</sup> Donde “ $S$ ” rangua sobre los miembros de una determinada comunidad lingüística, “ $p$ ” es reemplazada por una oración abierta que especifica rasgos contextuales y “ $s$ ” denota una oración ocasional.

motivada, pues, por la necesidad de saldar la brecha entre las que llamé oraciones-L y oraciones-V<sup>95</sup>.

Por su parte, la introducción del *principio de coherencia* parece responder a otra motivación, a saber: la teoría de la agencia racional que está en el corazón de la filosofía de la acción davidsoniana y que, según hemos dicho, se integrará con la teoría de la interpretación en el seno de la teoría unificada del pensamiento, el significado y la acción.

Una de las restricciones fundamentales que Davidson le impone a una teoría adecuada de la interpretación es que las interpretaciones de las oraciones y las asignaciones correspondientes de contenidos a sus creencias (y otras actitudes proposicionales) tornen a los hablantes *agentes racionales*. En su opinión, esto requiere que las acciones de los hablantes, incluyendo sus actos de habla, sean representados como mínimamente racionales en el sentido de que cada uno sea racionalizado por un par deseo-creencia, esto es, a la luz de algún fin del agente y alguna creencia acerca de cómo obtenerlo (cf. Davidson 1963/1980). Esta restricción nos permite empezar a avizorar el íntimo enlace que hay entre la *caridad* entendida como *racionalidad* y el rasgo del concepto de creencia que hemos analizado en el apartado anterior: el *holismo*. En efecto, podemos deslindar varios sentidos en que la *coherencia* o *consistencia* constituye la idea de *racionalidad*, a saber: tanto respecto de las creencias dadas otras creencias, en el sentido de que las creencias del hablante-agente deben ser consistentes (lo que hemos denominado “holismo *intra-actitud* o del contenido”)<sup>96</sup>, como también en el sentido de que la acción a ser explicada

---

<sup>95</sup>En este sentido, Lepore y Ludwig (2005) consideran tres formulaciones del principio de caridad sugeridas por distintos pasajes davidsonianos y que bautizan como: “Veracidad”, “Caridad” y “Acuerdo”. Según la primera, *ceteris paribus*, las actitudes de un hablante de sostener oraciones como verdaderas dirigidas hacia oraciones ocasionales son verdaderas. La segunda afirma que, *ceteris paribus*, las creencias de un hablante acerca de su entorno son de hecho verdaderas. La última sostiene que, *ceteris paribus*, el hablante y el intérprete están de acuerdo acerca del entorno del hablante. Allí Lepore y Ludwig consideran que la interpretación correcta es en términos de la segunda formulación, a saber: “Caridad”. No obstante, a su juicio, la Caridad resulta insuficiente para la tarea que le es asignada. Para una ampliación de este punto y del principio más fuerte (de la “Gracia”) que proponen en su reemplazo, véase Lepore y Ludwig (2005, cap.12).

<sup>96</sup>Recordemos que la idea es que los pensamientos tienen propiedades lógicas dados sus contenidos proposicionales: implican y son implicados por otros pensamientos. En virtud de que son las relaciones lógicas de un pensamiento las que parcialmente lo identifican como el pensamiento que es, los pensamientos no pueden ser totalmente incoherentes, porque si lo fueran, carecerían de toda posibilidad de ser identificados como un pensamiento en lugar de otro. El principio de caridad expresa esto diciendo: a menos que haya alguna coherencia en una mente, no hay pensamientos, y si un intérprete va a captar los pensamientos de otro, el intérprete debe descubrir un grado suficiente de coherencia para identificar esos pensamientos. Siguiendo esta línea Engel, P. (2003) señala un modo de comprender la estrecha relación entre la *caridad* y el rasgo del concepto de creencia que vimos en el apartado anterior: el *holismo*. Si el intérprete radical tiene que ser caritativo en el sentido de adscribir creencias racionales y coherentes a un hablante, entonces tiene que al menos adscribirle el dominio de ciertas inferencias básicas, tales como el *modus ponens*. Pero a los efectos de hacer inferencias uno debe tener más de una creencia, ha de tener

debe ser razonable a la luz de los deseos y creencias asignados y en el sentido de que los deseos y creencias atribuidos deben ajustarse entre sí (lo que hemos denominado “holismo *inter-actitud* o psicológico”). La idea subyacente es que nuestras creencias y preferencias deben ser racionales en el sentido de exhibir un patrón consistente que debe adecuarse a los axiomas de la teoría de la decisión. De manera que “es sólo contra el trasfondo de tal patrón que podemos identificar pensamientos” (Davidson 1975/1984). Así, según veremos, las normas de la agencia racional describen un patrón tal que si alguien se distanciara de él excesivamente, seríamos incapaces de describirlo como *agente* y, subsiguientemente, como *hablante* o *creyente*. La teoría de la agencia racional ubica, de este modo, restricciones *a priori* sobre los patrones de contenidos, actitudes, probabilidades subjetivas y órdenes de preferencia que podemos asignar a los hablantes y representa, así, una restricción poderosa sobre la interpretación. Tal es la íntima conexión que hay entre *caridad* y *holismo*.<sup>97</sup>

Ahora bien, habiendo esclarecido esta doble vertiente del principio de caridad, cabe preguntarnos: ¿qué estatus posee? Se ha sugerido que es posible distinguir entre diversas interpretaciones del principio de caridad davidsoniano<sup>98</sup>.

En primer lugar, encontramos la interpretación *fuerte*, según la cual el principio de caridad no sería un mero principio metodológico concerniente a la atribución de significados y creencias sino un principio *constitutivo* de los *significados* y las *creencias mismas*. Desde este punto de vista, se trata de un principio *ontológico* que podría formularse como sigue: el esquema conceptual del hablante *no puede ser* distinto del del intérprete: no puede haber dos esquemas

---

varias y aptas para ser relacionadas inferencialmente. Así, la racionalidad de la adscripción intencional obliga al intérprete a adscribir varias creencias relacionadas con una creencia dada (si es que el hablante ha de ser coherente).

<sup>97</sup> Como es sabido, Fodor y Lepore (1992) han examinado críticamente que la caridad implique el holismo de lo intencional. Como ya he señalado, mi intención en esta instancia no es la de una ponderación crítica, sino más bien exegética; es decir, no me voy a detener a evaluar la posibilidad del proyecto de la interpretación radical, sino que pretendo precisar los compromisos que asume Davidson para, luego, examinar sus implicaciones. Sin embargo, me parece relevante esbozar la crítica de Fodor y Lepore, dado que en ella aparecen algunas cuestiones conceptuales que voy a retomar. Ellos reconstruyen la argumentación davidsoniana a favor del holismo como una estrategia que posee un estilo trascendental; lo que la hace trascendental es que está destinada a mostrar que la interpretación radical sería metafísicamente imposible sin principios holistas, con la siguiente peculiaridad: tiene premisas que son simultáneamente epistemológicas y metafísicas. La idea básica es que es *metafísicamente constitutivo* de los hechos acerca del contenido que deben ser accesibles para alguien en la *situación epistemológica del intérprete radical* y que la interpretación radical es imposible a menos que se invoquen principios de caridad. El punto davidsoniano es (i) mostrar que ciertos principios de caridad son constitutivos de la adscripción de contenido; (ii) se sostiene que estos principios son intrínsecamente holistas y se sigue, así, el holismo de lo intencional. Por el contrario, para Fodor y Lepore (1992) no se sigue el holismo en virtud de que: (i) no se han dado buenas razones para sostener ni que la caridad sea un prerrequisito de la interpretación, (ii) ni a favor de que sea intrínsecamente holista.

<sup>98</sup> Sigo aquí las propuestas de Orlando, E. (1999) inspiradas, a su vez, en Moretti, A. (1993/2008).

conceptuales distintos. Por lo tanto, si deslindamos las dos tesis comprendidas en el principio de caridad, se estaría afirmando: por un lado, que *es constitutivo* del hecho de que una persona exprese oraciones significativas y tenga creencias el que esa persona sea (más o menos) *racional* de acuerdo con nuestros *standards*. Y, por otro lado, que *es constitutivo* del hecho de que una persona exprese oraciones significativas y tenga creencias el que esas emisiones y creencias sean *verdaderas* de acuerdo con nuestros *standards*, lo que implica que hay un sistema de creencias común<sup>99</sup> a hablantes e intérpretes y que la mayor parte de las creencias de ese sistema son verdaderas<sup>100</sup>.

En segundo lugar, cabe mencionar la interpretación *intermedia*. Desde esta perspectiva, la caridad es un principio que *regula* de la *atribución* de significados que, en tanto tal, señala una limitación esencial al intérprete. Ya no se trata de un principio ontológico acerca de los significados y las creencias, sino de un principio *gnoseológico* acerca del intérprete que podría formularse así: el intérprete *no puede creer* que el esquema conceptual del hablante sea distinto del suyo propio. Desde este enfoque, se sostiene que uno es absolutamente incapaz de comunicarse con quienes tuvieran un esquema conceptual distinto del propio. Asumir el rol de intérprete es asumir una restricción epistémica esencial: no es posible acceder a lo radicalmente distinto<sup>101</sup>.

---

<sup>99</sup> O, alternativamente, “esquema conceptual” común.

<sup>100</sup> Orlando, E. (1999:104-5) objeta que bajo esta interpretación, todo principio *gnoseológico*, concerniente al intérprete que realiza la atribución adquiere la categoría de un principio *ontológico*, concerniente a la *naturaleza* misma del significado o la creencia atribuidos. Ahora bien, en mi opinión, no cabe cuestionar este punto, al menos, al interior de la propuesta davidsoniana, dadas las razones siguientes. Tal como vimos a partir de Fodor y Lepore (1992) en la argumentación de Davidson hay un explícito solapamiento de aspectos metafísicos y epistemológicos involucrados en la perspectiva interpretacionista: es *metafísicamente constitutivo* del contenido el que un lenguaje (o un estado intencional) debe ser *radicalmente interpretable* (cf. Fodor y Lepore 1992, caps. 3 y 4). De este modo, ya no parece tener sentido escindir aspectos *gnoseológicos* de aspectos *ontológicos*. En otras palabras: dado que no hay significado o creencia independiente del proceso interpretativo, en el sentido de que es en virtud de que hay comunicación que hay creencia y significado (y no a la inversa), así como la otra de las restricciones impuestas por Davidson, el holismo de la atribución, implica el holismo semántico, que es *constitutivo* de la creencia y el significado mismos, la restricción que consideramos ahora, es decir, la caridad que debemos poner en juego para comprendernos mutuamente, implica que la racionalidad y la verdad son *constitutivas* de la naturaleza misma de la creencia o el significado. En suma: dado que, a mi juicio, no hay significado ni creencia sin proceso de atribución, es decir, toda creencia/significado es *constituída/o en el proceso mismo de (mutua) atribución*, ya no parece tener sentido deslindar tajantemente sus aspectos “gnoseológicos” de los “ontológicos”. Asimismo, para una consideración de la caridad como un principio a la vez epistémico y metafísico puede verse, *inter alia*, Glüer, K. (2006).

<sup>101</sup> En este punto, me gustaría introducir el siguiente reparo. Si nos tomamos en serio que no hay creencia ni significado por fuera del proceso de atribución, no parece lícito distinguir entre esta interpretación según la cual el principio de caridad sería meramente constitutivo de la atribución y la interpretación fuerte según la cual resulta constitutivo de la naturaleza misma de la creencia y el significado. Dado que toda creencia es creencia atribuida, no

Finalmente, encontramos la interpretación *débil* según la cual el principio de caridad simplemente concierne a la atribución de significados/creencias pero no es constitutivo de ésta: el principio de caridad es sólo uno entre los muchos principios que los intérpretes podrían elegir seguir<sup>102</sup>. De acuerdo con este sentido, el proceso de interpretación no involucra una unión profunda entre lenguaje y esquema conceptual sino un vínculo inicial y pasajero. Se trataría, en consecuencia, de un principio heurístico, puramente metodológico. Su formulación sería la siguiente: *Al empezar la interpretación*, el intérprete no puede creer que el esquema conceptual (tanto la lógica como el sistema de creencias, de modo de abarcar tanto a la caridad como racionalidad cuanto a la caridad como creencia verdadera) del hablante sea distinto del suyo propio.<sup>103</sup>

Ahora bien, ¿cuál es la interpretación que mejor se ajusta a las pretensiones davidsonianas? En mi opinión, la interpretación que es enfáticamente sugerida por el propio Davidson es la *fuerte*. En apoyo de que ello es así, cabe hacer referencia, por un lado, a lo que ya hemos señalado a propósito de la concepción davidsoniana acerca del holismo: no hay significados/ creencias independientemente de las prácticas de interpretación: la interpretación los constituye, esto es, no puede pensarse en significados/creencias como entidades previamente existentes cuyo intercambio entre hablantes e intérpretes hace posible la comunicación, sino que, a la inversa, *su naturaleza misma es posibilitada por la mutua práctica atributiva*. Así, leemos en Davidson: “el pensamiento depende de la comunicación” (1991/2001:209).

Por otro lado, cabe advertir el carácter *trascendental* del argumento mediante el cual introduce el principio de caridad. Dicho argumento puede reconstruirse como sigue:

- 1) La comunicación humana mediante el lenguaje es un hecho innegable.
- 2) La comunicación humana supone la interpretación radical.<sup>104</sup>

---

hay posibilidad de que el principio constitutivo de la atribución no permee a la naturaleza misma de la creencia y el significado. No obstante, me interesa retener este nivel de análisis a los efectos de tornar inteligible una de las acepciones de la *normatividad* que voy a proponer cuando, en el capítulo 4, analice su rol en el tratamiento davidsoniano del problema mente-cuerpo.

<sup>102</sup> Tal es la lectura propuesta, *inter alia*, por Vermazen, Wallace y Hacking. Desde esta perspectiva, Davidson es criticado por favorecer el principio de caridad sobre otros principios.

<sup>103</sup> La plausibilidad de esta interpretación reside, en opinión de Orlando (1999), en que podría ser conveniente empezar por asignar al hablante una lógica en común y verdad a todas sus emisiones; nada garantiza, en cambio, que lo asignado inicialmente se mantenga al finalizar el proceso de interpretación.

<sup>104</sup> Leemos en Davidson (1973/1984: 125): toda comprensión del habla de otro involucra interpretación radical.

- 3) Luego, existe un sistema de creencias universal y en su mayor parte verdadero que la hace posible.

De este modo, el mismo Davidson parece sostener explícitamente que el principio de caridad no es un principio que podamos optar entre otros principios, un principio empírico, descriptivo de la conducta lingüística, sino que lo introduce como un *principio a priori*<sup>105</sup> que constituye su condición de posibilidad.<sup>106</sup> En esta misma línea, Andrew Cutrofello (1999) subraya el carácter trascendental de la estrategia davidsoniana, que deslinda en el argumento “metodológico” que se presenta en “Radical Interpretation” (1973) y el argumento “ontológico” que se propone en “On the very idea of a conceptual scheme” (1974). El argumento “metodológico” mantiene que la interpretación no puede tener lugar a menos que los intérpretes traten de maximizar el acuerdo entre ellos y quienes interpretan. Por sí mismo, este argumento no puede –a juicio de Cutrofello– apoyar la extrapolación grandilocuente de Davidson de que todos los seres racionales deben compartir mayormente las mismas creencias. Después de todo, este argumento es estrictamente hipotético: dice sólo que *si* vamos a comprender a otros, *entonces* debemos interpretarlos de acuerdo con el principio de caridad. Cutrofello sugiere, pues, que resulta instructivo formular la cuestión en términos kantianos.

En su opinión, podemos decir que la justificación del principio de caridad carece de algo así como una *deducción trascendental* a los efectos de establecer su empleo *legítimo* en la

---

<sup>105</sup>A partir de enfatizar dicho carácter *a priori*, Ramberg advierte contra la confusión de los principios trascendentales de la interpretación radical con los principios de la interpretación efectiva, confusión que encuentra “fatal” (Ramberg: 1989:74).

<sup>106</sup>Fodor y Lepore (1992) también sostienen que el argumento de Davidson para la caridad tiene una forma *trascendental*. A su juicio, dicho argumento está fundado en la situación epistémica del intérprete radical (por tanto, para ellos será suficiente como objeción contra Davidson, socavar el supuesto de que la adscripción intencional deba entenderse por referencia a dicha situación) y puede reconstruirse como sigue:

- i. Si no se supone que el informante cree, en su mayoría, verdades, el aprendizaje del lenguaje desde la posición epistemológica del intérprete radical no sería posible.
- ii. El aprendizaje del lenguaje desde la posición epistemológica del intérprete radical es posible.
- iii. Debe suponerse que el informante cree, en su mayoría, verdades.

A juicio de Fodor y Lepore no hay una razón clara para aceptar la segunda premisa, y de este modo, el argumento trascendental para la caridad falla. A su vez, emprenden otra crítica a partir de la noción del intérprete omnisciente al que alude Davidson (1983/2001): si un ser omnisciente interpreta mis emisiones-creencias de manera tal que para él resultan verdaderas, entonces me malinterpretará toda vez que yo diga-crea algo falso. Esto ofrecería una reducción al absurdo de la idea de que la interpretación radical tendría que proceder mediante el ejercicio de la caridad. En breve: el intérprete omnisciente que emplea el método de la caridad para comprender a un hablante falible arribará a creencias falsas acerca de las creencias falsas del hablante. Omnisciencia y caridad son virtudes incompatibles.

experiencia efectiva. Recordemos que para Kant, demostrar que el entendimiento puede sólo pensar en términos de sus categorías *a priori* no es suficiente para mostrar que estas categorías son *legítimamente* aplicadas a la experiencia; la deducción *metafísica* tiene que ser complementada con la deducción *trascendental*. La invitación es a que pensemos, entonces, que el argumento metodológico de Davidson es como la deducción metafísica en el sentido de que muestra sólo cómo debemos interpretar; no muestra, en cambio, si interpretar del modo en que lo hacemos nos conducirá a hacer juicios adecuados acerca de los puntos de vista de otros. De este modo, la justificación metodológica del principio de caridad debe complementarse con algo análogo a la deducción trascendental.

Como sucede con la deducción trascendental kantiana, la tarea de este nuevo argumento es establecer el *quid juris* de la aplicabilidad del principio a nuestras experiencias efectivas –en este caso, nuestras experiencias de interpretar a otros-. Aquí la analogía entre Kant y Davidson resulta especialmente interesante porque la “deducción” de Davidson –lo que Cutrofello llama el “argumento ontológico”- es virtualmente idéntica a un argumento que Kant presenta en la “Estética Trascendental” de la primera *Crítica*. El argumento ontológico de Davidson trata de mostrar que todos debemos compartir la mayoría de las mismas creencias acerca del mundo, *porque hay sólo un mundo acerca del cual tenemos creencias*. Si sólo hay un mundo, todos los puntos de vista acerca del mundo son commensurables. Y si es así, se ha establecido el *quid juris* del principio de caridad<sup>107</sup>. No obstante, como ya he dicho, no es mi intención aquí evaluar si las pretensiones trascendentales del principio de caridad son adecuadamente argumentadas por Davidson. Mi interés ha sido precisar qué entiende por *caridad* para mostrar la manera en que, en mi opinión, en dicha noción se prefiguran, se anticipan conceptualmente, los diversos sentidos del próximo (y crucial) rasgo de la noción de creencia que paso a considerar: su *normatividad*.

---

<sup>107</sup> En efecto, en el argumento en contra del relativismo conceptual Davidson (1974) comienza por relacionar íntimamente los conceptos de *lenguaje, creencia, verdad y racionalidad* en el sentido siguiente: el intérprete radical no está –ni puede estar- en posición de creer que los hablantes que interpreta tienen creencias radicalmente diferentes de las suyas (ni masivamente erróneas) ni reglas inferenciales completamente inaccesibles. De ello se concluye que es imposible creer que existan múltiples esquemas conceptuales. Según esto, la tesis relativista extrema que sostiene la existencia de esquemas conceptuales inconmensurables es ininteligible. Más aún, se dirá que la idea misma de esquema conceptual lo es, en la medida en que supone una dicotomía concepto-objeto o esquema-contenido, que constituye para Davidson, el denominado “tercer dogma del empirismo”.

### 3. 2.2. Normatividad

Según veremos en la segunda parte de este trabajo, este rasgo del concepto de *creencia* resulta central en la contribución de Davidson a los problemas de la filosofía de la mente de los que voy a ocuparme. No obstante, en mi opinión, el tipo de *normatividad* que Davidson tiene en mente no ha recibido una articulación explícita y clara al interior de su propuesta. Recordemos que, según él mismo concede, lo ha “*dicho acerca de las normas que gobiernan nuestras teorías de la atribución intensional es, en los puntos cruciales, crudo, vago o demasiado rígido*” (1980b/2004:166, mis itálicas).

Como ya hemos dicho, frente a la “vaguedad”, la “crudeza” y la “rigidez” con las que el propio Davidson asume haber tratado a la *normatividad de lo mental*, se impone entonces la tarea de elucidar de un modo preciso este rasgo que porta la clave para captar la peculiaridad de un *nosotros humano*. En este sentido, cabe subrayar que a pesar de la vasta cantidad de trabajos publicados recientemente sobre este tópico, no hay un acuerdo razonable entre los autores acerca de cómo entender la *normatividad de lo mental*. Es por ello que, como ya hemos señalado, uno de los objetivos rectores del presente trabajo es el de clarificar en qué sentido atribuir estados intencionales es atribuir estados normativos y en qué medida responder a dicha pregunta involucra responder a la cuestión respecto de qué nos hace ser quienes somos.

A tal efecto, a continuación, voy a proponer un marco general en el que trazaré una distinción entre sentidos de *normatividad* que voy a retomar en la última parte del presente trabajo. De este modo, tras haber examinado en detalle el rol que desempeña la *normatividad de lo mental* en los argumentos que ofrece Davidson en torno a los tres problemas que vamos a considerar en la sección II (el problema mente-cuerpo, el problema de la relación entre lenguaje y pensamiento y el problema de las perspectivas de primera y tercera persona de la atribución mental) me abocaré a examinar bajo cuál/es de los sentidos de *normatividad* aquí propuestos resulta apropiado ubicar la contribución davidsoniana a la hora de arrojar luz sobre un *nosotros humano*.

Tradicionalmente, se considera que la esfera de la *normatividad* puede descomponerse en las dos dimensiones siguientes: la *deóntica* (recogida por nociones tales como “correcto”, “incorrecto”, “deber”, “obligación”, “permiso”, etc.) y la *evaluativa* (recogida por nociones tales

como “bueno”, “malo”, etc.)<sup>108</sup>. De este modo, se considera que la normatividad “es un rasgo común a ambos lados de la distinción deóntico/evaluativo” (Dancy 2000: vii).

Comencemos, entonces, por precisar la dimensión *deóntica* de la normatividad; para ello voy a valerme, en primer lugar, del célebre trabajo de Georg H. von Wright (1963, cap. I). En *Norm and Action*, él distingue entre tres tipos de normas, a saber: las prescripciones, las directivas y las reglas.

En primera instancia, von Wright caracteriza las *prescripciones* del modo siguiente: tienen su fuente en una *autoridad* que quiere que los sujetos adopten una cierta conducta; a los efectos de que su voluntad sea conocida, la autoridad promulga la norma y añade, en vistas de que sea efectiva, una sanción o amenaza de castigo. En términos generales, pues, las prescripciones son órdenes o permisos dados por alguien en posición de autoridad a alguien en posición de sujeto (von Wright cita a las leyes del estado como ejemplo paradigmático de prescripciones).

En segundo lugar, encontramos las *directivas o normas técnicas*. Este tipo de norma concierne a los *medios* que han de usarse para la obtención de un cierto fin. De este modo, resultan ilustraciones típicas de las directivas las así llamadas “instrucciones de uso”, en las que se presupone que la persona que las sigue apunta a la consecución de un fin o resultado. Su formulación estándar se da mediante oraciones condicionales en cuyo antecedente se menciona algo que se quiere y en cuyo consecuente figura algo que debe (o no debe) hacerse a tal efecto. Veamos un ejemplo de directiva: “Si Ud. quiere hacer habitable el refugio, debe calefaccionarlo”. ¿Se trata de una oración descriptiva o prescriptiva? se pregunta von Wright. Nos dice que, propiamente hablando, no es ni una cosa ni la otra y nos invita a que la comparemos con la oración que dice “Si el refugio va a hacerse habitable, debe calefaccionarse”. A esta última no duda en calificarla de “descriptiva”: dice que calefaccionar el refugio es una condición necesaria para hacerlo habitable y esto es verdadero (o no) con independencia de si alguien quiere hacer habitable el refugio y apunta a ello como un fin. De este modo, von Wright nos previene de una posible confusión: sería un error identificar las directivas con proposiciones que establecen que algo es (o no) condición necesaria para otra cosa.<sup>109</sup>

---

<sup>108</sup> Puede encontrarse esta presentación canónica, *inter alia*, en Dancy, J. (2000). Véase, también, Engel, P. (2001).

<sup>109</sup> von Wright reserva para estas últimas la calificación de “anancásticas”. Cf. (1963:10). No obstante, concede que hay una conexión esencial (lógica) entre ambas: al dar la directiva “si Ud. quiere hacer habitable el refugio, debe calefaccionarlo” está (lógicamente) presupuesto que si el refugio no es calefaccionado no se tornará habitable.

Finalmente, encontramos el tercer tipo de norma: las *reglas*. El ejemplo prototípico que nos ofrece von Wright son las reglas de un juego. Jugar un juego es una actividad humana que se lleva a cabo bajo patrones estandarizados que podemos llamar “movidas” del juego. La idea es que las reglas del juego *determinan* estas movidas y, mediante ellas, el juego mismo y la actividad de jugarlo. Desde la perspectiva del juego mismo, las reglas determinan, nos dice von Wright, cuáles son las movidas *correctas* y, desde la perspectiva de la actividad, cuáles son las movidas *permitidas* (de este modo, las movidas que no son correctas están *prohibidas* para los jugadores del juego y aquella movida que es la única correcta en una cierta situación, es la *obligada* cuando se está jugando el juego).

Pasemos, ahora, a considerar la dimensión *evaluativa* de la *normatividad*. A los efectos de precisar la relación entre dicha dimensión y la esfera deóntica de la normatividad, propongo que nos detengamos brevemente en el debate sobre normas y valores entablado entre Putnam y Habermas (2008). Si bien está precedido por un acercamiento progresivo entre las filosofías de ambos autores desde mediados de los años ochenta, el debate se abre propiamente cuando Putnam pronuncia en 1999 la conferencia “Valores y normas”. Allí expone su rechazo de la dicotomía que, a su juicio, Habermas postula entre los “valores” y las “normas”.

Según Habermas (1998), a la luz de las normas se puede decidir lo que estamos obligados a hacer, en tanto que en el horizonte de los valores podemos decidir qué comportamiento es más recomendable. Asimismo, mientras que las normas que merecen reconocimiento obligan a sus destinatarios sin excepciones y en igual medida, los valores expresan que merecen preferencia los bienes que en una comunidad determinada son considerados deseables. En tanto que las normas se cumplen en el sentido de una satisfacción de expectativas de comportamiento generalizadas, los valores sólo pueden realizarse mediante acciones orientadas a un fin. Además, las normas se nos presentan con una pretensión de validez binaria (y son o bien válidas o no lo son); por el contrario, los valores establecen relaciones de preferencia que significan que determinados bienes son más atractivos que otros. De este modo, la validez *deóntica* (*Sollgeltung*) de las normas tiene el sentido *absoluto de una obligación incondicional y universal*, mientras que el que unos valores sean atractivos tiene el sentido *relativo* de una estimación o evaluación de bienes adoptada en ciertas culturas y formas de vida. En suma: pareciera, dice Putnam, que el tratamiento habermasiano de las normas es “kantiano” en el sentido de que la fuerza vinculante de las normas es equiparada con la fuerza vinculante del

pensamiento racional, en tanto que los valores son productos sociales contingentes que varían conforme lo hacen los diferentes “mundos de la vida”.

Frente a esta posición, Putnam defiende la objetividad de las orientaciones de valor (tanto frente al no-cognitvismo como frente al relativismo) y sostiene que el que se acepte que los valores pueden discutirse racionalmente no es de ningún modo lo mismo que aceptar un autoritarismo con respecto a los valores. Aquí, encuentro un fuerte punto de contacto con la propuesta davidsoniana en torno a la objetividad de los valores.

En efecto, en “The objectivity of values” Davidson (1995c/2004) sostiene que lo mismo que rige para las creencias, rige para los valores. Antes de que podamos decir que dos personas están en desacuerdo acerca del valor de una acción o de un objeto, debemos estar seguros de que es la misma acción o el mismo objeto y son los mismos aspectos de esas acciones u objeto lo que tienen en mente: las consideraciones que prueban la disputa genuina revelarán también los criterios compartidos que determinan dónde yace la verdad. En este sentido, nos recuerda que decir cuál es el contenido de un determinado juicio moral es cuestión de interpretación y que las actitudes de una persona tienen contenido –son interpretables- sólo si esa persona está en comunicación con otras: esto es, sólo los intérpretes pueden ser interpretados. Ahora bien, ¿hasta qué punto estas consideraciones se aplican a la objetividad de los valores? Para Davidson, en el sentido de que la comprensión mutua se logra sobre la base de una convergencia de nuestros “valores básicos” – a los que también califica, sugerentemente, como “valores iluminados” (*enlightened values*)- (1995c/2004:49). La comprensión depende, nos dice, de encontrar un suelo común; dado suficiente suelo común, podemos comprender y explicar las diferencias, podemos criticar, comparar y persuadir. Si comprendemos sus palabras, “existe un suelo común, compartimos ya un modo de vida” (1995c/2004:51) La importancia de este trasfondo de creencias y valores compartidos es, en opinión de Davidson, que nos permite dotar de sentido a la idea de un estándar común de corrección e incorrección, de verdad y falsedad. De este modo, “una apreciación de qué constituye tal convergencia o acuerdo también muestra que los juicios de valor son tan verdaderos o falsos como lo son los juicios fácticos”. Y concluye, Davidson: “los valores están enraizados en las cosas” (1995c/2004:52).

Ahora bien, dado este marco general ¿bajo qué sentido de *normatividad* resultará adecuado explicitar, articular, el tipo de normatividad que, según Davidson, resulta distintivamente humana, esto es, *la normatividad de lo mental*? En rigor, dado el itinerario que

nos hemos impuesto, lo esperable es que nuestra respuesta surja tras considerar su impacto en los tres problemas de filosofía de la mente de los que vamos a ocuparnos. No obstante, cabe hacer algunas precisiones preliminares, para relegar hasta la última parte de este trabajo nuestra propuesta de elucidación.

Comencemos por el primer tipo de normas: *las prescripciones*. Estos candidatos no parecen proporcionar una clave adecuada para comprender la normatividad de lo mental al interior de la propuesta de Davidson. Según veremos en detalle en el capítulo 5, él rechaza explícitamente aquellos abordajes del lenguaje y la comunicación que reposan en la idea de que para nuestra comprensión mutua debemos seguir *prescripciones o convenciones* particulares ligadas a las palabras o a los conceptos que determinen su uso correcto.<sup>110</sup>

Así, por ejemplo, en “Communication and Convention” Davidson intentará hacernos ver –contra la idea de David Lewis– que no es en absoluto una trivialidad o un lugar común que hay –o deba haber– convenciones en el lenguaje, sino que se trata de una tesis filosófica cargada con contenido teórico. En primer lugar, recordemos que Davidson cambia el foco de atención del concepto de *un lenguaje* al de *comunicación mediante el lenguaje*. Lo que resulta significativo son las emisiones, esto es, acciones particulares de hablantes particulares en tiempos particulares. Más aún, la “noción de significado” depende “enteramente de casos exitosos de comunicación” (Davidson y Glüer 1995:8). El significado, por tanto, es de una naturaleza fundamentalmente intersubjetiva. En segundo lugar, para ser filosóficamente interesante la “trivialidad” relativa a las convenciones lingüísticas requiere ser investida con una cierta necesidad, esto es, ser leída como la tesis de que la comunicación mediante el lenguaje es *esencialmente* convencional. La pregunta filosóficamente interesante no es la cuestión empírica de si la gente de hecho habla entre sí de modos convencionales, sino si las convenciones son necesarias para la comunicación mediante el lenguaje.

Como es sabido, Davidson responde a esta cuestión argumentando que “la convención no ayuda a explicar qué es básico para la comunicación lingüística” (Davidson 1984b). Desde la perspectiva davidsoniana, está claro que no es necesario para la comunicación que haya convenciones compartidas que asignen significados a las palabras –y oraciones–; esto es, para

---

<sup>110</sup> La equiparación que estoy haciendo entre *prescripciones* y *convenciones* amerita, por lo menos, la siguiente aclaración. Las convenciones o hábitos, nos dice von Wright, se comportan como prescripciones en tanto influyen en la conducta ejerciendo “presión normativa” sobre los miembros de la comunidad. No obstante, en el caso de las convenciones sólo podemos hablar de “autoridad” en tanto nos referimos a la comunidad misma y, de esta suerte, las convenciones pueden caracterizarse como prescripciones “anónimas” (cf.1963:9)

que dos hablantes se comuniquen entre sí no es necesario que haya una convención que asigne el mismo significado a una expresión con independencia de cuál de los hablantes la emita. Todo lo que importa para la comunicación exitosa es que, en vistas a una emisión específica, el oyente asigne el significado que el hablante pretendió (*intended*). Kathrin Gluër (2001) subraya que para Davidson significar algo requiere que uno siga una práctica que puede ser entendida por otros. Si bien “práctica” es una palabra con carga wittgensteiniana, ella sostiene, no obstante, que todo lo que Davidson quiso decir es que es sin duda empíricamente necesario un grado considerable de consistencia en el uso de las palabras para que seamos comprendidos. La única norma relevante es práctica y se sigue del hecho de que usualmente queremos ser comprendidos.

Veamos, entonces, si el segundo tipo de normas, las *directivas*, permiten arrojar luz sobre la peculiar normatividad de lo mental que Davidson propone. Tomemos un ejemplo típico: las reglas para cocinar un determinado plato. Como hemos dicho, las directivas nos dicen por qué medios alcanzar ciertos fines, dados los hechos subyacentes de la naturaleza, esto es, dadas las subyacentes leyes naturales. Así, para la preparación de nuestro plato (un tiramisú, por caso) tendremos que atender, por ejemplo, a las condiciones del batido, circular y continuo, para que las claras adquieran “punto nieve”, respetar una determinada temperatura de cocción, etc. Como se advierte en von Wright (1963:9ss) las directivas deben, entonces, su validez, su carácter de guías para acción, a relaciones previas e independientes de cualesquiera de los imperativos hipotéticos derivados de ellas, esto es, deben su validez a las *leyes naturales*. Ahora bien, Davidson nos recuerda la lección de Wittgenstein respecto de que debe haber un modo de interpretar que no es cuestión de seguir directivas o instrucciones, en el sentido de que el tipo de conocimiento que describe una teoría del significado/creencia no es irrelevante pero nunca puede instruirnos acerca de cuándo aplicarla. No hay ninguna directiva en sentido estricto, como opuesta a meras máximas y generalidades metodológicas. De este modo, tampoco obtenemos una clarificación de la normatividad de lo mental que -según veremos en detalle en el capítulo 4- para Davidson abre, precisamente, un hiato respecto de las leyes naturales (que se suponen subyacentes a las normas entendidas como directivas).

Pasemos, entonces, a considerar nuestro tercer candidato: veamos si puede resultarnos de ayuda concebir la normatividad propia de lo mental en términos de *reglas*.

Como ya hemos señalado, los juegos son los casos paradigmáticos de actividades esencialmente constituidas por reglas. Sus reglas no dependen, como en el caso de las directivas,

de leyes naturales previas e independientes; por el contrario, las reglas de los juegos son claramente “creaciones humanas, hechas, no encontradas” (Baker y Hacker 1985:63). En este sentido, son *arbitrarias*. No hay algo tal como el juego correcto o el conjunto correcto de reglas; podemos crear incontables conjuntos de reglas, incontables juegos.

La segunda característica central de las actividades gobernadas por reglas, como jugar ajedrez, es que no sólo es imposible jugar ajedrez *sin reglas*, sino que es imposible jugar ajedrez sin las *reglas del ajedrez*. Tales actividades están *definidas* o *constituidas* por sus reglas: si cambiamos las reglas, automáticamente cambiamos la actividad.

Ahora bien, contempladas conjuntamente estas dos características, la *arbitrariedad* y la *constitutividad*, tornan complejo considerar la normatividad de lo mental como *las reglas de juego* de la práctica de nuestra comprensión mutua. La analogía entre las reglas de los juegos y las relaciones básicas constitutivas de nuestros conceptos parece quebrarse en aspectos decisivos: mientras que hay diferentes juegos que podemos jugar, hay un único sistema de conceptos. Recordemos que no tiene sentido hablar de sistemas conceptuales (radicalmente) diferentes (cf. Davidson, 1974). Tener conceptos es básicamente tener *nuestros* conceptos. Son estos mismos conceptos los que son una condición para tener creencias, deseos, intenciones y los medios para expresarlos a través del lenguaje. Son, por tanto, una condición para jugar o inventar todo juego, y una condición para tomar decisiones tales como la de jugar ajedrez ahora o dejar de jugarlo y jugar, en cambio, al truco.

Puesto en términos de “juegos”, aquí ya no parece tener sentido decir que podríamos jugar un juego distinto al *intencional*; este es el único “juego” que no podemos dejar de jugar, no importa qué hagamos. Su validez es condición para la posibilidad del significado y de la intencionalidad y como tal, es *sui generis*. En consecuencia, si se la piensa en términos de la analogía con los juegos, parece haber dificultades a los efectos de considerar la normatividad de lo mental como las reglas del juego de la práctica de nuestra mutua comprensión.<sup>111</sup> No obstante, tal vez podamos arrojar algo de luz sobre la *reglas* como normas a partir en de la clásica distinción trazada por Searle (1969, I, cap. 2). Como es sabido, en *Speech Acts* Searle intenta clarificar la distinción entre las reglas *regulativas* y las reglas *constitutivas*. Su idea es que las primeras *regulan* formas de conducta previamente existentes (ejemplo de ellas son las reglas de

---

<sup>111</sup> A propósito de la normatividad semántica, en su (2001) Kathrin Gluër hace uso de la analogía wittgensteiniana con los juegos para sugerir que el uso significativo de las expresiones lingüísticas carece de los rasgos característicos requeridos por las actividades esencialmente gobernadas por reglas.

etiqueta que gobiernan relaciones interpersonales que existen independientemente de tales reglas), en tanto que las segundas no regulan meramente sino que *constituyen* o *definen* nuevas formas de conducta (en este sentido, las reglas del fútbol, por ejemplo, crean la posibilidad misma de jugar ese juego). Quizás aquello que la normatividad de lo mental está subrayando es que se trata de reglas constitutivas de un tipo especial, en el sentido, de que su peculiaridad consiste en que sean *estas* reglas. Oportunamente, examinaré más detenidamente este punto.

Por su parte, Davidson (1998/2005: 326) nos dice cuál es su objeción central a la idea de concebir la normatividad en términos de convenciones o reglas: él no cree que haya normas “inherentes al lenguaje mismo”. Desde luego, concede que hay reglas en nuestra práctica lingüística, pero son “enteramente contingentes, dependientes de “*valores ulteriores* “(*further values*), no de valores intrínsecos al lenguaje. En este sentido, nos advierte que “las oraciones-T no son normativas en sí mismas: no nos dicen, por ejemplo, qué condiciones de verdad debemos asignar a una oración, ni nos dicen cuándo estaríamos en lo “correcto” al afirmarla, a menos que “correcto” aquí sólo signifique “verdadero”.<sup>112</sup> Ahora bien, si las normas que nos hacen singulares no proceden de la “superficie semántica” ¿de dónde vienen? ¿Tenemos que internarnos en la dimensión evaluativa de la normatividad para explicitar la dimensión deóntica peculiar de las reglas que constituyen el singular “juego” de nuestra mutua comprensión? ¿Cuáles son, entonces, esos valores ulteriores de los que depende la normatividad que advertimos en la “superficie” de dicha práctica?

En “Expressing evaluations” Davidson sostiene que hay un punto de partida: aquel que concierne a los *valores*, y que no sería errado decir que las actitudes evaluativas y las acciones que las revelan forman el “*fundamento de nuestra comprensión* del habla y la conducta de los otros” (1984c/2004: 35-6). El método, nos recuerda, supone una sociedad de dos: requiere que veamos a los otros como a nosotros mismos en cuestión de coherencia global y corrección, habitando un mundo como el nuestro. No se trata de una “política” (*policy*), en el sentido de ser

---

<sup>112</sup> A mi juicio, aquí es posible comenzar avizorar una pista para pensar que, precisamente, el tipo de normatividad que constituye el “giro esencial” (*essential twist*) en el que piensa Davidson como distintivo de lo mental procede de las *restricciones* que hay que imponerle a la estructura de tarskiana de la verdad. En este sentido, a mi juicio la normatividad de lo mental no surge bajo el signo de Tarski ( y contra McDowell (1998) no creo que la normatividad en sentido davidsoniano esté en la “superficie semántica”) sino a la vera del diálogo de Davidson con Quine y Wittgenstein que mencionamos en el cap. 2 : creo que hay que atender especialmente a la perspectiva del intérprete radical ( la operación por la que Davidson ubica en el espacio público al significado y la creencia) y, más específicamente, a la caridad entendida en términos del “principio de coherencia” para precisar el sentido de normatividad que nos hace peculiares.

una entre muchas posibles políticas exitosas. Es la única política disponible si queremos comprender otra gente. De este modo, cuando preguntamos de dónde vienen las normas que cada uno de nosotros aplica para comprender a los otros, la respuesta es que no pueden ser derivadas de una fuente externa a nosotros mismos, porque cualquier intento de considerar alternativas nos reconduce al proceso mismo de interpretación en el que necesariamente empleamos nuestras propias normas. No hay ningún modo de salirse de estas normas para comprobar si estamos en lo cierto, no más de lo que podemos comprobar si el cilindro de platino e iridio que se encuentra en la Oficina Internacional de Pesos y Medidas de Sèvres pesa un kilogramo. No podemos ponernos de acuerdo sobre la estructura de las oraciones o los pensamientos que usamos para explorar los pensamientos y los significados de los demás, pues el intento de alcanzar un acuerdo así simplemente nos remite al proceso mismo de interpretación del que depende todo acuerdo. No tiene ningún sentido cuestionar la adecuación de esta medida o buscar una norma o medida ulterior<sup>113</sup>. Aquí es donde, según veremos, llegamos a las fuentes últimas de la diferencia entre la comprensión de las mentes y la comprensión del mundo como algo físico.

A partir de examinar el rol que desempeña la *normatividad de lo mental* en las respuestas davidsonianas al problema mente-cuerpo, el problema de la relación entre el lenguaje y el pensamiento, y el problema de las perspectivas de la atribución mental, veremos, pues, si resulta necesario internarnos en su dimensión *evaluativa* para captar la peculiaridad de un *nosotros humano*. Quizás debamos atender al tipo de *valores* que han de resultar básicos como condición de posibilidad de la comprensión; tal vez podamos encontrar allí una clave adecuada para comprender el “giro especial” [*special twist*]” (1998/2005: 319) de la normatividad que es característicamente humana, aquello que hace que las reglas constitutivas del “juego” intencional resulten especiales. A tal efecto, creo que tendremos que tener presente la relevancia que Davidson le otorga al tercer rasgo del concepto de *creencia* que paso a considerar, la *triangulación*, dado que en mi opinión porta la clave para entender en qué sentido “una *comunidad* de mentes está en la base del conocimiento y proporciona la medida de todas las cosas” (1991/2001: 218). Pasemos, pues, a presentar su peculiar externismo.

---

<sup>113</sup> Davidson nos recuerda que: “la caridad no es una opción, sino una condición para tener una teoría practicable [...] La caridad nos es impuesta; nos guste o no, si queremos comprender a los demás debemos darlos por acertados en la mayor parte de los asuntos. Si somos capaces de producir una teoría que reconcilie la caridad y las condiciones formales para una teoría, hemos hecho todo lo que puede hacerse para asegurar la comunicación. No hay nada más que sea posible, ni hace falta nada más” (1974/1984:197).

### 3.3. El externismo de Davidson: la triangulación

Al comienzo de “Externalisms”, Davidson (2001c) sostiene que el *externismo* es la única alternativa al *subjetivismo*. La idea que está detrás de esta afirmación es que si no queremos ser escépticos acerca de la posibilidad del conocimiento del mundo externo o de las otras mentes, debemos rechazar la tesis de que todo el conocimiento del mundo depende de objetos que están directamente presentes en mentes individuales (esto es, objetos tales como *sense data*, impresiones, ideas o proposiciones, es decir, objetos que podrían ser exactamente como se nos presentan aún cuando el mundo fuera muy diferente). De lo que se trata, entonces, es de rechazar que tales objetos sean tomados como intermediarios epistémicos entre nuestras mentes y el resto del mundo. *Contrario sensu*, estaríamos ante teorías que “no sólo son individualistas sino también subjetivistas: el mundo de cada persona resulta construido a partir de material disponible en la conciencia, conectado sólo indirectamente con el mundo”.<sup>114</sup> Frente al subjetivismo, que se compromete con una conexión extrínseca entre el pensamiento y el mundo, el externismo davidsoniano se propone abrazar una conexión “intrínseca, no inferida, no construida, ni descubierta, sino que está allí desde el comienzo (*from the start*)” (Davidson 2001c:2). Examinemos, entonces, la “geometría” de la *triangulación*, su original tesis externista.

En “The conditions of thought” Davidson (1989c) se pregunta cuáles son las condiciones necesarias para la existencia del pensamiento; en particular, para la existencia de personas que tengan pensamientos. Allí sostiene que no podría haber pensamientos en una mente si no hubiese otras criaturas pensantes con las que dicha mente compartiese un mundo natural.

Su idea es que para tener una creencia no es suficiente discriminar entre aspectos del mundo, comportarse de maneras diferentes en circunstancias diferentes. Tener una creencia demanda, además, apreciar el contraste entre lo meramente creído y lo que es el caso, esto es, tener el concepto de verdad objetiva o de lo que es el caso independientemente de lo que se

---

<sup>114</sup> En el contexto de esta reflexión presenta al *empirismo* como una forma de *subjetivismo*, en la medida en que afirma que la evidencia última para las creencias acerca del mundo externo es algo no-conceptual que está dado directamente en la experiencia. No obstante, Davidson nos advierte que el dogma empirista no tiene que ser confundido con “la doctrina inofensiva de que la *fuerza* de todo el conocimiento empírico está en los sentidos” (2001c:2). En esta línea, él se opone a considerar las sensaciones como intermediarios epistémicos, pero las reivindica como meros intermediarios *causales*.

crea.<sup>115</sup> Pero ¿cuál es la fuente del concepto de verdad objetiva? A su juicio, es Wittgenstein quien nos dio la pista: la fuente del concepto de verdad objetiva es la comunicación interpersonal. El punto de Davidson es que no podemos acreditar a un individuo con la distinción entre lo que piensa que es el caso y lo que es el caso, a menos que la criatura tenga el estándar provisto por un lenguaje compartido. Ahora bien, sin compartir reacciones ante un estímulo común, nos dice, el lenguaje y el pensamiento no tendrían ningún contenido en absoluto: *requiere de dos puntos de vista* ubicar la *causa* de un pensamiento y definir así su *contenido*. De este modo, la forma de *triangulación* que propone como condición del pensamiento y del lenguaje involucra a *dos criaturas* en un *mundo compartido*, tal que cada criatura correlaciona sus propias reacciones a los fenómenos externos con las reacciones del otro. La siguiente pregunta es ¿cómo puede una persona determinar lo que está en otra mente? ¿Qué opera cómo “puente” entre ambas? Una parte de la respuesta depende de que, en los casos más simples, los eventos y objetos que *causan* una creencia *determinan* también los contenidos de la misma. Así, la creencia que es causada distintivamente y en condiciones normales por la presencia de algo amarillo, de la propia madre o de un tomate, es la creencia de que algo amarillo, la propia madre o un tomate están presentes. La idea no es, desde luego, que la naturaleza garantiza que nuestros juicios más llanos sean siempre correctos sino que la *historia causal* de tales juicios representa un importante rasgo constitutivo de sus contenidos, punto que lo emparenta con el *externismo* de Putnam (1975).

A su vez, su peculiar propuesta *externista*, aunque expresamente distinta de la de Burge (1979), pretende ubicar el *rol de la sociedad* dentro del nexo causal que incluye la interacción entre las personas y el resto de la naturaleza. La identificación de los objetos del pensamiento reposa así en una *base social*: sin una criatura para observar a la otra, la triangulación que ubica a los objetos relevantes en un espacio público no podría tener lugar. En este sentido, Davidson afirma que *no podemos resolver la cuestión de los contenidos de los estados mentales desde el punto de vista de una criatura única*.<sup>116</sup> Veamos, entonces, más detalladamente la dimensión social de su propuesta.

---

<sup>115</sup> Voy a volver, especialmente, sobre este punto al indagar la relación entre el pensamiento y el lenguaje en el capítulo 5.

<sup>116</sup> En “Three varieties of knowledge” Davidson (1991) se ocupa de extraer las consecuencias epistemológicas de su *externismo semántico*. Allí afirma que toda criatura racional tiene tres tipos de conocimiento: conocimiento del mundo objetivo, conocimiento de las mentes de los otros y conocimiento de los contenidos de su propia mente. Ninguno de estos tres tipos de conocimiento es reducible a ninguno de los otros dos o a una combinación de ellos.

Davidson dice encontrar interesante la forma de externismo social que S. Kripke (1982) promueve en nombre de Wittgenstein: la idea es que alguien captó las condiciones de verdad de una oración si usa la oración en situaciones relevantemente similares a las situaciones en que otros la usarían. Sin embargo, a su juicio, esta propuesta adolece de un defecto u omisión respecto de la comprensión de la objetividad.

Una razón es que Kripke ni siquiera intentó dar una explicación de los contenidos de pensamientos particulares; en el mejor de los casos, explica cuándo dos pensamientos son el mismo –o cuando dos emisiones tienen el mismo significado-. En parte, esto se debe a que él trata casi exclusivamente con ejemplos matemáticos y, así, no puede conectar los contenidos de los pensamientos con las circunstancias empíricas de la conducta de habla. Adicionalmente, dice Davidson, está la cuestión de cómo reconocemos que un contenido es el mismo desde una ocasión de emisión a otra, puesto que la respuesta de Kripke nos deja con otra pregunta. Su respuesta es que el agente continúa de la misma manera que antes, si continúa como lo harían otros. Pero Warren Goldfarb formula una dificultad ulterior: si no podemos decir de un único hablante cuándo sus respuestas constituyen su continuar del mismo modo, ¿cómo podemos decir cuándo sus respuestas son relevantemente como las de otro hablante? (Goldfarb 1985).

Davidson afirma que quizás podemos decir en defensa de Kripke que debemos dar por supuesta la noción de similitud. La cuestión no es, entonces, cuándo la respuesta de un hablante es como sus respuestas previas, sino qué es lo que hace a algunas respuestas *correctas* (y a otras no). Una serie de respuestas de un único hablante sólo muestra que esa persona tiene una u otra disposición; la noción de corrección o incorrección no tiene aplicación en una situación tal. Pero cuando agregamos otros hablantes, hemos añadido un elemento radicalmente nuevo y con ese agregado podemos decir que la corrección es definida como continuar como otros lo harían.

Sin embargo, esto también le resulta insatisfactorio a Davidson: ¿cómo puede introducir la distinción entre seguir una regla y sólo continuar de uno u otro modo el simple hecho de que dos o más personas hayan continuado de la misma manera? Todos los girasoles en el campo giran juntos hacia el sol: ¿están siguiendo una regla? No, en el sentido requerido: no tienen

---

La idea es que del hecho de que tengo cualquiera de estos tipos de conocimiento sé que tengo los otros dos, dado que el triángulo básico es condición del pensamiento, pero ninguno es temporal ni conceptualmente anterior a los otros. Si esto es así, el conocimiento de otra mente y el conocimiento del mundo resultan esenciales para el *conocimiento* de los contenidos proposicionales de la *propia mente* en el siguiente sentido: si no supiera lo que los otros con quienes comparto el mundo piensan, no tendría pensamientos propios y, por añadidura, no sabría lo que pienso. A la luz de esta reflexión, me parece que se advierte más cabalmente la afirmación con la que comencé este apartado, esto es, en qué sentido pretende que su externismo constituya una alternativa frente al subjetivismo.

conceptos y, por tanto, no pueden aplicarlos mal. El mero hecho de agregar otras criaturas con disposiciones idénticas no puede convertir las disposiciones en seguir una regla. Y Davidson nos interpela: ¿no es el propio Kripke quien advierte el punto cuando cita a Wittgenstein diciendo que ningún curso de acción podría ser determinado por una regla, dado que todo curso de acción puede hacerse acordar con una regla (Wittgenstein 1953:20)? Parece que la dificultad persiste sin importar cuántas sean las criaturas que siguen el mismo curso de acción. Lo que estaría faltando es algo que satisfaga lo que Kripke llama el “requisito básico” para una regla, a saber, que la regla le diga al hablante qué debe hacer en cada nueva instancia.

Pero ¿por qué *debo* continuar como los demás lo hacen? ¿Qué hace de su conducta mi norma? Según hemos visto en el apartado anterior, para Davidson no parece responder la pregunta decir que el maestro o la sociedad me lo dicen. El miedo al castigo en sí mismo no me da la idea de que hay algo más incorrecto con mi acción que el que a los demás no les guste. Lo que se está perdiendo de vista, para Davidson, es la idea de *comprensión*. Introducir más de una criatura agrega algo básico a la situación con una única criatura, porque con la posibilidad de que sus acciones *diverjan*, hemos introducido el hiato necesario para dotar de sentido a la idea de *error*. Pero la mera posibilidad de divergencia, aún cuando sea combinada con sanciones que alienten la conformidad, no introduce el tipo de norma necesaria para explicar el significado o la conceptualización. No obstante, a pesar de estas cuestiones irresueltas, Davidson nos dice que está persuadido de que la idea básica es correcta: “sólo la interacción social trae consigo el espacio en el que los conceptos de error, y así, de significado y pensamiento pueden tener aplicación. Un medio social es necesario, pero no suficiente, para el pensamiento objetivo” (Davidson 2001c:4).

Por su parte, el externismo *perceptual*, tal como es descrito por Burge (1988), suministra a juicio de Davidson, uno de los elementos faltantes en la reconstrucción kripkeana de la posición de Wittgenstein, en la medida en que sugiere cómo se puede asignar contenidos particulares a nuestras creencias perceptuales y explica, en parte, cómo es que el pensamiento y el lenguaje están anclados en el mundo. Sin embargo, para Davidson, la versión del externismo perceptual de Burge adolece de algunas “lagunas serias”.

En primer lugar, está el problema del *contenido de las creencias perceptuales*. Según Burge, es la “causa normal” de una creencia perceptual la que determina su contenido. ¿Pero cuál es la causa normal? Podría ser cualquier cosa desde la estimulación de las terminaciones

nerviosas hasta el Big Bang. Claramente, dice Davidson, necesitamos una manera de aislar la causa correcta. En segundo lugar, a juicio de Davidson, Burge no da una explicación seria del *error*, “y sin una explicación del error no hay manera de distinguir entre tener un concepto y, simplemente, tener una disposición” (Davidson 2001c:4). ¿Cómo podemos identificar un error como error? En otras palabras, en opinión de Davidson, Burge falla en resolver, o siquiera en advertir, el problema wittgensteiniano.

Ambos problemas, nos dice Davidson, son acerca de la *similitud relevante*. Las vacas de juguete son, en un sentido, relevantemente similares a las vacas reales (esto es por lo que cometemos errores). Pero en otro sentido, las vacas de juguete no son en absoluto como las vacas (no caen bajo el concepto vaca). Y, nos dice: es porque ocasionalmente confundimos cosas apócrifas con reales que puede decirse que tenemos conceptos, que clasificamos cosas. Distinguir vacas de otras cosas, sólo requiere de meras disposiciones. Pero “lo que es difícil de explicar es qué está pasando cuando alguien *piensa* que una vaca de juguete es una vaca, porque eso requiere tener el *concepto* de una vaca” (Davidson 2001c: 4).

El problema de la similitud relevante surge también en conexión con la idea de que el contenido de una creencia perceptual depende de su causa común o usual. Ahora bien, dado que cualquier conjunto de causas tendrá incontables propiedades en común, tenemos que buscar algún rasgo recurrente en quien las aglutina, alguna marca de que clasifica casos como similares. Este sólo puede ser algún rasgo o aspecto de sus reacciones (quizás de sus reacciones verbales), en cuyo caso tenemos que preguntar de nuevo: ¿qué hace a estas reacciones relevantemente similares entre sí? Reaparece así el problema de Wittgenstein.

No obstante, Davidson es consciente de que alguien podría decir que, desde luego, reaccionamos de maneras similares a ciertos estímulos. Es un hecho bruto que clasificamos, ponemos ciertas cosas juntas de este modo. Si vamos a describir algo, tenemos que usar el equipamiento que poseemos.

La observación le parece correcta a Davidson y cree que apunta en la dirección de la respuesta wittgensteiniana de Kripke al problema de seguir una regla. Es un hecho bruto que cada uno reacciona a los estímulos como lo hace; pero es un hecho sugerente, dice Davidson, que muchas de esas reacciones sean más o menos las mismas. ¿Qué lo hace sugerente? Que la mera conducta de la mayoría no explica en sí misma la conceptualización, dado que no puede explicar el *error*. Lo que sí puede es *hacer el espacio* para que algo pueda llamarse error: en la medida en

que el espacio para el error es creado por aquellos casos en los cuales un individuo se desvía del curso de acción ahí donde la mayoría no lo hace. Por ejemplo, cuando se presenta una vaca de juguete entre las vacas reales, la mayoría deja de decir “vaca” mientras que el individuo desviado persiste en decir “vaca”. Sin embargo, esto no es suficiente para justificar una atribución de error: todavía podemos decir que hay simplemente dos disposiciones diferentes en juego. Se hizo espacio para el error, pero el error en sí mismo todavía no ha sido explicado. Lo que a juicio de Davidson debe agregarse para explicar el error es algo que pueda contar como *reconocer o darse cuenta*, por parte de quienes comparten reacciones, de las reacciones del otro.

La situación básica, tal como la reformula Davidson, contiene un mínimo de tres elementos: dos criaturas y un mundo de objetos, propiedades y eventos que las criaturas pueden discriminar en la percepción. Por supuesto, el entorno social sería típicamente más numeroso, pero simplificaré la historia el que, por el momento, nos quedemos con esta “tríada irreducible” (Davidson 2001c: 5). Cada criatura lleva a cabo sus inducciones habituales, sean aprendidas o heredadas. Esto significa que para cada criatura hay estímulos que pueden ser puestos juntos en virtud de la similitud de las respuestas. Habrá casos que cuadran con esta descripción: la criatura A desarrolla respuestas similares a sus respuestas pasadas en los mismos lugares y en los mismos tiempos que la criatura B desarrolla respuestas similares a sus respuestas pasadas (las respuestas de A pueden o no ser similares a las de B). Podemos preguntar entonces qué estímulos causan estas respuestas repetidamente correlacionadas. Esta causa común será típicamente algún rasgo objetivo del entorno.

Nada en esta historia, dice Davidson, sugiere la necesidad de pensamiento, conceptualización o error. Sin duda, podemos decir, a partir de nuestra ventaja conceptual, que una criatura se equivoca cuando reacciona, quizás verbalmente, a vacas de juguete como lo hace respecto de vacas reales. El problema es poner a las criaturas en posición de pensar esto. Algo debe agregarse de modo que torne *disponible para las criaturas mismas una diferencia* entre los casos *en que reaccionan al unísono y los casos en que no lo hacen*. Esto puede hacerse, a juicio de Davidson, *añadiendo un elemento* de asociación o inducción natural: cada criatura asocia las respuestas de la otra con estímulos del mundo compartido. Supongamos entonces que hay dos criaturas, A y B, que a menudo se observan entre sí cuando ambas miran vacas. Supongamos, ahora, que la criatura A responde a las vacas emitiendo el sonido “Vaca”; la criatura B entonces asocia las vacas con las emisiones de “Vaca” por parte de A. La criatura B similarmente asocia

las vacas con las emisiones de “vache”. A y B, ahora, está cada una en posición de advertir ocasiones en las cuales las respuestas del otro difieren de sus respuestas usuales. Davidson se refiere a que las respuestas de A no son tomadas por B como las mismas respuestas previas de A, mientras que las respuestas de B son las mismas que sus respuestas previas.

Ahora, nos dice Davidson, se ha completado el triángulo: hay conexiones causales entre las vacas y las dos criaturas que reaccionan, por medio de la percepción, y hay conexiones causales entre las criaturas, también mediante la percepción. Davidson enuncia, entonces, su peculiar tesis de la siguiente manera: un triángulo interconectado de esa manera constituye una condición necesaria para la existencia de la conceptualización, el pensamiento y el lenguaje. Esto es lo que hace posible la *objetividad* de la creencia y las restantes actitudes proposicionales.

Así, dice Davidson, se llenan las lagunas del externismo social y del externismo perceptual debidos a Kripke y Burge, respectivamente<sup>117</sup>. Incorpora el elemento esencial de la solución de Kripke, que requiere que distingamos entre una mera disposición y seguir una regla. Ese elemento es provisto por la segunda criatura -que está en lugar de la sociedad-. El modelo simple que ha descrito no sugiere dónde yace la norma cuando las criaturas difieren en sus respuestas; tampoco el de Kripke lo explica. El punto no es identificar la norma, sino dotar de sentido que haya una norma, y esto se ha logrado, a juicio de Davidson, si podemos señalar la diferencia entre la preponderancia de casos en que las criaturas responden de manera similar y los casos desviados en que divergen. Él insiste en que no es suficiente que un tercero sea capaz de observar o describir estos dos casos; sino que es necesario que la existencia del contraste esté *disponible para las criaturas mismas*. La diferencia es la que tiene lugar “entre un comentarista externo deslizándose en sus categorías para dotar de sentido a una criatura aislada y un participante que observa a otro *participante*” (2001c:8; mis itálicas)<sup>118</sup>. Este elemento esencial entra en el triángulo cuando las criaturas observan sus reacciones mutuas hacia el mismo

---

<sup>117</sup> Asimismo, me parece interesante señalar la recuperación que en su (2001c) hace del “externismo” de Quine: su idea es que aunque la epistemología quineana siguió siendo resueltamente individualista, siempre ha habido un fuerte elemento externista en su teoría del significado. Aunque el significado de cada una de las emisiones de las oraciones observacionales es determinado por el estímulo proximal –esto es, patrones de estimulación sensorial- la única pista que Quine provee para ello es por medio de la comparación con los patrones de estimulación de otra gente. Hay, así, a juicio de Davidson, una cierta ambivalencia en Quine: por un lado, su epistemología es empirista, basada en “evidencia” privada; por otro, la teoría del significado es implementada mediante el proceso de traducción radical que “es básicamente social”.

<sup>118</sup> Volveré sobre esta noción de “participante” y la contrastaré con la de “espectador” desinteresado al examinar el problema de las perspectivas de la atribución mental en el capítulo 6.

fenómeno que ambas observan. El triángulo torna simple, así, lo que los ejemplos matemáticos de Kripke no pueden explicar: el rol que juega la percepción en el establecimiento social de una base para el pensamiento objetivo. Y lo hace de dos maneras: permite que las criaturas involucradas observen sus mutuas reacciones y dota –parcialmente- de sentido a la idea de que hay un único foco para las reacciones conjuntas.

De este modo, Davidson considera que su modelo triangular da un paso más allá para lidiar con otro rasgo problemático del externismo perceptual de Burge: la naturaleza indeterminada de los contenidos de las creencias perceptuales. Esta dificultad surge porque parece no haber manera de decidir la ubicación de los objetos y rasgos del mundo que constituyen el tema de las creencias perceptuales. Burge sólo nos dice que el contenido fue dado por la causa “usual” o “normal”; pero esto no nos ayuda a elegir entre estímulos proximales y distales. Al introducir otro ser percipiente, es posible ubicar la causa relevante: es la causa común a ambas criaturas, la causa que genera sus respuestas distintivas.

En suma, Davidson explicita las siguientes condiciones para que el pensamiento esté presente:

1-Con dos criaturas respondiendo en modos característicos a estímulos distales, podemos hablar del *foco* de sus respuestas, la causa común de los casos más frecuentes.

2-Si el pensamiento está presente, esto nos provee una indicación obvia del *contenido* del pensamiento.

3-Hay espacio para el concepto de *error*, que aparece cuando hay una divergencia en las reacciones normalmente similares.

4- Finalmente, y quizás lo más importante, dado que es lo que suele perderse de vista: nos hemos movido de la confianza en nuestra noción de similitud relevante al ámbito de los “intereses animales, adonde seguramente pertenece”. Lo hemos logrado al ver “la similitud relevante de las respuestas de un animal *a través de los ojos de otro* animal. Hemos abrazado la intuición wittgensteiniana de que la única fuente legítima de objetividad es la *intersubjetividad*” (2001c:13; mis itálicas).

Llegados a este punto, esto es, tras haber caracterizado los tres rasgos centrales del concepto de *creencia* según Davidson, ya es tiempo de que veamos qué rol cumplen en los

argumentos centrales de los tres problemas que vamos a considerar: el problema mente-cuerpo, el problema de la relación entre el lenguaje y el pensamiento y el problema de las perspectivas de primera y de tercera persona de la atribución mental.

## **SECCIÓN II**

**La filosofía de la mente de Donald Davidson**

## Capítulo 4: El problema mente-cuerpo

Nuestros cuerpos constituyen el vínculo esencial entre nuestras mentes y el resto de la naturaleza y el carácter de este vínculo suscita algunos de los problemas más arduos en filosofía. Si deseamos dar una explicación de la acción, la percepción, la memoria, o el conocimiento empírico en general, *es necesario que comprendamos las relaciones entre la mente y el cuerpo, lo mental y lo físico.* (Davidson 1993b/2005: 295, mis itálicas)

### 4.1. El anomalismo de lo mental

El argumento en contra de la existencia de leyes psicofísicas estrictas provee la clave de la peculiaridad de la psicología. (Davidson 1974c/1980:241)

Una de las revisiones más originales y decisivas de la agenda filosófica tradicional que nos ofrece Davidson tiene que ver, en mi opinión, con su contribución en filosofía de la mente. Comencemos, entonces, por precisar cómo concibe lo mental.

A mi juicio, la primera inflexión que hay en su consideración acerca lo mental está dada por el hecho de que, desde su ontología de eventos como particulares, como individuos irrepetibles, lo mental se exhibe como una cuestión de *vocabulario*. Su singular invitación es, entonces, a que dejemos de considerar lo mental como categoría ontológica, para pasar a verlo como categoría conceptual.<sup>119</sup> Recordemos que en “Mental Events” dispuesto a rigORIZAR qué quiere decir que un evento es mental, sostiene que un evento es mental “si y sólo si tiene una descripción mental” (1970/1980:211). El punto es que se trata de descripciones que involucran “verbos mentales”, esto es, aquellos que expresan actitudes proposicionales como *creer, desear, esperar*, etc.

---

<sup>119</sup> En este sentido, Ramberg B. (2000) nos propone ver la filosofía de la mente davidsoniana como “post-ontológica”. Allí, nos recuerda que la filosofía de la mente “ontológica” dice que nosotros tenemos algo que otras criaturas no tienen - una mente- y que este ingrediente es misterioso. A su juicio, una vez que dejamos de preguntarnos cómo están relacionados los dos ingredientes de los seres humanos y nos preguntamos, en cambio, por qué es útil tener –al menos- dos descripciones de la conducta humana, el misterio se desvanece.

En este sentido, su mirada conlleva una segunda inflexión respecto de la tradición cartesiana (que incluye como modos del pensamiento a la sensación y a la imaginación), para sostener una noción más *restringida*: en la concepción de Davidson, el rasgo distintivo de lo mental no es que sea privado, subjetivo, o inmaterial, sino que exhiba lo que Brentano llamó *intencionalidad*. De esta suerte, por “pensamiento” Davidson entiende “un estado mental con contenido especificable” (Davidson 1992).

Y, finalmente, el tercer giro respecto del tratamiento canónico de lo mental, tiene que ver, a mi juicio, con la original propuesta de su teoría de la interpretación radical, que hará que el vocabulario intencional sea siempre vocabulario atribuido, esto es, asignado desde la perspectiva del *intérprete*. Según veremos, son precisamente los principios que constriñen al intérprete en la empresa conjunta de asignación de significado y atribución de creencias aquello que, según Davidson, echa por tierra los intentos reduccionistas a propósito del problema mente-cuerpo. En efecto, en “Mental Events” él ofrece un tratamiento sumamente original del tradicional problema acerca de cómo se relacionan lo mental y lo físico, a partir de insertarlo en el marco seminal de su artículo de (1963) “Acciones, razones y causas”, esto es, bajo el signo de cómo conciliar la causalidad y la racionalidad de unas y las mismas acciones. Veamos, entonces, cómo lo tematiza.

En “Mental Events”, fiel a su estilo de redacción, Davidson comienza por presentarnos un problema que opera como “acicate” para lo que sigue. El problema en cuestión posee una inequívoca resonancia kantiana <sup>120</sup> (tal como lo evidencian las citas de la *Fundamentación de la Metafísica de las Costumbres* con las que da inicio y cierre al artículo) y consiste en la cuestión siguiente. Dado que los eventos mentales se resisten a ser capturados en la red nomológica de la

---

<sup>120</sup> Mi señalamiento no apunta a un mera cuestión histórica en busca de antecedentes filosóficos del monismo anómalo, sino a subrayar una preocupación conceptual que resulta central y recurrente en la obra de Davidson, y que lo revela, en mi opinión, como un filósofo inscripto en la tradición de la Ilustración: la preocupación acerca de cómo conciliar el determinismo mecanicista con la libertad humana. En esta misma línea puede verse Janke, T. (2008). A su juicio, la solución kantiana de la tercera antinomia prefigura algunos importantes desarrollos filosóficos contemporáneos, a saber: el “monismo anómalo” defendido por Donald Davidson y el “intencionalismo” propugnado por von Wright y Stoutland. En su opinión, la resolución kantiana de dicha antinomia entraña una posición sobre la agencia mucho más coherente y persuasiva de lo que suele admitirse. Janke lee la solución kantiana en la clave de Anscombe, quien es para mí el nexo con la propuesta de Davidson: se trata de uno y un mismo evento *bajo diferentes descripciones*. En la interpretación de Janke las preguntas acerca de la acción son preguntas *conceptuales*, mientras que las preguntas acerca de la ocurrencia de eventos son *empíricas*. Este es, a su juicio, otro modo de decir que las acciones caen bajo la “evaluación normativa”, mientras que los eventos estrictamente hablando no. Nos dice que las acciones difieren de los eventos en la medida en que las descripciones y explicaciones de las primeras son parte de lo que Wilfrid Sellars llamó “el espacio de las razones”: son *acciones* en la medida en que son parte de los intercambios sociales que involucran dar y pedir razones. Volveré sobre esta cuestión crucial a la hora de proponer mi elucidación de la *normatividad* de lo mental.

teoría física ¿cómo puede reconciliarse este hecho con el papel causal que poseen en el mundo físico? Así, Davidson parte del supuesto de que el anomalismo y la dependencia causal de los eventos mentales son hechos innegables y su desafío será explicar, frente a las aparentes dificultades, cómo es esto posible. Reconciliar la libertad con el determinismo causal, nos dice, ha de ser un caso especial de este problema<sup>121</sup>.

La “aparente contradicción” acerca de los eventos mentales que quiere discutir y, finalmente socavar, surge de tres principios. El primer principio es el de “interacción causal” y afirma que al menos algunos eventos mentales interactúan causalmente con eventos físicos. El segundo principio es el del “carácter nomológico de la causalidad” y sostiene que donde hay causalidad debe haber una ley: los eventos relacionados como causa y efecto caen bajo leyes deterministas estrictas. Finalmente, el tercer principio es el de la “anomalía de lo mental” y dice que no hay leyes deterministas estrictas a partir de las cuales puedan predecirse y explicarse los eventos mentales. La paradoja que desea discutir surge para quien se inclina a aceptar estos tres principios y piensa que son inconsistentes entre sí<sup>122</sup>. Muchos filósofos han aceptado, con o sin argumento, que los tres principios llevan a una contradicción; a Davidson, sin embargo, le parece que los tres principios son verdaderos, de modo que su reto será explicar y disipar la apariencia de contradicción.

Su estrategia va a consistir en mostrar la mutua consistencia de los tres principios, a partir de defender una tesis sobre la relación entre lo mental y lo físico, que no contenga contradicción interna y que los implique formalmente. Dicha tesis es una peculiar versión de la teoría de la identidad, que niega que pueda haber leyes estrictas que conecten lo mental y lo físico y que bautiza como “monismo anómalo”<sup>123</sup>. La tesis central de esta posición puede descomponerse del

---

<sup>121</sup> En la medida en que Davidson asume que el determinismo causal implica formalmente la captura en la red nomológica y la libertad requiere escapar de ella.

<sup>122</sup> La inconsistencia, por supuesto, no es formal, a menos que se añadan más premisas. No obstante, Davidson señala que es natural razonar que los primeros dos principios- el de interacción causal y el del carácter nomológico de la causalidad- juntos implican que al menos algunos eventos mentales pueden predecirse y explicarse en base a leyes, en tanto que el principio de anomalía de lo mental lo niega.

<sup>123</sup> Davidson intenta aclarar las pretensiones de su posición mediante la inserción del monismo anómalo en una clasificación de teorías acerca de la relación entre lo mental y lo físico. Parte de la novedad que introduce su planteo es enfatizar la independencia entre las tesis acerca de leyes y las tesis acerca de la identidad. La taxonomía que propone es, entonces, la siguiente: (1) El *monismo nomológico*: que afirma que hay leyes correlacionales y que los eventos correlacionados son uno (la teoría de la identidad de tipos pertenece a esta categoría); (2) el *dualismo nomológico*: que combina el dualismo ontológico con la existencia de leyes de correlación (comprende varias formas de paralelismo, interaccionismo y epifenomenalismo); (3) el *dualismo anómalo*: que combina el dualismo ontológico con el fracaso general de leyes que correlacionen lo mental y lo físico (el cartesianismo); y (4) finalmente, encontramos la posición que desea ocupar Davidson: el *monismo anómalo*, que paso a considerar.

modo siguiente: (I) las entidades mentales son entidades físicas (tal es el aspecto “monista” de su propuesta que, según veremos, lo compromete con una teoría de identidad de *casos*), pero (II) los conceptos mentales no son reducibles a los conceptos físicos (a) ni por definición, (b) ni por ley estricta (tal es el carácter “anómalo” de lo mental).

A los efectos de mostrar la manera en que se reconcilian los tres principios en cuestión, en el último apartado de “Mental Events” pretende inferir la verdad de (I), su original versión de la teoría de la identidad, mediante un argumento que puede reconstruirse del siguiente modo. Supongamos que *m*, un evento mental, causó *f* (por el principio de interacción causal); entonces, bajo alguna descripción, *m* y *f* instancian una ley estricta (por el principio del carácter nomológico de la causalidad). Esta ley sólo puede ser física (por el principio de anomalía de lo mental). Pero si *m* es subsumido por una ley física, tiene una descripción física, y por lo tanto, *es* un evento físico.

Ahora bien, creo que sin una adecuada explicitación de sus supuestos, se torna muy difícil comprender cabalmente el argumento davidsoniano que acabo de reconstruir. Pasemos, entonces, a ellos. En primer lugar, a partir de su análisis de la forma lógica de las oraciones de acción y de los enunciados causales singulares, Davidson se compromete con una ontología de eventos como particulares irrepetibles, como “individuos concretos” (1970/1980: 181 y 183). Así, de la misma manera en que podemos referirnos a un individuo mediante diferentes descripciones, todas verdaderas de él, podemos hacer referencia a un evento utilizando diferentes descripciones: tal como anticipé, un evento es un evento mental si y sólo si tiene una descripción mental (en la que figura, esencialmente, un verbo de actitud proposicional); en tanto que un evento será un evento físico en la medida en que lo seleccionemos mediante una descripción que sólo contenga “esencialmente, el vocabulario físico”.<sup>124</sup> Asimismo, en el argumento a favor de la teoría de la

---

<sup>124</sup> En varios lugares el propio Davidson se ocupa de subrayar la sugerente filiación que esta ontología de eventos “desnudos” (que sólo conforme a *descripciones* alternativas, serán ya mentales, ya físicos) parece guardar con una de corte spinocista (Cf. v. gr. 1970/1980 y 1993b/2005). A la vera de esta observación, creo que su posición ontológica puede concebirse como un monismo “neutro”. Desde el comienzo, Davidson reconoce explícitamente el parentesco que el monismo anómalo tiene con el monismo ontológico de Spinoza (1970/1980: 212). A su juicio, la posición de Spinoza acerca de la relación entre los modos de explicación físico y mental y acerca de la causación no sólo es consistente con la existencia de fallas de un análisis completo, sino que las hace inevitables y ofrece una explicación de por qué son inevitables. Tal análisis supone que haya leyes psicofísicas tan precisas como las de la física. Pero si Spinoza está en lo correcto, no hay tales leyes. Los sistemas de explicación a los que pertenecen lo mental y lo físico son irreduciblemente diferentes, a pesar del hecho de que lo mental y lo físico son sólo dos maneras diferentes de concebir –y por tanto, de clasificar– los mismos particulares. Las similitudes que Davidson señala explícitamente son las siguientes (cf. 1993b/2005: 308): (1) *Monismo ontológico*: no hay particulares que no puedan ser identificados mediante descripciones definidas formuladas exclusivamente en el lenguaje de la física y

identidad de casos entran en juego los dos supuestos siguientes: la tesis de la extensionalidad de la relación causal y de la identidad y la tesis del carácter lingüístico de las leyes. Según Davidson, la causalidad y la identidad son relaciones entre eventos individuales *no importa cómo se los describa*. Pero las leyes son lingüísticas, y de este modo, los eventos pueden instanciar leyes, y ser explicados y predichos a la luz de leyes, *sólo en tanto descriptos de cierto modo* (1970/1980:215). A partir de estos supuestos, volvamos, entonces, sobre el argumento davidsoniano para considerar cómo se reconciliarían los tres principios en cuestión.

Los eventos mentales están conectados causalmente con los eventos físicos sin importar cómo se los describa (puesto que el principio de interacción causal es “ciego a la dicotomía físico/mental”; Davidson 1970/1980:215) y en virtud del principio del carácter nomológico de la causalidad, cuando los eventos se relacionan como causa y efecto, tienen descripciones que instancian una ley.<sup>125</sup> Hasta aquí, Davidson estaría mostrando el *papel causal* de los eventos mentales en el mundo físico. Pero, puesto que los eventos son mentales sólo en tanto los seleccionemos mediante una descripción en la que figure un verbo de actitud proposicional, y tal descripción no instancia una ley del tipo que respalda las relaciones causales (esto es, una ley estricta<sup>126</sup>), a la vez que tendrían un papel causal, los eventos mentales se *escaparían de las redes*

---

mediante descripciones definidas que sean “mentales”; (2) *Dualismo conceptual*: los vocabularios físico y mental no son ni definicional ni nomológicamente reducibles en ninguna dirección; (3) *Dualismo explicativo*: aunque las ciencias naturales puedan arribar a un punto de explicación completa y por tanto predecir en teoría cualquier evento, ellas no pueden por principio predecir o explicar completamente ningún evento bajo una descripción mental. Aún más obviamente, tampoco una ciencia psicológica perfeccionada podría explicar o predecir ningún evento bajo una descripción física; (4) Se sigue que *no hay leyes deterministas estrictas* que conecten lo mental y lo físico; (5) Estos puntos no contradicen el reconocimiento de que hay relaciones causales entre lo mental y lo físico: las *relaciones* causales, tal como las concibe, tienen lugar entre eventos como sea que estén descriptos; las *explicaciones* causales, en cambio, dependen del vocabulario o de los conceptos usados para describir eventos y para formular leyes. Para una ampliación del posible “contrapunto” entre las obras de Davidson y Spinoza, puede consultarse el excelente estudio de van der Burg, F. (2007). Asimismo, para explorar una lectura de la posición de Davidson en términos de un “Minimal Materialism” cf. J.J.C. Smart (1987) y Davidson (1987/2004). Por su parte, Daniel Hutto remite a Buekens, F. (1997), para una interpretación de la posición metafísica de Davidson como una variante del monismo neutro sostenido por Bertrand Russell.

<sup>125</sup> Adviértase que este principio debe leerse con cuidado: no dice que cada enunciado singular causal verdadero instancia una ley, sino sólo que los eventos relacionados causalmente tienen alguna descripción que instancia una ley (para profundizar la curiosa concepción davidsoniana de la relación causal puede consultarse su (1967b/1980), donde intenta reconciliar la posición de Hume con la de Ducasse). En este sentido, creo que para comprender plenamente la estrategia de Davidson debemos tener presente la distinción que traza entre la dimensión *extensional* (propia de las conexiones entre *eventos* como *particulares concretos*) y la dimensión *intensional* (propia de sus *descripciones*; según veremos en breve, la idea de Davidson es que sólo las descripciones *físicas* instancian el tipo de ley que respalda las relaciones causales, esto es, las leyes estrictas).

<sup>126</sup> A juicio de Davidson sólo las leyes estrictas, esto es, sin excepciones, respaldan las conexiones causales y, adicionalmente, cree que sólo las leyes físicas satisfacen dicho requisito, en la medida en que son *homonómicas*. Aquí resulta pertinente recordar que él traza una importante distinción: por un lado, hay generalizaciones cuyas instancias positivas nos dan razón para creer que la generalización misma podría mejorarse añadiendo enmiendas y

*nomológicas de la teoría física*. De este modo, para ponerlo en los términos kantianos a los que Davidson apela: no existiría una verdadera contradicción entre la necesidad natural y la libertad de unas y las mismas acciones humanas.

Ahora bien, él es plenamente consciente de que la “prueba” que nos ofrece a favor de (I), esto es, del carácter monista de su posición, es a lo sumo “condicional”, dado que dos de sus premisas carecen de apoyo (el principio de interacción causal y el principio del carácter nomológico de la causalidad) y el argumento a favor de la tercera (el principio de anomalía de lo mental) puede parecer poco concluyente. Veamos, entonces, en qué consiste este último.

Para precisar (II), es decir, el aspecto no reductivo de su monismo, Davidson comienza por dar cuenta del carácter anómalo de las generalizaciones psicofísicas a partir de considerar la falla del reduccionismo *definicional* (en la versión del conductismo lógico). Su punto es que en la falla en cuestión hay un patrón: no importa cuánto ajustemos las condiciones no-mentales, siempre vamos a encontrar la necesidad de incorporar condiciones adicionales de carácter mental.<sup>127</sup> Ahora bien, ¿qué rasgo de lo mental subyace a esta argumentación en contra del reduccionismo definicional? En mi opinión, es el holismo de la atribución de los conceptos mentales, o lo que en la terminología introducida en el capítulo anterior llamamos “holismo psicológico” o “inter-actitudes”, esto es: la atribución de una actitud proposicional a un individuo tiene sentido sólo en tanto pueda ser ubicada dentro de una densa red de otras actitudes proposicionales. La idea es, entonces, que esta atribución en bloques interdependientes suscita en el *definiens* una inescapabilidad del vocabulario mentalista que hace imposible la reducción definicional.

A continuación, se dedica a argumentar en contra del reduccionismo *nomológico*. Veamos, entonces, cuál es su estrategia en contra de la existencia de leyes psicofísicas estrictas, y si ahí aparece en juego algún rasgo específico, peculiar de lo psicológico. Recordemos que

---

condiciones establecidas *en el mismo vocabulario* que la generalización original. Tales generalizaciones se denominan *homonómicas*. Por otro lado, hay generalizaciones que cuando se instancian pueden darnos razón para creer que hay una ley precisa trabajando, pero que puede establecerse sólo *cambiando a un vocabulario diferente*: son las que se llaman *heteronómicas*. El punto de Davidson es que las generalizaciones psicológicas y psicofísicas son de este último tipo y no califican, por tanto, para respaldar conexiones causales.

<sup>127</sup> Así, por ejemplo, si tratamos de decir sin usar términos mentales en qué consiste el hecho de que un hombre *crea* que hay vida en Marte, un intento podría ser: cuando un cierto sonido es producido en presencia del hombre (¿“Hay vida en Marte?”), él produce otro (“Sí”). Pero Davidson insiste en que esto mostraría que el sujeto en cuestión cree que hay vida en Marte, sólo si *comprende* el español, su producción del sonido fue *intencional*, etc. Es decir, para cada deficiencia descubierta, se va a agregar una enmienda que no puede sino ser mental.

Davidson considera que sus conclusiones no son nuevas. Su punto es que, si está diciendo algo nuevo sobre este tópico, es en los detalles de las razones que da para decir que las generalizaciones que combinan predicados psicológicos y físicos no son legaliformes (en el sentido fuerte en que las leyes completamente físicas, dice él, pueden serlo). Para indagar estas razones reconstruyamos, entonces, el argumento tal como lo presenta en “Mental Events”:

- (i) Dos ámbitos conceptuales cuyos principios constitutivos difieran radicalmente no pueden estar conectados mediante leyes.
- (ii) Lo mental y lo físico son ámbitos conceptuales cuyos principios constitutivos difieren radicalmente: en el primero rige el *principio de racionalidad*, ausente en el segundo.
- (iii) Por tanto, no hay leyes psicofísicas porque “los predicados mentales y físicos no están hechos el uno para el otro”<sup>128</sup>.

Más arriba hemos visto que el rechazo del reduccionismo *definicional* de los conceptos mentales a los físicos parece reposar en el *holismo psicológico* que Davidson abraza. Por su parte, el argumento que acabamos de reconstruir en contra de la existencia de leyes psicofísicas estrictas porta, a su juicio, “la clave de la peculiaridad de la psicología” (1974/1980:241) Cabe entonces preguntarnos: ¿qué rasgo propio de lo mental pone en juego Davidson en el argumento en contra del reduccionismo *nomológico*? ¿En qué consiste aquello que torna especiales a los conceptos mentales y que abre un dualismo profundo en relación con los conceptos físicos, al punto de imposibilitar por principio su conexión legaliforme?

Antes de examinar posibles candidatos para dar cuenta de la peculiaridad de lo mental creo que cabe advertir que si la condición para la reducción es tan exigente que demanda la existencia de leyes puente *estrictas* (tal como pretende Davidson), parece no haber mayor inviabilidad para la reducción de la psicología que la que se revela en otras ciencias especiales, como, por ejemplo, la geología o la aerodinámica.<sup>129</sup> En este sentido, la apelación al carácter heteronómico de las leyes, no permite capturar la peculiaridad de lo mental.

En “The significance of charity”, B. Ramberg (1999) sostiene que Davidson “sobredramatiza” la significación de la distinción que traza entre lo mental y lo físico, en el sentido de que el contraste mental/físico ofrecido por el argumento para la irreducibilidad no

---

<sup>128</sup> Cf. Davidson (1970/1980: 218). Para una exposición detallada del argumento davidsoniano en contra de las leyes psicofísicas puede consultarse Kim (1985).

<sup>129</sup> En este talante, Kim (1993:26) señala que la condición de reducción resulta “irrealistamente exigente”.

logra mostrar el punto de la distinción. A su juicio, el argumento para el monismo anómalo que Davidson ofrece en “Mental Events” y desarrolla en trabajos posteriores -tales como “Laws and cause” (1995d) y “Thinking causes” (1993c)- juega una parte importante en su proyecto, pero su rol es, no obstante, negativo. La oposición entre generalizaciones homonómicas y heteronómicas juega un rol dialéctico en el esfuerzo de establecer que el monismo es implicado por la irreducibilidad de los predicados mentales a los físicos. Pero esto no funciona para caracterizar el rasgo distintivo de lo mental. Así, en vena rortyana, sostiene que también la biología presenta generalizaciones heteronómicas que no son óbice para su reducción a la física.<sup>130</sup>

Por su parte, Fodor (1974) tampoco encuentra mayor “recalcitrancia” a la reducción en psicología que en otras ciencias especiales. La idea que subyace a su observación es que no habría argumentos específicos en contra de las leyes intencionales, sino que los argumentos que se les aplican, les están dirigidos en tanto leyes de las ciencias especiales. Así, su estrategia en (1991) para favorecer el estatus de la psicología como ciencia autónoma consiste en tratar de ofrecer una buena elucidación de las leyes de las ciencias especiales: esclarecida su naturaleza como leyes *ceteris paribus*, a su juicio no encontraremos obstáculos específicos para la psicología (en su proyecto, esto es, para aceptar leyes causales intencionales). En este sentido, la mera invocación de la irreducibilidad nomológica no parece funcionar para capturar el rasgo peculiar de lo mental<sup>131</sup>.

¿Qué sería peculiar, entonces, en el caso psicológico respecto de lo que sucede en las diversas ciencias especiales? ¿En qué sentido la irreducibilidad sería especialmente privativa de lo mental?

---

<sup>130</sup> A su juicio, el carácter *heteronómico por principio* de las leyes psicofísicas y psicológicas depende de una caracterización independiente de qué es lo que hace distintivo a lo mental: el principio de caridad. Según veremos en breve, la dimensión positiva de la noción davidsoniana de mente debe buscarse, pues, en su teoría de la interpretación, tal es la razón por la que en la primera parte la presenté como marco para mi indagación. Por su parte, Ramberg advierte que la tesis davidsoniana de que nuestro universo se divide en dos categorías fundamentales, esto es, su dualismo conceptual, emerge contra el trasfondo de su concepción de lo físico. Su propuesta es, pues, que debemos interpelar la comprensión de lo físico que se pone en juego. ¿Cómo debemos considerar aquello con lo que contrasta lo mental? Como observa Chomsky (1995) una dificultad con el problema mente-cuerpo es que nuestro concepto de lo físico no está realmente en mejor forma que nuestro concepto de lo mental.

<sup>131</sup> En mi opinión, esta estrategia tendría además el agravante de ser desembozadamente circular, esto es: el argumento pretende concluir la irreducibilidad y, por tanto, no puede invocar legítimamente este rasgo como aquél por el cual difieren radicalmente los ámbitos conceptuales de lo mental y lo físico. Asimismo, creo que vale la pena subrayar que el propio Davidson sostiene explícitamente en su “Reply a Ramberg” (Hahn 1999: 617) que en “Mental Events” no tenía un concepto claro de *reducción* en mente.

Consideremos, entonces, qué sucede con uno de los candidatos explícitamente propuestos por Davidson para dar cuenta del carácter peculiar de lo mental (1970/1980): la *indeterminación* de la interpretación. Para ello, a continuación voy a examinar una analogía a la que apela en varios artículos: la analogía entre las actitudes proposicionales y las actitudes numéricas. Veamos, pues, en qué consiste exactamente la analogía propuesta entre los sistemas empleados para adscribir estados mentales a los seres humanos y los sistemas de medición de objetos físicos macroscópicos, para ver si la apelación a la indeterminación de la interpretación nos permite o no captar la peculiaridad de lo mental.

#### **4.2. La analogía entre las actitudes proposicionales y las actitudes numéricas y la *indeterminación* de la interpretación**

Leemos en “Mental Events”:

Así como no podemos inteligiblemente asignar longitud a cualquier objeto a menos que una teoría comprensiva se sostenga con respecto a los objetos de tal tipo, tampoco podemos inteligiblemente atribuir ninguna actitud a un agente excepto dentro del marco de una teoría viable de sus creencias, deseos, intenciones y decisiones (1970/1980:221)

Esta analogía entre los sistemas de medición de objetos físicos macroscópicos y los sistemas empleados para adscribir estados mentales a los seres humanos, a la que Davidson recurre en diferentes contextos con diversos propósitos<sup>132</sup>, está motivada por varias coincidencias entre ambos.<sup>133</sup> En primera instancia, es posible advertir una semejanza estructural entre (i) oraciones por medio de las cuales se atribuye una medida a un cuerpo físico y (ii) las oraciones por medio de las cuales se atribuye una actitud proposicional a un ser humano. Por ejemplo, sean:

- (i) El objeto  $O$  tiene un peso (expresado en kilogramos) de  $n$ ;
- (ii) El individuo  $I$  tiene la creencia de que  $p$ <sup>134</sup>.

---

<sup>132</sup>Davidson también apela a esta analogía en (1977/1984) y en (1989/2001).

<sup>133</sup> Sigo en este punto la propuesta de Pérez, D. (1997).

<sup>134</sup> Dado el paralelo hasta aquí, hay una asimetría evidente entre ambos casos en virtud de que en (i) se añade entre paréntesis la frase “expresado en kilogramos” y en (ii) no hay nada que se corresponda con ésta, dado que las

(i) Es un caso de “actitud numérica” y (ii) es un caso de “actitud proposicional”. En ambos casos estamos frente a oraciones que parecen estructuralmente semejantes: un objeto concreto (un objeto físico o un ser humano, respectivamente) se encuentra en una relación (de peso, o de creencia) con una entidad abstracta (un número en el primer caso, una proposición en el segundo). En suma, podemos mencionar las siguientes coincidencias entre las actitudes numéricas y las actitudes proposicionales:

1- Se establecen relaciones cuyos sujetos son objetos físicos, en el primer caso y seres humanos, en el segundo; en ambos casos se trata de entidades concretas, situadas en el espacio-tiempo.

2-La relación se establece con entidades abstractas: números, en un caso, y proposiciones, en el otro.

3-Entre los objetos físicos se dan, efectivamente, ciertas relaciones como “ser más alto que...” o “estar tan frío como...”; y en el caso de los seres humanos, relaciones como “pensar lo mismo que...”, “desear lo contrario que...” Estas relaciones son las que se pretende reflejar al adoptar un sistema de medición, numérico, en el primer caso; proposicional, en el segundo.

4- Estas relaciones se pueden poner en correspondencia con ciertas relaciones existentes entre los objetos abstractos que se adoptan para la medición: relaciones como “ser mayor que...” en el caso de los números o relaciones de “implicación lógica y apoyo evidencial” en el caso de las proposiciones.

Pérez (1997) señala que a partir de la analogía en cuestión es posible extraer las siguientes consecuencias para una teoría de la mente. En primer lugar, dado que hay teorías físicas exitosas en las cuales se apela a actitudes numéricas, no hay razones para pensar que no podemos formular, en principio, teorías psicológicas que involucren estados con contenido proposicional, esto es, actitudes proposicionales. En segundo lugar, dado que no tenemos dificultades para aceptar que el peso (supongamos, de 80 kg.) del objeto *O* fue lo que causó que

---

proposiciones no pertenecen a ningún lenguaje. No obstante, la simetría entre los dos casos se restaura si se considera que las entidades con las que el verbo de actitud proposicional pone en relación al individuo no son proposiciones, sino algún tipo de entidad lingüística, como oraciones o emisiones (y, según veremos, es esto último lo que sostiene Davidson). Volveré sobre este punto, precisamente, al considerar la *indeterminación* de la interpretación.

se rompiera el cristalero sobre el que estaba apoyado, parece no haber razones, en principio, para no aceptar que, por ejemplo, la creencia (de que *p*) del individuo *I* causó cierta conducta que él realizó. No parece haber ningún problema especial para la psicología en lo relativo al papel causal de eventos que se identifican apelando a entidades abstractas.

Hasta aquí hemos presentado algunas coincidencias que parecen estar en la “superficie” de la analogía en cuestión. No obstante, la analogía tal como la propone Davidson es mucho más profunda que lo que hemos considerado hasta aquí. En efecto, es posible advertir coincidencias adicionales entre nuestro hablar acerca de la medida de objetos físicos y nuestro hablar acerca de las actitudes proposicionales de los seres humanos, que “permiten iluminar ciertas tesis davidsonianas acerca de lo mental”<sup>135</sup>, *inter alia*, la mentada tesis de la *indeterminación de la interpretación*. Veámoslo con más detalle.

En primer lugar, podría sostenerse que una coincidencia de fondo consiste en la *finalidad* por la cual adoptamos ambos lenguajes. En el caso de los objetos físicos macroscópicos adoptamos un sistema de medición para abreviar nuestros informes acerca de hechos concernientes a objetos concretos: esto es, por ejemplo, al decir cuánto pesa un objeto tratamos de informar acerca de la existencia de ciertas relaciones entre objetos (que algunos pesan más que otros, que dos de ellos pesan lo mismo, etc.). Paralelamente, podría sostenerse que adoptamos nuestro hablar en términos de actitudes proposicionales para no perder de vista ciertas propiedades y relaciones relevantes de los estados psicológicos de los individuos. En palabras de Davidson: “cuando hablamos acerca de cuánto pesan las cosas hablamos de una manera relacional: relacionamos objetos con números y, de este modo, unos objetos con otros” (Davidson 1989/2001). Si extrapolamos esta afirmación acerca de lo que permite un sistema de medición al caso de las actitudes proposicionales, obtenemos la observación siguiente: hablar acerca de las creencias es relacional, en el sentido de que relaciona personas con proposiciones y, así, unas personas con otras. El lenguaje intencional puede concebirse, pues, como una manera abreviada de describir ciertas (complejas) relaciones interpersonales. Como hemos visto en la primera parte de este trabajo al considerar los rasgos del concepto davidsoniano de *creencia*, la singular tesis externista que defiende, la triangulación, sostiene que los estados mentales se identifican, en parte, apelando a su historia causal e incluye las relaciones sociales como parte de esa cadena causal. Y estas relaciones con los demás y con el mundo son las que el lenguaje

---

<sup>135</sup> Pérez (1997: 34). A continuación, explicitaré varias coincidencias de fondo lúcidamente advertidas por ella.

intencional está llamado a preservar. De este modo, tanto en el caso de los sistemas de medida de los objetos físicos, como en los sistemas de adscripción de estados mentales a los seres humanos, la apelación a entidades abstractas puede entenderse como un mero recurso lingüístico, esto es, como la adopción de un lenguaje más simple para describir relaciones entre objetos concretos.

Un segundo aspecto en el que la analogía con los sistemas de medida de los objetos macroscópicos permite comprender mejor la posición de Davidson es atendiendo a las *consecuencias ontológicas* que se derivan de la adopción de uno y otro lenguaje (el que permite medir objetos físicos y el intencional). En efecto, introducir un patrón de medida u otro no altera las relaciones preexistentes entre los objetos ni nos obliga a agregar nuevas entidades a nuestra ontología: en términos de Davidson “no nos compromete con los *pesos* como objetos” (1989/2001:60) Es sólo una manera de relacionar un objeto con otros a los que se les aplica una misma escala. Análogamente, cuando se piensa y habla de las creencias de las personas no necesitamos suponer que existan entidades tales como las creencias. Ni tampoco tenemos que inventar objetos que sirvan como “los objetos de creencia” o como lo que está “ante la mente”, o en el cerebro.

Como advierte Pérez (1997:36), en este punto, Davidson parece coincidir con la estrategia de Dennett, quien muestra cómo es posible adoptar un hablar acerca de actitudes proposicionales, concibiéndolas como relaciones de un individuo con una proposición, y aceptando que tales actitudes desempeñan un rol causal sin que por ello debamos suponer que hay algo concreto en los cerebros de los individuos (como, por ejemplo, oraciones-caso del “lenguaje del pensamiento”) que intervenga en la historia causal en la que está involucrada la actitud proposicional, ni tampoco intermediarios psicológicos o epistémicos peculiares como los “objetos ante la mente” postulados en la modernidad<sup>136</sup>. Como hemos señalado ya en la primera parte de este trabajo, él rechaza tanto la existencia de objetos del pensamiento “concebidos como experiencias, sensaciones o datos de los sentidos no interpretados, como si se los concibe como proposiciones plenamente interpretadas” (Davidson 1988/2001: 46).

La analogía con los sistemas de medida de los objetos físicos abre un resquicio para que la tesis davidsoniana de la inexistencia de objetos ante la mente resulte, pues, sostenible. Los números no son intrínsecos a los objetos físicos que tienen peso, ni son parte de ellos, en tanto

---

<sup>136</sup> No obstante, en breve voy a introducir algunas diferencias que el propio Davidson advierte entre su posición y la de Dennett.

se trata simplemente de una manera de informar acerca de la existencia de ciertas relaciones entre objetos concretos. Del mismo modo, se puede afirmar que nuestro hablar en términos de actitudes proposicionales no nos compromete con la existencia de entidades intrínsecas a las personas.

En esta línea, Pérez (1997:37) sostiene que la analogía que nos ocupa muestra que es posible deslindar dos problemas: el problema lógico-semántico relativo al análisis que debemos adoptar para las oraciones de adscripción de actitudes proposicionales, y el problema metafísico acerca de la naturaleza de las actitudes proposicionales mismas, a saber: si resultan ser estados relacionales (diádicos o poliádicos) o si, por el contrario, resultan estados no relacionales.<sup>137</sup> Este divorcio le permite adoptar a Davidson un análisis de las oraciones de adscripción de estados mentales como relaciones entre individuos y emisiones, sin por ello comprometerse con la tesis metafísica de que los estados de actitud proposicional son relaciones diádicas entre un individuo y alguna otra entidad.<sup>138</sup> La analogía de los sistemas de medida con las atribuciones de actitudes proposicionales puede proponerse con dos finalidades diferentes: (i) para defender la idea de que concebir las actitudes proposicionales como relaciones con entidades abstractas no genera problemas especiales a la psicología, o bien (ii) para defender la idea de que el carácter relacional de la expresión lingüística “cree” no causa problemas a la psicología. Acuerdo con Pérez en que la actitud de Davidson es del segundo tipo. Y por ello, después de haber discutido profusamente la analogía en “What is present to the mind?”, reconoce no haberse pronunciado todavía acerca de cuáles son los objetos que se emplean para individuar el estado de una persona cuando se le aplica la palabra relacional “cree”. En este punto abraza las *emisiones* como el tipo

---

<sup>137</sup> En breve, veremos que una de las acusaciones que Davidson (1997b/2001) le hace a la lectura que Dennett efectúa de su posición es no haber advertido este deslinde. Ello, a su juicio, le impide ver que es posible defender la objetividad en nuestra atribución de creencias, sin comprometernos por ello con su estatus entitativo. Volveré sobre esta cuestión.

<sup>138</sup> La propuesta de Pérez (1997:38) al problema metafísico relativo a la naturaleza de las actitudes proposicionales adopta la siguiente dirección. No se trata ni de (i) relaciones genuinas entre un objeto concreto y una entidad abstracta, ni de (ii) una propiedad intrínseca de los objetos considerados, sino de (iii) una relación n-ádica que involucra al objeto concreto y a todos los objetos a los que se les puede aplicar la misma escala de medida, así como a los objetos del mundo compartido. Es, pues, algo así como una “propiedad super-relacional en tanto relaciona a un individuo con sus semejantes y con su medio”. Esta línea de interpretación permite eliminar el compromiso con entidades abstractas a favor de relaciones, más o menos complejas, entre objetos concretos. Por su parte, esta propuesta me interesa especialmente en la medida en que la encuentro consistente con mi interpretación del monismo anómalo como una posición “monista neutra” que se compromete sólo con eventos concretos particulares y descripciones de los eventos, pretendiendo que sólo con particulares concretos y expresiones lingüísticas alcanza para dar cuenta de nuestra vida mental.

de entidad con la cual el verbo de actitud proposicional pone en relación a los individuos. Así, afirma:

Si Ud. lo desea, hasta podría considerar que el objeto es una proposición. Dado que las emisiones, las oraciones y las proposiciones están íntimamente relacionadas, es probable que si una elección funciona, pueda hacerse que las otras también funcionen. Pero las emisiones tienen ciertas ventajas *prima facie*, ya que no son abstractas, y por eso están atadas a un hablante, un tiempo y un contexto. Asumiré, pues, que hemos escogido las emisiones, las mismas emisiones que se producen al atribuir actitudes, como los objetos que sirven para *individuuar* e identificar los diversos estados de la mente ( Davidson 1989/2001:63).

De este modo, son las emisiones de los que adscriben las actitudes proposicionales las que permiten individuar las creencias. Toda emisión pertenece a uno u otro lenguaje, por lo que la individuación de las creencias se vuelve relativa a unidades o a patrones, esto es, al lenguaje usado por el intérprete, y así las proposiciones no resultan constitutivas de las actitudes proposicionales. *Contrario sensu*, habría un modo independiente de todo lenguaje para determinar los estados mentales en los que se encuentran los individuos; pero esto, dada la estrecha relación postulada por Davidson entre pensamiento y lenguaje – que indagaremos en el capítulo próximo- resulta inconcebible en el marco de su filosofía. Una de las consecuencias centrales que tiene este planteo atañe, precisamente, al candidato que nos introdujo en este rodeo por la analogía entre actitudes numéricas y actitudes proposicionales, a saber: la *indeterminación de la interpretación*.

Del mismo modo en que es posible asignar diferentes números para la tarea única de medir el peso de objetos físicos (por ejemplo, se asignarán distintos números a un mismo objeto si se expresa su peso en libras y en kilogramos, o aún en kilogramos y en gramos), también es posible apelar a distintas emisiones para describir la misma creencia de un individuo.<sup>139</sup> Esto sucede, por ejemplo, si dos hablantes de distintos lenguajes naturales (digamos, español y francés) describen las creencias de un tercer hablante en un momento dado. Pero, aclara Davidson, del hecho de que se empleen distintos conjuntos de números para medir el peso de un cuerpo no se deduce que el peso no sea algo “real”, ni que la asignación de peso sea “relativista”. Análogamente, las emisiones de una persona pueden captar todos los rasgos significativos del pensamiento de otra persona de distintas maneras. Puede haber más de un conjunto de emisiones

---

<sup>139</sup> Con esto se anula la asimetría a la que hicimos alusión a propósito de la formulación original de la analogía.

que resulten igualmente exitosas para capturar los contenidos de la mente de una persona. Pero esto no amenaza la realidad de las actitudes.

En efecto, tal es la tesis que Davidson (1997b/2001) defiende explícitamente en “Indeterminism and Antirealism”. Una vez más, su posición es novedosa y rehúye de las clásicas antinomias.<sup>140</sup> Por un lado, caracteriza al antirrealismo como una “manifestación del impulso incontenible de la filosofía occidental de asegurar que lo real, sea lo que sea, puede ser conocido; el antirrealismo intenta conseguir esto considerando fuera de la existencia a aquello, sea lo que sea, que él decreta que está fuera del alcance del conocimiento humano”(1997b/2001:70). No obstante, declara que tampoco ve al realismo como alternativa al antirrealismo. Una caracterización común del realismo que trae, entonces, a consideración es la siguiente: hay algo en o acerca del mundo que hace que nuestras emisiones, aseveraciones o pensamientos sean verdaderos cuando lo son, tengamos o no el poder de determinar su verdad. El hecho es que, a su juicio, nadie ha sido capaz de decir de manera no trivial qué clase de cosa es lo que hace verdadera a una oración (o a otro portador de verdad). Como es sabido, en su opinión, la falla de las teorías de la correspondencia en dar sustento al concepto de verdad demuestra, al mismo tiempo, la vacuidad de las caracterizaciones del realismo en términos de correspondencia. El objetivo que persigue allí es, pues, el siguiente: discutir versiones del antirrealismo que suscitan dudas sobre la realidad de los estados y eventos mentales a la luz de la indeterminación de la interpretación. A su juicio, la *indeterminación* no muestra ni que las actitudes proposicionales no sean plenamente *reales*, ni que debemos modificar el concepto de *verdad* cuando hablamos de actitudes proposicionales. En otras palabras, muchas de nuestras creencias y enunciados sobre lo que las personas creen, pretenden, desean y esperan son verdaderas, y lo son porque las personas tienen esas actitudes. Se detiene, entonces, a analizar algunas consideraciones que parece que impiden aceptar esta tesis. La primera de ellas es (a) “cierta forma de cientificismo”, en tanto que la segunda es, precisamente, (b) la “indeterminación de la interpretación”.

Según (a), las actitudes proposicionales no parecen apropiadas para incorporarse a una concepción científica y unificada del mundo. Así, Davidson cita a Quine, quien ha dicho que “el lenguaje esencialmente dramático de las actitudes proposicionales” no tiene lugar en una ciencia

---

<sup>140</sup> Tal vez sea posible redoblar la apuesta “terapéutica” en este punto, para verlo como “disolviendo” la dicotomía realismo vs. antirrealismo. Para una lectura de su posición como “pragmatista” puede verse Orlando, E. (2001).

sería, o que nuestro vocabulario mentalista está “testarudamente en desacuerdo con los patrones científicos” (Quine 1960:219). Él considera que Quine está en lo cierto al mantener que el vocabulario mentalista no puede incorporarse o reducirse al vocabulario de la física, tal como lo atestiguan las líneas argumentativas que reconstruimos en el apartado anterior. Asimismo, subraya que Quine está de acuerdo con su versión del monismo de casos, que sostiene que cada objeto o evento que esté identificado en términos mentales es idéntico a un objeto o evento identificable en términos físicos, “aunque las *clases* configuradas mediante el vocabulario mental no puedan ser equiparadas por definición o mediante leyes estrictas a clases definibles con el vocabulario de la física” (Davidson 1997b/2001: 72). Esta posición que, tal como dijimos, él llama “monismo anómalo” se acerca, a su juicio, a reconciliar la mayor parte de lo que Quine ha afirmado acerca de las actitudes proposicionales, aunque no todo. El monismo anómalo da sentido a la afirmación de que las actitudes son disposiciones a comportarse de determinadas maneras, las cuales son a su vez estados fisiológicos, que finalmente son estados físicos, como también da sentido a la afirmación de que las descripciones intencionales no son reducibles a descripciones conductuales o físicas, y por ello no son apropiadas para ser incorporadas a un sistema estricto de leyes. El vocabulario mental “es práctico e indispensable, pero no está hecho para la ciencia más precisa” (Davidson 1997b/2001: 72). No obstante, dice Davidson, el monismo anómalo no sugiere que los eventos y estados mentales sean “meramente proyectados por el que atribuye algo a un agente; por el contrario, mantiene que los eventos mentales son tan reales como los eventos físicos, al ser idénticos a ellos, y que las atribuciones de estados son también objetivas. La descripción de Quine de las atribuciones de actitudes como retratos dramáticos no implica que no haya nada que retratar” (Davidson 1997b/2001: 72). A juicio de Davidson, pues, el hecho de que el vocabulario mental no encaje de manera que pueda ser incluido en las ciencias como la física o la fisiología no puede tomarse *per se* como una crítica a la realidad de los estados, eventos y objetos para cuya descripción se emplea tal vocabulario. Y agrega: “una física perfeccionada debe comprender cada objeto y evento, pero esto es un requisito ontológico y nomológico que define el propósito de la física: la física no dice nada de los intereses que exigen otras maneras de caracterizar las cosas” (Davidson 1997b/2001: 72).

Por su parte, veamos cuál es su posición con respecto a las implicaciones de (b) la *indeterminación* de la interpretación. Cabe recordar que en “Mental Events”, él sostiene que el carácter heteronómico de las generalizaciones psicofísicas se remonta a dicho rasgo. Tal es la

razón por la que he propuesto que nos internemos en la indeterminación de la interpretación para indagar si nos permite captar la peculiaridad de lo mental. En efecto, en su (1970) Davidson sostiene que en nuestra necesidad de hallar comprensible lo que dice una persona, trataremos de:

construir una teoría que la haga consistente, creyente de verdades y amante de lo bueno (...) Siendo la vida lo que es, no habrá ninguna teoría sencilla que satisfaga completamente estos requisitos. Muchas teorías efectuarán un compromiso más o menos aceptable, y entre estas teorías puede no haber bases objetivas para la elección. El carácter heteronómico de los enunciados generales que enlazan lo mental y lo físico se remonta a este papel central de la traducción en la descripción de todas las actitudes proposicionales, y a la *indeterminación* de la traducción (1970/1980:222; mis itálicas).

En (1997b) se pregunta, entonces, si la tesis de la indeterminación implica que las actitudes proposicionales no son reales. En esta línea, nos recuerda que la idea de que la indeterminación socava la realidad de los estados mentales es compartida por un cierto número de filósofos. Por ejemplo, Jerry Fodor ha pensado que esto constituye una reducción al absurdo de la indeterminación. En tanto que, por otro lado, a juicio de Davidson, Daniel Dennett acepta la indeterminación, pero está de acuerdo con que resta realidad a los estados mentales.<sup>141</sup> El punto

---

<sup>141</sup> Esta observación de Davidson amerita la siguiente aclaración: en rigor, Davidson (1997b/2001: 81-84) dirá luego que Dennett lo acusa de que, en su concepción, los estados mentales resultan *demasiado* reales (“Real Patterns”). En efecto, nos recuerda que la idea de Dennett es que lo que es real es el comportamiento y que los estados intencionales son patrones en el comportamiento. Estos patrones no se definen en términos de comportamiento, sino que son percibidos por un observador cuando éste adopta la “actitud intencional”. El valor de estos patrones es que reducen una vasta (y para nosotros indescritiblemente compleja) situación física a algo que podemos captar, algo sobre la base de lo cual podemos hacer predicciones aproximadas. Los patrones son reales en algún sentido, aunque sean abstractos, y personas distintas pueden percibir patrones distintos en el mismo terreno de comportamiento. Algunos de estos patrones pueden ser mejores que otros a la hora de predecir o entender algunos fenómenos. Todos estos patrones son “reales”, pero cuando son diferentes no tiene sentido decir cuál de ellos representa las actitudes reales del sujeto en cuestión. Ahora bien, en opinión de Davidson, no es fácil hacer un juicio sobre esta propuesta de Dennett y, todavía, dice, es más difícil captar cómo se relaciona con la cuestión del realismo. Ciertamente, no *todo* patrón que percibimos en el comportamiento es patrón de actitudes proposicionales. ¿Pero lo es alguno de ellos? Davidson lo interpela, pues, del modo siguiente. Dennett nos dice que los patrones son abstracciones como los centros de gravedad. Pero entonces ¿pueden ser creencias y deseos? Nos gusta pensar, dice Davidson –siguiendo su impulso original de (1963)- que las creencias y los deseos son estados de un cuerpo físico que pueden tener consecuencias causales; pero las abstracciones no tienen relaciones causales. Las fuerzas no actúan en los centros de gravedad sino en las cosas que tienen centros de gravedad, pero las fuerzas ciertamente alteran nuestras creencias e intenciones. ¿Percibimos patrones? A Davidson le parece que no; lo que percibimos es algo que *tiene* un determinado patrón y, con suerte (y la actitud apropiada) podemos percibir *que* tiene ese patrón. Así, pues, la cuestión no es si los patrones son reales. Davidson confiesa que al no ser un nominalista, cree que los patrones, al igual que las formas y los números son tan reales como puedan ser. Pero no ve cómo las actitudes proposicionales de una persona pueden ser patrones. Si nos preguntamos qué es lo que exhibe el patrón podemos decir que es el comportamiento observable de la persona. Pero en ambos casos *no está en juego el tema del rango ontológico* de las actitudes. A su juicio, Dennett ha confundido dos cosas. La primera es si las actitudes son *entidades*, y sobre eso Davidson piensa que no, a menos que se piense que los estados son entidades. Si no se piensa esto, se debería hablar simplemente de personas que tienen actitudes, lo que significaría que determinadas cosas que predicamos de ellas son verdad. La segunda es si hay una respuesta correcta a la pregunta de si alguien tiene una

de Davidson es, en cambio, discutir esta inferencia: pretende retener la indeterminación y, a la vez, dejar intacta la realidad de lo mental.<sup>142</sup>

Si uno niega, como hace Davidson, que al atribuir una creencia uno esté identificando un objeto con el cual el sujeto está en una relación peculiar de “captación”, no resulta necesario elegir una entre muchas interpretaciones exitosas como *el* contenido de la creencia.<sup>143</sup> Esto es así, porque como aclaré más arriba a partir de la analogía con los sistemas de medición, al adoptar un patrón de medida no se agregan objetos a la ontología ni se modifican las relaciones preexistentes entre los individuos y además porque, según sugerí, la lectura adecuada de Davidson consiste en pensar las atribuciones de creencias estableciendo relaciones entre unos individuos y otros y con el entorno, y no con objetos a los que tendríamos un oscuro modo de acceso.

En efecto, como concede Davidson, quizás sea natural suponer que la entidad a la que nos referimos para especificar el contenido de una creencia es el *objeto* de la creencia, una entidad que el sujeto de la creencia considera o capta, la *cosa* que el sujeto cree. Dada la semántica relacional de las oraciones sobre las creencias, y dada también la suposición natural de

---

determinada actitud. Esto lo interpreta no como una cuestión acerca de la vaguedad o de los casos dudosos, sino como una cuestión acerca de si hay fundamentos *objetivos* para escoger entre hipótesis rivales. Dennett ha afirmado que la respuesta a la segunda cuestión es que no existen tales fundamentos, pero Davidson no cree que haya dado ninguna razón para que aceptemos esta respuesta. A su juicio, tenemos que preguntarnos qué es lo que hace “objetivos” a los fundamentos. La fuente última (no el fundamento) de la objetividad es, nos dice, la intersubjetividad (1997:83). Tal como adelantamos en la primera parte de este trabajo, según su concepción, la posibilidad del pensamiento tanto como de la comunicación depende de un nexo triangular de relaciones causales que implican reacciones de dos (o más) criaturas respecto a cada uno y reacciones a los estímulos compartidos provenientes del mundo. Tal nexo es el que proporciona las condiciones necesarias para que tenga aplicación el concepto de error. Tal como anticipamos, Davidson sigue a Wittgenstein al afirmar que sin una segunda persona no hay base para juzgar que una reacción es errónea o, por lo tanto, que es correcta. En suma: frente a Dennett, defiende la *objetividad* en un sentido que pretende conciliar tanto con la *indeterminación* de la interpretación, así como con el carácter *no entitativo* de nuestras actitudes proposicionales. En esta línea, es que más arriba afirmé que en este punto Dennett parece no advertir el deslinde entre la dimensión semántica y la dimensión ontológica que introduce la analogía que estamos considerando. No obstante, dados los propósitos que persigo en el presente trabajo no me internaré en esta polémica. Sin embargo, en breve voy a volver –aunque muy sucintamente– sobre esta cuestión al tratar la respuesta de Davidson a Rorty (1997), esto es, cuando veamos que la indeterminación y la invariancia son, para él, dos caras de la misma moneda.

<sup>142</sup> Para una discusión reciente de este punto, puede verse el trabajo de Rawling P. (2003), quien sostiene –contra Davidson– que no es posible conciliar la indeterminación con la realidad de las actitudes proposicionales. En rigor, su posición es más fuerte aún: para Rawling la indeterminación produce como resultado – implica– que no hay de eso.

<sup>143</sup> Asimismo, esto explica por qué las atribuciones en primera persona no son infalibles. Las atribuciones de estados mentales que hago de mí misma son unas de entre las que hacen otros intérpretes. Puedo hacerlas por medio de una emisión externa –en voz alta– o interna– como cuando me digo a mí misma que estoy creyendo que *p* o deseando que *q*-. Las atribuciones que me hago serán correctas o no, dependiendo de exactamente los mismos parámetros con los que se juzgan las atribuciones que me hacen otros intérpretes. O todas son falibles, o ninguna lo es. Volveré sobre la peculiar concepción davidsoniana de la autoridad de primera persona en el capítulo 6.

que la entidad que identifica la creencia es el objeto que capta el sujeto de la creencia, nos vemos virtualmente forzados a concluir que la entidad designada debe ser única, puesto que si más de una entidad especificase el mismo estado de creencia igualmente bien, ¿qué entidad deberíamos decir que capta el sujeto? Este acertijo, a su juicio, puede resolverse desechando la idea de tratar los objetos gramaticales en las oraciones de creencia como términos que nombran objetos psicológicos reales. En este sentido afirma que “el único *objeto* que se requiere para la existencia de una creencia es la de un sujeto que crea algo” (1997b/2001: 74). Lo único que se necesita, entonces, para la verdad de la atribución de una actitud es que el predicado que se emplee sea verdad respecto de la persona que tiene esa actitud.

Precisamente, Davidson afirma que una de las virtudes de la analogía en cuestión es que aclara por qué diferentes asignaciones de objetos pueden capturar toda la información relevante acerca de una situación sin comprometer la verdad o “realidad” de la situación. Nadie piensa que el hecho de que podamos registrar el peso ya sea en libras o en kilogramos muestra que hay algo que no es real acerca del peso de un objeto; se pueden usar diferentes conjuntos de números para captar exactamente los mismos hechos. Supóngase que el hecho es que A pesa lo mismo que B; entonces el número que se asigna para medir los pesos de A y de B es el mismo número, ya sea que midamos el peso en libras o en kilogramos (...) Lo único que es necesario es que las entidades que usamos para captar cuánto pesan las cosas tengan una estructura en la cual puedan representarse ciertas características de los objetos con peso (1997b/2001: 75).<sup>144</sup>

¿Cuáles son, entonces, los objetos que podemos usar para captar las actitudes? Claramente, dice, deben constituir aproximadamente un campo tan complejamente configurado como lo son las propias actitudes. Los rasgos más conspicuos de las actitudes son sus estructuras básicamente racionales (si alguien cree que todo es blanco, esa persona tiene una creencia que implica que la nieve es blanca) y sus relaciones con el mundo (la creencia de que la nieve es

---

<sup>144</sup> Otra de las virtudes de esta analogía es, para Davidson, que aclara por qué Quine está justificado a insistir (contra Chomsky y otros) en que la indeterminación de la traducción es distinta de la subdeterminación de la teoría ante toda evidencia posible: “ciertamente es un problema el de si puede darse la subdeterminación en el sentido de Quine, puesto que ésta requiere que haya teorías empíricamente equivalentes pero incompatibles (aunque concede que Quine ha cambiado de opinión acerca de esta cuestión en más de una ocasión, cfr. *Pursuit of Truth*) La indeterminación no es así; las teorías empíricamente equivalentes que se aceptan como igualmente buenas para comprender a un agente no son incompatibles, no lo son más de lo que la medición del peso en libras o en kilos implica teorías incompatibles sobre el peso” (1997b/2001:76)

blanca es verdadera si y sólo si la nieve es blanca). Tal como anticipamos, a su juicio, nuestras *oraciones* o, alternativamente, nuestras *emisiones* son entidades que tienen este tipo de propiedades requeridas. Análogamente, entonces, podemos emplear distintos conjuntos de oraciones o emisiones para describir el mismo conjunto de relaciones interpersonales. Similitudes –supuestas u observadas- entre diversos estados en los que los humanos nos relacionamos unos con otros y con nuestro entorno quedan reflejadas por medio de las oraciones o emisiones con las que identificamos los estados mentales de los humanos involucrados, pero dichas emisiones no constituyen los estados mentales en cuestión, sólo resultan herramientas útiles para señalar similitudes objetivas entre diversos estados relacionales complejos.<sup>145</sup>

Ahora bien, llegados a este punto cabe preguntarnos si la *indeterminación* de la interpretación permite capturar el rasgo peculiar del vocabulario mental frente al físico, que abriría el hiato que supone el argumento en contra de la existencia de leyes psicofísicas estrictas, esto es, la brecha que subyace al “anomalismo” de lo mental.

Rorty (1997) ofrece dos reparos a propósito de la obra de Davidson. En la primera parte de este trabajo, dimos cuenta de uno de ellos, a saber, la propensión de Davidson a conectar todos sus argumentos con el proyecto de construir teorías T-para lenguajes naturales (en suma, con el proyecto interpretacionista). El segundo de sus reparos guarda relación con el punto que estamos considerando. A su juicio, Davidson debería abandonar la tesis de la indeterminación de la interpretación. Lo que él teme es que esta tesis implique que hay algo misterioso acerca de lo mental, tal como que no hay *hechos decisorios* acerca del significado o las actitudes proposicionales (*fact of the matter*), punto que a Rorty le parece un residuo del cientificismo y el reduccionismo que permeó al positivismo lógico.

En respuesta a esta imputación, Davidson (1998/2005:316) sostiene que en su concepción lo mental “no es más misterioso que la biología molecular o la cosmología...las actitudes proposicionales son tan reales como los átomos y los bates de baseball, y los hechos acerca de ellos son tan reales como los hechos acerca de cualquier otra cosa”. Al abrazar la tesis de la indeterminación, simplemente trató de atacar al platonismo del significado, pero “este punto

---

<sup>145</sup> Asimismo, cabe recordar que en “The Myth of the Subjective”, defiende la idea de que si bien las creencias de un individuo cualquiera se expresan en uno u otro lenguaje natural, esto no nos compromete con la idea de que las creencias sean relativas a esquemas, porque son las representaciones las que resultan relativas a esquemas (cf. 1974) y al adscribir a alguien una creencia no nos hemos de comprometer con una “representación mental”, es decir, con un “objeto ante la mente”.

negativo no implica que no haya –tal como le imputa Rorty- hechos decisivos: los hechos son las relaciones empíricas entre el hablante, sus oraciones y su entorno. Este patrón es invariante” (1998/2005:317).

Lo que me interesa señalar aquí es que en el seno de su respuesta a Rorty, el propio Davidson concede que en “Mental Events” incurrió en un error al sostener que la irreducibilidad de lo mental se debía, entre otras cosas, a la indeterminación de la interpretación.<sup>146</sup> El error, nos dice, es obvio: tal como admitió en “Three varieties of knowledge” (1991): la indeterminación es *endémica* en *todas* las disciplinas; tiene lugar toda vez que un vocabulario es lo suficientemente rico como para describir un fenómeno de más de una manera. Lo que importa es lo que es invariante. La indeterminación no es nada más que la otra cara de la invariancia. “Son las *ratios* las que son invariantes. Sólo las invariancias son los “hechos decisivos” (1998/2005:316).

La cuestión relevante para nuestros propósitos es, entonces, que la indeterminación de la interpretación no permite capturar el presunto carácter especial de los conceptos mentales.

---

<sup>146</sup> Es interesante advertir que muchos años después, en su (1997b), explicita algunas razones para la irreducibilidad de lo mental que no ofreció en “Mental Events”. Como veremos, una de ellas conecta directamente con su proyecto seminal de (1963): “Acciones, razones y causas”. En esta línea, sostiene que una propiedad poco advertida que separa el lenguaje de nuestros estados mentales del vocabulario de las ciencias avanzadas es su *dependencia de los conceptos causales*. El lenguaje físico ordinario, al igual que el lenguaje psicológico, está lleno de conceptos causales: la noción de catalizador es tan causal como la noción de acción intencional. La diferencia reside, a su juicio, en la promesa, intrínseca a la física, pero irrelevante para la psicología, de que con suficiente tiempo e investigación el concepto causal podrá ser sustituido por una explicación del mecanismo que explicará lo que la tosca noción causal meramente apuntaba. Si queremos, las leyes de la física pueden llamarse causales. La cuestión es que no emplean conceptos causales. La concentración de la psicología en el papel causal de las razones excluye toda esperanza de que los conceptos mentales básicos puedan acomodarse en un sistema cerrado de leyes. A su vez, señala otra razón novedosa respecto de “Mental Events”, en la que buscaré la clave para entender el rasgo peculiar de lo mental. Sostiene entonces que “hay otra consideración más debatida que milita en contra de la reducción – definicional y nomológica- de los conceptos mentales a los de la física, a saber, el hecho de que las actitudes proposicionales y otros estados y eventos relacionados se identifican en parte en términos de sus relaciones causales o de otro tipo de eventos ajenos en tiempo y lugar al agente que caracterizan. Así, pues, se sostiene correctamente que la historia de un aprendizaje y uso individual de las palabras y los conceptos es necesariamente, en casos centrales, un factor que determina lo que las palabras significan y los contenidos de los conceptos. La comunicación interpersonal desempeña un papel necesario en la posibilidad y naturaleza del pensamiento. Si los externismos de estas clases son también realmente rasgos dominantes e ineludibles de lo mental, es clara la imposibilidad de incorporar la psicología a una teoría científica unificada del mundo” (1997b/2001:71). En la última parte de este trabajo, precisamente, argumentaré en qué sentido este peculiar externismo, la *triangulación*, permite arrojar luz sobre el rasgo peculiar de lo mental que, como paso a sostener, hay que buscarlo en su especial *normatividad*.

En efecto, Davidson sostiene que John McDowell tiene en claro que nada de lo que él dijo torna misterioso a lo mental. El también está de acuerdo en que “hay una razón para enfatizar la irreducibilidad de lo mental, pero sólo porque la irreducibilidad surge de algo más interesante que la indeterminación de la interpretación. Él enfatiza correctamente el rol constitutivo de la *normatividad* en todos los conceptos mentales o psicológicos” (1998/2005:318). Y dice explícitamente que el vocabulario mental no es “privilegiado” porque es *irreducible* (como Rorty piensa que él cree) sino que “es irreducible porque es *normativo*” (1998/2005: 318; mis itálicas).<sup>147</sup>

Volvamos, entonces, sobre la analogía que estamos examinando para ver si podemos esclarecer la peculiaridad en cuestión a partir del candidato propuesto por Davidson: la *normatividad* de lo mental.

#### **4.3. La analogía entre las actitudes proposicionales y las actitudes numéricas y la *normatividad* de lo mental**

En la medida en que no logramos descubrir un patrón coherente y plausible en las actitudes y acciones de los otros, simplemente renunciamos a la posibilidad de tratarlos como *personas*.  
(Davidson 1970/1980:221-2; mis itálicas)

Según vimos en la primera parte de nuestro trabajo, una de las acepciones tradicionales de la *normatividad*, desde su dimensión *deóntica*, es la que se pone en juego en la noción de *regla*. Al considerarla, invocamos la distinción de Searle (1969) entre reglas regulativas y reglas constitutivas. Recordemos que, a su juicio, las primeras regulan formas de conducta previamente existentes, mientras que las segundas no regulan sino que *constituyen* o *definen* nuevas formas de conducta.

---

<sup>147</sup> En su “Reply a Rorty” (Hahn 1999: 599), subraya la diferencia entre el vocabulario intensional de la psicología y el resto. En ese contexto, distingue dos sentidos tradicionales de irreducibilidad (definicional y nomológica) y otro según el cual el vocabulario mental lo es en la medida en que no es una parte optativa de nuestros recursos conceptuales. En breve, veremos que la profundidad de su posición se pone en juego en el último sentido, es decir, en el carácter “inescapable” de nuestro vocabulario mental.

La analogía entre las actitudes proposicionales y las actitudes numéricas resulta particularmente importante para nuestros fines porque es, precisamente, al tratar de clarificar qué entiende por *reglas* o "*principios constitutivos*" de una teoría que Davidson la introduce en "Mental Events".

Siguiendo a Davidson, cuando nos enfrentamos con el conjunto de las leyes y postulados que rigen la medición de los objetos físicos macroscópicos no tiene sentido preguntarse si cada una de esas leyes o axiomas son analíticos o sintéticos, es decir, si son verdaderos en virtud del significado de los términos o si lo son por razones empíricas. Davidson adopta explícitamente la tesis quineana de acuerdo con la cual hay que rechazar una distinción tajante entre enunciados analíticos y sintéticos. Pero, entonces ¿qué estatus tienen desde esta perspectiva los postulados y leyes de la teoría de la medición? La respuesta de Davidson es: son *leyes constitutivas*, es decir, sintéticas *a priori*. En sus propios términos: "el conjunto total de axiomas, leyes, o postulados para la medida de longitud es parte constitutiva de la idea de un sistema de objetos macroscópicos, rígidos, físicos" (1970/1980:221). Esto quiere decir que el hecho de que un objeto deba tener algún largo, algún ancho y algún fondo es parte *constitutiva* de nuestro *concepto de objeto físico*. Paralelamente, en el caso del sistema de conceptos mentales, Davidson también considera que hay ciertos principios que los constituyen. Así, sostiene que "no podemos atribuir ninguna actitud proposicional a un agente excepto dentro del marco de una teoría viable de sus creencias, deseos, intenciones y decisiones" y este marco de creencias y deseos está gobernado por "el ideal *constitutivo* de la *racionalidad*" (1970/1980: 223). Este principio, según vimos en la primera parte de este trabajo, no es sino una de las dos caras del célebre principio de caridad, una de las restricciones fundamentales de la teoría de la interpretación a la hora de tornarla aplicable empíricamente. La racionalidad es un principio constitutivo de lo mental, esto es, de lo intencional, en el sentido de que abandonar la racionalidad de lo mental -es decir, fundamentalmente de las creencias, deseos y acciones humanas- sería como "cambiar de tema", esto es, no aceptar como criterio de lo mental el vocabulario de las actitudes proposicionales y, consecuentemente, abandonar nuestra concepción del hombre como animal racional.

En otras palabras, así como la satisfacción de las condiciones para medir longitud puede considerarse como constitutiva del rango de aplicación de las ciencias que emplean estas medidas, la satisfacción de las condiciones de consistencia y coherencia racional puede considerarse constitutiva del rango de aplicación de conceptos tales como los de creencia, deseo,

intención y acción. En consecuencia, a partir de la analogía que estamos analizando, esta acepción de la *normatividad* en términos de *principios constitutivos* no parece introducir ninguna diferencia radical entre los ámbitos conceptuales físico y mental, del tipo requerido por el argumento en contra de la existencia de leyes psicofísicas estrictas. El propio Davidson admite que así como hay elementos constitutivos en la aplicación de los conceptos psicológicos, los hay también para la aplicación de los conceptos físicos (1970/1980: 254).<sup>148</sup>

La dificultad que quiero subrayar es, pues, que no parece resultar suficiente entender la *normatividad* en términos de meros *principios constitutivos* a la hora de capturar lo que sería peculiar de lo mental. La idea que subyace a mi planteo es la siguiente: si en su esquema de eventos lo físico termina siendo de naturaleza tan lingüística o conceptual como lo mental<sup>149</sup> ¿cuáles son las razones para sostener que hay una diferencia constitutiva radical entre ambos, un dualismo profundo entre los conceptos mentales y los físicos? Esto es, dado que lo físico también se exhibe como conceptos en el esquema davidsoniano de eventos ¿en qué sentido los conceptos físicos se sustraerían al alcance de la consistencia y carecerían de condiciones de corrección? Así como la atribución de una actitud proposicional a alguien en un momento dado está constreñida por otras asignaciones de actitudes que mantengan consistencia, la asignación de una medida de longitud a un objeto particular debe respetar, por ejemplo, la ley de transitividad involucrada en la relación “más largo que” y, de este modo, está también constreñida por otras asignaciones de medida. En mi opinión, hasta aquí no habría nada “esencialmente normativo” en el primer caso, al menos nada que no esté presente en el segundo. En la medida en que son conceptos, tanto los físicos como los mentales tienen condiciones de corrección y están regidos por principios constitutivos de su rango posible de aplicación. Cuando en la primera parte del trabajo examinamos el principio de caridad como “antesala de la normatividad”, distinguimos

---

<sup>148</sup> En el caso de lo físico habrá que atender a los axiomas y postulados de la ciencia física; en el caso de lo mental, el problema de la racionalidad nos remite a las teorías existentes que buscan sistematizar las nociones de acción racional, creencia y preferencia, es decir, a la teoría de la decisión aludida en el capítulo 3. Respecto de esto último Levi, I. (Hahn 1999:533) lanzó duras críticas al planteo davidsoniano. En efecto, sostiene que una vez que con el monismo anómalo se abandonan los intentos de integración de la psicología con la biología y la teoría física y se admite que los principios de racionalidad son violados a menudo en la práctica, no llega a ver una gran ventaja en insistir en las virtudes predictivas o explicativas de los principios de la creencia racional, el deseo y la toma de decisiones. En esta línea, sostiene que el uso que Davidson hizo de las teorías formales de la racionalidad ha sido sumamente abstracto y especulativo y, más aún, que en lo concerniente al rol de la teoría de la decisión en la interpretación “ha sostenido puntos inconsistentes”.

<sup>149</sup> Recordemos que en la caracterización que ofrece en (1970) los eventos físicos son aquellos que seleccionamos mediante *descripciones u oraciones abiertas* que sólo contienen, esencialmente, el vocabulario físico (1970/1980:211).

tres interpretaciones del mismo. Dijimos que según la interpretación *intermedia*, la consistencia y la coherencia operan como un principio *gnoseológico* acerca del intérprete, constitutivo del *proceso de atribución*. En este sentido, podríamos pensar a propósito de la analogía que ahora estamos considerando, que la caridad, entendida en términos de coherencia, regula nuestra atribución de los conceptos físicos (tales como longitud o temperatura; cf. 1970/1980:254): si sostengo que el objeto A es más largo que el objeto B y que el objeto B es más largo que el C, estoy constreñida a sostener que el objeto A es más largo que el objeto C.

Ahora bien, Davidson insiste en que la “virtud” básica que segrega los conceptos mentales de los de las ciencias naturales es el “tipo especial de caridad” (1998/2005:319) requerida para comprender los pensamientos, el discurso y otras acciones de los agentes. Así afirma:

En el caso de la atribución de actitudes hay un *giro esencial* (*essential twist*): dado que las actitudes que atribuimos están constituidas en parte por consideraciones *racionales*, están también *normativamente* determinadas (2001b: 297; mis itálicas).

Cabe, pues, preguntarnos en qué consiste este tipo “especial” de caridad de los conceptos mentales, este “giro esencial” dado por su determinación *normativa*. Tal como adelantamos en la primera parte de nuestro trabajo, quizás aquello que la normatividad de lo mental está subrayando es que se trata de reglas constitutivas de un tipo especial, en el sentido de que su peculiaridad consiste en que sean *estas* reglas. Su singularidad habría que buscarla entonces, a mi juicio, en otro de los rasgos de la noción de creencia que tematizamos en el capítulo 3, a saber: su *holismo constitutivo* (que distinguimos del meramente *metodológico*).

En efecto, recordemos que para Davidson el argumento en contra de la existencia de leyes psicofísicas estrictas provee la clave de la peculiaridad de la psicología en la medida en que “conduce desde el carácter necesariamente *holista* de las interpretaciones de actitudes proposicionales al reconocimiento de un elemento irreduciblemente *normativo* en todas las atribuciones de actitud” (1974c/1980: 241; mis itálicas).

Así, pues, como más arriba vimos que la variante de su holismo que llamamos “inter-actitud” o “psicológica” hacía el trabajo en su argumento en contra del reduccionismo *definicional*, creo que es el holismo “del contenido” o “semántico” aquel que ahora desempeña el

trabajo central en su argumento en contra del reduccionismo *nomológico*, y aquel que permite comenzar a clarificar el carácter especial de las reglas constitutivas de lo mental.

Recordemos que según su *tesis holista del contenido*, la identidad de un pensamiento dado depende, en parte, de sus relaciones con otros pensamientos. En virtud de que esta variedad de holismo concierne a las relaciones entre varias creencias *dentro* de la categoría de creencia, Davidson la denomina “holismo *intra-actitud*”. La idea es que la individuación de lo creído depende de otras cosas creídas, esto es, no podemos tener un pensamiento aislado porque cada pensamiento supone un contenido y dicho contenido está determinado por su relación con los contenidos de otras creencias atribuidas. Ahora bien, la normatividad es un rasgo esencial de lo mental porque lo mental- a diferencia de lo físico- se constituye en un marco de actitudes que tienen contenidos *proposicionales* que guardan entre sí relaciones *lógicas*. En este sentido, no podemos tener cualquier colección arbitraria de pensamientos: la lógica nos restringe. Esto es, nuestros pensamientos están constreñidos por las relaciones que imponen la consistencia y la coherencia lógica propias del principio de racionalidad, que resulta una regla de tipo especial en tanto no tiene “eco en la teoría física” (1974c/1980:231). En otras palabras: la razón central en la que reposa la irreducibilidad de lo mental proviene del hecho de que cualquier explicación correcta de las creencias y otras actitudes de un agente debe reconocer un elemento *normativo* inherente en los contenidos cognitivos (1997b/2001:71). Una actitud tiene determinadas relaciones lógicas con otras; estas relaciones, aunque distorsionadas por los límites de nuestras capacidades sirven, sin embargo, para ubicar y de este modo identificar los contenidos de nuestros pensamientos. Cuando tratamos al mundo como algo sin mente, tal y como pasa con las ciencias naturales, nada se corresponde con esta dimensión de lo mental.

En función de lo expuesto, entonces, parece que el tipo “especial” de caridad que se pone en juego en la normatividad de lo mental amerita la interpretación fuerte que presentamos en el capítulo 3, según la cual el principio de caridad no sería un mero principio metodológico concerniente a la atribución de significados y creencias sino un principio *constitutivo de los significados y las creencias mismas*. En otras palabras: la caridad es “especial” en la medida en que no es meramente constitutiva del proceso de atribución, sino de *lo atribuido* mismo.

Tal es el tipo de peculiaridad que parece desempeñar un rol central en el argumento davidsoniano a favor del anomalismo, es decir, aquello que dota de “carnadura” a la premisa crucial (ii), que sostiene que los principios constitutivos de lo mental y lo físico difieren

*radicalmente*. El hiato se genera a partir de la dicotomía (y no de una mera distinción de grado) entre el carácter *normativo* de los conceptos mentales y el carácter *descriptivo* de los conceptos físicos. La consecuencia que se desprende de ello es un *profundo dualismo conceptual*.<sup>150</sup>

Dicho dualismo, a su vez, tiene la siguiente consecuencia: supone un *problema insuperable* para dar una descripción plena de la emergencia del pensamiento (Davidson 1997c/2001: 127; mis itálicas)<sup>151</sup>. Se trata de la dificultad de describir los primeros estadios en la maduración de la razón, los estadios que preceden a la situación en que conceptos como los de creencia, intención y deseo tienen una aplicación clara. Dice Davidson:

Tanto en la evolución del pensamiento en la historia de la humanidad como en la evolución del pensamiento de un individuo hay un estadio en el que no hay pensamiento y un estadio en el que sí lo hay. Describir la aparición del pensamiento sería describir el proceso que conduce del primero al segundo de esos estadios. Carecemos de un vocabulario satisfactorio para describir los estadios intermedios... Tenemos muchos vocabularios para describir la naturaleza cuando la consideramos no mental y tenemos un vocabulario mentalista para describir el pensamiento y la acción intencional; de lo que carecemos es de una manera de describir lo que ocurre entre estas dos cosas... Esto quiere decir que quizás haya un *problema insuperable* para dar una descripción plena de la aparición del pensamiento. Doy gracias de no estar en el campo de la psicología del desarrollo (1997c/2001: 127-128).

La consecuencia dramática que se sigue, pues, del dualismo conceptual profundo que tiene lugar a partir de la inscripción de la noción de *creencia* en el espacio “normativo” de las

---

<sup>150</sup> A juicio de Rorty (1999) esta diferencia categorial de Davidson es una “distinción filosófica espuria” y dice que no ve por qué el holismo de lo intencional habría de tener la consecuencia que acabamos de señalar. Él piensa en el holismo de la adscripción de predicados como una *cuestión de grado* y no ve por qué la complejidad creciente de sus condiciones de aplicación tendría que ser un síntoma de hiatos filosóficamente interesantes. Creo que parte de su “ceguera” consiste en no advertir las distintas variedades de holismo y las correlativas interpretaciones de la caridad que hemos tratado de distinguir. No obstante, considero que Rorty señala en la dirección correcta al sostener que detrás del planteo davidsoniano hay en juego una dicotomía –y no una mera distinción de grado-. Sólo que tal como acabo de anticipar, la dicotomía en cuestión se da, en mi opinión, entre los conceptos mentales que estarían *normativamente* determinados y los conceptos físicos que serían meramente *descriptivos*: en suma, la dicotomía *descriptivo/normativo*. Voy a detenerme, especialmente, sobre esta cuestión en el cap. 7.

<sup>151</sup> Cabe señalar que otra de las derivaciones cruciales que se le imputan es el *epifenomenalismo* o la *impotencia causal de lo mental*. En efecto, paradigmáticamente, Kim (1989) ha sostenido que si un evento causa otro lo hace *en virtud de poseer* la propiedad *física* que está involucrada en la ley, y no en virtud de sus propiedades mentales. De esta suerte, a su juicio se sigue que: (i) las propiedades mentales de un evento resultan causalmente impotentes; (ii) las propiedades mentales de un evento resultan explicativamente impotentes y (iii) el “monismo anómalo” no resulta plausible como posición fisicalista no reductiva (se torna peligrosamente cercano al eliminativismo). Por razones de pertinencia, no me voy a ocupar aquí de esta posible consecuencia del monismo anómalo. No obstante, para profundizarla pueden consultarse *inter alia*: Kim, J. (2000); Pérez, D. (1999) y Antony, L. (1994).

razones es la de detener, por principio, la discusión acerca de la *emergencia* del pensamiento. Se la clausura *a priori*.

Davidson considera que sería malinterpretar su posición resumirla diciendo que la psicología (la parte de la que él se ocupa) no es una ciencia; de este modo, sostiene que sus conclusiones (aunque no así sus argumentos) no son nuevas ni aptas para propiciar mucho debate: aquello que ha deseado subrayar es que la psicología se separa de otras ciencias “de una manera *interesante*” (1974c/1980: 240). La conclusión de la irreducibilidad nomológica de lo psicológico significa que no puede esperarse que las ciencias sociales se desarrollen de maneras exactamente paralelas a las ciencias físicas. El punto es que, dado que los fenómenos psicológicos no constituyen un sistema cerrado, no son pasibles de predicción precisa o subsunción bajo leyes deterministas. Davidson insiste en que el límite ubicado en las ciencias sociales es fijado, pues, por nosotros cuando consideramos a los hombres como agentes *racionales*.<sup>152</sup> Tal es el modo en que considera que el argumento en contra de la existencia de leyes psicofísicas estrictas provee la clave para la *peculiaridad de la psicología*.

Ahora bien, aún bajo la presunción optimista de que hemos logrado esclarecer el tipo peculiar de normatividad de lo mental a partir de su acepción *deóntica*, cabe preguntarnos si ello es suficiente para captar la profundidad de la posición de Davidson. En otros términos: ¿es posible escindir esta acepción de la normatividad de aquella otra que completa su tratamiento tradicional, esto es, su acepción *evaluativa*? Llegados a este punto cabe recordar lo que señalamos en el capítulo 3: Davidson (1998/2005:326) nos dice cuál es su objeción central a la idea de concebir la normatividad en términos de meras *reglas*: concede, desde luego, que hay reglas en nuestra práctica lingüística, pero subraya que son “enteramente contingentes, dependientes de “*valores ulteriores*” (*further values*), no de valores intrínsecos al lenguaje.

Tal como sostuve en la primera parte de este trabajo al detenerme en el debate sobre normas y valores entablado entre Putnam y Habermas (2008), Putnam expone allí su rechazo de la dicotomía que, a su juicio, Habermas postula entre las “normas” y los “valores”, entre la dimensiones *deóntica* y *evaluativa* de la normatividad. Recordemos rápidamente que para este

---

<sup>152</sup> En este sentido, y a propósito de lo que anticipamos en el cap. 2 (¿Comprender la comprensión?), Davidson parece volver sobre la discusión metodológica que suscitó el surgimiento de las ciencias del espíritu, para reivindicarlas desde un abordaje alternativo al explicacionista clásico. Así, por ejemplo, en (1997d/2005) se sitúa explícitamente en la tradición comprensivista. Al respecto, pueden verse Malpas, J. (2002b) y Ramberg, B. (2003).

último la validez *deóntica* de las normas tiene el sentido *absoluto de una obligación incondicional y universal*, mientras que el que unos valores sean atractivos tiene el sentido *relativo* de una estimación o evaluación de bienes adoptada en ciertas culturas y formas de vida. Frente a esta posición, Putnam defiende la objetividad de las orientaciones de valor (tanto frente al no-cognitivismo como frente al relativismo. Oportunamente, señalé que en esta defensa hay un fuerte punto de contacto con la propuesta davidsoniana en torno a la objetividad de los valores. En efecto, en “The objectivity of values” Davidson sostiene que lo mismo que rige para las creencias, rige para los valores: la comprensión mutua se logra sobre la base de una convergencia de nuestros “valores básicos” – a los que también califica, sugerentemente, como “valores iluminados” (*enlightened values*)- (1995c/2004: 49). Si comprendemos las palabras de otro, “existe un suelo común, compartimos ya un modo de vida” (1995c/2004:51).

Siguiendo a Putnam (2008), mi propuesta es que, por un lado, atendamos a la necesidad de dirigirnos a la dimensión evaluativa para comprender en toda su hondura la dimensión deóntica de la normatividad propia de lo mental y, por otro, que veamos a qué “valor esclarecido”, entre los que abraza Davidson, habría que invocar para ese fin.

Según acabamos de ver, las condiciones de coherencia y consistencia racional parecen crear la posibilidad misma de la interpretación. Recordemos, entonces, la cita con que dimos comienzo a este apartado: no podemos atribuir inteligiblemente creencias y deseos, o describir movimientos como conducta “a menos que estemos comprometidos a encontrar, en el patrón de conducta, creencia y deseo un amplio grado de racionalidad y consistencia, y si fallamos en descubrir un patrón coherente en las actitudes y acciones de los demás perdemos la oportunidad de tratarlos como *personas* (1970/1980:222; mis *itálicas*). En este punto, mi invitación es a que consideremos la posibilidad de interpretar la noción de “persona” desde una connotación *evaluativa*: la normatividad de lo mental que se exhibe en la interpretación resulta especial en la medida en que depende del siguiente “valor ulterior”: el de tratarnos mutuamente como *personas*<sup>153</sup>. Leemos en Davidson:

---

<sup>153</sup> Parece razonable preguntar en qué sentido la noción de persona entrañaría un “valor”. A los efectos de elaborar mi sugerencia, creo que puede resultarnos de ayuda la reflexión que Korsgaard (1996) hace en torno a la filosofía kantiana. Ella distingue entre el valor de los fines y el valor de la agencia: a su juicio, Kant advierte que consideramos cosas como importantes (*value of ends*) porque son importantes para nosotros y concluye, por tanto, que debemos tomarnos a nosotros mismos como importantes (*value of agency*). En esta línea, recordemos la segunda formulación del imperativo categórico, en la que se pone en juego el reconocimiento de la humanidad en la

[N]o podemos empezar a descifrar lo que dice una persona si no comprendemos sus actitudes en torno a sus oraciones, tales como sostener, desear o querer que sean verdaderas. Empezando por estas actitudes, debemos construir una teoría de lo que la persona quiere decir, dando así, simultáneamente, contenido a sus actitudes y a sus palabras. En nuestra necesidad de hallarla comprensible, trataremos de construir una teoría que la haga consistente, *creyente de verdades y amante de lo bueno* (1970/1980: 222; mis itálicas).

En un espíritu consonante con el que propongo aquí, Ramberg nos ayuda a ver que Davidson está tratando de quebrar la distinción entre la mente como *espectadora* y como *participante responsable* en las prácticas sociales<sup>154</sup>. En efecto, al combinar una teoría del significado con una teoría de la acción, Davidson entrelaza la distinción *mente-cuerpo* con la distinción *persona-cosa*<sup>155</sup>. Mientras que la noción de *mente* se vincula con una noción de racionalidad teórica entendida como *posesión de creencias y deseos*, la noción de *persona* se vincula con una noción de racionalidad práctica que involucra la idea de ser *uno de nosotros*, esto es, ser un miembro de nuestra comunidad, en la que estamos ligados por nuestras responsabilidades recíprocas. En este sentido, Davidson afirma que “cada mente se revela como parte de una comunidad de yoes libres” (Davidson, 1998b/2001: 91). A su vez, cabe recordar que en “Mental Events” afirma que la libertad implica la anomalía de lo mental (1970/1980:207). En este sentido, se ha sostenido que el monismo anómalo “no es sólo un principio de la filosofía de la mente, sino un principio ético-político” (Martin 1999: 348).

Es, justamente, el propio Davidson (1999) quien en su respuesta a Martin sostiene que hay cuatro áreas en las cuales ha tocado cuestiones relevantes para la ética, a saber: la naturaleza

---

propia persona y en la de los otros como *fin en sí mismo*. No obstante, no voy a alentar ningún compromiso sustantivo con la noción kantiana de persona, por razones que se advertirán más adelante.

<sup>154</sup> Cf. Ramberg B. (2000). Allí sostiene que en lugar de leer a Davidson bajo el signo de Brentano, esto es, en términos de la irreducibilidad de la intencionalidad, deberíamos pasar a leerlo en la clave hegeliana de nuestro mutuo reconocimiento (*Anerkennung*) como personas.

<sup>155</sup> Recordemos que, según vimos en el capítulo 3, la empresa de más largo aliento que persigue Davidson es la de proporcionar una teoría *unificada* del pensamiento, el significado y la acción. Lo que me interesa enfatizar aquí es que el carácter *unificado* de su teoría del pensamiento/significado y de la acción parece apuntar en la dirección señalada por Ramberg, esto es: el entrelazamiento de la racionalidad teórica (tener una mente/ser interpretable) y la racionalidad práctica (ser persona).

del razonamiento práctico, la cuestión de la objetividad de los juicios de valor, la autonomía de lo intencional y *los aspectos normativos de la interpretación*.

En el presente capítulo, he propuesto una elucidación de dichos aspectos a partir su dimensión deóntica. La idea que he intentado sostener es que la normatividad de lo mental, aquel rasgo que a juicio de Davidson porta la clave de lo humano (cf. 1982/2001), debe comprenderse en términos del principio de racionalidad como *regla constitutiva* de la práctica especial por la que nos tratamos mutuamente como *personas*. A partir de esta reflexión, creo que para comprender en toda su profundidad la peculiar normatividad de lo mental que propone Davidson, tenemos que internarnos, a su vez, en la dimensión *evaluativa*. El peso de la clarificación que persigo se traslada, entonces, a la noción de *persona*.<sup>156</sup> No obstante, voy a postergar esta cuestión hasta una vez que hayamos considerado los otros dos problemas que me propuse tratar aquí: el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de la primera y la tercera persona de la atribución mental.

---

<sup>156</sup> Para una interesante indagación de la conexión entre ser objeto de la interpretación y ser una persona, puede verse Carol Rovane (1999). Allí, ella sostiene explícitamente la relación entre ser interpretable y ser persona, a partir de lo que ella llama un “análisis normativo”: las condiciones que una cosa debe satisfacer para ser interpretable son condiciones de la cualidad de persona. Entre estas condiciones, la que realmente importa para su argumento es, precisamente, la condición normativa que Davidson postula con su principio de caridad.

## Capítulo 5: El problema de la relación entre el pensamiento y el lenguaje

Ni el lenguaje ni el pensamiento pueden explicarse plenamente en términos del otro, y ninguno de ambos tiene prioridad conceptual. Los dos, por cierto, están relacionados en el sentido de que cada uno requiere del otro para su comprensión, pero la relación no es tan completa como para que alguno de ambos baste, incluso cuando se lo refuerza razonablemente, para explicar al otro. (Davidson, 1975/1984: 156).

### 5.1. Una “geografía” del problema

Este capítulo está centrado en uno de los temas más transitados en la historia de la filosofía, en particular, en la metafísica, la filosofía del lenguaje y la filosofía de la mente, a saber: la naturaleza de las relaciones entre el pensamiento y el lenguaje. Para abordarlo, voy a valerme de la “geografía” de posiciones que ofrecen Carruthers, P. y Boucher, J. (1998). Lo que motiva mi adopción de este marco es que resulta interesante para mostrar, una vez más, la originalidad de la posición de Davidson. En efecto, a pesar de que proporcionan un trazado cuestionable de los bandos en pugna y una interpretación, a mi juicio, si no incorrecta, al menos parcial de la posición de Davidson, dividen las aguas de un modo que permite ver lo elusiva que es su mirada a la hora de querer “encorsetarla” en términos de dicotomías tradicionales.

¿Cuál es el lugar del lenguaje natural en el pensamiento y la cognición humana? Según Carruthers y Boucher (1998) las respuestas a esta pregunta pueden dividirse en dos grandes bandos – frente a los cuales, ellos pretenden abogar por una variedad de posiciones intermedias-. Por un lado, están aquellos que suscriben la idea de que el lenguaje tiene la función y el propósito exclusivos de *comunicar* el pensamiento (que resulta *independiente* de los medios de su transmisión). Esta concepción es la dominante en muchas áreas de las “ciencias cognitivas” (entendidas de un modo amplio, esto es, incluyendo la psicología cognitiva, la filosofía de la mente “empíricamente responsable”, la lingüística, la inteligencia artificial y la neurociencia cognitiva). En esta línea, Carruthers y Boucher, ubican a Locke (1690), Grice (1957,1969), Lewis (1969), Fodor (1978) y Searle (1983). De acuerdo con estos filósofos, la función central

del lenguaje es la de facilitar la comunicación, más que la de estar involucrado en el pensamiento –excepto indirectamente, por supuesto, posibilitando la adquisición de nuevas creencias-. El lenguaje funciona, así, completamente en el dominio público, interpersonal, más que en el dominio de la cognición individual. Desde luego, el lenguaje ha de ser representado y procesado dentro de la cognición de cada individuo, pero tal procesamiento existirá sólo para apoyar las funciones públicas del lenguaje, más que para tener un rol ejecutivo en el pensamiento y el razonamiento práctico del sujeto individual.

Los atractivos de esta concepción, dicen, han sido múltiples. En primer lugar, se puede argumentar que el lenguaje es empleado, simplemente, para propósitos de comunicación; y que es igualmente obvio que algunas formas de pensamiento, en particular, las que están involucradas en la imaginación visual, son independientes del lenguaje. Consideraciones de simplicidad sugieren, entonces, que *todo* pensamiento puede ser independiente del lenguaje, y que por tanto el lenguaje puede estar sólo para propósitos de comunicación. Por supuesto, Carruthers y Boucher advierten la debilidad de este argumento, dado que se requeriría una mínima evidencia en contrario para hacer zozobrar la apelación a una simplicidad de este tipo.

Un segundo argumento es de tipo semántico y se da de dos formas diferentes. En su versión histórica, empirista tradicional (Locke, 1690; Russell 1921) se formula del modo siguiente. Necesitamos una explicación de cómo es que algo- sea el lenguaje, o el pensamiento- puede *representar o significar* algo; en el caso de la imaginación visual se supone que es *transparente* cómo tienen lugar las representaciones -vía parecido-; pero no es en absoluto obvio cómo es que el lenguaje en sí mismo podría representar algo. Así, se obtiene una explicación satisfactoria si decimos que los poderes representacionales del lenguaje se derivan de los poderes representacionales del pensamiento (imágenes mentales), y si decimos que la función del lenguaje es, simplemente, comunicar pensamientos de una mente a otra. Este argumento clásico se tornó, no obstante, muy endeble al advertir que la teoría imaginista del pensamiento resulta inadecuada para abarcar todas las formas de pensamiento –v.gr. como explicación de cualquiera de nuestros pensamientos proposicionales-. Con todo, descendientes de este argumento sobreviven en la obra de Grice (1957), Lewis (1969) y Searle (1983). La idea es que podemos explicar las propiedades semánticas del lenguaje en términos de las propiedades semánticas de los pensamientos. En el caso particular de Searle (1983) la tesis es que *tenemos* que buscar una

explicación tal, dado que el contenido mental es simplemente más básico y primitivo que el significado lingüístico<sup>157</sup>.

De acuerdo con esta primera concepción de la relación entre el lenguaje y el pensamiento que presentan Carruthers y Boucher, el lenguaje es esencialmente sólo un dispositivo *input/output* para la cognición central. Se sostiene que el sistema del lenguaje natural es un mero conducto a través del cual se transmiten los pensamientos, desde y hacia los procesos cognitivos centrales de creer, pensar y razonar, *sin estar él mismo efectivamente implicado* en estas actividades.<sup>158</sup> Por tanto, desde esta perspectiva, si careciéramos de lenguaje natural, sin duda

---

<sup>157</sup> A juicio de Carruthers y Boucher, esta concepción del lenguaje como mero medio de transmisión del pensamiento puede rastrearse, asimismo, en la *psicología*. En particular, está prefigurada decimonómicamente en la “psicología de las facultades” de Gall. Hacia fines del siglo XIX, la afirmación de que habilidades particulares y, preeminentemente el lenguaje, están ubicadas en áreas particulares del cerebro donde funcionan de un modo relativamente independiente de otras habilidades, recibió apoyo a partir de la obra de neurofisiólogos y neurocirujanos como Broca, quien demostró que la habilidad para hablar depende de un área particular del cerebro, actualmente conocida como “área de Broca”. Ellos ven la noción de habilidades discretas, especiales, incluyendo la de una inteligencia lingüística, representada conspicuamente por la teoría contemporánea de las inteligencias múltiples (Gardner, 1983). Implícito en la noción de habilidades especiales está el supuesto de que el lenguaje se desarrolla, al menos parcialmente, con independencia de otros tipos de habilidad cognitiva, y no está involucrado en todos los tipos o clases de pensamiento consciente. No obstante, entre los psicólogos, dicen Carruthers y Boucher, no hay disputa acerca de que hay una “mediación del lenguaje natural” en el aprendizaje, la memoria, la planificación, la resolución de problemas, aunque persiste el disenso acerca de los detalles de su empleo (v.gr. Atkinson *et.al.* 1990). Aquellos psicólogos que argumentan que el lenguaje es una habilidad discreta que predominantemente subsirve la comunicación, aceptan pues que el lenguaje puede ser usado reflexivamente para facilitar el pensamiento (incluso Piaget (1932), quien consideró que el lenguaje es usado principalmente para la socialización, asume que el lenguaje también sirva para asistir en la formación y re-organización de estructuras de pensamiento). Aquí es importante distinguir entre dos posiciones. Por un lado, encontramos la tesis según la cual el lenguaje es el *medio* para algunos pensamientos y es parcialmente constitutivo de ellos (lo que parece una forma débil de la concepción que veremos a continuación, esto es, de la *constitución*). Por otro, encontramos la tesis según la cual el lenguaje *augmenta o facilita* algunas formas de pensamiento.

Por su parte, Carruthers y Boucher afirmar que en el ámbito de la *lingüística*, el argumento más radical y mejor conocido en favor de esta concepción proviene del célebre Noam Chomsky, quien en su temprano trabajo acerca de la gramática transformacional (1957) argumentó que en el proceso de la expresión lingüística el pensamiento tenía que ser traducido a lenguaje y, a su vez, traducido del lenguaje al “lenguaje del pensamiento” en el proceso de comprensión lingüística. Aunque, subsiguientemente, Chomsky abandonó la noción de gramática transformacional, su concepción del lenguaje como separado del pensamiento ha persistido y ha sido altamente influyente (Fodor, 1975, 1983; Levelt, 1989; Pinker, 1994). Asimismo, y en relación con su argumento de que el lenguaje es un sistema independiente, Chomsky ha argumentado –desde fines de los ‘50 y durante los ‘60- que la mejor explicación de los universales lingüísticos (rasgos gramaticales abstractos y comunes a todos los lenguajes naturales) es que los seres humanos poseemos una facultad del lenguaje innatamente estructurada y que el argumento de la pobreza del estímulo y los argumentos de la aprendibilidad del lenguaje implican que la adquisición de un lenguaje natural ha de concebirse en términos de madurativos, en lugar de en términos de aprendizaje (Chomsky, 1988). Las concepciones de Chomsky acerca del pensamiento y el lenguaje son por tanto consistentes con y, en efecto, han contribuido sustancialmente con, una pintura innatista y modularista de la mente.

<sup>158</sup> Por supuesto, todos acuerdan en que el lenguaje es importante para la cognición, dado que una vasta porción de las creencias normales y conceptos que poseemos de adultos sólo puede ser adquirida a partir de los otros, mediante la comunicación lingüística. Esto no está en cuestión.

creeríamos mucho menos de lo que actualmente creemos, pero nuestros procesos de pensamiento y nuestras capacidades intelectuales básicas permanecerían esencialmente sin cambios.

Ahora bien, en oposición a la concepción que acabamos de considerar, Carruthers y Boucher sostienen que históricamente ha habido una adhesión importante a la idea de que el lenguaje está *implicado crucialmente* en el pensamiento. Según esta segunda concepción, *pensamos en lenguaje natural*, de modo tal que las oraciones del lenguaje natural son el vehículo de nuestros pensamientos. Para discutir esta posición, Carruthers y Boucher nos invitan a deslindarla en las dos tesis siguientes. Por un lado, encontramos la tesis del *requisito*, que sostiene que el lenguaje es *requerido por* o es una *condición necesaria* del pensamiento, o ciertos tipos de pensamiento. Por otro lado, encontramos la tesis de la *constitución*, que afirma que el lenguaje está *constitutivamente involucrado* en los pensamientos o es el *medio* de los pensamientos. A juicio de Carruthers y Boucher, todo el mundo podría conceder que alguna forma de la tesis del requisito tenga, al menos, un grado limitado de aplicabilidad, dado que es obvio que los chicos adquieren muchas de sus creencias y conceptos sobre la base del lenguaje. En efecto, parece claro que una persona sin lenguaje no podría tener pensamientos acerca de, por ejemplo, electrones, neutrinos o división celular, dado que es sólo a través del lenguaje que podemos aprender estas cosas. Otra cuestión diferente, y más interesante, es la de si el lenguaje sigue implicado en las creencias así adquiridas -de manera tal que, por ejemplo, un pensamiento acerca de electrones sólo pueda tenerse por medio de la activación de una representación de la palabra “electrón” o algún equivalente-. De modo que para discutir esta segunda concepción acerca de las relaciones entre el lenguaje y el pensamiento, van a concentrarse en la tesis de la *constitución*. Lo que interesa a nuestros fines es que, a juicio de Carruthers y Boucher, la concepción según la cual el lenguaje es constitutivo del pensamiento ha sido abrazada por filósofos tan dispares como Leibniz (1704), Wittgenstein (1921, 1953), Dummett (1981,1991), Dennett (1991), McDowell (1994), Carruthers (1996) y el propio Davidson (1975, 1982).

Mientras que sus motivaciones han sido variopintas, ha habido, dicen, una línea argumentativa común en su favor, a saber: los seres humanos somos los únicos en el rango y la sofisticación de los pensamientos de los que somos capaces; los seres humanos somos también los únicos que poseemos lenguaje natural; la explicación más simple de la co-ocurrencia de estas dos características únicas es que el lenguaje natural es el que hace posible el pensamiento distintivamente humano. Este argumento parece tener, evidentemente, alguna fuerza. Adviértase,

sin embargo, que no apoya una versión de la tesis de la *constitución* según la cual *todo* pensamiento involucra lenguaje. Se trata de un argumento a favor de la tesis más débil según la cual aquellos de nuestros pensamientos que son distintivamente humanos *requieren* lenguaje.

A su juicio, muchos de los filósofos mencionados más arriba, específicamente, Wittgenstein, Dummett, McDowell y Davidson han defendido una tesis de la *constitución* del lenguaje en términos *universales y conceptuales*. Esto es, han sostenido que es conceptualmente necesario que *todo* pensamiento involucre lenguaje (la idea misma de pensamiento implica el lenguaje). De este modo, los animales y los niños pre-lingüísticos carecen de pensamientos (esto es, se supone que es lógicamente imposible para criaturas sin lenguaje tener actitudes proposicionales).

Los argumentos en esta línea, dicen, son básicamente epistemológicos. Como ejemplo paradigmático de ello ofrecen la siguiente reconstrucción de Davidson (1982). Las actitudes proposicionales tienen contenidos finamente discriminados, por ejemplo, el pensamiento de que Venus se ha puesto es distinto del pensamiento de que la estrella matutina se ha puesto, aunque Venus y la estrella matutina sean una y la misma cosa. Pero nada, excepto la actividad lingüística –por ejemplo, *afirmar* “Venus se ha puesto” mientras se *niega* “La estrella matutina se ha puesto” podría manifestar o establecer que una criatura tiene el primero de estos pensamientos como *opuesto* al segundo. Por tanto, sólo una criatura que use lenguaje puede *tener* un pensamiento tal.

En opinión de Carruthers y Boucher (desacertada, a mi juicio) este argumento depende, sin embargo, de un cuestionable verificacionismo o antirrealismo acerca de los estados mentales.<sup>159</sup> Porque sin una concepción semejante como premisa, simplemente no se sigue a partir del hecho de que no podamos *saber* qué piensa, precisamente, una criatura, que *no* esté pensando algo. Si, por el contrario, los pensamientos tienen una realidad que es distinta e independiente de nuestros modos de evidenciarlos, entonces es perfectamente posible que criaturas que no sean usuarias de un lenguaje tengan pensamientos completamente determinados, pero pensamientos que *nosotros* no podemos conocer. Para Carruthers y Boucher la evidencia de la etología sugiere fuertemente que los animales “son capaces de una variedad de clases de pensamiento y razonamiento, aún cuando no podamos saber precisamente *qué* pensamientos

---

<sup>159</sup> En el capítulo anterior, al considerar analogía entre las actitudes numéricas y las actitudes proposicionales, me he ocupado de mostrar en qué sentido Davidson pretende eludir el cargo de “antirrealista” acerca de los estados mentales.

finamente discriminados tienen en una ocasión dada” (Walker, 1983; Byrne y Whiten, 1988; Allen y Bekoff, 1997). Lo mismo se da, a su juicio, en la evidencia que ofrece el desarrollo infantil (Carey, 1985; Flavell, 1985; Sperber *et al.* 1995). De modo que, a su juicio, si pretende sostenerse alguna versión de la tesis de constitución, necesita ser más débil que la tesis *universal y conceptual*.<sup>160</sup>

Resumamos, entonces, la “geografía” de cuestiones que nos presentan Carruthers y Boucher acerca de la naturaleza de las relaciones entre el pensamiento y el lenguaje. En primer lugar, encontramos la concepción que defiende que el lenguaje natural tiene como función exclusiva *comunicar* el pensamiento, de modo que *no está constitutivamente involucrado* en él. Mientras que la versión fuerte sostiene que el lenguaje es un sistema cognitivo que simplemente produce *outputs* para el pensamiento, la versión débil defiende la idea de que el lenguaje natural *facilita o aumenta* algunas formas de pensamiento. En oposición a esta concepción, nos encontramos con aquella que sostiene que el lenguaje natural está *constitutivamente involucrado* en el pensamiento. Mientras que la versión fuerte de esta tesis sostiene que todo pensamiento involucra lenguaje natural por necesidad conceptual (paradigmáticamente, Davidson 1975, 1982), la versión débil defiende que el lenguaje natural está implicado sólo en cierta clase de pensamientos - los proposicionales conscientes - en virtud de una necesidad natural, i.e., en

---

<sup>160</sup> Dentro de la *psicología*, los proponentes más fuertes de una concepción *constitutiva* del lenguaje fueron un grupo de psicólogos rusos, quienes, como Sokolov (1972) sostuvieron que el “*inner speech*” está involucrado en muchos, si no todos, los aspectos del pensamiento. Otro psicólogo ruso, Luria (1959) describió al “*inner speech*” como “el sistema de segunda señalización”, y demostró su “rol en el control consciente de la acción motora, además de su rol en formas más intelectuales de pensamiento”. Aunque la metodología de estos estudios tempranos puede ser cuestionada, Carruthers y Boucher consideran que sugieren la tesis de que algunas clases de pensamiento involucran lenguaje natural. No obstante, aún cuando los psicólogos estén fuertemente comprometidos con la tesis de constitución, no adhieren a la idea de que sea imposible el pensamiento sin lenguaje. Vigotsky (1934), por ejemplo, argumentó a favor del involucramiento del lenguaje en el pensamiento, pero sostuvo asimismo que el lenguaje y el pensamiento tienen rutas diferentes e independientes y modos diferentes de operación en los chicos pre-lingüísticos. Desde la *lingüística*, una de las formas más radicales y mejor conocidas de la tesis de constitución proviene de Benjamin Lee Whorf (1956). De acuerdo con la hipótesis de la relatividad lingüística de Whorf, los conceptos humanos y, en efecto mucho de la estructura de la mente humana misma, son adquiridos por los chicos a partir de los adultos cuando les enseñan su lenguaje nativo; estos conceptos y estructuras difieren ampliamente, dependiendo de las estructuras y recursos conceptuales de los lenguajes naturales en cuestión. Esta concepción relativista-social de la estructuración de la mente todavía es la hegemónica en las ciencias sociales. Pinker (1994) se refiere a ella como “modelo estándar en ciencia social”. La línea del “modelo estándar en ciencia social” sostiene, a grandes trazos, lo siguiente. Encontramos variaciones amplias en las conductas y estructuras sociales de distintas culturas, y variaciones igualmente amplias entre los lenguajes; estas variaciones no son genéticas –los chicos que nacen en una cultura, pero se crían en otra, desarrollan sin dificultades las conductas y prácticas de un nativo-; así, la explicación natural es que la mente es inicialmente una suerte de *tabula rasa*, tal como supuso el empirismo tradicional; y toda la estructura impuesta a nuestra mente procede del lenguaje particular que la persona adquiere, y de la cultura particular en la que se halla inmersa.

virtud de las leyes de la naturaleza (tal es la posición que defiende, expresamente, Carruthers 1996)<sup>161</sup>.

Ahora bien, tal como veremos en este capítulo la tesis defendida por Davidson (especialmente, en 1975 y 1982) es la de una *interdependencia*: el *pensamiento* supone *lenguaje* y el *lenguaje* supone *pensamiento*. De este modo, y en la medida en que la filiación en alguno de los dos bandos trazados por Carruthers y Boucher (1998) supone priorizar uno de los elementos de la relación, en mi opinión la tesis de Davidson acerca de la naturaleza de la relación entre ambos no es legítimamente subsumible en ninguno.

---

<sup>161</sup> Como mera curiosidad, cabe señalar que el objetivo que persiguen Carruthers y Boucher (1998) es el de difuminar la dicotomía entre estas dos concepciones acerca de la naturaleza de las relaciones entre el pensamiento y el lenguaje; para ello intentan debilitar, alternativamente, a cada una de ellas. Por un lado, señalan que uno de los factores que condujo al abandono prematuro de la concepción de la *constitución* por parte de los científicos cognitivos fue que tenían en mente la forma más extrema de tal concepción, esto es, del tipo defendido por Wittgenstein, McDowell y Davidson. Carruthers y Boucher sostienen que estas versiones universales y conceptuales son altamente implausibles y que sobre la base de ellas se rechazó la concepción de la constitución *tout court*. No obstante, se la puede deflacionar en las dos direcciones siguientes: (i) pueden abandonarse las tesis *conceptuales o lógicas* acerca del carácter constitutivo del lenguaje respecto del pensamiento. Puede sostenerse, en cambio, que el lenguaje está implicado en -algunos o todos- los pensamientos como *cuestión de hecho* o como cuestión de necesidad *natural* resultante quizás de la estructura de la cognición humana: Bickerton (1990, 1995) Dennett (1991) Carruthers (1996) y (ii) puede debilitarse el compromiso con la *universalidad*. Así se puede permitir que algunos pensamientos (especialmente, los viso-espaciales y los disponibles para animales y niños pequeños) sean independientes del lenguaje y afirmar, no obstante, que otros pensamientos implican crucialmente el lenguaje. Esto requiere, entonces, que tracemos una distinción entre pensamientos que requieren lenguaje y pensamientos que no lo requieren. Hay, al menos, dos modos en que esto puede hacerse. Se puede trazar una división *horizontal*, en términos de tipos de contenido (Gómez, Carruthers y Segal, quienes exploran la idea de que los pensamientos acerca de estados mentales, en particular, involucrarían crucialmente lenguaje natural), se puede trazar una división *vertical*, en términos de los modos en que los tenemos. (Carruthers, Davies, Dennett) quienes discuten la idea de que el lenguaje natural puede estar implicado sólo en algunos *niveles* de pensamiento, p ej. los pensamientos conscientes como opuestos a no-conscientes o formados activamente como opuestos a pasivos (Franklish) o quizás explícitos como opuestos a implícitos (Perner). Por su parte, para debilitar la concepción según la cual la función central del lenguaje no es la de constituir sino *comunicar* el pensamiento consideran que es posible suavizarla sin alterar sus puntos fundamentales. Clark, por ejemplo, argumenta a favor de una concepción del lenguaje como una *herramienta cognitiva*, esto es, el lenguaje no sólo posibilita la comunicación, sino que aumenta los poderes cognitivos humanos (aunque el involucramiento del lenguaje sólo surge cuando nos comprometemos en un proceso extendido de pensamiento o razonamiento a través del tiempo).

En efecto, en “Seeing Through Language” introduce algunas precisiones que me parecen, especialmente, atendibles para ver la inadecuación de ubicarlo en la primera de las concepciones que presentamos, esto es, aquella según la cual el pensamiento está constituido con independencia del lenguaje, que resulta un mero medio para su transmisión o comunicación. Allí sostiene que vemos el mundo a través del lenguaje; pero es cauto a la hora de interpretar esta metáfora, puesto que puede ser sumamente inconducente. El lenguaje no media entre nosotros y el mundo: no vemos el mundo a través del lenguaje, más que como lo vemos a través de nuestros ojos. En este sentido, no vemos *a través de* sino *con* nuestros ojos. Explora, entonces, el paralelo entre nuestros sentidos y el lenguaje. Hay una analogía válida entre tener ojos y oídos y tener un lenguaje: “los tres son órganos con los que entramos en contacto directo con nuestro entorno” (Davidson 1997/2005:131).<sup>162</sup> Parece haber poca razón para dudar de que estamos genéticamente programados de modos bastante específicos para hablar como lo hacemos. Aparentemente todos los lenguajes comparten reglas estructurales a pesar de su variedad de superficie. La evidencia para esto, dice, reside parcialmente en el descubrimiento de restricciones universales sobre las gramáticas (en este sentido, acuerda explícitamente con Chomsky 1980).

Ahora bien, según Fodor y Pinker este lenguaje interno emerge como parte de nuestra herencia genética y es *previo a todo lenguaje hablado*. Pinker no tiene dudas sobre la prioridad del “lenguaje del pensamiento” o “mentalés”. Para Davidson, en cambio, esta concepción (que como hemos visto es aceptada por muchos lingüistas y científicos cognitivos) reposa en malos argumentos y conclusiones confusas. Él cree que vale la pena enfatizar que la aptitud para el lenguaje es parte de nuestro equipamiento natural; en este sentido, le gusta la analogía con los órganos sensoriales y, por tanto, la implicación de que el lenguaje no es algo que venga entre nosotros y la realidad: no puede venir en el medio, dado que es *parte de nosotros*. Pero la postulación de un lenguaje del pensamiento arruina el atractivo de esta postura: si el lenguaje del pensamiento es lo que es parte de nosotros, entonces nuestro lenguaje hablado es un intermediario entre el pensamiento y aquello acerca de lo cual es el pensamiento. En este espíritu

---

<sup>162</sup> No obstante, el habla no es un órgano más: es esencial a los otros sentidos si van a producir percepción proposicional. Percibir cómo son las cosas requiere pensamiento con contenido proposicional (y, según veremos, esta habilidad se desarrolla junto con el lenguaje). Asimismo, cabe aclarar que una vez que tenemos pensamiento proposicional, la percepción es directa y no mediada, en el sentido en que no hay intermediarios epistémicos sobre los que estén basadas las creencias perceptuales (cf. Davidson 1983/2001).

rechaza por débiles los argumentos a favor de la existencia de un lenguaje del pensamiento previo a, o independiente de, un lenguaje socialmente engendrado (1997/2005:133).

Por su parte, en mi opinión, sus afirmaciones en “Rational Animals” también toman inapropiada su ubicación estricta en el bando de los defensores de la tesis de la constitución, en el sentido en que Carruthers y Boucher (1998) la encuentran interesante (es decir, no en tanto se sostenga que el lenguaje es *requerido por* o es una *condición necesaria* del pensamiento, sino la de si la existencia de cada pensamiento depende de la existencia de una oración que lo exprese). En efecto, allí afirma que hay que distinguir la tesis que propugna de varias versiones relacionadas con ellas. Por ejemplo, no cree que se pueda reducir el pensar a la actividad lingüística. Tal como vimos en el capítulo anterior, dice que tampoco encuentra plausibilidad alguna en la idea de que los pensamientos puedan identificarse nomológicamente o ser correlacionados con fenómenos caracterizados en términos físicos o neurofisiológicos. Tampoco ve “ninguna razón para mantener que no podemos pensar sobre lo que no podemos hablar” (1982/2001: 100). Y afirma:

mi tesis no es entonces que la existencia de cada pensamiento dependa de la existencia de una oración que exprese ese pensamiento. Antes bien, mi tesis es que una criatura no puede tener un pensamiento a menos que tenga un lenguaje. Para ser una criatura racional pensante, esa criatura debe ser capaz de expresar muchos pensamientos y, sobre todo, ser capaz de interpretar el habla y los pensamientos de los demás (1982/2001: 100).

Finalmente, me interesa señalar que la *interdependencia* entre lenguaje y pensamiento que él defiende, y que encuentra su expresión paradigmática en la teoría de la interpretación radical como “gozne” entre su filosofía de la mente y su filosofía del lenguaje, diluye la asimetría necesaria para que la dicotomía entre las concepciones expuestas por Carruthers y Boucher (1998) resulte formulable. Pensamiento y lenguaje, están, pues, en pie de igualdad. En el capítulo anterior, nos hemos ocupado de la concepción davidsoniana de lo mental. Antes de avanzar sobre su relación con el lenguaje se impone, entonces, que tracemos algunas precisiones sobre la concepción davidsoniana de este último.

## 5.2. La realidad del lenguaje

Concluyo que no hay tal cosa como un lenguaje, si el lenguaje es algo como lo que han supuesto muchos filósofos y lingüistas.  
(Davidson 1986/2005:107)

En “A Nice Derangement of Epitaphs” Davidson (1986) declara que no hay “tal cosa como un lenguaje”, al menos no en el sentido en que muchos filósofos y lingüistas han querido sostener. Y, provocadoramente, afirma que él mismo había abrazado una perspectiva del tipo que busca desacreditar (como veremos, esta admisión de su parte no hizo más que echar “más leña al fuego” de las múltiples críticas que se le formularon)<sup>163</sup>.

A los efectos de ponderar el proyecto de Davidson en la teoría del significado resulta clave, pues, si su posición en “A Nice Derangement of Epitaphs” representa un alejamiento radical de su programa previo y, en particular, si constituye una amenaza capaz de minar su obra anterior. Lepore y Ludwig (2005, cap. 17) argumentan que una vez que se comprende exactamente cuál es la tesis que Davidson está atacando, es difícil ver cómo socavaría algo que haya mantenido previamente, y que en cierto sentido, es sólo una articulación ulterior de las

---

<sup>163</sup> Varios críticos han considerado que si el argumento de Davidson funciona, su propio programa en filosofía del lenguaje se vería seriamente socavado. Ian Hacking, por ejemplo, ha observado, que “verdadero-en-*L*” está en el corazón de la filosofía de Davidson, “pero ¿qué queda si no hay tal cosa como un *L*?” (1986: 447). Hacking reconoce que negar que haya lenguajes en el sentido de la tesis de Davidson, no muestra que no haya valores para “*L*”. Si los divorciamos del sentido problemático, podemos retener la idea de que hay sistemas formales que nos permiten generar recursivamente asignaciones de condiciones de verdad a las emisiones del hablante de oraciones de *L*. La objeción de Hacking se dirige, entonces, a la utilidad de contar incluso con tales *lenguajes* en la interpretación. El corazón de su crítica es que las observaciones de Davidson acerca de cómo divergen las teorías al paso de las normas públicas muestran que no construimos nada tal como una “teoría total” – estilo Tarski para un lenguaje *L* formalmente especificado-para la interpretación de un hablante en la conversación. El punto relevante es si esto contradice algún punto de la posición davidsoniana. Lepore y Ludwig creen que no. A su vez, Dorit Bar-On y Mark Risjord sostienen que a menos que la tesis radical de Davidson sea un alejamiento respecto de sus propios desarrollos, el programa davidsoniano parece desmoronarse. Para Lepore y Ludwig, Bar-On (1992) hace una lectura errada: en realidad las tesis de “The social aspect of language”(1994) y “A nice derangement...” (1986) no generan una ruptura respecto de los primeros escritos acerca de la interpretación radical, sino que son un desarrollo natural de ellos. Michael Dummett, por su parte, ha observado que la ocurrencia del fenómeno que le interesa a Davidson es incontrovertible: ¿cómo podríamos llegar a la conclusión de que no hay tal cosa como un lenguaje? (1986:465). No obstante, a juicio de Lepore y Ludwig (2005) tal vez haya menos desacuerdo entre ambos del que aparece a primera vista. Finalmente, entre quienes salieron en defensa de Davidson, cabe citar a Ramberg (1989) y Pietroski (1994).

consecuencias de su perspectiva metodológica básica. Adicionalmente, sostienen que no hay razón por la que Davidson tendría que haber suspendido su uso de la noción de *lenguaje* en un sentido perfectamente respetable (esto es, distinto de los que ataca) que sea adecuado a sus objetivos teóricos.

En la presente sección, entonces, vamos a examinar cuál es la *concepción del lenguaje* que Davidson quiere atacar, a considerar su *argumento* en contra de dicha posición y a tratar de clarificar del mejor modo posible su perspectiva. El interés adicional que para nosotros reviste esta cuestión, es que en el curso de su esclarecimiento, precisaremos también parte de lo que sostuvimos en el capítulo 3 de este trabajo a propósito de concepción davidsoniana de la *normatividad*: la impropiedad de pensarla bajo la primera de sus acepciones *deónticas*, a saber: en términos de *convenciones*.

La cuestión crucial para comprender el argumento de Davidson es qué quiere decir por “lenguaje”. Él se ocupa en primera instancia de dos preguntas relacionadas, a saber:

- a) ¿Cuál es el rol de la *convención* en la comunicación?
- b) ¿Cuál es el rol que juega en la comunicación el *conocimiento* que el intérprete trae a la interacción comunicativa?

De manera estándar se considera que el conocimiento previo de convenciones acerca del uso de las palabras en el lenguaje del hablante juega un rol importante en la comunicación. En un sentido, Davidson no niega esto. En efecto, él no niega que conocemos hechos acerca de convenciones según las cuales se usan las palabras, ni niega que este conocimiento, poseído previamente a la interpretación de otros en nuestra comunidad lingüística, pueda jugar un papel útil e importante al interpretarlos. Lo que niega, en cambio, es que tal conocimiento sea *esencial* al éxito interpretativo, es decir, niega que el conocimiento previo de convenciones acerca del uso de las palabras en el lenguaje del hablante sea necesario o suficiente para la interpretación. (Davidson 1994/2005). Su conclusión negativa se sigue de considerar los usos de las palabras que no coinciden con su significado usual (o de diccionario) en la comunidad lingüística del hablante, pero que, sin embargo, no impiden la comunicación. De hecho, el título de “A Nice Derangement of Epitaphs” está tomado de la obra *The Rivals* de Sheridan, dicho por cierta Sra. Malaprop, de cuyo apellido se deriva la expresión “malapropismo” (barbarismo). Como es sabido, un “malapropismo” es el uso incorrecto de una palabra, cuando dicho uso resulta de

confundir una palabra con otra similar en sonido a la que expresa el significado que el hablante tiene la intención de comunicar. Por esta razón, en general, es fácil determinar lo que el hablante tiene la intención de comunicar cuando incurre en un malapropismo, como con el intento de la Sra. Malaprop de comunicar que algo era una linda “sucesión de epítetos”. Este fenómeno lo lleva a Davidson a distinguir entre el significado literal, o lo que él llama “primer significado” y el significado convencional (de diccionario). Esta distinción entre primer significado y significado convencional puede considerarse como la clave de lo que subyace al rechazo de Davidson de una concepción del rol de la *convención* en la comunicación, y a una concepción del lenguaje, según la cual, el lenguaje consiste en portadores de significado convencional, cuyo conocimiento es necesario y suficiente para la comunicación.

El *primer significado* es aquel con que un hablante tiene la *intención* de que sean comprendidas sus palabras, como para formar la base de efectos subsiguientes que se logran al usar las palabras que emplea. Es el significado que él considera como “primero en el orden de la interpretación” (1986/2005:91). Diversas formas de significado no literal reposan siempre en el significado literal de una emisión. Así, entender una metáfora involucra entender primero lo que las palabras significan literalmente.

Usualmente damos por sentado que lo que las palabras significan literalmente está determinado por *normas públicas* que gobiernan su uso en la comunidad lingüística del hablante. Pero la distinción entre significado literal y significado del hablante sobrevive a usos no-estándar de las palabras y así la distinción misma debe ser relativizada a un hablante. Cuando un hablante usa sus palabras de acuerdo con normas públicas, sus significados *literales*, esto es, sus primeros significados, corresponderán con el significado del *diccionario*. Pero pueden darse por separado. Coinciden cuando el hablante competente tiene la intención de que se llegue a todo significado ulterior de sus palabras interpretando primero sus palabras de acuerdo con normas públicas. Divergen cuando tiene la intención de que las palabras que emite sean interpretadas en primera instancia de un modo que no corresponde al significado del diccionario, sea esto intencional o inadvertidamente (como en la metáfora o el barbarismo, respectivamente). La razón por la que es importante distinguir el *primer significado* del *significado convencional* es que el primero es esencial para toda comunicación lingüística (si no los distinguimos, supondremos que el significado convencional también es esencial para toda comunicación lingüística). Davidson no

traza explícitamente esta conexión, pero creemos que al considerarla se clarifica la idea que subyace a su argumento.

Ahora estamos en una mejor posición para establecer de modo más preciso el “blanco” del argumento de Davidson y para ver qué podría motivar la concepción a la que quiere oponerse. Dicha concepción es articulada a través de los siguientes tres “plausibles principios concernientes al primer significado en el lenguaje”:

(1) El primer significado es *sistemático*. Un hablante competente o un intérprete es capaz de interpretar emisiones, propias o de los demás, sobre la base de las propiedades semánticas de las partes, o palabras, en la emisión y la estructura de la emisión. Para que esto sea posible, debe haber relaciones sistemáticas entre los significados de las emisiones.

(2) Los primeros significados son compartidos. Para que hablante e intérprete se comuniquen exitosa y regularmente, deben *compartir un método de interpretación* del tipo descrito en (1).

(3) Los primeros significados están gobernados *por convenciones aprendidas o regularidades*. El conocimiento sistemático o la competencia del hablante/ intérprete es aprendido/a previo a las ocasiones de interpretación y es convencional en carácter.

El blanco de Davidson es, en particular (3), en la medida en que lo interpretemos como diciendo que los primeros significados están esencialmente gobernados por convenciones aprendidas o regularidades - (3) es equivalente a identificar el primer significado con el significado convencional del tipo expresado por las definiciones de diccionario.

Ahora bien, ¿cómo está conectado el rechazo de (3) con la tesis de que no hay tal cosa como un lenguaje? Primero, se sostiene que para ser capaces de interpretar a un hablante debemos saber qué lenguaje está usando en la ocasión de interpretación. Segundo, se identifica el lenguaje que habla como determinado por los primeros significados, esto es, aquellos significados con los que él tiene la intención de que interpretemos sus palabras en ese momento. Tercero, identificamos los primeros significados con significados convencionales de las palabras en su comunidad lingüística.

Así, arribamos a una concepción del lenguaje como algo esencial para la interpretación, porque el conocimiento de los primeros significados del hablante es esencial para la

comunicación y tal conocimiento consiste en el *dominio de convenciones determinadas por prácticas de la comunidad*, dominio que *debe necesariamente ser adquirido antes de la interpretación*. Desde esta concepción, *un lenguaje es (a) un vocabulario y un conjunto de reglas determinadas por convenciones de una comunidad lingüística, (b) dominado por los miembros de la comunidad y cuyo dominio es (c) necesario y (d) suficiente para interpretar a sus hablantes*. Es evidente que argumentar en contra de la existencia de algo llamado “lenguaje” en este sentido no compromete con que no haya lenguajes en ningún sentido respetable.

Pasemos, entonces, a considerar el argumento. Éste necesita establecer las dos cuestiones siguientes:

- I. El conocimiento de los significados convencionales es *insuficiente* para interpretar al hablante.

Esta tesis puede entenderse en sentido débil y en sentido fuerte. (i) En sentido *débil*, es la afirmación de que el conocimiento de los significados convencionales no garantiza la correcta interpretación *en la ocasión en cuestión*; deben aducirse hechos adicionales para arribar a ella. (ii) En sentido *fuerte*, es la afirmación de que *siempre* debemos apelar a otro conocimiento además del conocimiento del significado convencional para interpretar correctamente al hablante. Lepore y Ludwig (2005) sugieren que el “blanco” de Davidson es (i), aunque consignan que Pietroski cree que es (ii) -tesis que, desde luego, implica (i)-. Lepore y Ludwig encuentran plausible que todo hablante al menos una vez haya empleado incorrectamente una palabra en relación con las normas de su comunidad, y no obstante, haya sido comprendido. Comprender a la gente a pesar de sus deslices lingüísticos es un ejercicio rutinario de caridad.

Lepore y Ludwig (2005) proponen, entonces, reformular la discusión en términos de la distinción que Davidson introduce entre *teorías previas* y *teorías al paso*<sup>164</sup>. Hacen explícito que

---

<sup>164</sup>Si aceptamos esta distinción y concedemos también la explicación de la competencia lingüística dada por los principios 1 y 2, surge el siguiente problema. De acuerdo con dicha explicación, cada intérprete (y esto incluye a los hablantes, dado que los hablantes deben ser intérpretes) llega a un intercambio lingüístico exitoso preparado con una “teoría” que constituye su competencia lingüística básica y que comparte con aquellos con quienes se comunica. Dado que cada parte tiene tal teoría compartida y sabe que otros comparten su teoría, y sabe que otros saben que él sabe (etc.), algunos podrían decir que el conocimiento o las habilidades que constituyen la teoría pueden ser llamadas *convenciones*. Davidson piensa que la distinción entre teoría previa/teoría al paso, si se toma seriamente, socava esta explicación comúnmente aceptada de la competencia lingüística y la comunicación. Lo que se debe compartir para que la comunicación sea exitosa es la teoría al paso, dado que la teoría al paso es la que el intérprete efectivamente usa para interpretar una emisión y es la teoría que el hablante tiene la intención de que use el

no debemos suponer que estas teorías son efectivamente sostenidas por el intérprete; se trata de teorías que el *teórico* usa para caracterizar las disposiciones del intérprete para comprender a un hablante particular. En un intercambio comunicativo, un intérprete siempre está listo para modificar cómo está dispuesto a entender a un hablante a la luz de información provista por el contexto y por lo que el hablante ya ha dicho. La *teoría previa* para el intérprete es aquella que caracteriza su disposición para interpretar al hablante previo a su intercambio comunicativo. La *teoría al paso* para el intérprete es aquella que caracteriza sus disposiciones para interpretar las emisiones del hablante en medio del intercambio comunicativo. Las teorías previas y al paso para los intérpretes están siempre relativizadas a hablantes y tiempos particulares. Asimismo, las teorías previa y al paso para el hablante están concebidas como relativizadas a hablantes y tiempos particulares. La teoría previa del hablante es “lo que el hablante *cre*e que es la teoría previa del intérprete, mientras que su teoría al paso es aquella que tiene la *intención* de que use el intérprete” (Davidson 1986/2005: 101). La teoría al paso es aquella que expresa, entonces, lo que las palabras en su boca quieren decir mientras habla con el intérprete.

El otro punto que necesita establecer el argumento es el siguiente:

II. El conocimiento de los significados convencionales es *innecesario* para interpretar al hablante.

Esta tesis también puede entenderse en sentido débil y en sentido fuerte. (i) En sentido débil, es la afirmación de que *en principio* puede descartarse o prescindirse del conocimiento de los significados convencionales. (ii) En sentido fuerte, es la afirmación de que como una *cuestión de hecho*, siendo como son los hablantes, no se requiere conocimiento de los significados convencionales para interpretarlos exitosamente. Esta tesis según la cual el significado

---

intérprete. Sólo si las teorías al paso coinciden, hay comprensión. Ni la teoría previa ni la teoría al paso describe lo que llamaríamos el lenguaje que una persona sabe y ninguna caracteriza una competencia lingüística del hablante o el intérprete. Lo que dos personas necesitan –si van a entenderse mediante el habla– es la habilidad de convergir en sus teorías al paso de emisión a emisión: lo que hablante e intérprete comparten no es aprendido, y así no es un lenguaje gobernado por *reglas o convenciones* conocidas por adelantado. Y, así, lo que conocen por adelantado no es (necesariamente) compartido.

convencional es innecesario para la comunicación es presentada claramente en “Communication and Convention”. Allí afirma:

El conocimiento de convenciones lingüísticas es ...un bastón práctico (practical crutch) para la interpretación, un bastón sin el cual no nos la podemos haber en la práctica, pero un bastón que bajo condiciones óptimas para la comunicación, podemos finalmente arrojar y del que, en teoría, podríamos prescindir desde el comienzo (Davidson 1984b/1984:279).

Abordemos, primero, la cuestión de si el conocimiento compartido de convenciones previas es *en principio* necesario para el éxito interpretativo. Al examinar este punto es importante tener presente que no se requiere que supongamos que el hablante y el intérprete emplean los mismos portadores de significado convencional. Todo lo que se requiere para que la tesis sea verdadera es que compartan los significados convencionales de las palabras que usan, sean o no las mismas palabras. Uno podría hablar en mandarín y otro en francés y comprenderse entre sí perfectamente. Esta cuestión (dejando de lado las limitaciones naturales del conocimiento y las habilidades cognitivas) es equivalente a preguntar si hay cuestiones independientes de las convenciones lingüísticas que determinen –o que podrían determinar- lo que un hablante significa con sus palabras. El punto es que si la perspectiva metodológica básica adoptada por Davidson es correcta, la respuesta a esta pregunta es, tal como sostienen Lepore y Ludwig (2005), claramente afirmativa, según vimos en la primera parte de este trabajo: lo que fija el significado son las disposiciones del hablante para usar las palabras, tal como está dispuesto a hacerlo en su entorno. Si se asume que el intérprete conoce las disposiciones del hablante para usar las palabras en su entorno, conoce todo lo que debe conocer a los efectos de interpretar correctamente las palabras del hablante. El conocimiento en cuestión es, entonces, insuficiente en la práctica e innecesario en teoría. No obstante, se podría insistir en que las *convenciones*, en un sentido, son necesarias para la comunicación. Esto, desde luego, depende de cómo entendamos qué es *participar de una convención*<sup>165</sup>.

---

<sup>165</sup> En “Communication and Convention” las conclusiones de Davidson derivan de una cierta concepción de la convención en la obra de David Lewis (básicamente, como regularidad). La noción que proponen Lepore y Ludwig (2005) difiere de la de Lewis. Para ellos, dado que el hablante y el intérprete necesitan convergir en una teoría al paso para la comunicación exitosa, necesitan convergir en un conjunto común de reglas compartidas, en convenciones compartidas (en un acuerdo explícito). En este sentido, las convenciones son necesarias para el éxito comunicativo. Esto, no obstante, no socava, a juicio de ellos, la conclusión de Davison. Por el contrario, la caracterización que proponen Lepore y Ludwig - que permite que exista en ausencia de regularidades antecedentes en la comunidad- hace más tratable la siguiente afirmación de Davidson: “los filósofos que hacen de la convención

¿Qué uso podemos encontrar, pues, para el concepto de lenguaje? Aquello que recoge es la *habilidad* para interpretar o hablar con otra persona, que consiste en la habilidad que le permite construir una teoría al paso correcta, esto es, convergente para las transacciones lingüísticas con esa persona. La habilidad para comunicarse por medio del habla consiste en la habilidad de hacerse comprender y comprender. Debemos abandonar, entonces, la idea de una *estructura compartida claramente definida* que los usuarios del lenguaje adquieren y luego aplican a casos. Debemos abandonar el intento de apelar a *convenciones* para iluminar cómo nos comunicamos. Dice Davidson:

Cuando hablamos de *reglas* del lenguaje, usualmente pensamos en las descripciones que los gramáticos o los lingüistas hacen en la práctica (generalizada e idealizada), o en las *prescripciones* que los gramáticos desean que se sigan. Las reglas pueden servir de ayuda para aprender una lengua, pero sólo podemos disponer de esa ayuda en la adquisición de una segunda lengua. La mayor parte del aprendizaje de cómo usar palabras se logra sin aprender explícitamente ningún tipo de reglas. (1992b/2001:113; mis itálicas).

Lo que hemos abandonado no es sólo la noción ordinaria de lenguaje, sino que hemos borrado el límite entre: saber un lenguaje (*knowing a language*) y saber cómo lidiar con el mundo (*knowing our way around in the world*). Y, dado que no hay reglas para arribar a las teorías al paso, *no hay reglas en sentido estricto*, como opuestas a máximas y generalidades metodológicas.

Llegados a este punto, creo que ahora podemos entender más cabalmente aquella observación que anticipamos en el capítulo 3 a propósito de la insuficiencia de la dimensión *deóntica* a la hora de capturar la normatividad propia de lo mental/lingüístico, esto es la insuficiencia de pensarla en términos de “convenciones” o “reglas”, puesto que él no cree que haya normas “inherentes al lenguaje mismo”. Las únicas que hay en nuestra práctica lingüística, son “enteramente contingentes, dependientes de “valores ulteriores “(*further values*), no de valores intrínsecos al lenguaje (1998/2005:326).

Por su parte, Davidson (1994) se pregunta qué es conceptualmente primario: ¿el idiolecto o el lenguaje? Si se afirma que el idiolecto, la aparente ausencia de norma social hace difícil

---

un elemento necesario para el lenguaje, ponen la cuestión al revés. La verdad es que el lenguaje es condición para tener convenciones” (1984b/1984: 280).

explicar el éxito en la comunicación; si se adhiere a que es el lenguaje, el peligro es que la norma no tiene una relación clara con la práctica. Dummett piensa que al promover la primacía del idiolecto, Davidson iría en contra la posición de Wittgenstein acerca del lenguaje privado; en opinión de Davidson, en cambio, Dummett no aprecia adecuadamente el elemento social esencial en la conducta lingüística. En este artículo se ocupa, entonces, de clarificar las cuestiones involucradas.

Tanto Davidson como Dummett insisten en que la conducta verbal es necesariamente social: no podría haber nada como un lenguaje sin más de una persona. Quizás incluso están de acuerdo, dice Davidson, en la razón wittgensteiniana subyacente: sin un medio social nada podría contar como aplicar incorrectamente las palabras. En donde se distancian es en cómo piensan que el medio social hace su contribución esencial. En este contexto, vuelve sobre el carácter trascendental de su empresa, que consideramos en el capítulo 2: su interés no fue describir una práctica efectiva, sino decidir lo que es necesario para la comunicación lingüística. Y aquí es donde ve razones claras para dudar de que el lenguaje sea esencial (si se va a entender como implicando modos compartidos de hablar). Las mismas dudas se aplican a la noción de *seguir una regla*, comprometerse en una *práctica*, o adecuarse a *convenciones*, si van a considerarse como implicando tal *compartir* (1994/2005:123) Davidson sostiene, en concordancia con Dummett y Kripke (y quizás Wittgenstein), que la respuesta a la pregunta de qué es continuar como antes requiere hacer referencia a la *interacción social*.

En la última parte de este trabajo, exploraremos si a partir del singular tipo de interacción social que Davidson pone en juego, es posible clarificar el tipo peculiar de normatividad *evaluativa* que, a mi juicio, se pone en juego en la interpretación.

### 5.3. La tesis de la conexión conceptual entre el pensamiento y el lenguaje<sup>166</sup>

¿Qué animales son racionales? Tal es la pregunta con la que nos interpela Davidson en “Rational Animals”. Él aclara de inmediato que su propósito no es nombrar especies o grupos; esto es, su motivación no es decidir si los delfines, los embriones humanos o los políticos son racionales. Su interés principal reside, en cambio, en elucidar las condiciones de posibilidad de la racionalidad, en establecer qué es lo que hace que un animal (o cualquier otra cosa, en principio) sea racional. A tal efecto, comienza por ofrecer su criterio de racionalidad: ser racional es tener actitudes proposicionales. Surge, entonces, la siguiente pregunta: ¿cuándo podemos decir que una criatura tiene actitudes proposicionales? ¿Qué evidencia es relevante para decidirlo? Davidson enfatiza que debe observarse un patrón muy complejo de conducta para justificar la atribución de pensamiento. Y sostiene que hay tal patrón sólo si el agente tiene lenguaje. En efecto, en “Rational Animals” (1982) defiende la tesis según la cual sólo pueden tener pensamientos aquellos individuos que se comunican entre sí por medio de un lenguaje hablado.

Como señalan Lepore y Ludwig (2005), el supuesto de que el lenguaje es necesario para el pensamiento está a la base de varios argumentos davidsonianos. Por ejemplo, en el argumento *contra la posibilidad de una pluralidad de esquemas conceptuales radicalmente diferentes* se asume que un esquema conceptual podría ser identificado con un conjunto de lenguajes intertraducibles, lo que supone que sólo los seres lingüísticos tienen esquemas conceptuales. En el argumento para *la imposibilidad del error masivo*, la conclusión de que la creencia es verídica por naturaleza reposa en el supuesto de que *la creencia es por su naturaleza lingüística*. A su vez, la explicación de la *autoridad de primera persona* se sustenta en la apelación a una asimetría en el conocimiento de los significados: esta no puede ser la explicación correcta de por qué los pensadores tienen un conocimiento especial de sus propios pensamientos, a menos que todos los pensadores sean hablantes de un lenguaje. En efecto, el supuesto de que sólo los hablantes de un lenguaje son pensadores *debe* estar en el trasfondo de todo argumento que pretenda extraer conclusiones acerca de la naturaleza de las actitudes proposicionales a partir de

---

<sup>166</sup> Como señalé más arriba la tesis davidsoniana sostiene una *interdependencia*: esto es, el lenguaje requiere del pensamiento y el pensamiento requiere del lenguaje. No obstante, por razones de pertinencia expositiva, voy a restringirme a considerar especialmente la última. Asimismo, para una defensa de una versión relacional lingüística de la intencionalidad inspirada en esta propuesta davidsoniana, puede verse Brandom, R. (1994), especialmente, caps. 3 y 8.

los supuestos que tiene que hacer un intérprete radical. Si pudiera haber seres no lingüísticos con actitudes proposicionales, las conclusiones alcanzadas sobre la base de lo que debe suponerse acerca de seres lingüísticos a los efectos de interpretarlos no se extenderían a seres no lingüísticos y por tanto no podrían ser caracterizadas como revelando nada acerca de la naturaleza de las creencias como tales. Además, si las conclusiones alcanzadas pudieran ser vistas -en el mejor de los casos- como restringidas a seres lingüísticos, las conclusiones mismas estarían en riesgo de ser socavadas. Surgiría la pregunta de por qué si nuestra concepción de la creencia y de otras actitudes proposicionales es independiente de qué es hablar un lenguaje, y así independiente de la interpretación, tendríamos que concebir la perspectiva de la interpretación como la única perspectiva desde la cual comprendemos nuestras creencias, y otras actitudes. ¿Por qué insistir tan fuertemente, por ejemplo, sobre la posibilidad de ser interpretado a partir de la perspectiva del intérprete radical, que subyace a los fuertes resultados que alcanza Davidson acerca de la naturaleza de nuestras creencias? En suma, pues, al abordar la tesis acerca de la dependencia conceptual del pensamiento respecto del lenguaje, estamos rozando un punto sumamente “sensible” de la filosofía davidsoniana. La necesidad del lenguaje para el pensamiento detenta un rol central: los argumentos acerca de la *naturaleza* de la creencia a partir de lo que debe suponerse en la *interpretación* serían socavados si la creencia no estuviera esencialmente ligada a ser interpretable.

A mi juicio, Davidson tiene -al menos - dos grandes líneas argumentativas para sostener la tesis de que los seres no lingüísticos son incapaces de pensamiento. En la primera de ellas pone en juego uno de los rasgos del concepto de creencia que presentamos en el capítulo 3, a saber: el *holismo* de lo mental. En efecto, el corazón de su estrategia se concentra en el contraste entre la *naturaleza holística* de la atribución de actitudes y el tipo de evidencia disponible para la atribución de actitudes en ausencia de conducta verbal. Davidson considera el caso del perro de Malcolm, al que tendemos a adscribirle la *creencia* de que el gato subió al roble, luego de que vemos que lo siguió y que le ladra al árbol. Y dice:

Solamente cuando podemos ubicar los pensamientos dentro de una densa red de creencias relacionadas, identificamos pensamientos, hacemos distinciones entre ellos y los describimos según lo que son. Si realmente queremos adscribir inteligiblemente una creencia a un perro, *debemos ser capaces de imaginar cómo decidiríamos si el perro tiene muchas otras creencias del tipo necesario para dar sentido a la primera*. Me parece que empecemos por donde empecemos, muy pronto llegaremos a creencias sobre las cuales no tendremos idea de si

decir que el perro las tiene, y que sin embargo son tales que sin ellas, nuestra confiada atribución inicial parece inestable (1982/2001:98; mis itálicas).

Dicha argumentación puede reconstruirse del modo siguiente.

**A. El argumento a partir del holismo de lo mental:**

A1. Las creencias y otras actitudes proposicionales son adscribibles sólo en densas redes.

A2. La atribución de una densa red de actitudes proposicionales a otros requiere en su apoyo de un patrón rico de conducta que le dé sustancia a tales atribuciones. (Las premisas A1 Y A2 son centarles a la concepción de Davidson acerca de la naturaleza de las actitudes tal como se presenta en “Rational Animals” (1982/2001) y “The Method of Truth in Metaphysics” (1977b/2001).

A3. El patrón de conducta requerido no puede ser exhibido en ausencia de conducta verbal interpretable como actos de habla (1982/2001)

A4. Por tanto, sólo los animales lingüísticos pueden tener actitudes proposicionales.

El argumento es inconcluyente, como reconoce el propio Davidson. El problema yace en conferirle sustento a A3. Aquí hay lugar para aplicar la analogía a la que nos referimos en el capítulo anterior, esto es, entre las actitudes proposicionales y las actitudes numéricas a la que Davidson apela, entre otras cosas, para favorecer la tesis de la *indeterminación de la interpretación*. Recordemos que así como podemos emplear distintas escalas para medir la temperatura, las distintas oraciones de creencia que empleamos para rastrear la conducta del hablante capturan igualmente bien el fenómeno que nos interesa, lo que es diferente de decir que no haya tal fenómeno. Según vimos en el capítulo 4, lo que sugiere la indeterminación es que los conceptos particulares que empleamos al atribuir creencias son más ricos en contenido y estructura que aquellos que podríamos atribuir a un perro u otro ser no lingüístico. Dado que no tenemos palabras que expresen conceptos del grano requerido, empleamos los nuestros y, así sobredescribimos los contenidos de las actitudes del perro. Pero, podría objetarse –tal como vimos antes que hacen Carruthers y Boucher (1998), en una línea también adoptada por Jeffrey

(1985)- que esto no ha de hacernos suponer que no hay actitudes en absoluto, sino sólo que nuestras oraciones no están hechas para capturar exactamente sus condiciones de verdad.

En otro el argumento al que apela Davidson (1982) para favorecer la tesis de que el pensamiento requiere de lenguaje, veremos que desempeñan un rol crucial los otros dos rasgos del concepto de creencia que presentamos en el capítulo 3, a saber: **normatividad** y **triangulación**. Pasemos, entonces, a considerarlo.

### **B. El argumento a partir del *concepto* de creencia**

B1. Se puede tener actitudes proposicionales sólo si se tiene creencias.

B2. Se puede tener creencias sólo si se tiene el concepto de creencia.

B3. Se puede tener el concepto de creencia sólo si se tiene un lenguaje.

B4. Por tanto, se puede tener actitudes proposicionales sólo si se tiene un lenguaje.

Según vimos en el capítulo 3, a (B1) subyace el holismo del *contenido*, según el cual atribuir cualquier actitud requiere atribuir creencias y atribuir cualquier creencia requiere atribuir un rango adicional de otras creencias con contenidos relacionados. Allí señalamos que de acuerdo con Davidson, sería impropio insistir en una lista particular o fija de creencias asociadas (en este sentido, dijimos que no se trata de un holismo “rígido”, dado que siempre puede haber explicaciones disponibles de por qué aparece una creencia y no otra, pero el punto es que sin tal patrón apropiado, no podemos sostener nuestra atribución). Esta variedad de holismo, dijimos, es la posición que dota de “carnadura” a la tesis de la “centralidad de la *creencia*” (Fodor y Lepore, 1992: 18). Así el resto del argumento está centrado en la creencia.

A propósito de (B2), la premisa más cuestionable del argumento, Davidson traza algunas aclaraciones preliminares, a saber: no se compromete con que todo pensamiento es autoconsciente o que cuando pensamos que *p* debemos creer que creemos que *p*. Su posición es, en cambio, que para tener cualquier actitud proposicional en absoluto, es necesario “tener el concepto de creencia, tener una creencia sobre una creencia” (1982:104). Según veremos, en este punto va a desempeñar un papel central el rasgo peculiar de lo mental: su **normatividad**. La idea

básica es esta: dado que toda creencia implica otras creencias y otras actitudes proposicionales, y que toda creencia es pasible de ser revisada a la luz de otras creencias, tener una creencia es estar preparada para cambiar nuestras creencias a la luz de nueva información, esto es, tenemos que estar preparados para distinguir lo meramente creído de lo que es el caso. A su juicio, el fenómeno de la *sorpresa* torna claro el caso. En efecto, Davidson trata de conferirle apoyo a (B2) mediante:

**El argumento (C) a partir de la *sorpresa*.**

C1. Se puede tener un *stock* general de creencias sólo si se puede ser sorprendido.

C2. Se puede ser sorprendido (esto es darse cuenta de que una o más de nuestras expectativas no fueron *correctas*) sólo si se tiene el concepto de creencia.

C3. Por tanto se puede tener creencias sólo si se tiene el concepto de creencia.

Frente a (C1) podríamos imaginar un ser omnisciente: por hipótesis tiene creencias, pero es obviamente incapaz de ser sorprendido (cf. Lepore y Ludwig (2005) No obstante, parece claro que la premisa especialmente cuestionable es (C2). Veamos, más detenidamente cómo ilustra el punto Davidson.

Supongamos que creo que tengo una moneda en el bolsillo. Vacío mi bolsillo y no encuentro ninguna. Me sorprende. Claramente, no podría sorprenderme de no haber tenido, en primer lugar, una creencia (la de que tenía una moneda en el bolsillo). Ahora bien, la idea de Davidson es que no es suficiente que primero tenga la creencia en cuestión y que luego de vaciar mi bolsillo ya no la tenga, sino que la sorpresa involucra un paso ulterior. Requiere que sea consciente del contraste entre lo que creía y lo que llego a creer. Tal consciencia es, a su juicio, una creencia acerca de una creencia: si estoy sorprendida, entonces *creo* que mi *creencia* original era falsa. En suma, la idea es que tener el concepto de creencia supone el *dominio del contraste entre lo creído y lo que es el caso*, i.e. la noción de un mundo objetivo independiente. El punto que él pretende establecer por esta vía es que “no se puede tener un stock general de creencias, del tipo necesario para tener cualquier creencia sin más, sin estar sujetos a sorpresas que involucran **creencias acerca de la corrección de las propias creencias**” (1982/2001:104).

Tal es el sentido en el que considero que aquí desempeña un papel central la **normatividad** de lo mental, en su dimensión *deóntica*. Recordemos que Davidson afirma que por “pensamiento” entiende un estado mental con contenido especificable. En tanto dicho contenido tiene condiciones de satisfacción que se juegan en la dirección de ajuste al mundo, lo mental tiene **condiciones de corrección** que involucran una dimensión normativa: el estar preparada para ajustar mi creencia está sujeto a que dicha revisión satisfaga condiciones de consistencia y coherencia racional. De este modo: tener una creencia es estar en **un estado regido por el principio constitutivo de la racionalidad** que introduce, según vimos en el capítulo anterior, un carácter “irreduciblemente normativo” en la esfera de lo mental.

Ahora bien, esta estrategia argumentativa a favor de (B2), la premisa crucial que afirma que se puede tener creencias sólo si se tiene el *concepto* de creencia, es pasible de ser “blanco” de la siguiente crítica. Se podría hacer caso omiso del carácter conceptual del ejercicio que nos propone Davidson e intentar socavarlo desde consideraciones *empíricas*.

En efecto, la noción davidsoniana de creencia parece resultar incompatible con aquella que está a la base de estudios empíricos experimentales tales como los que se vienen llevando a cabo en psicología del desarrollo sobre *teoría de la mente*. Pese a que no hay una caracterización unívoca de “teoría de la mente”, aludo a ella bajo el sentido que le fue conferido originalmente por David Premack y G. Woodruff (1978), esto es, como la habilidad para explicar, predecir e interpretar tanto la conducta propia como la de terceros mediante la atribución de creencias y deseos. Cuando un individuo tiene una teoría de la mente, comprende que otros pueden tener creencias que difieran de las suyas, dada la naturaleza representacional de la mente. Se sostiene que, en general, esto no sucede hasta alrededor de los 4 años. El test estándar que se emplea para defender esta afirmación es conocido como la “tarea de la falsa creencia”.<sup>167</sup> La idea es que para resolver el test el sujeto debe ser capaz de considerar pensamientos del tipo *ella cree que p*. Para ser capaz de considerar pensamientos de esta forma, el sujeto debe poseer el *concepto* de creencia. El punto aquí es, precisamente, contrario al espíritu davidsoniano: la idea es que tener creencias no es asimilable a tener el concepto de creencia. De hecho, se concede que los chicos

---

<sup>167</sup> Brevemente, la tarea consiste en lo siguiente: al sujeto se le cuenta una historia acerca de Sally, una chica que ubica una barra de chocolate en una canasta y luego deja la habitación. Mientras Sally está afuera, Anne entra en la habitación, pone la barra de chocolate en una caja y se va. Una vez que Sally vuelve a la habitación, se le pregunta al sujeto dónde va a buscar Sally la barra de chocolate. Los chicos por debajo de los 4 años en general dicen que Sally va a buscarla en la caja (su ubicación efectiva), mientras que los chicos más grandes indican la canasta (que es el lugar donde Sally dejó el chocolate) y resuelven exitosamente la tarea.

por debajo de 4 años tendrían *creencias* (acerca del mundo, aunque no acerca de otras creencias) y se explica su falla en dar la ubicación donde Sally vio la barra de chocolate por última vez en virtud de la carencia de una teoría de la mente y, por tanto, del *concepto de creencia*. De modo que frente a la noción davidsoniana que nos llevaría a decir muy contraintuitivamente que los chicos por debajo de 4 años no tienen creencias, el paradigma experimental de la tarea de la falsa creencia invita a pensar en una noción “deflacionaria” a la vera de la cual *tener creencias no requiere tener el concepto de creencia*.

En un espíritu análogo, Lepore y Ludwig (2005) citan a Gopnik y Astington (1988) y Perner (*et. al.* 1987) para abonar que hay evidencia según la cual los chicos que no pasan el test en cuestión, y que no tendrían el concepto de creencia, no obstante, son capaces de ser sorprendidos (pese a que, si Davidson está en lo correcto, no tendrían actitudes proposicionales en absoluto).

Por su parte, Pérez, D. (2005: 186) pone en jaque las pretensiones del holismo, que en el capítulo 3 llamé “psicológico” o “inter-actitud”. A su juicio hay “evidencia convergente acerca de que los chicos no adquieren todos los conceptos mentales al mismo tiempo”. La evidencia sugiere que los chicos normales de 2 años usualmente comprenden conceptos como “desear”, los de 3 conceptos como “saber” y “pensar” y los de 4 el de “creencia”, esto es, comprenden situaciones en las que la gente puede tener creencias diferentes acerca del mismo estado de cosas. Dado que median dos años entre la adquisición del concepto de “deseo” y el de “creencia” parece natural poner en cuestión el holismo extremo de lo mental que propugna Davidson. Desde luego, Pérez contempla el recurso de decir que el concepto de “deseo” que el chico tiene previo a los 4 años es diferente del que tiene luego, pero lo encuentra como una mera movida *ad hoc* para retener el holismo en cuestión. Este señalamiento, en su opinión, socava la premisa (B2): si no hay creencias son el concepto de creencia, y no hay deseos (como un caso especial de actitud proposicional) sin creencias, se sigue que no hay deseos sin el concepto de creencia. Sin embargo, tal es precisamente el caso en los niños pequeños según la psicología del desarrollo: entre los 2 y los 4 años, los chicos tienen el concepto de deseo, pero no tienen el concepto de creencia.

Sin embargo, estas críticas podrían objetarse dado su talante empírico “externo” a la propuesta conceptual de Davidson e, incluso podrían cuestionarse al “interior” de la misma, en la medida en que está concebida para dar cuenta de la situación *adulto* normal, y en consecuencia,

las consideraciones tomadas del ámbito de la psicología del desarrollo no tendrían por qué hacerle mella. Con todo, creo las dificultades persisten aún si nos restringimos a la estricta esfera conceptual.

Según acabamos de ver, la normatividad resulta un rasgo central de la noción de creencia bajo una de sus acepciones *deónticas*, a saber, en términos de la *regla* o *principio constitutivo* de la *racionalidad* (no así como *convención*, según vimos en la sección anterior). Ahora bien, se trata de un rasgo que, según vimos en el capítulo 4, *segrega radicalmente* los conceptos mentales de los conceptos físicos.

La consecuencia que dicho dualismo conceptual tiene para el problema que estamos considerando es la siguiente: nos compromete con la admisión de un salto de “prestidigitación” desde una situación triangular no proposicional, meramente causal o física, a la situación triangular proposicional inscrita en el espacio “normativo” de las razones, que tiene lugar entre adultos normales.

Para comprender, cabalmente, esta dificultad pasemos, entonces, a considerar la estrategia argumentativa a favor de (B3), la premisa que afirma que se puede tener el concepto de creencia sólo si se tiene un lenguaje. Aquí resultará crucial el papel que Davidson le asigna al otro de los rasgos de la noción de creencia, a saber: la **triangulación**.

A los efectos de apoyar (B3), el primero de los argumentos que ofrece Davidson es **(D) a partir del error**, y puede reconstruirse así:<sup>168</sup>

D1. Para tener el concepto de creencia, se debe tener el concepto de error, o lo que es lo mismo, de verdad objetiva.

D2. La afirmación de que una criatura posee el concepto de error, o de verdad objetiva, requiere de fundamentación, esto es, necesitamos dar una explicación de cómo una criatura es capaz de tener un concepto tal, qué condiciones deben estar satisfechas.

D3. Podemos comprender cómo una criatura que está en comunicación con otras podría tener el concepto de error, como una herramienta usada en la interpretación para lograr un mejor ajuste

---

<sup>168</sup> Aquí sigo los lineamientos centrales de la propuesta de Lepore y Ludwig (2005). Sus fuentes principales son: Davidson 1975, 1982, 1988 y 1991.

racional de la conducta de un hablante con la evidencia que tenemos para sus creencias y significados.

D4. Hay un campo para la aplicación del concepto de verdad objetiva en la conducta o experiencia de una criatura sólo si está –o ha estado- en comunicación con otras.

D5. Por tanto (de D2-D4), para tener el concepto de error se debe estar o haber estado en comunicación con otros.

D6. Por tanto, de (D1-D5) para tener el concepto de creencia se debe tener un lenguaje y estar o haber estado en comunicación con otros.

Según vimos, para Davidson, sólo podemos tener creencias si tenemos la noción de un mundo objetivo independiente de nuestras creencias, noción a la que sólo se puede llegar, si se tiene *lenguaje*. Una criatura puede interactuar con el mundo en modos complejos sin tener ninguna proposición; puede discriminar entre colores, sabores, sonidos y formas. No obstante, ninguna de estas cosas, por exitosas que sean, muestran para Davidson que la criatura pueda dominar el contraste en cuestión. ¿Qué mostraría que se maneja el contraste? Claramente para Davidson la *comunicación lingüística* es condición suficiente. La comunicación depende de que cada comunicador tenga y piense correctamente que el otro tiene el concepto de un mundo compartido. Davidson sugiere, entonces, que el concepto de verdad intersubjetiva es suficiente como base para la creencia y, por tanto, para los pensamientos en general. Y le parece plausible que tener el concepto de verdad intersubjetiva dependa de la comunicación en sentido lingüístico. No obstante, asume que para completar su argumento todavía tendría que mostrar que el lenguaje es condición necesaria, esto es, tendría que mostrar que es el único modo en que se puede llegar a tener el contraste entre lo creído y lo que es el caso. Confiesa que no sabe cómo mostrar esto último, aunque afirma al mismo tiempo que tampoco tiene idea respecto de qué otra forma podría arribarse al concepto de mundo objetivo. Apela, entonces, a una línea persuasiva en la que en tercero de los rasgos de la noción de creencia que consideramos en el capítulo tres, la situación de *triangulación*, cobra un rol protagónico.

Finalmente, encontramos el **argumento (E) a partir de la triangulación**.

E1. Podemos dotar de sentido a que haya un objeto determinado de pensamiento para una criatura sólo si podemos verla como un hablante que triangula con otro con el que está en comunicación acerca de un objeto común de pensamiento.

E2. Nada puede tener pensamientos a menos que pueda haber objetos determinados de sus pensamientos.

E3. Por tanto, nada es un pensador a menos que sea un hablante que está en –o ha estado en– comunicación con otro hablante.

Como adelantamos en el capítulo 3, la idea central aquí es que nuestro sentido de la objetividad es la consecuencia de un tipo de **triangulación** que requiere de dos criaturas, cada una de las cuales interactúa con un objeto, pero en la que lo que le da a cada una el concepto del modo en que las cosas son *objetivamente* es la línea de base formada entre las criaturas por medio del *lenguaje*. Veamos este punto más detenidamente. Sin la situación de triangulación no podríamos explicar la objetividad del pensamiento por lo siguiente. El pensamiento proposicional es objetivo en el sentido de que tiene un contenido que es verdadero o falso independientemente de la existencia del pensador. Además, este es un hecho del que debe ser consciente un pensador: no se puede creer algo sin saber que lo que se cree puede ser verdadero o falso y que se puede estar equivocado. ¿De dónde obtenemos esta idea de que podemos estar equivocados? Según vimos, en la situación de triangulación cada una de las criaturas correlaciona sus propias reacciones a fenómenos externos, con las reacciones del otro. Una vez que estas correlaciones están establecidas, cada criatura está en posición de esperar el fenómeno externo cuando percibe la reacción asociada del otro. Lo que introduce la *posibilidad de error* es la falla ocasional de la expectativa, las reacciones que no se correlacionan. Ahora bien, a menos que la línea de base del triángulo -entre los dos agentes- sea reforzada al punto de que pueda implementar la comunicación de contenidos proposicionales, no hay manera de que los agentes puedan hacer uso de la situación de triangulación para formar juicios acerca del mundo. Sólo

cuando el *lenguaje* está en su lugar las criaturas pueden apreciar el concepto de verdad objetiva (1997c/2001:130). Leemos en Davidson:

La única forma de saber que el segundo ángulo del triángulo –la segunda criatura o persona- está reaccionando al mismo objeto que uno es saber que la otra persona tiene en mente el mismo objeto. Pero entonces la segunda persona debe también saber que la primera persona constituye un ángulo del mismo triángulo en el cual el otro ángulo lo ocupa la segunda persona. La razón es que para que dos personas sepan que ambas están relacionadas de esta manera, que sus pensamientos están relacionados así, se requiere que estén en comunicación. Cada una de ellas debe hablar a la otra y ser entendida por la otra. Como he dicho no tienen que significar la misma cosa mediante las mismas palabras, pero cada una debe ser intérprete de la otra (1992/2001:121).

A los efectos de precisar, entonces, en qué consiste tener un lenguaje resulta pertinente remitirnos brevemente a “Thought and Talk” (1975). Allí sostiene que ser hablante de un lenguaje supone ser *intérprete* de los demás. Recordemos que el problema básico que debe afrontar la teoría davidsoniana de la interpretación es que no se puede asignar significado a las emisiones de un hablante sin saber lo que el hablante cree, en tanto que, a su vez, no se pueden identificar creencias sin saber lo que significan las emisiones del hablante. De esto resulta que debe proveerse una teoría de la creencia y una teoría del significado al mismo tiempo. Como vimos en el capítulo 2, Davidson sostiene que el modo de lograrlo es vía la aplicación del principio de caridad que involucra optimizar el acuerdo entre nosotros y aquellos a quienes interpretamos, esto es, interpretar a los hablantes como sosteniendo creencias verdaderas toda vez que sea plausible hacerlo. Si suponemos que las creencias del hablante están de acuerdo con las nuestras –al menos en los casos más simples y básicos- y resultan así en su mayoría verdaderas, podemos usar nuestras creencias acerca del mundo como guía hacia las creencias del hablante. Y dado que podemos identificar emisiones asertóricas simples por parte del hablante (recordemos que esto es así en virtud de que podemos identificar la actitud de sostenerlas como verdaderas) la interconexión entre creencia y significado nos permite usar nuestras creencias como guías hacia los significados de las emisiones del hablante. En suma: la interpretación es considerada por Davidson como una empresa conjunta de asignación de significado y atribución de creencias. De esta suerte, para ser un *hablante* hay que ser un *intérprete*.

Una vez más, cabe recordar que Davidson hace explícito que él simplemente describe “rasgos de ciertos conceptos”. En este sentido, su abordaje de la relación entre pensamiento y lenguaje a muchos les ha resultado un ejemplo paradigmático de las limitaciones que un enfoque exclusivamente conceptual puede importar en filosofía de la mente. Tal es el caso de Kristin Andrews (2002). Para ella, la falta de interés por los desarrollos empíricos, lo ha llevado a elaborar una teoría de la mente que, aunque resulte altamente coherente, esté desconectada del mundo tal como es comprendido por medio de las ciencias. En este espíritu, veamos algunas dificultades involucradas en la conclusión derivada de “Thought and Talk”, según la cual todos los hablantes son *intérpretes*. A su juicio, estudios recientes sobre casos de autismo severo podrían jugar como contraejemplos a dicha tesis. Esto es, algunos casos de autismo parecen ponernos ante hablantes que, pese a tener algunas dificultades semánticas y pragmáticas, calificarían no obstante como *hablantes* competentes y que no podrían considerarse *intérpretes* dado, precisamente, su déficit severo en teoría de la mente (esto es, su déficit para atribuir actitudes proposicionales). El punto es el siguiente. Como ya se ha señalado, la interpretación es considerada por Davidson como una empresa conjunta de asignación de significado y *atribución de creencias*. Ahora bien, hay evidencia que indicaría que la mayoría de los individuos autistas no puede satisfacer este requisito de atribución intencional. La mayoría de ellos es incapaz de pasar la tarea de la falsa creencia, cuya resolución es asumida como criterio para determinar la presencia de una teoría de la mente, esto es, como índice de la habilidad de explicar y predecir la conducta propia y ajena mediante la *atribución de creencias* y deseos. Sin embargo, se han reportado casos de autistas con déficit severo en teoría de la mente que logran, con todo, desempeñarse como hablantes competentes. En tales casos se registran algunos problemas tales como escasa sensibilidad a rasgos prosódicos, dificultad en el uso de artículos, respuestas literales a pedidos indirectos; dificultades que, no obstante, no impedirían calificar a los sujetos en cuestión como hablantes competentes. Ahora bien, a mi juicio, Davidson podría poner serios reparos a este tipo de objeciones. Además de volver a defender el carácter conceptual de su empresa, podría sostener, asimismo, que el contraejemplo propuesto no afecta a su análisis en la medida en que éste pretende ocuparse sólo de personas adultas *normales*, por lo que el caso autista resultaría inadecuado.

No obstante, está el otro frente crítico que presentamos más arriba y que, a mi juicio, resulta más relevante en tanto hace justicia al carácter conceptual de la propuesta davidsoniana.

Según vimos, el triángulo que Davidson señala es esencial para la existencia, y por tanto, para la emergencia del pensamiento. Dijimos que sin él hay dos aspectos del pensamiento que no podemos explicar: la objetividad y el contenido empírico de los pensamientos acerca del mundo. Ahora bien, aquí hay al menos dos señalamientos que resultan relevantes.

Por un lado, él defiende que la relación triangular es necesaria para el pensamiento, pero no es suficiente, tal como se muestra por el hecho de que puede existir en animales a los cuales no les atribuimos juicios. En efecto, esta interacción triangular entre criaturas y un entorno compartido que no requiere pensamiento o lenguaje ocurre con gran frecuencia entre animales que ni piensan ni hablan. El punto crucial es que se trata de una insuficiencia por principio: nunca obtendremos condiciones suficientes mediante la añadidura de más detalles a esta descripción, dado que si ello fuera posible, el pensamiento sería reducible a lo físico, cuestión que tal como vimos en el capítulo anterior, Davidson clausura *a priori*: la situación triangular en la que hay intercambio de contenidos proposicionales está atravesada por una irreducible normatividad.

Por otro lado, cabe señalar que la idea de “*triangulación*” es usada por Davidson para diferentes fines en distintos contextos.<sup>169</sup> Como señala el propio Davidson, la triangulación no es una explicación empírica del proceso de adquisición del lenguaje, sino una elucidación de las condiciones de posibilidad del pensamiento, el lenguaje, la ostensión y el contenido perceptual. Él admite tres aplicaciones diferentes de esta idea (2001b): primero, como una descripción de las condiciones necesarias (completamente no-intencionales) para el pensamiento; segundo, como una descripción de la situación de aprendizaje en la que una criatura aprende su primer lenguaje (la situación en la que el pensamiento y el lenguaje “emergen”); y tercero, la situación en la que se determinan los contenidos de nuestras creencias perceptuales.

Hay una situación triangular adicional, la original en el pensamiento de Davidson: la situación de interpretación radical, esto es cuando dos personas no comparten el mismo lenguaje (no obstante, cabe recordar aquí que las situaciones ordinarias de interpretación no difieren en ningún punto esencial de la situación de interpretación radical: “toda comprensión del habla de otro involucra interpretación radical” (1973/1984:125). El punto que me interesa señalar aquí es que, dado el rasgo peculiar que atraviesa esta dimensión propiamente proposicional, su *normatividad*, no podemos conectarla - y ello por principio- con el primer triángulo

---

<sup>169</sup> Sigo aquí varios señalamientos hechos por Pérez (2005: 185).

(causal/físico), tal es el sustento del “anomalismo” de lo mental que vimos en el capítulo anterior. Los conceptos mentales y físicos “no están hechos el uno para el otro”. *Contrario sensu*, la reducción psicofísica sería posible. Por su parte, tampoco podemos conectar –y ello también por principio- el triángulo causal/físico con aquel que se pone en juego entre el niño y el adulto que le enseña la primera lengua: en la interacción con el adulto el vínculo está atravesado inexorablemente por la normatividad propia de los conceptos mentales que éste “inescapablemente” emplea en su trato con el niño. En otras palabras: hay una mediación intersubjetiva de la causalidad que introduce inevitablemente la normatividad. Recordemos que desde la perspectiva interpretativa, la normatividad es siempre normatividad *atribuida*. De este modo, a partir de interactuar con un adulto intérprete que no puede sino proyectar las restricciones estructurales del juego interpretativo, el niño que con él triangula, se vincula con los objetos y eventos del mundo mediante lazos causales que son inescindibles de la dimensión normativa. De este modo, habría que comprometerse con el paso de “prestidigitación” que mencionamos más arriba entre la situación pre- lingüística, pre-conceptual y la cabalmente intencional.

Frente a este estado de la cuestión, mi invitación es a que volvamos sobre la peculiar normatividad de lo mental y el trasfondo en el que se inscribe en el pensamiento davidsoniano: la dicotomía descriptivo/normativo, que resuena desde la dicotomía seminal en “Acciones, razones y causas” (1963) entre causalidad y racionalidad. Hemos señalado que Davidson parece apuntar a una dimensión no meramente *deóntica* de la normatividad sino a los *valores ulteriores* en los que reposa la singular práctica de nuestra comprensión mutua.

La conclusión de sus consideraciones en “Rational Animals” es que “la racionalidad es un *rasgo social*. Solamente la tienen quienes son capaces de comunicarse” (1982/2001: 105; mis *itálicas*). Todas estas consideraciones, dice, no se aplican solamente al lenguaje, sino que se aplican igualmente al pensamiento en general. A su juicio,

la creencia, la intención y las otras actitudes proposicionales son todas *sociales* en el sentido de que son estados en que una criatura no puede estar sin tener el concepto de verdad intersubjetiva, y este es un concepto que no se puede tener sin compartir y saber que uno comparte *con alguien más* un mundo y una manera de pensar acerca del mismo” (1992/2001:121; mis *itálicas*).

Según vimos en la sección anterior, Davidson afirma que aquellos que insisten en que las prácticas compartidas son esenciales están parcialmente en lo correcto: debe haber un *grupo interactivo* para que emerja el pensamiento proposicional (1994/2005:125). Mi sugerencia es, pues, que profundicemos en este señalamiento que Davidson no exploró en toda su hondura: quizás encontremos el *valor ulterior* en el que reposa la comprensión mutua en un tipo peculiar de *interacción social*.

La idea que pretendo sostener es que la normatividad de lo mental, aquel rasgo que a juicio de Davidson porta la clave de lo humano, debe comprenderse en términos del principio de racionalidad como regla constitutiva de la práctica especial por la que nos tratamos mutuamente como *personas*. Tal como anticipé en el capítulo anterior, el peso de la clarificación que persigo se traslada, entonces, a la noción de *persona*. La motivación que me guía es que, tal vez, a la vera del tipo de interacción social en la que ponemos en juego en este valor, ni el problema mente-cuerpo ni el problema de la relación entre el pensamiento y el lenguaje queden confinados a que, por principio, detengamos nuestra discusión.

Tal como vimos, Davidson adopta, en la senda de Quine, la perspectiva de *tercera persona* (la del intérprete) como metodológicamente básica. Asimismo, sostiene que su propuesta teórica es conciliable con el reconocimiento de la autoridad de la *primera persona*. Y según acabamos de ver, finalmente, afirma que sin una *segunda persona* el pensamiento y el lenguaje no serían posibles. Consideremos, pues, cuál es el tratamiento que ofrece de cada una de estas perspectivas, a ver si podemos clarificar en qué sentido en la noción de *persona* se pone en juego un valor central de nuestra concepción de un *nosotros humano*.

## Capítulo 6: El problema de las perspectivas de primera y tercera persona de la atribución mental. Subjetivo-Intersubjetivo-Objetivo.

Un intérprete no es un *espectador ocioso*: es un *jugador esencial* en una actuación que requiere interacciones causales complejas entre la gente y el mundo.  
(Davidson, 1995c2004: 49; mis itálicas).

Hemos señalado que aquella concepción de lo mental que Davidson pretende minar es la que está a la base de los dualismos modernos, a saber: el dualismo subjetivo/objetivo y el dualismo esquema/contenido, esto es, un concepto de la mente con estados y objetos privados. En este sentido, hemos considerado que el desafío central que asume su filosofía de la mente es el de cómo superar la concepción cartesiana. Frente a dicha concepción, según dijimos, su propuesta ha de ser la de una mente *social e interactiva* (1988/2001: 52).

Como es sabido, a la vera de la concepción cartesiana se abrió una brecha profunda entre lo subjetivo y lo objetivo, entre la mente humana y el resto de la naturaleza. Así, el problema general acerca de cómo combinar la perspectiva subjetiva, esto es, la de una persona particular que se halla dentro del mundo, con una concepción objetiva de ese mismo mundo (incluidos la persona y su perspectiva) se tornó acuciante.

Se trata de un problema que resulta relevante en la medida en que nuestra respuesta (o nuestra carencia de ella) a esta cuestión determinará sustancialmente nuestra concepción del mundo y de nosotros mismos, nuestra actitud hacia nuestras vidas, nuestras acciones y nuestras relaciones con los demás. Tal como sostiene Thomas Nagel (1986), la tensión entre lo subjetivo y lo objetivo penetra en la vida humana y predomina, en particular, en la generación de problemas filosóficos en torno a la teoría del conocimiento, el libre albedrío y la metafísica de la mente.

En el ámbito específico de la filosofía de la mente contemporánea, el problema consiste en cómo conciliar la subjetividad (o la denominada “perspectiva de primera persona”) con la

objetividad (o “perspectiva de tercera persona”), de modo de cerrar la brecha abierta entre ambas en la modernidad. Se trata de un “problema” en la medida en que si nos ceñimos a la perspectiva de tercera persona, no podemos dar cuenta de la autoridad especial que parece poseer la primera persona, y si nos atenemos a la perspectiva de primera persona, se suscita el problema escéptico relativo a las otras mentes.

En particular, en la filosofía de la mente de Donald Davidson nos encontramos con el ambicioso intento de superar la concepción cartesiana de la mente y de retener ambas perspectivas, esto es, tanto la subjetiva como la objetiva, confiriéndoles un carácter igualmente genuino e irreducible.

En la primera parte de este trabajo, hemos señalado que la elección de la perspectiva de la *tercera persona* como aquella desde la cual se ha de investigar el pensamiento, el lenguaje, y su relación, está motivada por la idea de que el lenguaje es esencial y centralmente un fenómeno social y debe ser comprendido desde una dimensión pública. En ese espíritu, hemos subrayado que Davidson le atribuye el punto básico a Quine: él revolucionó nuestra comprensión de la comunicación al tomar seriamente el hecho de que no puede haber nada más para el significado que lo que una persona adecuadamente equipada puede aprender y observar. En otros términos: hemos visto que aquello que liga la empresa interpretativa de Davidson con el proyecto quineano es que, en virtud de que el lenguaje es por su naturaleza un instrumento para la comunicación, el significado debe ser accesible desde un punto de vista público: los datos relevantes se restringen, pues, a lo disponible para el intérprete radical. El intérprete radical debe construir su teoría acerca de los pensamientos y palabras de los demás y su conexión con la realidad a partir de los recursos públicos disponibles para cualquier sujeto cognoscente. En este sentido, la empresa interpretativa incluye la dimensión *objetiva*. Hasta aquí, nos hemos ocupado de su tratamiento de esta dimensión.

No obstante, Davidson pretende que la asunción de la perspectiva de tercera persona, como metodológicamente básica, deje lugar para el reconocimiento de la autoridad de la *primera persona*, esto es, de la dimensión *subjetiva*. En efecto, a su juicio, cuando un hablante afirma que tiene una creencia, se presume que no está equivocado, presunción que no se da respecto de las adscripciones de estados mentales similares cuando se atribuyen a otros. ¿Por qué existe esta asimetría entre las atribuciones de actitudes a nosotros mismos en el presente y las atribuciones de las mismas actitudes a los demás? ¿Qué es lo que da cuenta de la “autoridad que va unida a

las afirmaciones en la primera persona del presente y que se les niega a las afirmaciones de segunda y tercera personas?” (1984d/2001:3) ¿Qué concepción de la subjetividad pretende retener el planteo davidsoniano?

En “The Myth of the Subjective” dice conservar dos rasgos de la subjetividad tal como fue concebida de manera estándar, aunque resignificados del modo siguiente. Por un lado, insiste en que conserva *la asimetría entre la primera y la tercera persona*, pero en tanto inserta en su concepción interpretativa de la mente. Como hemos anticipado, si el intérprete no hace una presunción de verdad respecto de las autoatribuciones del hablante, no habría nada que interpretar. Por otro lado, subraya que retiene la *privacidad*: los pensamientos son privados en el sentido “obvio pero importante en que la propiedad puede ser privada, a saber, *pertenecen a una persona*” (1988/2001: 52).

En el presente capítulo, me voy a ocupar de su tratamiento de la dimensión subjetiva, esto es, de la perspectiva de primera persona, desde el marco provisto por la teoría de la interpretación radical. A partir de estas reflexiones, en la última sección, voy a ponderar cómo juegan las perspectivas consideradas hasta aquí, a saber, de primera y tercera persona, en un debate en el que Davidson no intervino expresamente, pero en el que clásicamente se lo interpreta de manera “teoricista”: el debate acerca de la atribución mental. Intentaré argumentar en contra de dicha interpretación, en una dirección que señalará allí donde Davidson, a mi juicio, no terminó de explorar: el rol de la *intersubjetividad* en su concepción “social e interactiva” de la mente. Según veremos, frente a las perspectivas de primera y tercera persona, en mi planteo cobrará una relevancia central la perspectiva de *segunda persona* de la atribución mental tal como se la introdujo recientemente en la filosofía de la mente.

### **6.1. La primera persona en el marco de la teoría de la interpretación radical: la autoridad de primera persona. Una “geografía del problema”.**

La primacía del punto de vista que hemos explorado hasta aquí, esto es, el de la tercera persona, a la vera del cual surge una concepción de la mente como un resultado del proceso de interpretación, no puede dar cuenta fácilmente del hecho de que, por lo que respecta a nosotros mismos, normalmente no nos atribuimos estados mentales a través de la interpretación de nuestra

propia conducta. En nuestro propio caso, los estados mentales parecen tener una realidad inmediata e independiente de su posible papel en el proceso de explicación e interpretación de la conducta y del lenguaje. En este hecho se ha fundado la concepción cartesiana de la mente, y los cartesianos actuales se basan en él para tratar de mostrar la insuficiencia de las concepciones contextuales o relacionales de la vida mental, entre las que se incluye paradigmáticamente el tratamiento interpretacionista de Davidson. A fin de cuentas, la concepción representacional de la mente, de raíz cartesiana, surgió, en parte, como un intento de explicar la certeza indudable que acompaña los enunciados en que un sujeto se atribuye a sí mismo, sinceramente, un estado mental. Davidson rechaza, a mi juicio con muy buenas razones, esta concepción representacional, pero nos resta ver si logra proporcionar una explicación alternativa satisfactoria de ese hecho fundamental acerca de la mente. Yo sé lo que deseo sin esperar a ver cómo actúo y sin tener en cuenta las relaciones de coherencia entre ese deseo y otros estados mentales o acciones. Pareciera que una teoría adecuada de la vida mental humana debe dar cuenta de ello.

En “First Person Authority”, Davidson (1984d) pretende explicar la autoridad de la primera persona como un supuesto esencial de la interpretación, es decir, como un aspecto necesario del conocimiento de tercera persona acerca de la mente de otros y del significado de sus palabras. Lo que debe aceptarse es que, en general, el hablante sabe lo que quiere decir con sus emisiones. Ciertamente, el proceso de interpretación pierde sentido si no se admite tal cosa. Por lo tanto, este supuesto debe ser respetado, y de él se sigue la autoridad de primera persona. Al mismo tiempo, el hablante, si desea ser entendido, ha de proporcionar a su oyente las claves necesarias, empleando coherentemente los sonidos ante objetos y situaciones que considera presentes tanto para él como para su oyente. Así, aunque el hablante no determina los contenidos de sus propios estados mentales observando las relaciones causales que median entre él y su entorno o las relaciones de coherencia que guardan con otros estados mentales, su autoridad sobre su propia mente deriva de y presupone el contexto público de la interpretación recíproca, por lo que no puede, en general, entrar en conflicto con él.

Antes de pasar a considerar la particular concepción davidsoniana de la autoridad de primera persona, me parece pertinente dar cuenta brevemente del marco general del debate en torno al autoconocimiento, para poder apreciar el carácter- un vez más-innovador de su propuesta.

Recordemos, en primer lugar, que el término “autoconocimiento” es usado comúnmente en filosofía para referir al conocimiento de los propios estados mentales particulares. Lo cierto es que, al menos desde Descartes, la mayoría de los filósofos ha creído que el autoconocimiento es diferente de un modo importante respecto de otros tipos de conocimiento: esto es, del mundo externo y de las otras mentes. Pero hay poco consenso acerca de qué es lo que distingue al autoconocimiento respecto del conocimiento en otros campos. ¿Qué es lo especial del autoconocimiento? Si se cree que hay algo especial en él - en el sentido de que podemos saber lo que pensamos, creemos, deseamos, y demás, de una manera distinta de cómo conocemos los estados mentales de los otros-<sup>170</sup> las posiciones que intentan dar cuenta de esta asimetría entre la primera y la tercera persona suelen dividirse, en la literatura filosófica, en dos grandes vertientes.

Por un lado, están aquellas posiciones que postulan un *acceso epistémico privilegiado* a nuestros propios estados mentales y, por otro, encontramos aquellas que consideran que el individuo posee una *autoridad especial de primera persona* con respecto a las autoadcripciones. Según la primera perspectiva, el acceso epistémico puede tomar la forma o bien de cierta seguridad epistémica (Descartes, 1641)<sup>171</sup> o bien de algún método especial<sup>172</sup> (como el “mirar hacia adentro”, Armstrong, 1981; Lycan, 1987; o el “mirar hacia afuera”, Evans, 1982; Dretske, 1994). A su vez, entre las posturas que sostienen que poseemos una autoridad de primera persona, la idea es que las emisiones acerca de los propios estados portan una autoridad especial o una presunción de verdad. Esta perspectiva se aleja de la ofrecida por los dos candidatos anteriores en tanto considera que lo que es especial de las autoatribuciones no es, en absoluto, un factor epistémico sino uno *pragmático*. Estrictamente hablando, esta posición parece que no se ocupa primariamente de lo que es especial acerca del auto-*conocimiento* sino, en cambio, del rasgo distintivo de las auto-*atribuciones* (se suele filiar en esta vertiente, *inter alia*, a Wittgenstein 1953; Moran 1997; Bilgrami 1998).

---

<sup>170</sup> Esto descarta posturas que consideran que no hay nada especial en el autoconocimiento porque conocemos nuestra propia mente del mismo modo que conocemos la mente de los demás, infiriendo los estados mentales a partir de la conducta observable. Según veremos en breve, Ryle (1949) es un conspicuo representante de esta línea.

<sup>171</sup> Es decir, se identifica el rasgo distintivo del autoconocimiento en términos del estatus epistémico especial que tendría una cierta clase de creencias.

<sup>172</sup> La idea es que lo que hace peculiar al autoconocimiento es el empleo de un método único, singular, para determinar los propios estados mentales.

Dado este marco, en mi opinión resulta clara la inscripción de Davidson en esta última alternativa, la opción pragmática. Lo que él intenta hacer en "First person authority" es legitimar la autoridad de primera persona, entendida a la manera de una presunción a favor del hablante. La asimetría que él tiene en mente es la siguiente: hay una presunción de que el hablante está en lo correcto cuando se atribuye a sí mismo -a su yo presente- una creencia, deseo o intención, presunción que no resulta apropiada cuando otros le hacen atribuciones similares. Y Davidson entiende que su contribución es explicar dicha asimetría en términos del proceso de interpretación. Así, mediante un argumento trascendental concibe la autoridad de primera persona como un rasgo esencial de dicho proceso, y sostiene que a menos que supongamos que el hablante sabe lo que quiere decir, el intérprete no tiene nada que interpretar. Veamos, entonces, en qué consiste su propuesta<sup>173</sup>.

¿Qué es lo que da cuenta de la autoridad que va unida a las afirmaciones en la primera persona del presente y que se les niega a las afirmaciones de segunda y de tercera personas? Davidson sostiene que la misma pregunta se puede formular tanto para el lenguaje como para la epistemología, y da por supuesto que si se logra explicar la autoridad de primera persona en el caso del hablante, habremos hecho mucho - si no es que todo- lo que necesitamos para caracterizar una teoría epistemológica (1984d).

A su juicio, la conexión entre el problema de la autoridad de primera persona y el problema tradicional de las otras mentes es evidente, pero tal y como plantea el primer problema existen las siguientes diferencias. La autoridad de primera persona es un problema más restringido, puesto que lo tratará sólo en tanto aplicado a actitudes proposicionales. No analizará lo que, a menudo, se considera que es central en el problema de las otras mentes, a saber: los dolores, las sensaciones, y otros estados psicológicos dirigidos a objetos no-proposicionales. A su juicio, lo que se mantenga para las actitudes proposicionales ha de ser relevante para las sensaciones y lo restante, pero no se ocupará de explorar dicha conexión.

Todas las actitudes proposicionales exhiben, en opinión de Davidson, el rasgo de la autoridad de primera persona, pero en diversos grados y clases. La creencia y el deseo son ejemplos relativamente claros y simples mientras que la intención, la percepción, la memoria y el conocimiento son, de una manera u otra, más complejos. Así pues, al evaluar la afirmación de

---

<sup>173</sup> Para ello voy a seguir sus lineamientos centrales en "First Person Authority" (1984d) y "Knowing one's own mind" (1987b).

alguien que ha notado que la casa está ardiendo hay al menos tres cosas a considerar: si la casa está ardiendo, si el hablante cree que la casa está ardiendo y cómo el fuego causó esa creencia. Con respecto a lo primero, el hablante no tiene una autoridad especial; con respecto a lo segundo sí que la tiene; y con respecto a lo tercero, la responsabilidad es mixta y compleja. La autoridad especial se encuentra directamente en afirmaciones acerca del deseo y la creencia y, menos directamente, en afirmaciones acerca de la conexión causal necesaria. Estas diferencias entre las maneras en que la autoridad de primera persona puede aplicarse a las actitudes proposicionales son importantes y merecen ser exploradas. En cada caso, la autoridad de primera persona es relevante y lo que considerará Davidson es el caso general, a saber: se concentrará, un vez más, en el caso de la *creencia*<sup>174</sup>.

Vale aclarar que aunque haya autoridad de primera persona respecto a las creencias y otras actitudes proposicionales, el *error es posible*; y esto es una consecuencia del hecho de que las actitudes son disposiciones que se manifiestan de diversas maneras y en intervalos de tiempo. El error es posible como también lo es la duda. Así pues, no siempre tenemos un conocimiento indubitable o cierto de nuestras propias actitudes. Ni *tampoco* son *incorregibles* nuestras afirmaciones sobre ellas. Es posible que otros accedan a evidencias que les permitan refutar los juicios sobre uno mismo.

Nos acercamos, dice Davidson, a una caracterización de la autoridad de primera persona al hacer notar que la persona que se autoatribuye algo no basa normalmente sus afirmaciones en evidencias o en la observación, ni tampoco normalmente tiene sentido preguntarle a la persona que se autoatribuye algo por qué cree que tiene las creencias, los deseos o las intenciones que afirma que tiene<sup>175</sup>. Wittgenstein señaló este rasgo de las autoatribuciones: “¿Cuál es el criterio

---

<sup>174</sup> Puesto que en todos los ejemplos que considera la autoridad de primera persona recae –al menos parcialmente– en un componente de creencia.

<sup>175</sup> En “Knowing one’s own mind” Davidson (1987b) insiste en que no hay ningún secreto sobre la naturaleza de las evidencias que utilizamos para decidir lo que otras personas piensan. Observamos sus actos, leemos los emails que nos envían, estudiamos su expresión, escuchamos sus palabras, nos enteramos de sus vidas y observamos sus relaciones con la sociedad. Otra cosa es la manera en que somos capaces de ensamblar todo este material para formarnos una imagen convincente de una mente; *sabemos cómo hacerlo sin saber necesariamente cómo lo hacemos*. Davidson dice: “A veces me entero de lo que creo de la misma manera en que lo hacen los demás, es decir, observando lo que digo y hago. Puede haber veces en que ésta sea la única manera de acceder a mis propios pensamientos. Según Graham Wallas, “Aquella niña tenía madera de poeta, pues al decirle que se asegurara de lo que quería decir antes de hablar dijo: “¿Cómo puedo saber lo que pienso hasta que no veo lo que digo?” Robert Motherwell expresó un pensamiento similar: “Diría que los mejores pintores no saben lo que piensan hasta que lo pintan”. Como es sabido, Gilbert Ryle estaba completamente de acuerdo con el poeta y el pintor en este asunto; mantuvo tenazmente que conocemos nuestras propias mentes exactamente de la misma manera en que conocemos las mentes de los demás, es decir, observando lo que decimos, hacemos o pintamos. Según veremos a continuación,

para atribuir rojez a una imagen? Para mí, cuando es la imagen de otro la que se considera: lo que dice o hace. Para mí mismo, cuando es mi propia imagen la que considero: nada (*Philosophical Investigations*, # 377) La mayoría de los filósofos ha seguido a Wittgenstein en esto y, tal como veremos, han extendido este criterio a las actitudes proposicionales. En suma: las autoatribuciones no se basan en evidencias. Veamos cuál es la dificultad que Davidson encuentra aquí.

Normalmente se sostiene que lo que se consideran evidencias para la aplicación de un concepto ayuda a definir el concepto, o al menos restringe su individuación. Si dos conceptos dependen regularmente para su aplicación de criterios distintos o del apoyo de diferentes tipos de evidencias, tales conceptos deben ser diferentes. De modo que si lo que aparentemente es la misma expresión, a veces se utiliza correctamente sobre la base de un cierto tipo de evidencias y a veces sobre la base de otro tipo de evidencias (o ninguno), la conclusión obvia parece que sería que la expresión es ambigua. Si es ambigua, entonces no hay razón para suponer que tiene el mismo significado cuando se aplica a uno mismo que el que tiene cuando se aplica a otra persona. Si aceptamos que el carácter necesariamente público e interpersonal del lenguaje garantiza que, de manera frecuente, aplicamos estos predicados a los demás y que, por tanto, frecuentemente sabemos lo que *los demás* piensan, entonces la cuestión que se debe plantear es la de qué fundamento tiene cada uno de nosotros para pensar que sabe –en el mismo sentido- lo que él mismo piensa. En suma, a juicio de Davidson, la respuesta de estilo wittgensteiniano

---

Ryle, a juicio de Davidson, estaba equivocado. Rara vez sucede que necesitemos o apelemos a evidencias o a la observación para averiguar lo que creemos; normalmente sé lo que creo antes de hablar o actuar. Incluso cuando tengo evidencias, rara vez hago uso de ellas. Puedo estar equivocada acerca de mis propios pensamientos, por lo que no es irrelevante recurrir a lo que se puede determinar públicamente. Pero la posibilidad de que se pueda estar equivocada sobre sus propios pensamientos no puede acabar con la suposición dominante de que una persona sabe lo que cree; en general, la creencia de que tenemos un pensamiento es suficiente para justificar esa creencia. Pero aunque esto sea verdad e incluso obvio para la mayoría de nosotros, este hecho, tal y como Davidson lo ve, no tiene una explicación fácil. Mientras que está bastante claro, al menos a grandes rasgos, lo que tenemos que hacer para intentar desentrañar cuáles son los pensamientos de los demás, no está nada claro por qué, por lo que respecta a nosotros mismos, podemos saber tan frecuentemente lo que pensamos sin apelar a evidencias o sin recurrir a la observación.

Dado que usualmente sabemos lo que creemos sin necesidad de recurrir a las evidencias –incluso cuando están a nuestro alcance- nuestras declaraciones sinceras respecto de nuestros estados mentales presentes no están sujetas a los defectos de las conclusiones que se basan en evidencias. Así pues, las afirmaciones sinceras sobre pensamientos en primera persona del presente, aunque no son infalibles ni incorregibles, tienen una autoridad que no puede tener ninguna afirmación ni en segunda ni en tercera. No obstante, reconocer este hecho, como veremos, no significa explicarlo.

puede solucionar el *problema de las otras mentes*, pero crea el concomitante *problema sobre el conocimiento de nuestra propia mente*. No obstante, la correspondencia no es completa.

El problema original de las otras mentes provocaba la pregunta de cómo sabe uno que los demás tienen mente. El problema al que ahora nos enfrenta Davidson se puede especificar del modo siguiente: sé dónde buscar para atribuir pensamientos a los demás. Me aplico a mí misma los mismos predicados usando criterios totalmente distintos (o ninguno); de este modo, la pregunta escéptica que se plantea es: por qué debería pensar que son *pensamientos* los que me estoy atribuyendo a mí misma.

Así pues, el rasgo de la autoridad de primera persona que Wittgenstein pone en el centro de la escena, aunque puede ser sugerente, no ayuda a explicar esa autoridad. La razón principal es, simplemente, que las afirmaciones que no se basan en evidencias no tienen, en general, más autoridad que las afirmaciones que están basadas en evidencias, ni son probablemente correctas. En opinión de Davidson, los filósofos contemporáneos que han tratado la cuestión de la autoridad de la primera persona han hecho escasos intentos para contestar la pregunta de por qué las autoatribuciones resultan privilegiadas. Hace tiempo que ha pasado de moda, dice, explicar el autoconocimiento sobre la base de la introspección y es fácil ver el por qué, ya que esta explicación conduce sólo a la pregunta de por qué habríamos de ver mejor cuando inspeccionamos nuestras propias mentes que cuando inspeccionamos las mentes de los demás.

## **6.2. La autoridad de primera persona y el externismo davidsoniano**

A juicio de Davidson, pocos filósofos han negado la existencia de la asimetría en cuestión. Ryle (1949) es un tenaz ejemplo de la negación. En *El concepto de lo mental*, afirma que lo que tomamos por “acceso privilegiado” no se debe más que al hecho de que generalmente estamos en mejor posición de observarnos a nosotros mismos de lo que lo están los demás. Escribe Ryle: “en principio, a diferencia de la práctica, las maneras que tiene John Doe de saber cosas sobre John Doe son las mismas maneras que tiene John Doe de saber algo acerca de Richard Doe”. Y prosigue así:

Las diferencias son diferencias de grado, no de tipo. La superioridad del conocimiento que posee el hablante de lo que él hace respecto al oyente no indica que el primero tenga un Acceso Privilegiado a hechos que sean de un tipo que los haga inevitablemente inaccesibles al oyente, sino solamente que está en una muy buena posición para saber aquello que el oyente frecuentemente está en muy mala posición para saber. (Ryle 1949: 179)

Davidson está de acuerdo con Ryle en que cualquier intento de explicar la asimetría entre las afirmaciones de primera persona del presente acerca de actitudes y las afirmaciones de otras personas y otros tiempos haciendo referencia a una manera especial de *conocer* o a una *clase especial de conocimiento* conducen por fuerza al escepticismo.

En cierto momento, Ayer (1963) tomó una ruta similar a la de Ryle. En *The Concept of a Person* enfatiza que las adscripciones de primera persona pueden ser erróneas y admite que estas adscripciones son privilegiadas (Ayer 1963: "Privacy"). No obstante, cuando quiere describir la autoridad de las autoadscripciones, las compara con la autoridad que le otorgamos a un testigo presencial cuando se lo compara con los informes de segunda mano. Esta analogía resulta insatisfactoria para Davidson por dos razones. La primera es que no nos dice por qué una persona es como un testigo presencial respecto de sus propios estados y eventos mentales, mientras que los otros no lo son. La segunda es que no ofrece una descripción detallada de qué es la autoridad de primera persona, ya que las atribuciones de primera persona no se basan en evidencias mejores, sino que frecuentemente no se basan en ningún tipo de evidencias. La autoridad de los testigos presenciales, en el mejor de los casos, se basa en probabilidades inductivas que fácilmente se ven superadas en determinados casos: se desacredita a un testigo presencial y las evidencias que ofrece no se toman en cuenta si es un observador que se destaca por ser poco fiable, con prejuicios o miope. Sin embargo, una persona nunca pierde su derecho a estar en lo cierto respecto a sus propias actitudes, incluso cuando su afirmación se pone en cuestión o resulta refutada.

Joseph Agassi ha mantenido que conocemos los estados y eventos mentales de otras mentes mejor que los de las nuestras. Agassi distingue entre el acceso privilegiado y la verdad de sentido común según la cual "cada persona accede a alguna información disponible solamente a esa persona y esto involucra a uno mismo, al menos como un testigo presencial". Continúa:

La doctrina del acceso privilegiado es que yo soy la autoridad para todas mis propias experiencias...La tesis fue refutada por Freud (conozco tus sueños mejor que tú), Duhem (conozco tus métodos de descubrimiento científico mejor que tú), Malinowski (conozco tus costumbres y hábitos mejor que tú) y los teóricos de la percepción (puedo hacer que veas cosas que no están ahí y describir tus percepciones mejor de lo que puedes hacerlo tú) (Agassi, 1975: 120).

Con excepción del caso de Freud, hay poco en este texto que amenace la autoridad de primera persona, en opinión de Davidson. Freud al extender los conceptos de intención, creencia y deseo para incluir el inconsciente, implica que con respecto a algunas actitudes proposicionales una persona pierde la autoridad directa. De hecho, la pérdida de autoridad es el rasgo principal que caracteriza a los estados mentales inconscientes. Por supuesto, las actitudes prefreudianas siguen estando sujetas a la autoridad de primera persona. Lo que es más interesante es el hecho de que en la práctica psicoanalítica, la recuperación de la autoridad sobre una actitud se considera frecuentemente como la única evidencia sólida de que la actitud estaba allí antes de ser apreciada no –inferencialmente por su poseedor. Por ello, los casos de estados mentales inconscientes que de un modo no sistemático se reconocía que existían antes de Freud son incluidos indirectamente por el psicoanálisis dentro del ámbito de la autoridad de primera persona. Así pues, no cree que la existencia de las actitudes inconscientes amenace la importancia de la autoridad de primera persona.

A continuación, Davidson considera filósofos que han supuesto que existe la autoridad de primera persona y han aceptado la descripción wittgensteiniana de la diferencia entre las atribuciones de primera y de tercera persona.

Strawson analiza la autoridad de primera persona en el contexto de intentar responder al *escepticismo sobre otras mentes*. De acuerdo con Strawson, si el escéptico comprende su propia pregunta (“¿Cómo alguien conoce lo que pasa en la mente de otra persona?”), sabe la respuesta, puesto que si el escéptico conoce lo que es una mente, sabe que debe estar en un cuerpo que tiene pensamientos. También sabe que atribuimos pensamientos a los demás sobre la base del comportamiento observado, pero lo hacemos a nosotros mismos sin una base tal. Dice Strawson:

Para *tener* este tipo de concepto [el de una propiedad mental], debemos ser tanto un autoadscriptor como un adscriptor-a-otros de tales predicados [los que adscriben propiedades mentales], y hemos de ver a los demás como autoadscriptores. Para *comprender* este tipo de concepto, debemos darnos cuenta de que hay una clase de predicado

que es adscribible de manera adecuada y no ambigua *tanto* sobre la base de observaciones de los sujetos de tales predicados *como* sin esa base, es decir, independientemente de las observaciones del sujeto (Strawson1959:108)<sup>176</sup>

Para Davidson, ésta no es una respuesta satisfactoria para el escéptico, puesto que el escéptico podría replicar que aunque Strawson puede haber descrito correctamente la asimetría entre las adscripciones de predicados mentales en primera persona y en las otras personas, no ha hecho nada para explicarla. En ausencia de explicación, el escéptico seguramente está justificado, dice Davidson, en preguntar cómo sabemos que la descripción es correcta. En particular ¿por qué habríamos de pensar que no es ambiguo un predicado que a veces se aplica sobre la base de la observación y a veces no? Esta cuestión, que Strawson no ha tratado, constituye una fuente importante del escepticismo respecto a las otras mentes.

Por su parte, Richard Rorty ha intentado ofrecer una explicación tal. Se nos pide que imaginemos que, originalmente, las autoadscripciones se hicieron sobre la base de la misma clase de observaciones o evidencias de comportamiento que las adscripciones a otros. Se cayó entonces en la cuenta de que las personas podían adscribirse propiedades mentales sin hacer observaciones y sin usar evidencias de comportamiento y que, a la larga, resultaba que las autoadscripciones ofrecían mejores explicaciones del comportamiento que las adscripciones de tercera persona. De ese modo, tratar las autoadscripciones como privilegiadas se convirtió en una convención lingüística: “se convirtió en una condición limitativa de las explicaciones de la conducta el que deberían encajar todos los pensamientos o sensaciones sobre los que se informaba en la explicación global ofrecida” (Rorty 1970:416).<sup>177</sup> Esta teoría, dice Davidson, no pretende que se la tome seriamente como un fragmento de antropología de sentido común, sino que lo que se pretende es que parezca razonable tratar las autoadscripciones como algo que tiene una autoridad especial. Pero la cuestión sigue ahí: ¿cuál es la razón que ofrece Rorty para mostrar que las autoadscripciones que no se basan en evidencias tienen que ver con los mismos estados y eventos que las adscripciones de los mismos predicados mentales que están basadas en observaciones o evidencias? Hay una diferencia claramente discernible en los modos en que se hacen los dos tipos de adscripciones, y la manera en que explican el comportamiento también es

---

<sup>176</sup> Anita Avramides (1999) ha señalado que Strawson ha contribuido más a esta problemática de lo que Davidson concede.

<sup>177</sup> Tal como señala Davidson, la teoría de Rorty se deriva, a su vez, de la de Sellars, W. (1956).

diferente. Lo que Rorty describe como el descubrimiento de que las autoadscripciones que no se basan en evidencias explican mejor el comportamiento, el escéptico lo describe como el hecho de que lo que se adscribe es, en cada caso, según todas las apariencias, distinto.

Puede que sea sorprendente el darse cuenta de que los filósofos que ha considerado, en realidad, a juicio de Davidson, no han tratado el viejo problema del escepticismo respecto al conocimiento de las otras mentes; no obstante, él cree que esto tiene una explicación fácil. Históricamente, el problema se ha enfocado desde los puntos de vista cartesiano o empirista y ambos admiten que cada persona conoce lo que está en su propia mente. El problema ha parecido ser, entonces, el de ofrecer una base para el conocimiento de las otras mentes (y, por supuesto, del mundo externo). Ahora, los filósofos se han dado cuenta de que parte de nuestra comprensión de los conceptos (o predicados) mentales consiste en saber qué tipo de comportamiento observable justifica la adscripción de estos conceptos a los demás. Pero este tipo de respuesta al escéptico, no explica la autoridad de primera persona o la asimetría entre las autoadscripciones y las adscripciones a otros. Aún podemos preguntar por qué creemos que estos dos tipos de adscripciones conciernen a lo mismo. Esta es una buena pregunta, reconozcamos o no su tradicional estirpe escéptica.

Quizás, debiera señalarse la siguiente peculiaridad acerca de los conceptos mentales: pese a que muchos conceptos pueden aplicarse sobre la base de múltiples criterios, los conceptos mentales son los únicos con respecto a los cuales quienes realizan la adscripción *deben*, en determinadas ocasiones, utilizar criterios distintos (1984d/2001: 33). Si queremos explicar esta anomalía y a la vez evitar el escepticismo, la explicación debe señalar una asimetría natural entre los otros observadores y nosotros mismos, una asimetría que no haya sido simplemente inventada para resolver el problema.

A juicio de Davidson, el primer paso para la solución depende de poner en claro *cuáles son las entidades* a las que se aplica la autoridad de primera persona. Pasa, entonces, a una formulación de Sidney Shoemaker que hace una mención explícita del lenguaje: “Entre los enunciados incorregibles encontramos enunciados acerca de...eventos mentales, por ejemplo...informes sobre pensamientos...Tales enunciados son incorregibles en el sentido de que si una persona asevera sinceramente tal enunciado no tiene sentido suponer que está equivocada, esto es, que lo que él dice es falso, y no podría aceptarse nada que mostrara tal cosa” (Shoemaker 1963: 215-16).

Davidson deja a un lado la condición de incorregibilidad y la sustituye por algo menos exigente –algo que, nos dice, equivalga a la autoridad de primera persona-. Lo que es importante para los fines de Davidson es que Shoemaker asigna la presunción de corrección no a una *clase de conocimiento* sino a una *clase de preferencias*. Esta idea podría conducirnos a una explicación de la autoridad de primera persona si la clase de esas preferencias pudiera especificarse en términos sintácticos. Desafortunadamente, esto no puede hacerse. No se ha ofrecido todavía una explicación satisfactoria de la asimetría entre las atribuciones en primera persona y las atribuciones a los otros de actitudes, aunque centrarse en las oraciones y preferencias, más que en proposiciones o en los significados, es un paso importante hacia una *nueva dirección* promisoria. La razón de ello es la siguiente. Mientras planteemos el problema en términos de la clase de justificación o autoridad que alguien tiene con respecto a las afirmaciones acerca de la actitud de un agente hacia una proposición, estamos constreñidos a dar cuenta de las diferencias simplemente postulando diferentes tipos de fuentes de información. Pero esta vía no hace más que volver a plantear el problema y, con ello, se invita al escepticismo acerca del conocimiento de las mentes de los demás (o de nuestra propia mente). Sin embargo, a juicio de Davidson, si planteamos el problema en términos de las relaciones entre agentes y preferencias podemos evitar el punto muerto.

En esta instancia, resulta necesario distinguir dos asimetrías diferentes pero relacionadas. Por un lado, encontramos la diferencia que nos resulta familiar entre autoatribuciones y atribuciones a otros de la misma actitud a la misma persona: el sostener por mi parte que yo creo que Wagner murió feliz y el sostener por su parte que yo creo que Wagner murió feliz. Por otro lado, podemos considerar mi preferencia de la oración “creo que Wagner murió feliz” y después contrastar la justificación que tengo para pensar que he dicho algo verdadero, con la justificación que usted tiene para pensar que yo he dicho algo verdadero.

Por razones que en breve se harán evidentes, Davidson considera que debemos tratar la segunda versión de la asimetría. La cuestión se convierte, pues, en la siguiente: ¿qué es lo que explica la diferencia en la clase de seguridad que usted tiene en que estoy en lo cierto cuando digo “creo que Wagner murió feliz” y la clase de seguridad que yo misma tengo? Por ahora, sabemos que no nos sirve de ayuda decir que yo tengo acceso a una manera de conocer mis propias creencias que usted no tiene, ni tampoco que usamos distintos criterios al aplicar el concepto de creencia.

Los supuestos que tenemos hasta aquí son los siguientes. Usted y yo sabemos que sostenía que la oración “Wagner murió feliz” era una oración verdadera cuando la he proferido; y que sabía lo que la oración significaba cuando esta emisión tuvo lugar. Así pues, es esta diferencia entre nosotros lo que tiene que explicarse: según estos supuestos, sé lo que creo, mientras que usted puede que no. La diferencia se sigue del hecho de que el supuesto de que sé lo que quiero decir necesariamente me da a mí, pero no a usted, conocimiento de la creencia que he expresado mediante mi proferencia. Todavía queda por mostrarse por qué debe existir la presunción de que los hablantes, pero no sus intérpretes, no están equivocados acerca de lo que significan sus palabras. Y esto es así porque esa presunción es esencial a la naturaleza de la interpretación. Planteando la cuestión en su versión más simple: no puede haber ninguna garantía de carácter general de que el oyente esté interpretando correctamente a un hablante; no importa lo fácil, automática, irreflexiva y exitosamente que un oyente entienda a un hablante, ese oyente bien puede cometer un serio error. En este sentido especial, siempre se puede considerar que interpreta al hablante. El hablante no puede, en el mismo sentido, interpretar sus propias palabras.

Un oyente interpreta –normalmente, sin pensarlo o hacer una pausa- sobre la base de muchas pistas: las acciones y otras palabras del hablante, lo que supone acerca de la educación, lugar de nacimiento, inteligencia y profesión del hablante, la relación del hablante con los objetos que se encuentran cerca y lejos, y así sucesivamente. El hablante, aunque debe tener muchas de estas cosas en mente cuando habla, en la medida en que depende de él intentar ser comprendido, no puede estar considerando si, de manera general, lo que dice se corresponde con lo que quiere decir. Desde luego, el hablante puede estar equivocado acerca de lo que sus propias palabras significan. Ésta es una de las razones por las que la autoridad de primera persona no es infalible. Sin embargo, la posibilidad de error no elimina la asimetría. La asimetría descansa en el hecho de que al interpretar al hablante, el intérprete debe apoyarse en lo que, si se hiciera explícito, sería una *inferencia difícil*, mientras que el hablante no.

Ni el hablante ni el oyente conocen de una manera especial o misteriosa lo que significan las palabras del hablante y ambos pueden estar equivocados. Pero hay una diferencia. El hablante, después de aplicar cualquier conocimiento y habilidad que tenga a la tarea de decir lo que sus palabras significan, no puede mejorar el siguiente tipo de formulación: “Mi emisión de la oración “Wagner murió feliz” es verdadera si y sólo si Wagner murió feliz”. Un intérprete no

tiene ninguna razón para suponer que ésta será *su* mejor manera de enunciar las condiciones de verdad de la preferencia del hablante. La mejor manera de apreciar esta diferencia es volviendo sobre la teoría de la interpretación radical que presentamos en el capítulo 2, esto es, imaginando una situación en la que a dos personas que hablan lenguas no emparentadas se las deja solas para que aprendan a comunicarse entre sí. Descifrar una nueva lengua no es como aprender una primera lengua, puesto que un genuino principiante no tiene el poder de razonamiento ni el surtido de conceptos que los participantes de la situación imaginada tienen a su disposición. Imaginemos ahora que una de las personas del par habla mientras que la otra trata de comprenderlo. Lo mejor que puede hacer el hablante es ser *interpretable*, esto es, usar una cantidad finita de sonidos distinguibles aplicados consistentemente a objetos y situaciones que crea que resultan evidentes para su oyente. Obviamente, el hablante puede fracasar en este proyecto de vez en cuando; en ese caso podemos decir, si queremos, que no sabe lo que significan sus palabras. Es igualmente obvio que el intérprete no tiene nada más en que apoyarse que la configuración de sonidos que el hablante muestra en conjunción con otros eventos (incluyendo acciones posteriores por parte de ambos). No tiene sentido que en esta situación nos preguntemos si el hablante está haciendo las cosas mal. Su comportamiento puede muy bien no ser simplemente interpretable, pero si lo es, entonces, lo que significan sus palabras – generalmente- es lo que pretende que ellas signifiquen. Se presume –presunción que inevitablemente forma parte de la naturaleza de la interpretación- que el hablante sabe usualmente lo que quiere decir. Así pues, se presume que si sabe que piensa que una oración es verdadera, sabe lo que cree.

La idea que subyace, pues, al planteo davidsoniano es que si logramos abandonar una imagen defectuosa de la mente, la autoridad de primera persona dejará de ser un problema; incluso puede suceder que la autoridad de primera persona *dependa de y se explique por*, los factores *sociales y públicos* que se suponía que socavaban esa autoridad. De este modo, Davidson pretende que en su planteo haya lugar para el reconocimiento de la autoridad de primera persona y, a la vez, se diluya el problema acerca del conocimiento de las otras mentes.

En efecto, considera que dicho problema también se transforma cuando reconocemos que las creencias, los deseos y demás actitudes proposicionales “no son subjetivas del modo en que pensábamos que lo eran. El problema es cómo una persona conoce lo que está en la mente de otra”. Sostiene que quizás es obvio que si la explicación que bosquejé acerca de nuestra

comprensión del lenguaje y su conexión con los contenidos del pensamiento es correcta, “la accesibilidad a las mentes de otros está asegurada desde el comienzo. El escepticismo acerca de la posibilidad de conocer otras mentes es así excluido” (1988/2001: 46).

De esta suerte, aunque el último tribunal de apelación sea personal, no se sigue que nuestros juicios sean arbitrarios o subjetivos, dado que:

fueron formados en un *nexo social* que asegura la objetividad, si no la corrección de mis creencias. La *intersubjetividad* es la raíz de la objetividad, no porque aquello sobre lo que la gente acuerde es necesariamente verdadero, sino porque la intersubjetividad depende de la interacción con el mundo (1998b/2001: 91; mis itálicas).

Como hemos adelantado, a juicio de Davidson, hay una imagen de la mente que ha llegado a estar tan profundamente arraigada en nuestra tradición filosófica que es casi imposible escapar a su influencia, incluso cuando se reconocen y se rechazan sus defectos. Según vimos, una versión muy tosca pero familiar es la siguiente: la mente es un teatro en el que el yo consciente contempla escenas (las sombras en la pared). Las escenas constan de “apariencias”, datos sensoriales, *qualia*, lo que la experiencia “da”. Lo que aparece en el escenario no son los objetos comunes del mundo que “el ojo externo registra y el corazón ama”, sino sus pretendidos representantes. Cualquier cosa que sepamos sobre el mundo externo depende de lo que recojamos a través de pistas internas.

Desde el principio, la dificultad manifiesta que tiene esta descripción de lo mental es ver cómo es posible trazar una trayectoria fiable desde lo interior a lo exterior. Otra dificultad destacable, aunque quizás haya sido menos apreciada, es la de dónde ubicar el yo en la imagen, puesto que el yo parece que, por una parte, incluye el teatro, el escenario, los actores y la audiencia, pero, por otra, lo que se conoce y se registra pertenece sólo a la audiencia<sup>178</sup>. Este segundo problema podría plantearse también como el de la *ubicación de los objetos de la mente*: ¿están *en* la mente o, simplemente se ven *mediante* ella?

En “Knowing one’s own mind”, Davidson (1987b) vuelve sobre esta cuestión de la que ya nos ocupamos, especialmente, en el capítulo 4 al considerar la analogía que propone entre las

---

<sup>178</sup> Cabe recordar que Davidson defiende una tesis de irreducibilidad del concepto de yo. En efecto en “Irreducibility of the Concept of the Self”, sostiene que el pronombre personal “yo” nos permite expresar un conocimiento que no puede ser expresado de ninguna otra manera. Claramente hay algo “irreducible, irremplazable, encarnado en el uso de los pronombres de primera persona...No hay sustituto para esta manera de ubicarme en el espacio público” (1998b/2001: 86).

actitudes proposicionales y las actitudes numéricas. El interés que la cuestión reviste ahora, para nosotros, es que mediante una nueva consideración de la misma, Davidson intenta conciliar la *autoridad de primera persona* con uno de los rasgos centrales de la noción de creencia, que tematizamos en el capítulo 3: su peculiar externismo, a saber, la *triangulación*.

En efecto, en “Knowing one’s own mind” Davidson afirma que si adherimos al externismo y reconocemos, a la vez, la autoridad de primera persona, estamos ante un auténtico enigma. La dificultad surge porque el admitir que los factores externos desempeñan un rol de identificación e individuación del contenido parece llevar a la conclusión de que nuestros pensamientos podrían no ser conocidos por nosotros mismos. Desde la perspectiva de Davidson, quienes aceptan la tesis del externismo parecen tener un problema similar al del escéptico, pero *desplazado al conocimiento de nuestras propias mentes*: dado que la relación externa no está determinada subjetivamente, el sujeto no tendría autoridad acerca de ella; una tercera persona bien podría estar en una mejor posición que el sujeto en cuestión para saber en qué objeto está pensando.

Le resulta inquietante que quienes suscriben el externismo no se hayan ocupado de esta consecuencia aparente de su punto de vista. Pero, ¿el externismo socava, efectivamente, la autoridad de primera persona?

La respuesta parece depender, para Davidson, del tipo de externismo que se abrace, del modo en que se piense que la identificación de contenidos mentales depende de factores externos. A su juicio, la consecuencia en cuestión se sigue para cualquier teoría que sostenga que las actitudes proposicionales son identificadas por objetos que están “ante” la mente y que contengan o incorporen como “ingredientes” objetos o eventos fuera del sujeto. De esta suerte, la estrategia de Davidson para tornar compatible el externismo con “nuestra fuerte intuición de que la autoridad de primera persona existe” se articula del modo siguiente. Por un lado, va a desmontar la pretendida base para la inferencia: esto es, va a criticar los supuestos que subyacen al externismo estándar para mostrar que es posible defender una versión alternativa que no comprometa necesariamente la autoridad de primera persona. Y por otro lado, va a ofrecer una explicación de la autoridad de primera persona sobre la base de su peculiar propuesta externista.

Davidson no se ocupa aquí de objetos de la mente como los datos señoriales, sino de sus “primos enjuiciables”, los presuntos objetos de las actitudes proposicionales, ya se conciban como proposiciones, ejemplares de proposiciones, representaciones o fragmentos de “mentales”.

La idea central que va a atacar es que éstas son entidades que la mente pueda “abrigar”, “asir”, “tener ante ella”, o “estar en contacto con”.

Una vez más, afirma que es fácil ver cómo el descubrimiento de que los hechos externos contribuyen a la individuación de los estados de la mente altera la imagen de la mente que ha estado describiendo, ya que si estar en un estado mental consiste en que la mente esté en alguna relación con un objeto, entonces cualquier cosa que ayude a determinar qué objeto es ése debe ser igualmente captada si es que la mente conoce en qué estado se encuentra. Esto es particularmente evidente si un objeto externo es un “ingrediente” del objeto que se encuentra ante la mente. Pero en cualquier caso, la persona que está en ese estado mental puede ser que no conozca en qué estado mental está.

Es este el punto en que *el concepto de lo subjetivo* - de un estado mental- comienza a desintegrarse. Por una parte, tenemos a los verdaderos estados internos con respecto a los cuales la mente conserva su autoridad; por otra, tenemos los estados usuales de creencia, deseo, intención y significado que están contaminados por sus conexiones necesarias con el mundo social y público.

Una vez más, la solución davidsoniana consiste en deshacerse de la *metáfora de los objetos ante la mente*. La mayoría de nosotros, dice Davidson, abandonamos hace tiempo la idea de que las percepciones, los datos sensoriales o el flujo de la experiencia son cosas “dadas” a la mente; deberíamos hacer lo mismo con los objetos proposicionales. Es evidente que las personas tienen creencias, deseos, dudas, etc., pero admitir esto no significa sugerir que las creencias, los deseos, las dudas son *entidades* en o ante la mente, o que estar en tales estados requiera que haya los objetos mentales correspondientes.

Davidson hace explícito que los escrúpulos ontológicos no forman parte de lo que le interesa. Siempre necesitaremos una fuente infinita de objetos que nos ayuden a describir e identificar actitudes como la creencia; no está sugiriendo ni por un momento que las oraciones de creencia y las oraciones que atribuyen otras actitudes no sean de naturaleza relacional. Ya hemos visto en el capítulo 4 que lo que está diciendo es que los objetos con los que relacionamos a las personas para describir sus actitudes *no necesitan ser objetos psicológicos* en ningún sentido, que puedan ser captados, conocidos o abrigados por la persona cuyas actitudes se están describiendo. Esto es algo conocido.

El origen del problema es el dogma de que tener un pensamiento es tener un objeto ante la mente. Putnam<sup>179</sup> y Fodor –y muchos otros- han distinguido dos clases de objetos: aquellos que son verdaderamente internos y entonces están “ante la mente” o ella “los capta”, y aquellos que identifican el pensamiento a la manera usual. Davidson está de acuerdo en que no hay objetos que puedan servir a ambos propósitos. Putnam piensa que la dificultad proviene del hecho de que un objeto parcialmente identificado en términos de relaciones externas no puede considerarse que coincide con un objeto ante la mente porque la mente puede ignorar la relación externa en cuestión. Quizás esto sea así, pero de ello no se sigue que podamos encontrar *otros* objetos que nos aseguraran la deseada coincidencia, puesto que si el objeto *no está* conectado con el mundo, nunca podemos llegar a saber cosas sobre el mundo teniendo a ese objeto ante la mente; y por razones recíprocas, sería imposible detectar tal pensamiento en otro. En consecuencia, parece que lo que está ante la mente no puede incluir las conexiones externas –su semántica-. Por otro lado, si el objeto *está* conectado con el mundo, entonces no puede estar enteramente ante la mente en el sentido relevante. Y, con todo, a menos que un objeto *semántico* pueda estar ante la mente *en su aspecto semántico*, el pensamiento, concebido en términos de tales objetos, no puede escapar al destino de los datos sensoriales.

La dificultad básica es simple: si tener un pensamiento es tener un objeto ante la mente, y la identidad del objeto determina de qué pensamiento se trata, entonces siempre debe ser posible estar equivocado acerca de lo que uno está pensando, puesto que a menos que uno lo sepa *todo* sobre ese objeto, siempre habrá sentidos en que uno no sabe de qué objeto se trata. Se han hecho muchos intentos de encontrar una relación entre una persona y un objeto que se dé en todos los contextos si y solamente si puede decirse intuitivamente que la persona sabe de qué objeto se trata. Sin embargo, ninguno de esos intentos ha tenido éxito y Davidson piensa que la razón es clara. El único objeto que podría satisfacer ambos requisitos, el de estar ante la mente y también el de que determine el contenido de un pensamiento, debería, como las ideas e impresiones de

---

<sup>179</sup> La sugerencia de Davidson es que la dificultad surge de los dos siguientes supuestos no cuestionados del externismo estándar (que él identifica básicamente con la posición de Putnam en “The meaning of “meaning””):

- (i) Si un pensamiento es identificado por una relación con algo que está fuera de la cabeza, no está totalmente en la cabeza.
- (ii) Si un pensamiento no está totalmente en la cabeza, no puede ser “captado” por la mente del modo requerido por la autoridad de primera persona.

Davidson desafía ambos supuestos. Por razones de pertinencia aquí voy a concentrarme en su crítica al segundo de ellos. En la que desempeñará un rol central su crítica a los “objetos ante la mente”.

Hume, “ser lo que parece y parecer lo que es”. No existen tales objetos, públicos o privados, abstractos o concretos.

Los argumentos de Burge, Putnam, Dennett, Fodor, Stich, Kaplan, Evans y muchos otros para mostrar que las proposiciones no pueden *a la vez* determinar los contenidos de nuestro pensamiento y asegurarse subjetivamente son, en opinión de Davidson, variantes del argumento simple y general que ha bosquejado. No son solamente las proposiciones las que no pueden desempeñar esa tarea; ningún objeto podría desempeñarla.

Cuando *nos liberamos del supuesto de que los pensamientos deben tener misteriosos objetos*, podemos ver cómo el hecho de que los estados mentales tal y como usualmente los concebimos se identifiquen en parte por su *historia natural* no sólo habla a favor del carácter interno de tales estados ni pone en peligro la autoridad de primera persona, sino que *abre el camino a una explicación* de la autoridad de primera persona.

La explicación se obtiene al caer en la cuenta de que lo que las palabras de una persona significan depende, en los casos más básicos, de la clase de objetos y eventos que hayan causado que la persona considere que las palabras se aplican; lo mismo vale para aquello sobre lo que versan los pensamientos de una persona. Un intérprete de las palabras y de los pensamientos de otra persona debe depender de informaciones dispersas, de un adiestramiento afortunado y de conjeturas imaginativas para llegar a comprender a esa persona. Sin embargo, el propio agente no está en posición de preguntarse si de manera general está usando sus propias palabras para aplicarlas a los objetos eventos correctos, puesto que aquello a lo que las aplique regularmente, sea lo que fuere, proporciona a sus palabras el significado que tienen y a sus pensamientos los contenidos que éstos poseen. Por supuesto, en cualquier caso particular puede estar equivocado con respecto a lo que cree acerca del mundo; lo que es imposible es que pueda estar equivocado la mayor parte del tiempo. La razón es, para Davidson, patente: a menos que haya una presunción de que el hablante sabe lo que quiere decir o significa, es decir, a menos que acierte con su lenguaje, no habría nada que el intérprete pudiera interpretar. Para decirlo de otro modo, nada podría ser considerado como un caso en el que alguien regularmente aplica mal sus propias palabras. Una vez que hemos abandonado el mito de lo subjetivo –la idea de que los pensamientos requieren objetos mentales– es natural pensar que *la autoridad de primera persona, el carácter social del pensamiento y los determinantes externos del pensamiento y el significado se dan conjuntamente*. Dice Davidson:

Si puedo pensar, sé que hay otros con mentes como yo, y que habitamos un tiempo y un espacio públicos con objetos y eventos, muchos de los cuales son conocidos por los otros. En particular, yo, como toda otra criatura racional, tengo tres clases de conocimiento: conocimiento del mundo objetivo (sin numerosas ostensiones exitosas, no tendría pensamientos), conocimiento de las mentes de otros; y conocimiento de los contenidos de mi propia mente. Ninguno de estos tres tipos de conocimiento es reducible a cualquiera de los otros o a una combinación de ambos. En efecto, se *sigue* del hecho de que tengo alguno de estos tipos de conocimiento, que tengo los otros, dado que el triángulo básico es condición del pensamiento, pero ninguno es conceptual o temporalmente previo a los otros (1998b/2001:87).

Así, pues, su peculiar externismo, la triangulación, y la autoridad de primera persona se concilian en el seno de su concepción interpretativa de la mente: así como los supuestos de la interpretación obligan a concebir el *contenido* de nuestras creencias como un *evento u objeto público* - y no como una entidad intermedia entre el sujeto y el mundo- obligan por su parte a hacer *una presunción en favor del hablante*. Externismo y autoridad de primera persona se compatibilizan en tanto condiciones del juego interpretativo. Ahora bien, si esto es así, la asimetría entre la primera y la tercera persona parece resultar un rasgo primitivo de la “gramática” de nuestro juego atributivo cotidiano. Quienes consideren que, en rigor, se trata de un rasgo que requiere de una explicación ulterior quedarán, pues, insatisfechos con esta propuesta<sup>180</sup>.

Como es sabido, la concepción filosófica de la subjetividad está cargada de una historia y un conjunto de supuestos desde los cuales se sostiene que lo subjetivo es previo a lo objetivo, que hay un mundo subjetivo previo al conocimiento de la realidad externa. La pintura davidsoniana del pensamiento y el significado no deja lugar para tal prioridad dado que, como hemos visto, predica el autoconocimiento (resignificado del modo que acabamos de ver) sobre la base del conocimiento de las otras mentes y del mundo. Lo *objetivo* y lo *intersubjetivo* son así

---

<sup>180</sup> Por razones de pertinencia, no voy a ocuparme aquí de problematizar este punto. Sin embargo, cabe decir que el abordaje davidsoniano ha sido “blanco” de muchas críticas. Entre otras, se ha señalado que (i) el estatus de la teoría no es claro, dado que la teoría no resulta fenomenológicamente adecuada; (ii) la teoría, en el mejor de los casos, explicaría sólo aquella parte del fenómeno de la autoridad de la primera persona que se debe al hecho de que no hay dos hablantes que hablen exactamente el mismo idiolecto, pero la autoridad podría ser un fenómeno más amplio que éste; y (iii) el argumento de Davidson depende de la afirmación de que “no equivocarse en el uso de las propias palabras” equivale a “saber qué significan las propias palabras”, puntos que no parecen ser idénticos.

esenciales a cualquier cosa que podamos llamar *subjetividad* y constituyen el contexto en que toma forma.

### **6.3. Primera y tercera persona en el marco de la atribución mental: el debate entre la teoría-teoría y la teoría de la simulación. Subjetivo-intersubjetivo-objetivo.**

Hasta aquí, hemos visto que Davidson logra ofrecer una elaboración de las nociones de *objetividad* y *subjetividad*, identificándolas respectivamente con las *perspectivas de tercera y de primera persona*, esto es, considerándolas como perspectivas distintas y genuinas. Sin embargo, creo que no se puede decir lo mismo respecto de la noción de *intersubjetividad*.

En efecto, con su peculiar propuesta de externismo semántico, la *triangulación*, Davidson introduce la *intersubjetividad* como una dimensión que parece, en principio, un recurso conceptual promisorio para cerrar la brecha abierta por el cartesianismo entre lo subjetivo y lo objetivo. Como vimos, se trata de una propuesta en la que la identificación de los objetos del pensamiento reposa en una base *social*: sin una criatura para observar a la otra, la triangulación que ubica a los objetos relevantes en un espacio público no podría tener lugar, y con ello, no sería posible nuestra mutua comprensión.

En este sentido, la perspectiva correlativa en la que Davidson parece cifrar su esperanza para elucidar la noción de *intersubjetividad* es la de *segunda persona*. No podemos resolver la cuestión de los contenidos de los estados mentales desde el punto de vista de una criatura única. Es necesaria una “*segunda persona*” (1992b/2001). No obstante, más allá de “bautizarla”, no termina de articularla conceptualmente como una perspectiva auténtica y diferente de las dos tradicionales. En otras palabras, en mi opinión, su *segunda persona* sigue siendo una *tercera persona*. Esto es, la “segunda” persona con la que interactuamos en el proceso de triangulación, según la posición explícita en Davidson, parece ser un mero *espectador* u observador no comprometido (ubicado en la perspectiva propia del intérprete, que construye una teoría a partir de la tercera persona) y no un *participante* en las prácticas sociales de los interpretados. En este sentido, creo que Davidson vislumbró la riqueza de introducir la dimensión *intersubjetiva* en nuestra reflexión acerca del pensamiento y el lenguaje, pero no terminó de elaborarla en toda su

hondura, esto es, como auténticamente distinta y, a la vez, conciliadora de lo subjetivo y lo objetivo. Lo que propongo es, entonces, que tratemos de profundizar esta suerte de hallazgo inconcluso en Davidson. A tal efecto, creo que puede resultar fructífero apelar a la perspectiva de *segunda persona* tal como se la introdujo recientemente en la filosofía de la mente a propósito del debate acerca de la atribución mental (Gomila: 2001, 2002, 2003; Scotto: 2002).

En la presente sección, me propongo abordar los términos dicho debate planteado en torno a la denominada “teoría de la mente”, esto es, la capacidad de atribución mental, de entender a los demás, y a nosotros mismos, en términos de estados mentales tales como: creencias, deseos, emociones, intenciones. Mi interés reside en examinar cuál sería la ubicación de la propuesta Davidson. Tal como dije al principio del presente capítulo, si bien él no intervino expresamente en esta disputa, se lo suele interpretar tradicionalmente de un modo que pretendo cuestionar aquí, en un sentido que tal vez nos permita empezar a avizorar de qué manera clarificar el tipo de normatividad peculiar de lo mental.

La expresión “teoría de la mente” apareció en el artículo de Premack y Woodruff (1978) “¿Tienen los chimpancés una teoría de la mente?”, donde se plantea si la chimpancé Sarah entendía la conducta de sus cuidadores humanos como intencional, propositiva. La discusión que acompañó a ese artículo sirvió para poner de manifiesto qué criterio resulta adecuado para establecer la capacidad de comprensión intencional de los chimpancés, o cualquier otro sistema cognitivo, a saber: cuando alguien es capaz de distinguir sus propios estados de los de los demás, por ejemplo, cuando se es capaz de atribuir a otros creencias diferentes de las propias. Esto es lo que ocurre en los casos de engaño o de atribución de ignorancia, y por ello, puede decirse que para que sea posible que alguien entienda el engaño o sea capaz de engañar, es preciso que cuente con una “teoría de la mente”. A ello siguieron desarrollos interesantes de la investigación (tales como los que consideramos en el capítulo anterior al referirnos al “Test de la falsa creencia”): destacadamente, el diseño de varias sencillas pruebas experimentales para valorar si un niño es capaz de entender el engaño (Wimmer y Perner, 1983; Gopnik y Astington, 1983); la búsqueda en la investigación comparada de evidencias de engaño entre los chimpancés, tanto en estado natural (Withen y Byrne, 1988) como en estudios de laboratorio (con esfuerzos para establecer un equivalente no verbal de la prueba de Wimmer y Perner: Tomasello, Call y Hare, 2003); y por supuesto, la hipótesis de que el autismo podría deberse a un déficit específico en la teoría de la mente (Baron-Cohen, Leslie, Frith, 1985).

Paralelamente a estos esfuerzos, se planteó también la discusión sobre si era apropiado denominar “teoría” a esa capacidad. En este plano, quienes defendían esa noción asumían como trasfondo el planteo funcionalista en sentido amplio, integrando tanto al funcionalismo mental, según el cual cada tipo de estado se caracteriza por su rol en la interacción con el resto, como el funcionalismo teórico, según el cual cada concepto teórico se caracteriza por sus relaciones con el resto de los términos de la teoría. De alguna manera, el éxito del *funcionalismo* como filosofía de la mente se manifestaba en esta comprensión de los términos de estado mental como términos *teóricos* y de los propios estados mentales, como estados funcionales representacionales, prescindiendo, por tanto, de consideraciones fenoménicas en su caracterización, por ser ponderadas como incompatibles con ese marco (Fodor y Block, 1972)<sup>181</sup>.

Frente a esta posición teoricista, Gordon (1986, 1992) y después Goldman (1989, 1992) recuperaron la vieja idea de la *empatía* como modo de entender a los demás, y la desarrollaron como la capacidad de *simular* el estar en la posición del otro, como mecanismo por el cual se realiza la atribución mental. Básicamente, su motivación consistía en evitar el *intelectualismo* que suponía la “teoría de la teoría”, al no requerir competencia conceptual alguna como medio para la atribución, sino simplemente la capacidad para hacer funcionar “off line” el propio sistema intencional e imputar el resultado de ese mecanismo al otro. Ese funcionamiento “off line” se concibe como una simulación, del mismo modo que, en el ámbito de la visión o del control motor se ha propuesto que la imaginación visual consiste en “simular”, en hacer funcionar “off line”, el propio sistema visual (Kosslyn, 1994), y la imaginación motora, el sistema motor (Jeannerod, 2001).

Surgió así una paradójica contraposición entre quienes proponían la “teoría” de que tener una “teoría de la mente” consistía en tener una *teoría* y quienes sostenían la “teoría” de que la capacidad de teoría de la mente no se basa en tener una teoría (a pesar del nombre), sino en el ejercicio de la capacidad de “ponerse en lugar del otro”, de *empatizar*, a través de simular sus

---

<sup>181</sup> En suma: la teoría da cuenta de las capacidades de atribución mental, denominadas “teoría de la mente”, como la posesión de una teoría –que cambia en el curso del desarrollo, en algunas versiones; modularista e innata, en otras -. De este modo, en algunas concepciones poseer una teoría de la mente sería como tener una teoría científica que es posible aplicar a casos particulares, esto es: observando la conducta y postulando estados mentales (términos teóricos) es posible explicar y predecir la conducta futura. En versiones alternativas, tal teoría se parece más a la gramática universal chomskiana y la pretensión es que la lectura de mentes revelaría un módulo de la mente humana.

estados, es decir, de atribuirle los estados que alcanza nuestro propio sistema intencional, al activarse, no por las circunstancias, sino por la consideración de las circunstancias ajenas<sup>182</sup>.

Ahora bien, dado este marco ¿dónde resulta adecuado ubicar la propuesta de Davidson de la teoría de la interpretación radical, una vez que se recuerda que “toda comprensión del habla de otro involucra interpretación radical” (Davidson 1973/1984:125)?

Recordemos que en “The second person” Davidson (1992b) intenta mostrar el carácter básicamente social de la atribución de pensamientos y significados en un contexto de interacción comunicativa, en la que hablante y oyente se vinculan por medio de patrones comunes de respuesta reconocidos ante un mismo objeto saliente en el entorno.

Ahora bien, se ha señalado que el enfoque davidsoniano está “demasiado tensionado por una *visión teoricista* (atenuada en algún sentido por tratarse de “teorías momentáneas” que el intérprete revisa constantemente (y por una muy estrecha concepción acerca de la interdependencia de intencionalidad mental e intencionalidad lingüística, de tal suerte que sólo los sujetos con competencia lingüística puede decirse que tengan pensamientos y creencias” (Scotto 2002:139; mis *itálicas*)<sup>183</sup>.

Cabe preguntarse, pues, si esta ubicación de Davidson en el lado *teoricista* del debate en torno a la atribución mental, hace justicia a su propuesta. Para ello, una vez más, resultará iluminador volver a su disputa con Rorty (Davidson 1998/2005).

Allí Davidson introduce precisiones sobre lo que quiso decir en un ensayo que consideramos en el capítulo anterior: “A Nice Derangement of Epitaphs”. Considera que se lo malinterpretó al pensar que sostuvo que hablantes e intérpretes tienen y forman *teorías* del significado sobre la base de lo que hablan e interpretan. Stroud pregunta, en ese contexto, por qué Davidson desea hablar de “teoría” en el caso de la Sra. Malaprop y Rorty, directamente, mina la idea de que la gente opere con *teorías* del significado. Leemos en Rorty:

---

<sup>182</sup> Se suele considerar que el desarrollo reciente del debate ha favorecido a los simulacionistas por su interpretación del descubrimiento de las “neuronas espejo” en el córtex premotor como el mecanismo neuronal de la simulación de bajo nivel, es decir, de atribución de intenciones básicas y emociones (antes que de actitudes proposicionales cf. (Goldman, 2006). Sin embargo, y aparte de las dificultades intrínsecas que tiene la interpretación funcional de las “neuronas espejo” (Gomila 2007), Gomila cree que sigue siendo válida la afirmación de que este planteo *pasa por alto el carácter distintivo de las interacciones cara a cara*. En el capítulo 8 me ocuparé centralmente de esta omisión, al considerar los fenómenos de atribución mental que son recogidos desde la- así llamada- perspectiva de *segunda persona* de la atribución mental.

<sup>183</sup> Cabe señalar que también Rabossi (2000) considera a Davidson comprometido con una perspectiva teoricista.

Por qué no tratar el trabajo de los gramáticos y los lexicógrafos (o de sus contrapartes ideales, los sostenedores de definiciones de verdad tarskianas...) como manteniendo con el hablante ordinario la misma relación que el físico mantiene con el ciclista ordinario? (citado por Davidson 1998/2005:324)

En efecto, Davidson dice haber seguido este consejo rortyano mucho antes de que se lo hubieran dado. Afirma que desde los albores de su presentación de la teoría de la interpretación radical sostuvo que “aquellos que usan lenguaje normalmente no tienen una *teoría*” (1998/2005:324). Todo lo que sostuvo de una teoría satisfactoria en este sentido fue que *si* alguien tenía una teoría tal para un hablante, en un momento, esa teoría bastaría para comprender una emisión arbitraria de ese hablante en ese momento (y remite, en este espíritu al primer párrafo de “Radical Interpretation”).

Asimismo, en (1974b), hace expreso que usa la teoría de la interpretación para describir un aspecto de la competencia del intérprete en la comprensión de lo que se dice.<sup>184</sup> Pero, insiste en que la teoría de la interpretación es un trabajo común del lingüista, del psicólogo y del filósofo. En este sentido, la teoría puede usarse para describir lo que el intérprete sabe, esto es, un subconjunto especificable de las verdades de la teoría, pero de ningún modo se sugiere que el intérprete conozca la teoría en cuestión.

Por su parte en (1976), Davidson sostiene que alguien está en posición de interpretar las emisiones de los hablantes de un lenguaje L si ese alguien posee un determinado cuerpo de conocimiento implicado por una teoría de la verdad para L y sabe que ese conocimiento está implicado por dicha teoría. Mi error, dice Davidson, no fue como parece sugerir Foster, suponer que toda teoría que diera correctamente condiciones de verdad serviría para la interpretación: mi error consistió en pasar por alto el hecho de que alguien podría conocer una teoría lo suficientemente peculiar sin saber que lo es. Imaginé, afirma, que quien conociera la teoría sería alguien que la habría construido a partir de la evidencia, y tal persona no podía pasar por alto que su teoría satisfacía las restricciones. Tenemos un paralelo preciso si preguntamos qué datos debe conocer alguien para ser un físico. Una rápida respuesta: las leyes de la física. Pero Foster diría, y Davidson acuerda, que esto no basta. El físico debe conocer también que esas leyes son leyes.

---

<sup>184</sup> Allí afirma que también podemos sostener, si lo deseamos, que hay un mecanismo en el intérprete que corresponde a la teoría. Si esto significa solamente que hay alguno u otro mecanismo que ejecuta esa tarea, resulta difícil ver cómo la afirmación podría dejar de ser verdadera. De donde, desde luego, no se sigue en absoluto que le confiera realidad psicológica a dicha teoría.

El intérprete, dice Davidson, sabe que su conocimiento consiste en lo que está enunciado por una teoría T, una teoría que satisface la Convención T. Pero no hay razón para suponer que el intérprete puede expresar su conocimiento en alguna forma lingüística específica, y mucho menos en un lenguaje particular. Quizás deberíamos insistir en que una teoría es una oración o un conjunto de oraciones de algún lenguaje. Pero para conocer una teoría no es ni necesario ni suficiente conocer que estas oraciones son verdaderas. No es suficiente ya que esto podría conocerlo alguien que no tuviera idea de lo que las oraciones significan, y no es necesario pues alcanza con conocer las verdades que expresan las oraciones de la teoría y esto no requiere el conocimiento del lenguaje de la teoría.

Tal como sugerí en la primera parte de este trabajo, Davidson apela a una “teoría” para, desde un planteo trascendental, hacer explícitas las condiciones de posibilidad de la interpretación que está involucrada toda vez que comprendemos el discurso de otro. La *teoría* de la interpretación es su explicitación de la competencia lingüística del intérprete. En este sentido, la llama “teoría”, sólo porque una descripción de la competencia del intérprete requiere una explicación recursiva.

Más recientemente, tal como nos recuerda en su (1998/2005) enfatizó que es el *filósofo*, tratando de comprender la comprensión, quien necesita una *teoría* a los efectos de decir qué es lo que el intérprete sabe si comprende a un hablante.

Un intérprete ordinario está preparado, dice Davidson, para interpretar cualquier cosa que diga un hablante –una infinidad potencial de emisiones-. Así, el intérprete conoce lo que porta cada oración-T implicada por una teoría de la verdad para el lenguaje del hablante, y sólo una teoría puede especificar esta totalidad en términos finitos. Pero considera haber sido siempre muy claro respecto de que esto no sugiere que el *intérprete* conozca una *teoría tal*.

¿Tiene sentido, pues, hablar de una teoría en el caso de la Sra. Malaprop? La idea que Davidson tenía en mente, dice, era simple. Un intérprete en una situación conversacional particular está preparado con un conjunto general de expectativas (lo que describe la “teoría previa”). Cuando la expectativa se frustra, lo que es nuevo se acomoda (usualmente, de manera automática; por ejemplo, leer “epítetos” por “epitafios”). Si el hablante continúa así, estas sustituciones en todos los contextos posibles producen un nuevo lenguaje, que puede ser delineado por una “teoría al paso”. Los resquicios en la “teoría previa” se llenan de maneras novedosas. En este sentido, el aparato conceptual de la “teoría previa” y de la “teoría al paso”

fue, dice autocríticamente, una manera innecesariamente problemática de expresar esta idea. (1998/2005:325). Y afirma:

Saber un lenguaje es, en algunos aspectos, como saber andar en bicicleta. En ambos casos, como Rorty señala, hablamos de *saber cómo*, y en ninguno es necesario o común saber una *teoría* que explique lo que hacemos. Pero también hay diferencias notables. Hay incontables cosas que un hablante o intérprete debe saber: las condiciones de verdad que un intérprete probablemente tomará como las que tienen las emisiones del hablante, las condiciones de verdad que la mayoría de las oraciones que escuche tendrán, relaciones de implicación, contradicción, y apoyo evidencial entre oraciones. Andar en bicicleta no requiere en absoluto conocimiento proposicional. A los animales no lingüísticos se les puede enseñar a andar en bicicleta. (1998/2005:325; mis itálicas).

Tal es la manera en la que, según indicamos ya en la primera parte de este trabajo, Davidson considera que su interés persistente en la semántica tarskiana puede tornarse compatible con la agenda wittgensteiniana en que lo inscribe Rorty.

A mi juicio, las consideraciones que formula el propio Davidson acerca del estatus “teórico” de su propuesta, ponen en jaque la interpretación “teoricista” en la que clásicamente se lo ha ubicado en el debate acerca de la atribución mental.

Ahora bien, podría juzgarse que los términos en los cuales se plantea dicho debate resultan profundamente insatisfactorios, y ello por dos razones principales (Gomila 1993). Por un lado, porque basa la discusión en una visión cuando menos sesgada de la atribución mental, limitada exclusivamente a la atribución de *actitudes proposicionales*, sin tener en cuenta los diversos modos y niveles de interacción social psicológicamente mediada. Y, por otro, porque supone la *reedición del viejo debate de principios de siglo XX*, en el momento fundacional de las ciencias sociales, sobre la naturaleza de la explicación psicológica, cuando se enfrentó el psicologismo de la escuela de la empatía al enfoque racionalista de Max Weber. Además, se comparte el supuesto de que la función de la atribución es la de *explicar o predecir* la conducta de los demás, desde un plano observacional distanciado de su acción.

Por ello, la estrategia de Gomila consiste en mostrar lo insatisfactorio tanto de los teóricos de la teoría, como de los teóricos de la simulación, para dar cuenta de la complejidad de nuestras relaciones intersubjetivas, basadas inicialmente en la *interacción espontánea*, en la

*participación (engagement)*, y no en una actitud distanciada, observadora, puramente interesada en la predicción de la conducta ajena. Su objetivo primordial consiste, así, en poner de relieve que en los casos de interacción intencional cara a cara, la atribución intencional que media la interacción se realiza implícitamente y se caracteriza por la *intersubjetividad*, por la experiencia de reciprocidad en tal atribución.

En este mismo espíritu, Scotto (1992: 137) señala que en el debate en cuestión ha pasado inadvertida la consideración de una perspectiva de atribución de mental: la de *segunda persona*. Ella sugiere dos hipótesis para explicar esta inadvertencia. La primera es que los argumentos a favor y en contra, tanto de un lado (teoricistas) como de otro (simulacionistas), han intentado explicar la atribución mental como parte del intento de fundamentar la explicación psicológica de carácter científico, sin distinguir adecuadamente los problemas propios de la atribución mental ordinaria de aquellos vinculados con la explicación psicológica. El problema de la atribución intencional, dice, adquiere una luz diferente cuando lo que se intenta describir y/o explicar es, justamente, una competencia ordinaria, distinguiendo esta tarea del propósito de segundo orden de responder a la cuestión de si una psicología científica debería hacer referencia a estados con contenido y cómo debería explicarse su atribución. Piensa que esta omisión puede deberse también a una “característica deformación del análisis y la especulación filosóficos”, demasiado dependientes de la selección de casos paradigmáticos; esto es, que ejemplifican fenómenos más localizados o especializados que los que se deberían tomar en consideración si se quiere proponer generalizaciones sobre un dominio dado. Estos casos paradigmáticos son, además, frecuentemente descriptos en abstracción de los contextos prácticos y las variaciones graduales que de hecho exhiben. Wittgenstein, dice Scotto, nos enseñó a evitar ambos defectos. Por un lado, se puso énfasis en casos paradigmáticos de “pensamientos-yo” de un sujeto, autoconsciente, reflexivo y lingüísticamente competente, o piénsese en el papel concedido a fenómenos tales como el “acceso privilegiado” o la “conciencia fenoménica”, por un lado. Por otro, se puso énfasis en los cálculos predictivos acerca de la conducta ajena que realizamos mediante el recurso a la observación y a la hipótesis acerca de su vida mental o a ciertos casos complejos de comunicación intencional<sup>185</sup>. El diagnóstico de Scotto es que, de este modo, ambas perspectivas han distorsionado los rasgos de la atribución intencional, al subrayar dimensiones

---

<sup>185</sup> En este espíritu, de manera tradicional, se liga al bando “simulacionista” del debate con la perspectiva de primera persona, y al bando “teoricista” con la perspectiva de tercera persona de la atribución mental.

distintas (la subjetividad o la objetividad) y restringirse a dar cuenta de ciertos propósitos especiales de la atribución mental, especialmente los epistémicos.

Frente a este estado de la cuestión, su invitación es que atendamos a la perspectiva de *segunda persona de la atribución mental*. Dicha perspectiva, en cambio, ubica en el centro de la escena a la *intersubjetividad* e implica un "*compartir intencional*", una *sintonía simultánea* que constituye una experiencia psicológica genuina, diferente de las que proporcionan las perspectivas de primera y tercera persona, y que resulta clave en el proceso de desarrollo ontogenético (Hobson, 1995,2005).

Hemos visto que, en virtud de las precisiones que el mismo Davidson ofrece, no parece lícito enrolarlo sin más en el bando "teoricista" del debate en cuestión. Asimismo, al principio de esta sección, afirmé que, en mi opinión, logra ofrecer una elaboración precisa de las nociones de *objetividad* y *subjetividad*, identificándolas respectivamente con las *perspectivas de tercera y de primera persona* (esto es, considerándolas como perspectivas distintas y genuinas) pero que, sin embargo, no se puede decir lo mismo respecto de la noción de *intersubjetividad*. Más allá de "bautizar", la noción que le sería correlativa, a saber: la *segunda persona*, no termina de articularla conceptualmente como una perspectiva auténtica y diferente de las dos tradicionales. En otras palabras, en mi opinión, su *segunda persona* sigue siendo una *tercera persona*. Esto es, la "segunda" persona con la que interactuamos en el proceso de triangulación, según la posición explícita en Davidson, parece ser un mero *espectador* u observador no comprometido (ubicado en la perspectiva propia del intérprete) y no necesariamente un *participante* en las prácticas sociales de los interpretados. No obstante, hemos visto que Davidson equipara en muchos aspectos el saber un lenguaje con el saber andar en bicicleta. Si nos tomamos en serio esta idea, que él mismo alienta, creo que podemos encontrar una manera de profundizar su "segunda persona", esta suerte de hallazgo inconcluso en Davidson, a partir del modo cómo se la tematiza actualmente desde la filosofía de la mente (Gomila: 2001, 2002, 2003; Scotto: 2002).

Sin embargo, antes de embarcarnos en esta propuesta, se impone que nos desembaracemos del "corset" en el que se confina la filosofía de la mente davidsoniana. Según vimos en los capítulos 4 y 5, tanto su respuesta al problema mente-cuerpo, como su posición en torno al problema de la relación entre el pensamiento y el lenguaje reposan en una dicotomía (y no una mera distinción de grado) entre la dimensión descriptiva y la dimensión normativa. Según mostramos, dicha dicotomía detiene por principio nuestra discusión al respecto.

La idea que voy a pasar a considerar en el capítulo próximo es la siguiente: si logramos deshacernos de dicha dicotomía, se abre un espacio conceptual en el que la perspectiva de *segunda persona* de la atribución mental puede darle sustento a la concepción de la intersubjetividad que pone en juego Davidson, de un modo que nos permita arrojar luz sobre la peculiar normatividad de lo mental y continuar nuestra discusión, ahí donde Davidson la cancela *a priori*.

**SECCIÓN III**

**En busca de nuevos horizontes**

## Capítulo 7: Davidson y la dicotomía descriptivo/normativo: ¿el “cuarto dogma del empirismo”?

Estoy sorprendido al encontrar que Rorty, quien se opone tanto a distinguir el vocabulario mental del vocabulario de la ciencia natural, compe la vieja distinción positivista entre el “vocabulario *descriptivo* de la intencionalidad” y el “vocabulario *prescriptivo* de la normatividad”.  
(Davidson, 1998/ 2005: 320; mis itálicas).

### 7.1. Tipos de debate y dicotomías. De la “discusión” a la “controversia”

En su artículo “Dichotomies and types of debate”, Marcelo Dascal (2007) comienza por recordarnos que las dicotomías son “ubicuas”: están presentes en el pensamiento deliberativo, en la toma de decisiones y al argumentar en todas las esferas de la vida. En particular, las dicotomías se tornan evidentes, en su opinión, en los debates, esto es, en los intercambios dialógicos argumentativos “caracterizados por su naturaleza agonística”. Los “protagonistas” en un debate deben su nombre a que sostienen posiciones contrapuestas —o que asumen como tales—. En algunos casos, esto puede conducir a una polarización del debate, es decir, a una “dicotomización”.

Pasa a explorar, entonces, la relación entre tipos de debate y sus estrategias características. En la taxonomía que propone encontramos, por un lado, las *discusiones* y *disputas* que reposan, a su juicio, en una estrategia del tipo mencionado, es decir, de “dicotomización” y, por otro, las *controversias*, a las que subyace una estrategia de “desdicotomización”. En el primer caso, se radicaliza el debate, al punto de tornar compleja, y aún imposible su resolución, mientras que en el segundo caso, se abre la posibilidad de una solución.

El primer tipo de estrategia, la dicotomización, puede retrotraerse a la tradición platónica. Recordemos que la vía propuesta para la definición de las ideas, la *diaeresis*, es un método que procede a través de una sucesión de divisiones dicotómicas de los conceptos. Para una primera aproximación a la noción de dicotomía, veamos entonces cuál es su caracterización

lógica. Una dicotomía puede ser definida como una operación por medio de la cual un concepto, A, es dividido en otros dos, B y C, cada uno de los cuales excluye al otro, cubriendo complementariamente el dominio completo del concepto original. Esto es, si  $x$  pertenece a A, entonces  $x$  pertenece a B o  $x$  pertenece a C, y si  $x$  pertenece a B entonces no pertenece a C y viceversa, lo que implica que  $C = \sim B$  y  $B = \sim C$ . Los términos de la dicotomía se revelan, pues, como exclusivos y excluyentes.

Dascal rechaza esta caracterización que pone el fundamento de la dicotomía en la relación lógica de exclusión porque, entre otras cosas, entiende que implica un compromiso con la idea de que la oposición de que se trate resultaría insuperable en tanto constituiría una clase formal. De este modo, frente a la aproximación platónica de las dicotomías en términos que involucran un compromiso con su realidad conceptual, con su estatus como entidades, Dascal reivindica su caracterización en términos que denomina “constructivos y pragmáticos” y que le confieren a las dicotomías el estatus de “estrategias”.

Desde su perspectiva, el que determinadas oposiciones sean consideradas dicotómicas depende de una variedad de factores tales como: el marco socio-cultural, las creencias básicas, los intereses personales, las necesidades y propósitos argumentativos, etc. Desde su enfoque, entonces, la deconstrucción de cualquier dicotomía es siempre posible. Es, precisamente, este segundo tipo de estrategia (la des-dicotomización) aquella que está a la base de lo que denomina “controversias”.

En suma, según la propuesta de Dascal, hay dos tipos de estrategia diferentes que configuran distintos tipos de debate. Por un lado, la estrategia de *dicotomización*, que consiste en radicalizar una polaridad, esto es, en enfatizar la incompatibilidad de los polos y sostener la inexistencia de alternativas intermedias, configura el espacio conceptual para las *discusiones* y *disputas* filosóficas. Por otro lado, la estrategia de *des-dicotomización*, que consiste en mostrar que la oposición entre los polos puede admitir alternativas intermedias, configura el espacio conceptual para las *controversias*.

En lo que sigue, intentaré proponer el pasaje de una *discusión* a una *controversia*. Trataré de mostrar que la peculiar respuesta de Davidson a los tres problemas que hemos considerado en la segunda parte de este trabajo, a saber: el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema las perspectivas de primera y tercera persona de la atribución mental, se inserta en una *discusión* que reposa en una dicotomía

que hunde sus raíces en la historia de la filosofía: la dicotomía entre una dimensión normativa y una descriptiva. Si podemos deshacernos de ella, estaremos ante una *controversia* que abre un horizonte conceptual para contemplar alternativas intermedias y continuar el diálogo.

## 7.2. La dicotomía descriptivo/normativo: su historia y su “colapso”.

Entre los ejemplos célebres que considera Dascal, se concentra especialmente en la dicotomía hecho/valor, esto es, la dicotomía entre una dimensión descriptiva y una dimensión normativa. Dascal encuentra en la obra de Putnam (2002) no sólo una crítica específica a la dicotomía en cuestión, sino una objeción de talante más general a los excesos de la dicotomización como estrategia, sugiriendo cómo evitarla sin incurrir en el extremo opuesto consistente en suprimir, junto con ella, distinciones que parece interesante retener.

En *The collapse of the fact/value dichotomy*, Putnam (2002) se ubica como continuador de una tradición crítica de las dicotomías y retoma el propósito central que lo animó en *Razón, verdad e historia*, esto es: terminar con la presión asfixiante que unas cuantas dicotomías parecen ejercer tanto sobre el pensamiento filosófico como sobre el de sentido común. Allí asume el desafío de repensar el dogma que consiste en comprometerse con que los *hechos* son objetivos, los *valores* son subjetivos y no hay encuentro posible entre ambos. Veamos, pues, en qué consiste la dicotomía en cuestión.

A juicio de Putnam, el asunto de los hechos y los valores, formulado de un modo suficientemente amplio, nos concierne a todos. En esto se diferencia claramente de muchos problemas filosóficos. Señala que desde un punto de vista mayoritario, los problemas de la filosofía del lenguaje, de la epistemología y hasta de la metafísica pueden parecer asuntos interesantes, aunque de elección opcional. Pero el problema de los hechos y los valores parece de elección forzosa: cualquier persona reflexiva ha de tener una opinión frente a él. Ocurre que una particular solución a este problema se ha arrogado el estatus de institución cultural, y es la siguiente: hechos y valores pertenecen a esferas totalmente distintas, la dicotomía entre *enunciados de hecho* o *enunciados de valor* es absoluta y, concomitantemente, no hay base objetiva para decidir si las cosas son buenas o malas, mejores o peores. La pregunta que está

detrás de esta reflexión es la de si los “fines”, los valores, pueden ser discutidos racionalmente o no; esto es, si hay una noción de racionalidad aplicable a cuestiones normativas.

En su libro de (2002), Putnam vuelve a arremeter contra esta dicotomía, para detenerse especialmente en su historia. Parte de su interés reside en mostrarnos la manera en que la práctica de ciertos filósofos ha convertido en dicotomías absolutas lo que parecen ser distinciones inocuas. Así, pues, el punto de vista con respecto a la relación entre “hechos” y “valores” que defiende en el libro en cuestión es el que John Dewey sostuvo prácticamente a lo largo de toda su vida. Lo que atacaba Dewey, dice Putnam, no era la idea de que, para ciertos propósitos, pueda ser útil trazar una distinción (entre, digamos, “hechos” y “valores”). El blanco de su crítica era más bien lo que él llamaba el “dualismo” hechos/valores, uno de los muchos dualismos filosóficos que Dewey estaba interesado en identificar, diagnosticar y exorcizar de nuestro pensamiento. Como cuestión preliminar, Putnam traza entonces la siguiente aclaración terminológica: una *dicotomía* es caracterizada como una distinción tajante y omnipresente que pretende poder aplicarse a todo enunciado significativo en absolutamente toda área; mientras que una *distinción* ordinaria posee rangos de aplicación y no nos sorprende si no se aplica siempre.

Recordemos, pues, en primera instancia, en qué consiste la dicotomía hecho/valor a partir de la caracterización “en espejo” que Putnam ofrece de ella desde la célebre dicotomía analítico/sintético. En efecto, Putnam pretende mostrar que estas dos dicotomías “han pervertido nuestro pensamiento tanto en lo que concierne al razonamiento ético como a la descripción del mundo, impidiéndonos advertir que *evaluación y descripción están entrelazadas y son interdependientes*” (2002/2004: 16). Como dice en el capítulo 1, su idea es que pensar sin los dogmas hecho/valor y sintético/analítico (dogmas que, según veremos, hunden sus raíces en dos dicotomías fundacionales del empirismo clásico) es entrar en “una genuina posmodernidad, entrar en un nuevo campo de posibilidades intelectuales”. En mi opinión, parte de la originalidad de su planteo reside en sostener que ambas dicotomías están prefiguradas por una dicotomía humeana y que colapsan de manera análoga. En otras palabras: tienen una génesis y un destino común. La primera, en la medida en que, a juicio de Putnam, ambas reposan en una dicotomía humeana. El segundo, en el sentido de que ambas colapsan como dicotomías, pero, a su vez, vale la pena retenerlas como distinciones útiles. De este modo, Putnam sostiene que la historia de

la dicotomía hecho/valor es en ciertos aspectos paralela a la de la dicotomía analítico/sintético. Veamos esta cuestión más detenidamente.

Putnam sostiene, por un lado, que la dicotomía analítico/sintético está prefigurada por la distinción humeana entre “cuestiones de hecho” (*matters of fact*) y “relaciones de ideas” (*relations of ideas*). Por otro, afirma que la dicotomía hecho/valor está prefigurada por el *dictum* humeano de que un “debe” (*ought*) nunca puede ser derivado de un “es” (*is*). El punto en el que dichas dicotomías se “hermanan” es, para él, la noción de “hecho” que subyace a ambas. En efecto, su esfuerzo reside en mostrarnos en qué sentido la noción humeana de “hecho” resulta extremadamente “estrecha” y signa, desde el comienzo, el destino de estas célebres dicotomías construidas sobre ella.

La transición crucial desde una distinción inocente a un dualismo metafísico -de la que Dewey trató de alertarnos- puede observarse ya, dice Putnam, en la manera en que Kant concibió la distinción entre juicios “analíticos” y juicios “sintéticos”, prefigurada por la distinción humeana entre “cuestiones de hecho” (*matters of fact*) y “relaciones de ideas” (*relations of ideas*). De lo que se trata, para Putnam, es que Kant forzó la pregunta: “las verdades matemáticas ¿son analíticas o son sintéticas?”. Como es sabido, Kant consideraba que eran a la vez sintéticas y *a priori*, tesis que ha resultado un anatema para los empiristas. La réplica de los positivistas lógicos fue que los principios de las matemáticas son efectivamente necesarios (como pensaba Kant), pero no sintéticos, sino analíticos. Sin embargo, a fin de dar esta respuesta, los positivistas lógicos “forzaron la noción de analiticidad hasta hacerla añicos” (2002/2004:25).

Célebremente Quine (1951) criticó a los positivistas, en primer lugar, por su fracaso en apreciar que el elemento de convención es un rasgo de las hipótesis científicas de manera general -más que una propiedad de ciertas proposiciones científicas particulares y no de otras- y, en segundo lugar, por su fracaso en apreciar que el hecho de que una proposición sea inicialmente aceptada como una convención no implica que no haya que someterla luego al tribunal de la experiencia, al igual que el resto de las oraciones de la teoría.

No obstante, a juicio de Putnam, Quine fue demasiado lejos en su ataque inicial a la distinción en “Dos dogmas del empirismo”, donde perdía de vista el punto principal al negar que tenga sentido alguno distinguir una clase de verdades analíticas de las verdades sujetas a contrastación observacional. En opinión de Putnam, puede aceptarse la intuición de Quine, según la cual hay amplias categorías de proposiciones que no pueden clasificarse simplemente como

verdades analíticas o como enunciados acerca de hechos observables, pero a la vez no abandonar la “modesta idea” de que hay también casos que caen en uno u otro lado de la distinción. En otras palabras: para Putnam, reconocer una diferencia entre enunciados de un cierto lenguaje que son trivialmente verdaderos en virtud del significado de sus palabras y enunciados que no lo son, no implica que todos los demás tipos de enunciados se subsuman bajo una única clase de “enunciados acerca de cuestiones de hecho” (Hume) o de “enunciados sintéticos” (Kant). En resumen: domesticada de este modo, la noción de enunciado analítico resulta ser una noción modesta y en ocasiones útil, pero ya no esa poderosa arma filosófica capaz de “funciones tan maravillosas como la de explicar por qué las verdades matemáticas no representan ningún problema para el empirismo” (2002/2004: 27).

Por su parte, según adelantamos, la dicotomía hecho/valor está implícita en la doctrina humeana de que no puede inferirse un “debe” (*ought*) a partir de un “es” (*is*). Aquello que, en opinión de Putnam, Hume quiso señalar no es una mera cuestión sobre los cánones de la inferencia formal, sino que cuando un juicio de “es” describe una “cuestión de hecho” no puede derivarse ningún juicio de “debe”. La idea de Putnam es que lo que constituye el fundamento de la alegada inderivabilidad de “debe” a partir de “es” es, una vez más, la dicotomía *metafísica* humeana entre “cuestiones de hecho” (*matters of fact*) y “relaciones de ideas” (*relations of ideas*).

El criterio de Hume para las “cuestiones de hecho” presupone lo que se ha denominado “semántica pictórica”. En la teoría de la mente de Hume, los conceptos son un tipo de “idea” y las “ideas” son ellas mismas figurativas; el único modo en que pueden representar cualquier “cuestión de hecho” es *asemejándose* a ella. Las ideas también tienen, no obstante, propiedades no figurativas: pueden entrañar o estar asociadas a sentimientos, emociones. Hume no dice tan sólo que no podemos inferir un “debe” de un “es”; afirma, en un sentido más amplio, que no hay “cuestión de hecho” acerca de lo correcto ni “cuestión de hecho” acerca de la virtud. La razón de ello es que, si hubiese cuestiones de hecho acerca de la virtud y el vicio, entonces tendría que ser el caso -si se acepta su semántica pictórica- que la virtud fuera *figurable* en el modo en que lo es la propiedad de ser una manzana. Por su parte, los positivistas lógicos radicalizaron la idea de Hume de que los juicios éticos no son enunciados de hecho, sino expresiones de sentimiento o bien imperativos disfrazados. Esta exclusión de los enunciados éticos del dominio del discurso racional recibió una expresión vehemente por parte de Carnap en *The Unity of Science*.

Al decir de Putnam, la confianza de los positivistas lógicos en que podrían expulsar la ética del dominio de lo racionalmente discutible derivaba en parte del modo en que, en sus manos, los dualismos analítico/sintético y hecho/valor se reforzaban mutuamente. Según los positivistas, para construir conocimiento, las “proposiciones” éticas deberían ser o bien analíticas -cosa que manifiestamente no son- o bien “fácticas”. Y el convencimiento de los positivistas de que no podían ser fácticas, al igual que el de Hume de que “el crimen de la ingratitud no es ningún hecho individual”, derivaba del convencimiento de que sabían exactamente qué es un *hecho*. A juicio de Putnam, pues, en los escritos de los positivistas, tanto en el caso del dualismo de enunciados analíticos y fácticos como en el del dualismo de juicios éticos y fácticos, es la *concepción de lo fáctico* la que hace todo el trabajo filosófico.

La estrategia general de Putnam (2002) puede, entonces, resumirse del modo siguiente: explícitamente reticente a efectuar una “operación” sobre los juicios de valor mediante un intento de “naturalización”, Putnam centra su atención, en cambio, en los juicios de hecho y pretende acercar los juicios de valor a los juicios de hecho mostrando que estos últimos no tienen las propiedades “sacrosantas” que se suponía que tenían. De esta suerte, ataca al empirismo en su nudo gordiano: el diagnóstico de Putnam es que el gran error se cometió respecto de los juicios de *hecho* -al suponer que son *objetivos* en tanto garantizables más allá de toda discusión ulterior- y cree que, correlativamente, se incurrió en un error que es parasitario de aquél: creer que los juicios de *valor* son ineludiblemente *subjetivos*. La posición de Putnam puede articularse, a mi juicio, en dos instancias: una parte crítica o negativa destinada a mostrar que la dicotomía está sustentada en argumentos indefendibles y una parte constructiva o positiva a los efectos de revelar la imbricación (*entanglement*) que habría entre lo fáctico y lo valorativo. Consideremos brevemente cada una de ellas.

Putnam va a dirigir su ataque, especialmente, a la “pobreza de la concepción del lenguaje de los positivistas lógicos”. La ciencia ha cambiado radicalmente desde los días de Hume, y los positivistas se vieron cada vez más forzados a abandonar su noción inicial de “hecho”, en cierto modo familiar a la de Hume, con el fin de hacer justicia a la revolución científica de la primera mitad del siglo XX. El punto de Putnam es que, al revisar su noción de “hecho” destruyeron el mismísimo fundamento sobre el cual habían erigido la dicotomía hecho/valor.

En tiempos de Hume, todavía era razonable mantener que no había predicados científicamente indispensables referidos a entidades no observables por los sentidos humanos<sup>186</sup>. En efecto, según vimos, la noción humeana de “hecho” es, simplemente, la de algo de lo que puede haber una “impresión sensorial”. La situación era, sin embargo, muy distinta en la época en que se constituyó el Círculo de Viena. A partir de 1939, los positivistas lógicos empezaron a liberalizar su famoso “criterio de significatividad cognoscitiva”, afirmando que el lenguaje cognitivamente significativo puede contener no sólo términos observacionales, sino también los llamados “términos teóricos”, términos referidos a inobservables e introducidos por sistemas de postulados, los postulados de las diversas teorías científicas. El criterio liberalizado de significatividad cognoscitiva resultante quedaba resumido así: en la medida en que el sistema en su conjunto nos permite predecir mejor nuestras experiencias con tales predicados (teóricos) que sin ellos, esos predicados tienen que aceptarse como “empíricamente significativos”. Pero, observa Putnam, “predecir” algo significa –para los positivistas lógicos- deducir enunciados observacionales a partir de una teoría. Y para deducir algo a partir de un conjunto de postulados empíricos se necesitan no sólo estos postulados, sino también los *axiomas de las matemáticas y la lógica*. Según los positivistas lógicos, estos axiomas no enuncian “hecho” alguno. Son “analíticos” y, por tanto, “vacíos de contenido fáctico”. En suma, “pertenecer al lenguaje de la ciencia es (desde el punto de vista de los positivistas lógicos) un criterio de significatividad científica, pero no todo lo científicamente significativo es un “enunciado de hecho”; entre los enunciados científicamente significativos hay, según los positivistas lógicos, enunciados *analíticos* además de los *sintéticos* (es decir, fácticos). Así, la búsqueda de una demarcación satisfactoria de lo “fáctico” se convirtió en la búsqueda de un modo satisfactorio de trazar “la distinción analítico/sintético”.

Como hemos señalado, en la lectura de Putnam, Quine derribó la noción (metafísicamente exagerada) de lo “analítico”, para satisfacción de la mayoría de los filósofos. Sin embargo, no sugirió que todo enunciado del lenguaje de la ciencia deba ser considerado un enunciado de hecho (esto es, “sintético”); más bien observó que la idea entera de clasificar todo

---

<sup>186</sup> Desde luego, Putnam es consciente de que los “átomos” figuraban en una parte importante de la especulación científica de la época –principalmente la británica-. No obstante, Locke sostenía que nunca podremos saber nada sobre ellos, y Berkeley y Hume estaban absolutamente determinados a descartar tales especulaciones por ininteligibles, del mismo modo que consideraban –filosóficamente-ininteligible hablar de puntos en geometría o de infinitesimales en cálculo.

enunciado -incluidos los de las matemáticas puras- o bien como “fácticos” o bien como “convencionales” (lo que para los positivistas lógicos equivalía a “analíticos”) era un atolladero insoluble. Pero, prosigue Putnam, si la idea de que hay una noción *clara* de lo que es un *hecho* se derrumba junto con la irremediablemente restrictiva imagen empirista que le dio lugar ¿qué ocurre con la *dicotomía hecho/valor*? Como ha escrito el filósofo economista Vivian Walsh, “tomando prestada y adaptando la elocuente imagen de Quine, si una teoría puede ser negra por el hecho y blanca por la convención, bien podría ser -en lo que alcanzan a decir los empiristas lógicos-roja por los valores. Dado que para ellos la confirmación o la falsación tienen que ser propiedades de una teoría en su totalidad, no tienen manera de desenredar toda la madeja” (Walsh 1987).

Así, señala Walsh (y antes de él, Morton White, amigo de Quine) que, tras abandonar Carnap -entre 1936 y 1939- la caracterización de las proposiciones “fácticas” como susceptibles de ser confrontadas una a una con la experiencia sensorial (que era justamente, como hemos visto, la caracterización que hacía el empirismo tradicional) y tras la crítica de Quine a la concepción positivista del lenguaje de la ciencia claramente dividido en una “parte fáctica” y otra “analítica”, el argumento *en pro de la dicotomía clásica entre hecho y valor se viene abajo*.

A su vez, en la parte positiva de su estrategia, Putnam ancla su análisis en nuestro lenguaje ordinario a los efectos de mostrar el carácter difuso de la distinción hecho/valor. Su idea es que si reparamos en el vocabulario de nuestro lenguaje como un *todo* – y no meramente en la parte que fue supuesta por el positivismo lógico como suficiente para la descripción de los “hechos”- vamos a encontrar que hay una *imbricación profunda de hecho y valor* incluso en el nivel de los predicados individuales. Dicha “imbricación” se patentiza, a su juicio, al considerar términos éticos “densos” (*thick*) como “cruel”. Veamos, brevemente, de qué se trata. El núcleo de su propuesta consiste en señalar el “*entretelado*” entre *descripción* y *evaluación* que estaría a la base del uso competente de tales términos. Dicho de otro modo: sostiene que lo característico, tanto de descripciones “negativas” (como “cruel”) como de descripciones “positivas” (como “valiente”), es que para poder usarlas competentemente se debe ser capaz de identificarse imaginariamente con un punto de vista *evaluativo*. Esta es la razón por la que, por ejemplo, quien piensa que “valiente” significa “no temeroso de arriesgar su vida” no podrá comprender la distinción socrática entre la simple *temeridad* y la genuina *valentía*.

Al señalar la dependencia del uso descriptivo respecto de la evaluación, Putnam distingue explícitamente su posición de aquellas que: o bien no advirtieron el problema que suscitan los términos éticos “densos” o bien no pudieron darle una resolución adecuada. Por un lado, como ejemplo paradigmático de “ceguera” frente al problema, presenta la posición de Mackie (1978) según la cual los conceptos éticos “densos” no son, en rigor, éticos o normativos sino simplemente *fácticos*. Por otro lado, encuentra en la posición de Hare (1981) un abordaje más sofisticado, pero incorrecto. Hare parece favorecer un análisis de los que denomina términos “secundariamente evaluativos” -en lugar de “éticos densos”- en un componente descriptivo y otro prescriptivo. Ahora bien, si la propuesta de Hare consiste en que es posible aislar cada uno de estos componentes, de modo que, aislando el componente descriptivo se pueda recoger- independientemente de toda evaluación- un genuino rasgo del mundo, entonces, sostiene Putnam, se trata de una empresa inviable. No obstante, Hare parece reconocer que “cruel” ha sido considerado como un término *inseparablemente* descriptivo y evaluativo pero, a juicio de Putnam, la presentación que ofrece de lo que sostienen los proponentes de la *imbricación (entanglement)* está distorsionada por la proyección de sus propios puntos de vista. Básicamente, Putnam rechaza el abordaje que Hare hace de la *imbricación* porque cree incorrecto su tratamiento del componente evaluativo como motivacional. El punto es que según la lectura motivacional de Hare, estos términos se comportarían a la manera emotivista (expresarían actitudes e intentarían despertar actitudes similares en el interlocutor) y es – precisamente- este compromiso con una teoría metaética subjetivista lo que Putnam rechaza de la reconstrucción ofrecida por Hare.

A propósito de la dicotomía hecho/valor, Putnam sostiene entonces que el problema no es sólo que la noción empirista de “hecho” fue demasiado estrecha desde el principio; un problema más profundo es que desde Hume en adelante los empiristas fallaron, en su opinión, en apreciar los modos en que la *descripción* y la *evaluación* pueden y deben estar *entrettejidos*. Su idea es que el advertir que mucho de nuestro lenguaje descriptivo es un contraejemplo viviente, tanto de la pintura empirista clásica como de la positivista lógica del campo de los “hechos”, debería sacudir la confianza de quien supone que hay una noción de *hecho* que *contrasta absolutamente* con la noción de *valor*.

En resumen, a juicio de Putnam (2002), la dicotomía hecho/valor colapsa de manera análoga a como colapsó la famosa dicotomía analítico/sintético: su punto es que persistiría la

distinción, pero no sería dicotómica porque si se toma la clase de todos los enunciados no se puede establecer una línea divisoria clara y universal. En otros términos: se puede aceptar que hay grandes rangos de enunciados que no pueden ser simplemente clasificados como verdades analíticas o enunciados sintéticos, esto es, habilitar una zona de “grises” y retener, no obstante, la idea más modesta de que hay también casos que caen en uno de los lados de la distinción: enunciados de un lenguaje que son trivialmente verdaderos en virtud de su significado y otros que no lo son. De esta suerte, Putnam asume una rehabilitación moderada de la distinción en vistas de que la noción de enunciado analítico puede ser una noción ocasionalmente útil. Análogamente, respecto de la dicotomía hecho/valor, la idea de Putnam es que si la “deflacionamos” lo que obtenemos es que hay una distinción a trazarse entre juicios éticos y otros tipos de juicio que es útil en algunos contextos –así como hay una distinción útil a trazarse entre juicios químicos y juicios que no pertenecen al campo de la química- pero *no se seguiría nada metafísico* de la existencia de la distinción hecho/valor en este sentido modesto.

Su invitación es, pues, a que reparemos en nuestro lenguaje para advertir una imbricación que no se limita a los tipos de “hecho” que reconocieron los positivistas lógicos y a valores meramente cognitivos. Así, el espíritu de fondo que atraviesa *The collapse of the fact/value dichotomy and other essays*, es la reivindicación del *dictum* de los pragmatistas clásicos según el cual los valores y la normatividad permean toda la experiencia.

Ahora bien ¿qué tipo de valores tiene en mente Putnam? Y ¿cuál es su relación con las normas? Putnam nos recuerda que los pragmatistas clásicos, Peirce, James y Mead, no se referían sólo al género de los juicios normativos que llamamos “éticos” o “morales”, sino que también los juicios sobre “coherencia”, “plausibilidad”, “razonabilidad” y lo que, en la célebre expresión de Dirac, es la belleza de una hipótesis, son todos ellos juicios normativos en el sentido de Charles Peirce, juicios sobre “lo que debe ser en el caso del razonamiento”. Frente a esta concepción, en el capítulo 7 se concentra en los intentos frustrados por parte de los diversos positivistas lógicos de evitar reconocer que la selección de teorías siempre presupone valores. Pero, observa Putnam, del mismo modo que muchos filósofos empiristas estaban determinados a mantener los ojos cerrados ante el hecho de que la ciencia presupone juicios sobre coherencia, simplicidad, belleza, naturalidad y demás, muchos autores contemporáneos que se refieren a los valores como puramente “subjetivos” y a la ciencia como puramente “objetiva”, siguen con los ojos cerrados ante el mismo hecho. Y, sin embargo, “la coherencia, la simplicidad y similares

son valores” (2002/2004:46). En este sentido, le resulta claro que los valores epistémicos son también valores. Asimismo, considera que resulta obvio que esta afirmación no equivale a negar que haya diferencia entre los valores epistémicos y los éticos. De hecho, hay diferencias incluso entre los diversos valores éticos mismos; en el Talmud, por ejemplo, la diferencia entre la *din* (justicia) y la *hesed* (compasión) divinas es descripta a veces como un conflicto en el interior del mismo ser de Dios. La preocupación que está conectada con los valores que nos guían al escoger entre diferentes hipótesis –coherencia, simplicidad, etc.- es la preocupación por “la descripción correcta del mundo”, que para muchos –dice Putnam- ha resultado ser lo mismo que la “objetividad”. Si esto fuese cierto, no sólo entrañaría que los valores éticos y los epistémicos están conectados con preocupaciones distintas, sino también que los valores éticos no están conectados en absoluto con la objetividad. Pero esto, afirma Putnam, es un error.

Su intención es dismantelar la ilusión de quien intente justificar los valores epistémicos mediante la apelación externa a una “descripción correcta del mundo”, ayuna de esos valores. Cuando afirmamos que estamos más próximos a la verdad al elegir una teoría que exhibe, por ejemplo, coherencia, simplicidad y fertilidad predictiva, en lugar de una basada en el método de la autoridad, lo que hacemos- a su juicio- es comprometernos con una hipótesis empírica compleja que elegimos guiados, a su vez, por esos mismos valores epistémicos. Con ello, no pretende expresar una posición escéptica respecto de la superioridad de estos criterios sino enfatizar que si estos valores epistémicos nos permiten describir correctamente el mundo, “ello es algo que vemos a través de los cristales de esos mismos valores”. Aunque pueda ser una justificación circular “aún es lo suficientemente justificativa para la mayoría de nosotros” (2002/2004:48). De este modo, la escisión que pretendía trazarse (desde una ecuación entre *descripción correcta del mundo* y *objetividad*) entre valores epistémicos y valores éticos ya no puede sostenerse. Esta línea argumentativa parece hacerse eco de lo ya sostenido en *Razón, Verdad e Historia*. Recordemos que allí Putnam afirma que el único criterio para decidir qué es un *hecho* es lo que es *racional aceptar* y que la concepción de la *racionalidad* es parte de la noción de *florecimiento humano*, noción propia de la dimensión ética. En otros términos, su idea es que:

si consideramos el ideal de aceptabilidad racional que se revela al examinar cuáles son las teorías científicas que tanto los científicos como la gente de la calle consideran racionalmente aceptables, observaremos que el propósito en nuestra actividad científica es construir una representación del mundo que tenga las siguientes

características: eficacia instrumental, coherencia, alcance comprensivo y simplicidad funcional. Pero ¿por qué razón? Yo respondería que la razón por la que queremos este tipo de representación [...] es que este tipo de representación forma parte de nuestra idea de florecimiento cognitivo humano y, por tanto, parte de nuestra idea de florecimiento humano total o *eudaemonia* (Putnam 1981: 137-8).

De esta suerte, su tesis de imbricación alcanza no sólo a la relación entre hechos y valores, sino, a su vez, a la relación entre los valores epistémicos y los valores éticos. Los valores epistémicos son “acondicionados” a la luz de valores éticos. Ahora bien, ¿cuál es, a su vez, la relación de los valores con las normas? En la primera parte de este trabajo, a los efectos de esclarecer el marco desde el cual considerar la normatividad de lo mental, hicimos alusión a su polémica con Habermas, a la que también vuelve en *The collapse of the fact/value dichotomy*. Veamos en qué términos lo hace.

En los capítulos 6 y 7 de su (2002), Putnam aborda dos cuestiones que tienen que ver con el fundamento de la ética. Precisamente, en el último de ellos examina la posición de Habermas, que traza –según vimos ya- una distinción tajante entre “normas” éticas universales y valores no universalizables y sostiene que sólo las primeras- las “normas”- son objetivas. La tesis de Putnam es que allí subyace, una vez más, una dicotomía indefendible y que, de hecho, la objetividad que Habermas atribuye a las normas *presupone la objetividad de por lo menos algunos valores*.

Tal es el tema del capítulo 6; allí considera la cuestión de cómo sostener la creencia en la *objetividad* del juicio ético si uno se niega (tal como él lo hace) a postular cualquier tipo de reino “platónico” especial las propiedades éticas. Sostiene que puede hallarse el fundamento para una explicación completamente opuesta a la platónica en los escritos de John Dewey.

Asimismo, esa es la empresa que acomete Putnam (2004b) en *Ethics without ontology*. Su motivación allí es defender la objetividad de los juicios morales eludiendo el compromiso con tres estrategias metafísicas. En primer lugar, desecha la que denomina “inflacionaria” y que encuentra representada paradigmáticamente por el platonismo y el intuicionismo mooreano –ie la postulación de propiedades morales *qua* propiedades no-naturales captables por una peculiar intuición-. Luego, se diferencia de dos estrategias “deflacionarias”: la “reduccionista” que establece identidad entre propiedades morales y propiedades naturales, y la “eliminativista” que sostiene que nuestro discurso moral ordinario es tan errado como el discurso alquimista o el discurso acerca del flogisto. Su compromiso es, en cambio, con lo que denomina un “pluralismo

pragmático” según el cual no es un accidente que en nuestro lenguaje cotidiano empleemos muchos tipos de discurso, sujetos a diferentes rasgos lógicos y gramaticales -diferentes “juegos de lenguaje” en el sentido wittgensteiniano-.

Una evaluación respecto de si Putnam tiene éxito en esta empresa, esto es, la de dar cuenta de la objetividad ética vía su “pluralismo pragmático” excede completamente los límites de este trabajo. Simplemente, me interesa señalar que Putnam nos recuerda que si abandonamos la idea de disputa ética “racionalmente irresoluble” no nos comprometemos por ello con el proyecto de la resolución efectiva de todos nuestros desacuerdos éticos, sino con la idea de que siempre hay posibilidad de discusión y examen ulterior. En este sentido, el aspecto más deleznable de la dicotomía hecho/valor, de la dicotomía entre una dimensión descriptiva y una dimensión normativa (en sentido amplio, esto es, tanto deóntica como evaluativa, dado que según vimos normas y valores están entretreídos), es que funciona como un punto en el que se detienen la discusión y el pensamiento.

### **7.3. La dicotomía descriptivo/normativo en la filosofía de la mente de Davidson: ¿el “cuarto dogma del empirismo”?**

Al considerar la peculiar respuesta de Davidson al problema mente-cuerpo nos detuvimos en su argumento a favor del anomalismo de lo mental. Allí advertimos que la premisa clave es aquella que reza que lo mental y lo físico son ámbitos conceptuales cuyos principios constitutivos difieren *radicalmente*. Recordémoslo brevemente:

- (i) Dos ámbitos conceptuales cuyos principios constitutivos difieren radicalmente no pueden estar conectados mediante leyes.
- (ii) Lo mental y lo físico son ámbitos conceptuales cuyos principios constitutivos difieren radicalmente: en el primero rige en *principio de racionalidad*, ausente en el segundo.
- (iii) Por tanto, no hay leyes psicofísicas porque “los predicados mentales y físicos no están hechos el uno para el otro.

A los efectos de dotar de sustento a la idea de una “diferencia radical” entre ambos dominios exploramos los distintos candidatos ofrecidos por Davidson para capturar la peculiaridad de lo mental, a saber: la irreducibilidad, la indeterminación de la interpretación y la *normatividad*. Al respecto, argumentamos que es el último rasgo, esto es, la normatividad, aquello que torna singular al vocabulario mental. En términos del propio Davidson: “el vocabulario mental no es “privilegiado” porque es irreducible sino que es irreducible porque es *normativo*” (1998/2005:318; mis itálicas). A partir del marco que presentamos en la primera parte de este trabajo, distinguimos entre dos sentidos de normatividad: un sentido *deóntico* y un sentido *evaluativo*. Hemos visto que a la vera de esta distinción, es posible deslindar las siguientes dimensiones en la normatividad de lo mental. Por un lado, nos encontramos con las condiciones de *corrección* del contenido mental y con la racionalidad como *regla o principio constitutivo* de lo mental. En efecto, dado que el contenido mental tiene *condiciones de corrección* que se juegan en la dirección de ajuste al mundo, encontramos aquí una primera dimensión de tipo deóntico de la normatividad. A su vez, hemos visto que no podemos atribuir ninguna actitud proposicional a un agente excepto dentro del marco de una teoría viable de sus creencias, deseos e intenciones, y dicho marco está gobernado por la *regla o principio constitutivo* de la racionalidad. Encontramos aquí una segunda dimensión de tipo deóntico de la normatividad.

En su intento por clarificar el rasgo normativo de lo mental, Pascal Engel (1999) apela a una comparación entre las posiciones de Davidson y Wittgenstein. Para Wittgenstein hay reglas de lenguaje, que registran el uso de las palabras por los hablantes, y por tanto hechos o regularidades acerca de estos hablantes. Pero no hay hechos adicionales acerca de lo que hace correctas a las reglas. Por su parte, si bien -como vimos en el capítulo acerca de la relación entre lenguaje y pensamiento- Davidson no siente atracción por hablar de reglas y de normas específicas del lenguaje su posición acerca del estatus de las normas de racionalidad es, a juicio de Engel, bastante similar a la de Wittgenstein: para Davidson hay una manera correcta –a pesar de la ineliminable indeterminación– de interpretar a un hablante y de hacer inteligible lo que piensa o hace, pero no hay ningún hecho acerca de lo que hace correcta a esta interpretación, en tanto no hay modo de codificar los patrones racionales que uno tiene que atribuir a un hablante en una circunstancia particular. Así, para Engel, Wittgenstein y Davidson están entre los filósofos que más han insistido en afirmar la existencia de un tipo particular de normas: las

normas interpretativas que son distintas de los tipos usuales de normas –prácticas, epistémicas o estéticas-. Ambos sostuvieron que tales normas son distintivas y que no pueden ser reducidas a hechos. Pero, frente a Engel, yo considero que si nos quedamos en la consideración de la mera normatividad deóntica, no se logra dotar cabalmente de sentido al pretendido carácter *distintivo* de las normas interpretativas, que está supuesto en la premisa (ii) del argumento en cuestión. En otras palabras, no se logra romper la analogía del modo requerido: los principios psicológicos son ellos mismos normativos en un sentido en que los principios físicos no lo son; la racionalidad involucra un *giro esencial* ausente en el caso de los principios constitutivos del ámbito físico.

En “The norms of the mental” Engel se detiene en lo que le parece un rasgo inquietante de la posición de Davidson: las normas de lo mental resultan *elusivas*. En lo personal, acuerdo con este diagnóstico inicial de Engel. También en mi opinión resultan *elusivas* o *escurridizas*, pero en un nivel diferente al considerado por Engel y por otras razones. Veamos, primero, cuál es el sentido de “elusivas” que captura Engel y que, lejos de reprobarnos, valida como necesario. Engel dice que se ha objetado que las normas de lo mental en el esquema davidsoniano son elusivas en la medida en que no pueden ser explicadas o caracterizadas de modo preciso; varios críticos se han quejado de que son demasiado inespecíficas para servir como principios efectivos de interpretación (si la mayoría de las creencias de alguien tienen que ser ciertas, esto no nos dice, como intérpretes, cuáles, y si un agente tiene que ser racional, no nos dice cómo). Pero para Engel así es como debe ser. A su juicio, Davidson a veces habla como si las normas generales interpretativas estuvieran codificadas en las teorías normativas que usamos<sup>187</sup>; a veces también alude a principios epistémicos específicos<sup>188</sup>. En tales casos, dice Engel, es importante ver cuán inespecíficos y generales son tales principios y por qué deben ser así. Si las normas de racionalidad fueran específicas en el sentido de estar “codificadas” en un conjunto de principios de los cuales pudiera derivarse la predicción estricta acerca de la conducta o pensamiento de los agentes, podrían ser aplicadas como leyes estrictas a cualquier caso particular y nos dirían qué hacer o pensar en una situación particular. De esta suerte, Engel sostiene que si su abordaje de la concepción davidsoniana de las normas de racionalidad es correcto, no tendríamos que esperar ninguna definición ni ningún tipo de especificación de las normas de lo mental. Su idea es que

---

<sup>187</sup> La lógica, la teoría de la decisión, o principios generales o leyes usados por tales teorías: la ley de no contradicción, o el principio de maximización de la utilidad esperada.

<sup>188</sup> Como el requisito de Carnap de “evidencia total” para el razonamiento inductivo, o su contraparte para el razonamiento práctico.

pensemos en las normas de la racionalidad como estructuras indefinidas en las que debe reposar una teoría de la interpretación. Así aunque no podemos decir cuáles son las normas, podemos decir qué hacen.

Ahora bien, en mi opinión, Engel parece defender el necesario carácter *elusivo* o *inespecífico* de las normas de lo mental considerando que si así no fuera, colapsaría el anomalismo de lo mental, esto es, contaríamos con leyes estrictas que posibilitarían la predicción precisa de la conducta o el pensamiento. Frente a esto, creo que Engel da por supuesto aquello que un defensor de la irreducibilidad de lo mental como Davidson tiene, precisamente, que fundamentar, a saber: su anomalismo. Decir que las normas de lo mental tienen que ser inespecíficas porque de otro modo habría reducción, no ofrece ninguna elucidación del sentido normativo peculiar que fundaría el anomalismo y que sería la razón en la que el propio Davidson ve su aporte en el debate. Más aún, incluso si se concediera la plausibilidad de interpretarlo a la manera de Engel, dado que en el mejor de los casos sólo podríamos decir *qué hacen* las normas de lo mental, no podríamos ir más allá de la dimensión deóntica de la normatividad que alude a la función constitutiva de los principios interpretativos. Pero, entonces, a mi juicio lo que se *elude* o se *escurre* es el paso conceptualmente necesario, en el nivel de la elucidación filosófica, para entender la asimetría en la que se funda la peculiaridad de lo mental. Es decir, en mi opinión lo que se *escurre* es el necesario dar cuenta del *giro esencial* que involucrarían los principios interpretativos y que haría de los conceptos mentales algo singular. Veamos más detenidamente cuáles son mis razones.

En primer lugar, recordemos que en la caracterización davidsoniana, un evento es un evento mental si y sólo si tiene una descripción mental, esto es, una descripción en la que figura esencialmente un verbo de actitud proposicional y que los eventos físicos son aquellos que seleccionamos mediante descripciones u oraciones abiertas que sólo contienen, esencialmente, el vocabulario físico (Davidson, 1970/1980: 211). Según sostuve, Davidson asume así un compromiso ontológico con eventos *desnudos*, es decir, con eventos que sólo según descripciones alternativas serán mentales o físicos. El ámbito extensional de los eventos resulta, de esta suerte, “ciego a la dicotomía mental-físico” (Davidson, 1970/1980: 215). Ahora bien, si en su esquema de eventos lo físico termina siendo de naturaleza tan lingüística o conceptual como lo mental ¿cuáles son las razones para sostener que hay una diferencia constitutiva radical entre ambos, un dualismo profundo entre los conceptos mentales y los físicos? Esto es, dado que

lo físico también se exhibe como conceptos en el esquema davidsoniano de eventos: ¿en qué sentido carecerían de *condiciones de corrección*? Y ¿en qué sentido los conceptos físicos se sustraerían al alcance del *principio constitutivo* de la consistencia? Para volver sobre la analogía que consideramos en el capítulo 4, así como la atribución de una actitud proposicional a alguien en un momento dado está constreñida por otras asignaciones de actitudes que mantengan consistencia, la asignación de una medida de longitud a un objeto particular debe respetar, por ejemplo, la ley de transitividad involucrada en la relación “más largo que” y, de este modo, está también constreñida por otras asignaciones de medida. No hay nada *esencialmente normativo* en el primer caso, al menos no hay nada que no esté presente en el segundo. En la medida en que son conceptos, tanto los físicos como los mentales tienen *condiciones de corrección* y *principios que constituyen* su rango posible de aplicación. ¿Cuál es, entonces, la explicación de la pretendida peculiaridad normativa de los conceptos mentales?

En el capítulo 4 anticipé mi respuesta: hay que elucidarla desde una dimensión *evaluativa* de la normatividad: la peculiaridad de lo mental reside en el *contenido* de su especial principio constitutivo, a partir del cual la racionalidad introduce el *valor* de considerarnos como *personas*. Recordemos la conexión conceptual íntima que Davidson establece entre *racionalidad* y *agencia*, conexión que se hace manifiesta en la teoría de la decisión racional que abrazó a lo largo de los años.<sup>189</sup> Así, pues, parece que Davidson pone en juego la noción de una *racionalidad práctica*. En esta línea, tal vez no sea excesivamente aventurado decir que Davidson está revitalizando la tesis kantiana de la primacía de la razón práctica, según la cual los valores últimos no son los cognitivos o epistémicos sino los éticos: la racionalidad está inextricablemente unida al valor de la libertad, y sería esta unión la que abriría el hiato entre esta dimensión y la de la mera necesidad natural.<sup>190</sup> Su punto es que “no habría pensamiento si los individuos no jugaran el rol indispensable e inevitablemente creativo de *árbitro final*” (Davidson, 1998b/2001: 91).

Como es sabido, Kant trató de separar la distinción *mente-cuerpo* de la distinción *persona-cosa*, argumentando que la primera está dentro del ámbito del *entendimiento* y la última dentro del ámbito de la *razón práctica*. No obstante, aún el propio uso kantiano del término “racional”

---

<sup>189</sup> En los '70 con el énfasis puesto, en mi opinión, en mostrar la peculiaridad de la psicología como ciencia y en los '80 y '90 con el énfasis puesto en el entramado atributivo cotidiano al que, a su juicio, subyace dicha teoría.

<sup>190</sup> En este sentido, vale la pena recordar que al comienzo de “Mental Events” Davidson hace expresa su simpatía para con la formulación kantiana del problema que se propone tratar: el de dar cuenta de la ausencia de una verdadera contradicción entre la libertad y la necesidad natural en las mismas acciones humanas. Cf. 1970/1980:207. Voy a volver sobre este punto en el capítulo próximo.

nos tienta a volver a juntar las dos distinciones. Davidson sigue esta línea y a menudo usa “racionalidad”, “normatividad”, “intencionalidad” y “agencia” como si fueran co-extensivos.

A mi juicio, pues, la clave para comprender la relación mente-cuerpo no está en comprender la irreducibilidad de lo mental a lo físico, sino en comprender la *inescapabilidad* de la *normatividad*. En contraste, ningún vocabulario meramente *descriptivo* es inescapable, excepto quizás el vocabulario de sentido común –por razones kantianas/strawsonianas- en que describimos los movimientos de los objetos macrofísicos perceptibles por los órganos sensoriales humanos. Podríamos haber hecho un montón de predicción y control aún si no se nos hubiera ocurrido indagar en un nivel microestructural, aún si la ciencia hubiera seguido siendo aristotélica. Pero no hubiéramos podido siquiera hablar -en términos de sentido común o aristotélicos o corpusculares- a menos que no nos hubiéramos tratado ya mutuamente como *personas*, que tienen deberes de responder a ciertas situaciones con ciertas palabras en lugar de otras.

En efecto, en el capítulo en el que me ocupé de la respuesta davidsoniana al problema mente-cuerpo, argumenté que la dimensión deóntica propia del principio constitutivo de la racionalidad sólo se comprende cabalmente a partir de contemplar la dimensión evaluativa en la que se engarza: en nuestra necesidad de hallar comprensible a alguien, trataremos de construir una teoría que “la haga consistente, *creyente de verdades y amante de lo bueno*” (1970/1980: 222; mis *itálicas*). De lo que se trata es de advertir, tal como hace Bill Martin (1999), que el monismo anómalo no es sólo un principio de la filosofía de la mente, sino un principio ético-político. Sostuve, entonces, que el valor que pone en juego la racionalidad es el de tratarnos mutuamente como *personas*.

Ahora bien, si asumimos la plausibilidad de llenar la laguna conceptual de Davidson en torno al nivel normativo peculiar de lo mental desde una concepción de la racionalidad como introduciendo el valor de considerarnos como *personas*, todavía resulta necesario adherir a un supuesto adicional para defender el carácter “*irreduciblemente* normativo” de lo mental (Davidson, 1974c/1980: 241). En efecto, podría argüirse, tal como sostiene Putnam (2002) que las normas son tales a la luz de valores y que éstos se encuentran entretreídos con los hechos, de modo que el reconocimiento de una dimensión normativa y una descriptiva no conduciría *per se* necesariamente a un dualismo conceptual profundo, del tipo que Davidson pretende establecer entre lo mental y lo físico. Por tanto, la diferencia radical entre los principios constitutivos de lo

mental y de lo físico se sigue en la medida en que repose en el compromiso con una dicotomía (y no una mera distinción) entre la dimensión descriptiva y la evaluativa. Efectivamente, en varios pasajes Davidson asume un compromiso explícito con esta dicotomía.<sup>191</sup>

En suma: en su peculiar respuesta al problema mente-cuerpo, a los efectos de respaldar un dualismo profundo entre los conceptos mentales y los conceptos físicos, Davidson sostiene una diferencia tajante entre la dimensión evaluativa propia de la racionalidad y la dimensión descriptiva del mundo de la necesidad natural, posición que resulta difícil de conciliar con su inscripción general en la tradición crítica de las dicotomías.

Por su parte, cuando consideramos el problema de la relación entre el lenguaje y el pensamiento, nos detuvimos en las aplicaciones que ofrece Davidson de la noción de *triangulación*. Primero, como una descripción de las condiciones necesarias (completamente no intencionales) para el pensamiento; segundo, como una descripción de la situación de aprendizaje en la que una criatura aprende su primer lenguaje (la situación en la que el pensamiento y el lenguaje “emergen”); y tercero, la situación en la que se determinan los contenidos de nuestras creencias perceptuales. Señalamos, asimismo, que hay una situación triangular adicional, la original en Davidson: la situación de interpretación radical, es decir, cuando dos personas no comparten el mismo lenguaje. El punto que, entonces, advertimos es que dado el rasgo peculiar que atraviesa esta dimensión propiamente proposicional, esto es, su *normatividad*, no podemos conectarla (y cabe recordar que se trata de una imposibilidad por principio) con el primer triángulo (causal/físico), tal es el sustento del anomalismo de lo mental que acabamos de considerar. Caso contrario, la reducción sería, en principio, posible. Ahora bien, en este escenario, los dos participantes ya tienen los conceptos de contenido objetivo y de verdad; saben lo que es un objeto físico y en qué consiste que una creencia sea verdadera o falsa. Por su parte, vimos que tampoco podemos conectar conceptualmente el triángulo causal/físico con aquel que se pone en juego entre el niño y el adulto que le enseña la primera lengua. Una vez más, aquí nos topamos con la *normatividad* como aquel rasgo que abre el hiato con la dimensión meramente descriptiva del vocabulario causal/físico. Veamos en qué sentido.

---

<sup>191</sup> Así, por ejemplo, se vale de ella en “Deception and Division” a instancias de discernir entre pensamiento desiderativo y autoengaño, por un lado, y debilidad de la justificación, por otro. En consonancia con esto emplea la dicotomía para trazar una distinción entre tener una razón (evaluativa) para creer cierta proposición y tener una evidencia (cognitiva) a cuya luz es razonable considerar verdadera la proposición. Cf. Davidson, 1986b/2004.

A juicio de Davidson, la interacción entre el adulto y el niño en la situación del aprendizaje ostensivo provee las condiciones necesarias para la emergencia del lenguaje y el pensamiento proposicional, al crear un espacio en el que el *error* sea posible. Leemos en él:

Es difícil exagerar la magnitud del paso desde la disposición innata o aprendida a responder a estímulos de un cierto tipo, a emplear un concepto con consciencia de la posibilidad de error. Es el paso desde reaccionar a estímulos proximales a pensar acerca de objetos y eventos distales, el paso desde la mera respuesta condicionada hasta lo que Wittgenstein llamó “seguir una regla (1997e/2005:14).

La idea que está detrás de esta afirmación es que uno no puede fallar al actuar de acuerdo con una disposición, pero uno puede fallar al seguir una regla. Estamos suponiendo, dice Davidson que:

el que aprende tiene muchos conceptos a su alcance, ¿pero cuál es el *correcto*? Solamente quien le enseña puede determinarlo. Se necesitan dos para triangular la ubicación del estímulo distal, dos para una prueba objetiva de corrección o error. Si esto es así, debe serlo desde el comienzo, antes de que quien aprende tenga la idea de objetividad cuando está aprendiendo, por medio de ostensiones lo que para empezar no son más que maneras de establecer asociaciones, qué es *juzgar* que una respuesta es correcta o es errónea. La posibilidad del pensamiento viene en compañía (1998b/2001:88).

Sin un intérprete, pues, no se puede dar contenido a la afirmación de que el hablante se ha equivocado. Ahora bien, llegados a este punto cabe recordar sus diferencias con Kripkestein, consideradas en la primera parte de este trabajo. Al respecto se pregunta Davidson:

Pero al eliminar la condición de que el hablante debe seguir como el intérprete (u otros) lo haría(n), ¿no hemos destruido al mismo tiempo inadvertidamente toda oportunidad de caracterizar el error lingüístico? ¿Si no hay práctica social con la cual comparar la actuación del hablante, no ocurrirá que sea lo que sea lo que diga el hablante éste estará, como señala Wittgenstein, en concordancia con alguna regla (esto es, en concordancia con alguna lengua)? (1992b/2001: 116)

En otras palabras, si el comportamiento lingüístico de los demás no proporciona la *norma* al hablante, ¿qué lo puede hacer? Al considerar el problema de la relación entre el lenguaje y el pensamiento, hemos visto que Davidson se opone explícitamente a apelar a la dimensión deóntica de las *convenciones o reglas* a la hora de ofrecer su elucidación de qué es hablar una lengua. En efecto, hemos advertido que afirma que:

Deberíamos preguntarnos sobre la adecuación del concepto de seguir una regla para describir lo que se supone que es hablar una lengua. Cuando hablamos de reglas del lenguaje, usualmente pensamos en las descripciones que los gramáticos o los lingüistas hacen de la práctica (generalizada e idealizada), o en las prescripciones que los gramáticos desean que se sigan. Las reglas pueden servir de ayuda para aprender una lengua, pero sólo podemos disponer de esa ayuda en la adquisición de una segunda lengua. La mayor parte del aprendizaje de cómo usar las palabras se logra sin aprender explícitamente ningún tipo de regla (1992b/2001:113).

Hasta aquí, hemos visto que aquello que introduce el “paso”, cuya magnitud “no puede exagerarse”, entre el primer triángulo y aquel en el que interactúan el adulto que enseña y el niño que aprende es la dimensión *deóntica* de la normatividad que se pone en juego en el concepto de *corrección* como la otra cara del *error* posible. En la primera etapa del aprendizaje ostensivo, el error no tiene sentido para el que aprende, porque no hay nada acerca de lo cual estar equivoocado, y “donde el error no tiene sentido, no hay un concepto o pensamiento” (1997e/2005:14). Ahora bien, a la hora de precisar en qué sentido la interacción con el adulto provee una “norma” que permita percatarse de la distinción entre lo meramente creído y lo que es el caso, Davidson rechaza abiertamente, la otra dimensión deóntica que hemos considerado, a saber, que se pueda elucidar en términos de “reglas”. Hemos señalado ya, que parece apuntar a una dimensión no meramente deóntica de la normatividad, sino a los *valores ulteriores* en los que reposa nuestra comprensión mutua. Según vimos en el capítulo dedicado al problema de la relación entre el lenguaje y el pensamiento, Davidson afirma que quienes insisten en que las prácticas compartidas son esenciales están parcialmente en lo correcto: debe haber un *grupo interactivo* para que emerjan el lenguaje y el pensamiento proposicional. Y dice: “dado un terreno suficientemente común podemos comprender y explicar las diferencias; podemos

criticar, comparar y persuadir. El punto central es que encontrar el terreno común no es resultado de la comprensión, sino una condición de ella: si comprendemos sus palabras, existe un terreno común, un “modo de vida compartido” (*shared “way of life”*; Davidson 1984c/2004:37). Y sostiene: “así como con la creencia, no hay espacio para relativizar los valores, o para preguntar si las comparaciones interpersonales de valores son posibles. La única manera que tenemos para saber cuáles son los valores de otro es la que –como en el caso de la creencia- se constituye en un marco común (*common framework*; Davidson 1995c/2004:50). De este modo, “las criaturas racionales son necesariamente habitantes del mismo mundo objetivo que nosotros y necesariamente *comparten sus principales valores (leading values)* con nosotros” (1984c/2004:36). Ahora bien, en la primera parte de este trabajo, argumenté que uno de los puntos de contacto entre las posiciones de Putnam y Davidson consiste en su defensa de la objetividad de los valores. Al respecto Davidson aclara que no está diciendo que “los valores son objetivos porque hay más acuerdo que el que “encuentra el ojo” y ciertamente tampoco está diciendo “que aquello en lo que acordamos es por tanto verdadero. La importancia de un trasfondo de creencias y valores compartidos (*background of shared beliefs and values*) es que tal trasfondo nos permite dotar de sentido a la idea de un *estándar común de lo correcto y lo erróneo*, lo verdadero y lo falso” (1995c/2004:51; mis itálicas). El error puede ser inteligible “sólo si hay un adecuado trasfondo de acuerdo (*adequate background of agreement*)” (1993d/2005:43).

En consecuencia, según argumentamos oportunamente, resulta que la dimensión *deóntica* de la normatividad, que se exhibe en el argumento que ofrecimos en el capítulo correspondiente al problema de la relación entre pensamiento y lenguaje, esto es, el argumento a partir del error (según el cual para tener el concepto de creencia se debe tener el concepto de error y para ello se debe estar en comunicación con otros) supone una dimensión *evaluativa* de la normatividad en el siguiente sentido: sin un trasfondo de valores compartidos que constituyen un terreno común, la comunicación no podría tener lugar. Por tanto, la estrategia davidsoniana para defender la “magnitud” nunca “exagerada” del paso desde el triángulo inicial, a aquel en el que emergen el pensamiento y el lenguaje, deviene un “salto”, un “paso de prestidigitación” en la medida en que, una vez más, la argumentación reposa en una dicotomía (y no una mera distinción) entre la dimensión *descriptiva* de los conceptos físico-causales propios del primer triángulo y la

dimensión *evaluativa* que posibilita la emergencia del pensamiento y el lenguaje en el segundo triángulo (niño-adulto) y que se exhibe cabalmente en el tercero (entre adultos normales).

Ahora bien ¿cuáles son los valores que constituyen el trasfondo requerido por la comprensión? Davidson se encarga de enfatizar que hay un “lugar de partida y el lugar concierne a los valores o deseos. No sería errado decir que las actitudes evaluativas y las acciones que las revelan, forman el fundamento (*foundation*) de nuestra comprensión del habla y la conducta de los otros” (1984c/2004:35). Enfatiza que “no hay una lista fija de estándares, ni una jerarquía eterna de valores a los que todos debamos suscribir” (1993/2005:169), pero al mismo tiempo subraya que “tendríamos que esperar que los valores iluminados (*enlightened values*) converjan...tendríamos que esperar que la gente iluminada y que se comprende completamente entre sí acuerde en sus *valores básicos*” (1995c/2004:49). En el capítulo relativo al problema de la relación entre el pensamiento y el lenguaje, argumenté que el valor básico que subyace al tipo de acuerdo (*agreement*) que posibilita la comprensión (*understanding*) es el de considerarnos como *personas*. En efecto, el propio Davidson alude al valor en cuestión al referirse a “nuestros valores y creencias fundamentales” entre los que cuenta –además de la noción de persona- al “conocimiento”, la “virtud” y la “honestidad” (cf.1994b/2005:259).

Ahora bien, si también a propósito del problema de la relación entre el pensamiento y el lenguaje asumimos la plausibilidad de llenar la laguna conceptual de Davidson en torno al nivel normativo peculiar de lo mental desde una concepción de la racionalidad como introduciendo el valor de considerarnos como *personas*, todavía resulta necesario adherir a un supuesto adicional para defender el carácter “*irreduciblemente* normativo de lo mental” y sostener que hay un problema conceptual quizás “insuperable” a la hora de describir la aparición del lenguaje y el pensamiento: la existencia de una dicotomía (y no una mera distinción) entre la dimensión descriptiva del primer triángulo y la normativa que se exhibe cabalmente en el que tiene lugar entre adultos normales que nos tratamos como *personas*.

Una vez más, nos topamos aquí con una dicotomía que es difícil de compatibilizar con la posición crítica de Davidson frente a las dicotomías tradicionales.

Por su parte, al abordar el problema de las perspectivas de primera y tercera personas en la atribución mental, hicimos manifiesto que si bien Davidson no interviene explícitamente en el debate, su posición suele interpretarse en términos “teoricistas” que lo ubican defendiendo una perspectiva de “tercera persona”. Frente a esta lectura estándar de su posición, nos detuvimos en

precisiones que él mismo ofrece (especialmente en Davidson 1986/2005 y 1998/2005) y que, según vimos, desalientan ese enrolamiento. En vistas de ello, concluimos que allí donde logra ofrecer una articulación precisa de las nociones de “subjetividad” y “objetividad”, no puede decirse lo mismo de su tratamiento de la “intersubjetividad”. En efecto, en mi opinión, Davidson no logra darle sustento a una afirmación crucial como la siguiente:

Así al introducir explícitamente al intérprete completamos en un microcosmos la situación social que da contenido a la idea de estar *en lo correcto o equivocado* acerca de un mundo público compartido. Un intérprete no es un *espectador ocioso (idle bystander)*: es un *jugador esencial (essential player)* en una actuación (*performance*) que requiere interacciones causales complejas entre la gente y el mundo (1995c2004: 49; mis itálicas).

En este pasaje, Davidson conecta la normatividad entendida desde la dimensión *deóntica* concerniente al concepto de *corrección* (contracara del concepto de error) con la introducción de un intérprete que no sea un mero “espectador ocioso” o “desinteresado”, esto es, en términos de la polémica que vimos en el capítulo anterior, con la introducción de una perspectiva que no sería propia de la “tercera persona”. Tal como anticipé, la idea de un “*participante*” o “jugador esencial” requiere abandonar el planteo tradicional en que se estableció el debate en torno a la atribución mental, esto es, como restringido a las perspectivas de primera y tercera persona, para dar cabida a una genuina *perspectiva de segunda persona*. Sólo entonces la normatividad deóntica a la que aludimos más arriba quedará adecuadamente caracterizada. Recordemos, brevemente, la manera en que dicha perspectiva pretende diferenciarse de las dos canónicas.

Por un lado, para establecer la diferencia entre las perspectivas de primera y segunda persona, se suele acudir a la distinción entre la *empatía* y la *simpatía*. El fenómeno de la simpatía (como reseña Wispé (1986) en su ya clásico artículo, donde intenta distinguir ambos conceptos) consiste en la conciencia y *participación* en el sufrimiento o, más ampliamente, en las emociones de otra persona. Se distingue de la empatía que es el intento de comprender estados positivos o negativos del otro. La etimología de estas dos expresiones sugiere una diferencia de significado – a pesar de que en la literatura se emplean con cierta frecuencia de manera intercambiable–: el prefijo “sim” significa “con”, “junto con”, mientras que el prefijo “em” significa “en”. Wispé

describe la diferencia entre empatía y simpatía del siguiente modo: “En la empatía el yo es el vehículo para la comprensión, y nunca pierde su identidad. La simpatía, por otra parte, tiene más que ver con la comunión que con la certeza (...) En la empatía uno se sustituye a uno mismo por otra persona; en la simpatía uno sustituye a otros por uno mismo. Saber cómo se sentiría ser esa persona es empatía. Conocer lo que sería ser esa persona es simpatía. En la empatía uno actúa “como si” uno fuera la otra persona (...) el objeto de la empatía es la comprensión. El objeto de la simpatía es el bienestar de la otra persona. En suma, la empatía es un modo de conocimiento; la simpatía es un modo de relacionarse.

La empatía es, como se sabe, uno de los mecanismos básicos que, según los teóricos de la simulación, explican el funcionamiento de la atribución mental en la *folk psychology*, oponiéndola a los procesos inferenciales postulados por la Teoría de la Teoría. Se trata de un mecanismo de primera persona consistente en la capacidad para comprender a otro y no se limita a una experiencia afectiva, aunque puede incluirla. La simpatía, en cambio, implica el experimentar el estado mental o emocional del otro, por lo tanto, es compartir una experiencia con alguien más. Al constituir un comportamiento social básico antecede a la empatía y no al revés (aunque, desde luego, no todos coinciden en explicar las cosas de este modo). Goldie (2000) asemeja la simpatía a la empatía por presuponer ambas alguna clase de comprensión, negando en cambio este rasgo al contagio emocional, el que no podría incluso ser descripto tampoco como una respuesta en algún sentido *apropiada* a las conductas de los demás. Goldie avanza aún más sosteniendo que para comprender las emociones de los demás no se requiere ser capaz de contagiarse emocionalmente (pp. 177-219). La empatía y otros procesos implicados en la adscripción mental que discuten los simulacionistas son mecanismos y habilidades básicamente intelectuales, no emocionales. Lo cierto es que, podría alegarse, que la simpatía es un fenómeno intencional más básico y primitivo, como bien reconocía Wittgenstein: “Lo que es esencial para nosotros es, después de todo, el *acuerdo* espontáneo, la *simpatía* espontánea (*Investigaciones Filosóficas*, # 699; mis itálicas). Creo que no es aventurado pensar que, en esta misma línea (es decir, apelando a una perspectiva de segunda persona) podría ofrecerse una articulación del tipo de *acuerdo* que, a juicio de Davidson, subyace a la comprensión y que deja sin precisar, en la medida en que sólo se atiende a la perspectiva de “tercera persona” (pese a vislumbrar la riqueza de introducir una alternativa que sólo “bautiza” sin explorar).

Por su parte, y respecto de diferenciar la perspectiva de segunda persona de la de tercera, recordemos que Wittgenstein cuestionó también los intentos de interpretar las competencias ordinarias para la comprensión mental como basadas en alguna clase de *teoría* (en el fragmento #223 de *Zettel* sostiene explícitamente que "...el "lenguaje ingenuo", nuestro modo de hablar ingenuo, normal, no contiene teoría alguna (...) acerca de ningún fenómeno psicológico, "no muestra ningún tipo de teoría", (tampoco una "teoría popular") sino tan sólo conceptos mediante los cuales hablamos acerca de ellos). Como señalamos en el capítulo anterior, para estos enfoques demasiado intelectualistas de la comprensión mental, Wittgenstein proponía un remedio no-teórico: " (...) uno quisiera decir: Vive un tiempo entre nosotros y lo entenderás " (*Observaciones sobre la Filosofía de la Psicología, II, # 630*). Frente a una concepción como la davidsoniana, para muchos demasiado tensionada por una visión "teoricista", la atribución mental desde la perspectiva de segunda persona, es un conjunto de habilidades o una competencia compleja para la comprensión recíproca, cuyo desarrollo y expresión se da en contextos interactivos, a la vez públicos, sociales y prácticos. Como ya advertimos, es evidente que la noción misma de "segunda persona" tiene su clara inscripción en la gramática de los pronombres personales (que como señala Scotto es un primer indicador vago pero correcto de la noción que se quiere considerar), esto es, es la categoría con la cual nos referimos a otro que está *próximo*, frente a quien se está en cierta relación y puede constituirse en un "jugador esencial" o "participante", distinguiéndose así de la "tercera persona" o "espectador ocioso" y de uno mismo, el "yo" o "primera persona". La atribución intencional, desde la perspectiva de segunda persona, ocurre así en un marco interactivo o incluso, en sus manifestaciones más simples, es ella misma una clase de interacción. A la vera de esta noción se abre, entonces, también la posibilidad de precisar el tipo de interacción que a juicio de Davidson es propia del grupo (*interacting group*, 1994/2005: 125) que está a la base de la comprensión.

Ahora bien, según vimos en el capítulo anterior, pese a que Davidson parece avizorar la riqueza conceptual que ofrece la perspectiva de *segunda persona*, no la despliega. Asimismo, advertimos que a diferencia de los otros dos problemas considerados (el problema mente-cuerpo y el problema de la relación entre el pensamiento y el lenguaje) Davidson no interviene explícitamente en el debate en torno a la atribución mental. Sin embargo, a partir de lo que consideramos hasta aquí podemos aventurar una hipótesis plausible acerca de su "ceguera" frente a la capacidad explicativa de esta perspectiva. Las dimensiones subjetiva y objetiva, esto es, las

perspectivas de primera y tercera persona que él explora, son dos caras de la misma “moneda” que se pone en juego en la interpretación que tiene lugar entre adultos normales, esto es, ahí donde los conceptos de actitud proposicional tienen una cabal aplicación y poseen la *normatividad* como rasgo constitutivo. En otras palabras: tanto la presunción a favor de la verdad del hablante, como la adopción de la perspectiva del intérprete, se revelan como condiciones necesarias de la interpretación. Dado su carácter “irreduciblemente normativo” se clausura por principio toda posibilidad de dar cuenta de cómo emergen. Recordemos que dado que “no tenemos una manera de hacer corresponder los eventos descritos con el vocabulario de lo físico con eventos descritos con el vocabulario mental y no podemos esperar encontrarla...Esto quiere decir que quizás haya un problema insuperable en dar una descripción plena de la aparición del pensamiento. Doy gracias de no estar en el campo de la psicología del desarrollo” (1997c/2001:128). La imposibilidad de vislumbrar la riqueza conceptual de la perspectiva de segunda persona tiene que ver, en mi opinión, con que recoge un tipo de interacción que relaciona aquello que, según Davidson, por principio no podría conectarse: las relaciones causales con otro próximo posibles de ser recogidas por un vocabulario meramente *descriptivo* con las relaciones *normativas* que tienen lugar en la instancia interpretativa normal entre adultos. Aquello que impide, pues, la posibilidad de elucidar la normatividad de lo mental en el sentido pretendido por Davidson, esto es, mediante su conexión con un “jugador” o “participante esencial” (y no un mero observador o espectador desinteresado u ocioso) es, una vez más, el compromiso de Davidson con la dicotomía (y no una mera distinción) entre una dimensión descriptiva y una normativa.

Como podrá apreciarse a esta altura, la consecuencia más perniciosa que, en mi opinión, posee su adhesión a dicha dicotomía es que (como observa Putnam 2002) detiene la discusión. En efecto, vimos que a propósito del problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de la primera y la tercera persona de la atribución mental, la discusión se clausura *por principio*. Esta consecuencia, constituye a mi juicio, una buena razón para abandonar la dicotomía en cuestión. Veamos, más detenidamente, sobre qué pilares se la puede rechazar y qué horizonte conceptual se abre en caso de hacerlo.

Como cuestión preliminar, cabe señalar que una estrategia posible es la de quienes apuntan a socavar la dicotomía descriptivo-normativo *per se*, esto es, en términos absolutos, tal como es el caso de Putnam (2002), cuyos lineamientos centrales reconstruimos en la sección anterior, y

cuyo propósito es más ambicioso que el que persigo en el presente trabajo. En efecto, mi argumentación en contra de la dicotomía descriptivo-normativo es, en cambio, relativa a la contribución de Davidson en filosofía de la mente. En este sentido, creo que mi indicación debe leerse con cautela. Como señalé al principio de este capítulo, tradicionalmente se considera que una dicotomía es exclusiva y excluyente. Mi sugerencia a propósito de despojarnos de la dicotomía descriptivo/normativo presentará las dos siguientes salvedades: por un lado, pretendo que alcance sólo a las singulares respuestas que ofrece Davidson a los tres problemas aquí considerados (esto es, no pretendo hacerla extensiva a otros dominios de reflexión, al menos no en el presente trabajo); por otro lado, me interesa retener los polos de la distinción pero prescindiendo, simplemente, de su carácter excluyente, es decir, habilitando un espacio conceptual en que lo descriptivo y lo normativo puedan entrelazarse (en consecuencia, no cuestionaré su carácter exclusivo y no propondré, por tanto, una tercera alternativa).

En primer lugar, creo que hay una interpretación plausible de la posición de Davidson que hace que su adhesión a la dicotomía descriptivo-normativo se revele como inconsistente – o al menos, como entrando en una fuerte tensión- con el resto de su sistema. En efecto, creo que es posible ofrecer una lectura de su contribución en clave pragmatista. En la primera parte de este trabajo, señalamos que *-inter alia-* Glock propone una interpretación en esa dirección. Sostiene que Quine y Davidson pueden considerarse “*pragmatistas lógicos*”: han estado influidos por el positivismo lógico, por un lado, y por el pragmatismo americano, por otro. Lo que conecta sus posiciones es, como hemos dicho, una antropología filosófica, una concepción singular del ser humano y de la acción humana. A su vez, al rechazar el dualismo esquema-contenido que el empirismo lógico daba por supuesto, Davidson, dice Glock, mantuvo la lógica pero se desembarazó del empirismo, mantuvo la atención en el lenguaje y se deshizo de la epistemología. Nos permitió usar las intuiciones de Frege para confirmar las doctrinas holista y pragmatista de Dewey. Su trabajo hace posible el tipo de síntesis de pragmatismo y positivismo que Morton White previó como posible “*reunion in philosophy*”. Por su parte, Rorty (1986) caracteriza al pragmatismo como un movimiento que se ha especializado en dismantelar dualismos y en disolver los problemas tradicionales creados por esos dualismos. A su juicio, la cercana filiación de la obra de Davidson en la de Quine y, a su vez, la de Quine en la obra de Dewey, torna tentador verlo a Davidson como perteneciente a la tradición pragmatista americana. En este sentido, bajo la acepción propuesta por Rorty, creo que es plausible pensar

en Davidson como un “pragmatista”<sup>192</sup>, que ha adherido a las críticas quineanas hacia el reduccionismo neopositivista y la dicotomía analítico-sintético y ha denunciado la dicotomía esquema-contenido. Así leemos:

En vez de decir que es la dicotomía esquema-contenido la que ha dominado y definido los problemas de la filosofía moderna, podríamos entonces también decir que es la manera en que se ha concebido el dualismo entre lo objetivo y lo subjetivo. La razón de ello es que ambos dualismos tienen un origen común: un concepto de mente con sus estados y objetos privados.

He llegado al punto al que me estaba dirigiendo puesto que me parece que el cambio más promisorio e interesante que ocurre en filosofía actualmente es que estos dualismos se están poniendo en cuestión de nuevas maneras o que se están redefiniendo radicalmente. Hay una buena oportunidad para que sean abandonados, al menos en su forma presente. Estamos asistiendo a la aparición de una visión revisada de la relación entre la mente y el mundo (1988/2001:43).

En la primera parte de este trabajo, he sostenido el carácter crucial que posee su crítica a la concepción cartesiana de la mente. Al dismantelarla, Davidson pretende poner en jaque los dualismos modernos que reposan en ella. De este modo, no sólo comparte la crítica quineana al dualismo analítico/sintético (y, con ella, el repudio a la distinción entre creencia y significado y entre teoría integral y lenguaje; cf. 1994c/2005:81), sino que va más allá y denuncia el dualismo entre esquema y contenido como el “tercer dogma del empirismo” (1974/1984: 189). Ahora bien, si la lectura que propongo es correcta, la peculiar concepción de la mente “social e interactiva” que propone sobre la cartesiana tradicional reposa, a su vez, en la dicotomía entre la dimensión descriptiva de los conceptos físicos y la dimensión normativa “esencial” de los conceptos mentales. Dada la historia en la que dicha dicotomía hunde sus raíces -según el relato de Putnam (2002) que reconstruí en el apartado anterior- podríamos bautizarla como el “cuarto dogma del empirismo”.

---

<sup>192</sup> Soy plenamente consciente de que Davidson ha negado de manera explícita que su ruptura con la tradición empirista lo convierta en pragmatista (1984: xviii). El concibe el pragmatismo en términos de una identificación de verdad con asertabilidad bajo condiciones ideales. Pero, como advierte Rorty, si tal identificación es esencial al pragmatismo, en efecto, Davidson es tanto anti-pragmatista como anti-empirista. No obstante, bajo la acepción más acotada que propone Rorty, coincido con él en que es tentador interpretar a Davidson como un pragmatista.

Mi invitación es, entonces, a que despojemos su filosofía de la mente del dualismo en cuestión. En mi opinión, al desembarazarlo de la dicotomía entre una dimensión descriptiva y una dimensión normativa creo que tornamos más “davidsoniano” a Davidson al hacerlo más “consistente” con su propia inscripción en la tradición crítica de las dicotomías tradicionales. Luego de los tres célebres “dogmas” del empirismo, a saber: el reduccionismo neopositivista, la distinción analítico-sintético y el dualismo esquema-contenido, la dicotomía descriptivo/normativo podría concebirse como un “cuarto dogma” que tiene la consecuencia indeseable, a mi juicio, de detener -por principio- la discusión acerca de cómo aparece en nosotros lo que nos hace peculiares. Así, si logramos deshacernos de dicha dicotomía, desmantelamos la condición de posibilidad para que surja el “problema conceptual” sobre el que nos advierte Davidson, y abrimos con ello un nuevo horizonte conceptual.

Con todo, considero que más allá de estas razones “externas” que reposan en la posible filiación de la filosofía de Davidson en la tradición pragmatista americana, hay razones “internas” a su propio planteo que instan al abandono de la dicotomía descriptivo/normativo. Veámoslo más detenidamente.

En la primera parte de este trabajo vimos que un rasgo central de la posición defendida por Davidson es que el recurso al principio de caridad no es una opción disponible al intérprete que éste pueda tomar o dejar de lado, sino que la asunción de una perspectiva caritativa es condición de posibilidad de toda interpretación. La interdependencia entre creencia y significado impone el principio de caridad al intérprete, y la única alternativa real a la aplicación del mismo es el colapso total de la empresa interpretativa. No resulta adecuado, entonces, pensar la atribución caritativa de creencias verdaderas y coherentes (según el juicio del intérprete) como una suposición o una hipótesis, ya que de ese modo sugeriríamos que, al igual que con cualquier otra hipótesis, podríamos eventualmente contrastarla o al menos concebir las circunstancias en que podríamos, en principio, llegar a reconocer su falsedad o su adecuación. En el caso de la caridad, sin embargo, esta posibilidad se encuentra por principio clausurada, ya que para reconocer la falsedad de las atribuciones que la caridad demanda tendríamos que poder contar con alguna medida contra la cual poder evaluar su corrección, o al menos concebir la posibilidad de tal medida. El problema radica en que esta medida sólo podríamos encontrarla al interior de alguna interpretación, y la conclusión del análisis ofrecido por Davidson es precisamente que

seguir los dictados de la caridad es, ante todo, condición de posibilidad de toda interpretación. Según afirma Davidson:

La política de acomodamiento racional o caridad en la interpretación no es una política en el sentido de ser una entre muchas políticas exitosas posibles. Es la única política disponible si queremos entender a los demás. De este modo, en lugar de llamarla una 'política', haríamos mejor en concebirla como una forma de expresar el hecho de que las criaturas con pensamientos, valores y habla deben ser criaturas racionales, deben habitar necesariamente el mismo mundo objetivo que nosotros habitamos, y deben compartir necesariamente los valores que nos guían (1984c/2004: 36).

Creo que se imponen dos comentarios, al menos, para aclarar el sentido preciso de esta observación. En primer lugar, como observamos oportunamente, que el intérprete deba proceder bajo la suposición del acuerdo entre sus creencias y las del hablante no implica que haya ningún asunto particular sobre el que no sea concebible la diferencia de opinión. La tesis de Davidson es que toda diferencia local, dondequiera que se encuentre, deberá ubicarse dentro de un panorama de acuerdo global (so pena de ininteligibilidad), aún cuando no sea posible ni necesario señalar *a priori* ningún conjunto determinado de creencias sobre el que ambos deban coincidir.

En segundo lugar, el marco de acuerdo prescripto por la caridad debe alcanzar a *todo* el conjunto de las actitudes proposicionales del hablante. Según la formulación de esta idea en "Mental Events", el principio de caridad no nos demanda únicamente encontrar que los hablantes interpretados son (mayormente) consistentes y que sus creencias son (mayormente) verdaderas, sino que exige igualmente que sean, a juicio del intérprete, "*amantes de lo bueno*" (1970/1980: 222). Esta extensión del alcance de la caridad resulta inevitable si creemos, siguiendo a Putnam (2002), que el colapso de la dicotomía entre lo descriptivo y lo valorativo está dado por el holismo resultante de rechazar la dicotomía entre lo analítico y lo sintético. En el caso Davidson, la conclusión no es requerida sólo por su adhesión al holismo semántico (el 'holismo intra-actitudinal' que examinamos en la primera parte de este trabajo), sino que es al mismo tiempo reforzada por su tesis de la imbricación entre las diferentes actitudes que asumimos ante esos contenidos (el 'holismo inter-actitudinal', que denominamos también como "holismo psicológico"), que vuelve dramática la diferencia entre tener una actitud proposicional y no tener

ninguna. Este doble aspecto del holismo propio de lo mental bloquea todo intento de restricción de la aplicación de la caridad a un subconjunto señalado de actitudes proposicionales. Si la interpretación ha de ser posible, entonces, el intérprete deberá proceder según el principio normativo de la caridad, y si ha de interpretar caritativamente, ha de suponer que las emisiones del hablante expresan un conjunto ampliamente coherente de pensamientos, que sus creencias son mayormente verdaderas y que sus *valores son mayormente los correctos*, según sus propios estándares. Hay, pues, un derrotero conceptual al interior de la propia propuesta davidsoniana que parece conducir desde el rechazo de la dicotomía analítico-sintético a la adhesión de un compromiso holista que hace colapsar el trazado de una distinción tajante entre elementos que serían meramente descriptivos y elementos valorativos.

Adicionalmente, en polémica con Rorty el propio Davidson considera que la dicotomía descriptivo/normativo es un lastre del viejo positivismo:

Estoy sorprendido al encontrar que Rorty, quien se opone tanto a distinguir el vocabulario mental del vocabulario de la ciencia natural, compe la vieja distinción positivista entre el “vocabulario *descriptivo* de la intencionalidad” y el “vocabulario *prescriptivo* de la normatividad” (1998/ 2005: 320; mis itálicas).

Frente a esta dicotomía él afirma que lo mismo que dijo acerca de las creencias lo sostiene respecto de los valores morales:

antes de que podamos decir que dos personas están en desacuerdo acerca del valor de una acción u objeto, debemos estar seguros de que es la misma acción u objeto y que son los mismos aspectos de esas acciones u objetos los que tienen en mente. Las consideraciones que prueban la disputa genuina –las consideraciones que conducen a la interpretación correcta–revelarán también los criterios compartidos que determinan dónde yace la verdad (Davidson 1995c/2004:47).

Y desafiando abiertamente la dicotomía descriptivo/normativo, propia del empirismo clásico y lógico, afirma explícitamente la siguiente tesis de “imbricación”: “los valores están enraizados en las cosas” (*value are rooted to things*; 1995c/2004: 52).

Ante este estado de la cuestión mi propuesta es, pues, la siguiente. Si el trazado de las dimensiones descriptiva y normativa resulta insostenible como dicotomía incluso al interior de la propia filosofía de la mente de Davidson, podemos conservarlo, no obstante, como una distinción más modesta - de grado- para retener la normatividad de lo mental, pero sin pagar el precio que paga Davidson: el de un dualismo profundo que impide, por principio, la posibilidad de contar una historia acerca de cómo pasamos de lo pre-conceptual a lo conceptual, acerca de qué nos hace ser quienes somos.

En una línea consonante con la que propongo aquí, Ramberg (2000) ofrece su iluminadora lectura de la filosofía de la mente davidsoniana. A su juicio, hay una precedencia de la dimensión normativa, que permea toda dimensión descriptiva. En este sentido, al igual que pretendo con mi planteo, se conservan los polos normativo y descriptivo de la distinción (esto es, su carácter exclusivo), pero se socava su carácter excluyente, es decir, se habilita la posibilidad conceptual de que las dimensiones se “entrelacen”. Todo vocabulario descriptivo presupone un vocabulario normativo, no en virtud de relaciones inferenciales entre oraciones de un vocabulario y oraciones del otro, sino pragmáticamente: no podemos desplegar el vocabulario descriptivo a menos que podamos desarrollar el normativo (así como no podríamos emplear un destornillador si no tuviéramos manos). En opinión de Ramberg, si Davidson está en lo correcto, *describir* algo es una habilidad que tenemos *sólo porque es posible para otros vernos como personas* como en una conformidad general a las normas que encarnan los predicados de la agencia. La agencia, entendida como la habilidad de ofrecer descripciones en lugar de ruido, *sólo aparece si un vocabulario normativo ya está siendo usado*: las descripciones emergen como descripciones de cualquier tipo sólo contra un trasfondo supuesto de conducta propositiva y, por tanto normativamente describible por parte de los comunicadores involucrados.

Ahora bien ¿por qué nos tienta juntar los conceptos de *mente* y de *persona* y de racionalidad como *intencionalidad* y racionalidad como tener *responsabilidades*? ¿Por qué está tentación fue lo suficientemente fuerte como para hacer caer a Kant en la metafísica, al sostener que la libertad sólo es posible si hay un tipo de realidad no espacio-temporal? Según Rorty, la respuesta que Ramberg ofrece a las dos preguntas es que hay una yuxtaposición considerable entre seres que hablamos *acerca de (talk about)* usando el vocabulario intencional y seres *entre quienes nos hablamos (talk to)* usando el vocabulario normativo. Hay muchos vocabularios descriptivos (maneras de conferir “sapiencia” a diferentes patrones causales en el mundo) y

muchas comunidades de usuarios de lenguaje, pero siempre debemos, a la vez, recoger tal patrón y pertenecer a alguna comunidad. No podemos parar de prescribir y sólo describir. De esta suerte, los intereses predictivos que están expresados en el vocabulario de la agencia son de un tipo muy peculiar: se activan revelando los tipos de rasgos que nos permiten reconocernos a nosotros mismos en aquello acerca de lo cual estamos hablando, y traer a la luz aquellas complicadas consideraciones en torno a la *noción moral de persona*<sup>193</sup>.

En una línea afín, leemos en Bill Martin:

Davidson acepta no sólo la posibilidad de lo ético, sino, más aún la inevitabilidad de la posibilidad de lo ético. Para Davidson, participar de un lenguaje es depender de la posibilidad de lo ético, dado que es estar comprometido en y, en efecto, ser formado por un despliegue nunca completo ... de lenguaje y evaluación. (En Hahn 1999: 345).

Mi invitación es, pues, la siguiente. Una vez que nos hemos despojado de la dicotomía descriptivo/normativo, se abre un horizonte conceptual nuevo para elucidar la *normatividad de lo mental*, aquello que nos hace peculiares, a partir del rol que desempeña la perspectiva de *segunda persona* de la atribución mental. Tal es el sentido en el que planteo que pasemos de una “discusión” a una “controversia”. Veamos, a continuación, mi propuesta.

---

<sup>193</sup> En este espíritu, Ramberg (2000) propone que en lugar de leer a Davidson bajo el signo de Brentano, esto es, en términos de la irreducibilidad de la intencionalidad, pasemos a leerlo en la clave hegeliana de nuestro mutuo reconocimiento (*Anerkennung*) como personas. A su juicio, Davidson ha comprendido mejor que Rorty que reconocer a algunos seres como obedientes-seguidores de normas, reconocerlos como miembros de una comunidad, es un requisito para usar un lenguaje, para la habilidad de desplegar un vocabulario descriptivo. El reconocimiento establece una comunidad de usuarios de herramientas. Los distintos vocabularios descriptivos son herramientas en el “kit” de nuestra comunidad. Sin caja de herramientas, no hay comunidad –si no describiéramos no tendríamos críticas que ofrecer a las descripciones de otros-. Pero sin comunidad, no habría caja de herramientas –si no nos criticáramos, nuestras descripciones no serían descripciones-.

## Capítulo 8: La normatividad de lo mental y el rol de la segunda persona

El descubrimiento que el niño hace de sí mismo como persona es también su descubrimiento de sí mismo como miembro de un mundo de personas...El descubrimiento de mí mismo como persona es el descubrimiento de que puedo hablar y de que soy una *persona* o hablante; al hablar soy tanto hablante como oyente, y puesto que el descubrimiento de mí mismo como persona es también el descubrimiento de otras personas a mi alrededor, es el descubrimiento de hablantes y oyentes distintos a mí.

(Collingwood, citado por Davidson, 1991/2001: 219).

El desafío que afrontamos en este capítulo es el siguiente. En la primera parte de nuestro trabajo presentamos la teoría de la interpretación radical como marco de nuestra indagación. En efecto, ella constituyó nuestro “hilo de Ariadna” para mostrar la figura que surge del fascinante “rompecabezas” conceptual davidsoniano: su contribución antropológica. Para ello, intentamos acomodar las “piezas” en juego de modo de iluminar las relaciones conceptuales que hay al interior de su filosofía de la mente y examinar sus consecuencias. El recorrido que propusimos, a tal efecto, fue el siguiente. Nos concentramos en los tres rasgos de su noción de *creencia*, a saber: *holismo*, *normatividad* y *triangulación*. En cada caso, distinguimos variedades y propusimos elucidaciones, ahí donde a nuestro juicio Davidson dejó inarticulado su planteo. En la segunda parte de nuestro trabajo, indagamos el rol que desempeñan dichos rasgos en las respuestas de Davidson a tres problemas cruciales en el área de la que nos ocupamos: el *problema mente-cuerpo*, el *problema de la relación entre el pensamiento y el lenguaje* y el *problema de las perspectivas de primera y tercera persona de la atribución mental*. A partir de su consideración, en el capítulo anterior advertimos un supuesto en el que, a nuestro juicio, reposa su filosofía de la mente, y que la compromete con consecuencias indeseables: la dicotomía descriptivo/normativo. He defendido la idea de que si despojamos su filosofía de la mente de este “cuarto dogma” del empirismo, no sólo habremos de tornar más “davidsoniano” a Davidson, sino que abriremos un horizonte conceptual para ofrecer una elucidación de aquel

rasgo central de su planteo, que él mismo asume haber dejado vago: la *normatividad de lo mental*. He argumentado que si nos restringimos a la dimensión meramente “deóntica” de su propuesta, no la capturamos en toda su peculiaridad y hondura. En tal sentido, sostuve la necesidad de profundizarla en su dimensión “evaluativa”. En el curso de la segunda parte de este trabajo, alegué que el valor “ulterior” que Davidson pone en juego a la hora de esclarecer qué nos hace ser quienes somos es el de tratarnos mutuamente como *personas*. Asimismo, argumenté que dado su tratamiento de las perspectivas de primera, segunda y tercera persona, dicho valor no logra captarse en la singularidad requerida por el propio planteo davidsoniano. Mi invitación es, pues, a hacerlo desde la perspectiva que él simplemente “bautizó”, pero no exploró en toda su profundidad: la perspectiva de *segunda persona* de la atribución mental. Ha llegado, pues, el momento de mostrar en qué sentido dicha perspectiva resulta relevante a la hora de esclarecer en qué consiste tratarnos mutuamente como *personas*. Tal es el desafío que abordamos aquí. Comencemos, entonces, por ofrecer una caracterización precisa de la perspectiva de segunda persona de la atribución mental.

### **8.1. La perspectiva de segunda persona: repensando la atribución mental.**

#### **El regreso de las emociones<sup>194</sup>**

Como ya anticipé, hasta aquí hemos visto el papel que la perspectiva de *tercera persona* - que aparece claramente relacionada con la noción de *objetividad* - desempeña en la teoría de la interpretación radical. Asimismo, hemos advertido que dicho marco deja lugar para el reconocimiento de la perspectiva de *primera persona*, en un sentido que ressignifica la concepción tradicional de *subjetividad*. Por su parte, la perspectiva de *segunda persona* parece introducir una dimensión que, como dijimos, sugiere la posibilidad de cerrar la brecha entre lo objetivo y lo subjetivo abierta en la modernidad, esto es: la *intersubjetividad*. Ahora bien, he señalado que allí donde considero que Davidson logró ofrecer una elaboración precisa de las

---

<sup>194</sup> En esta sección voy a seguir los lineamientos centrales de las excelentes presentaciones de Scotto, C. (2002) y Gomila Benejam A. (2002).

perspectivas de *primera y tercera persona*, esto es, como perspectivas genuinas y distintas, no logró lo propio con la perspectiva de *segunda persona*, a la que más allá de “bautizarla” no terminó de articular conceptualmente como una perspectiva “auténtica” y diferente de las dos tradicionales. Tal como argumenté en el capítulo 6: su *segunda persona* sigue siendo una *tercera persona*. Esto es, la “segunda” persona con la que interactuamos en el proceso de triangulación, parece ser un mero observador no comprometido y no un participante en las prácticas sociales de los interpretados. En este sentido sostuve que, a mi juicio, Davidson vislumbró la riqueza de introducir la dimensión intersubjetiva en nuestra reflexión acerca del pensamiento y el lenguaje, pero no terminó de elaborarla en toda su hondura, esto es, como auténticamente distinta de lo subjetivo y lo objetivo.

A los efectos de profundizar esta suerte de hallazgo inconcluso en Davidson, creo que puede sernos de ayuda apelar a la perspectiva de *segunda persona* tal como se la introdujo recientemente en la filosofía de la mente. Hasta donde sé, ha sido Antoni Gomila Benejam (2000) quien identificó el dominio de la perspectiva de segunda persona de la atribución mental<sup>195</sup> y justificó su relevancia, observando con acierto que resulta “sorprendente” que en el debate acerca de las perspectivas, no se haya advertido la posibilidad de otra perspectiva genuina y diferente de las clásicas de primera y tercera persona.<sup>196</sup>

---

<sup>195</sup> La primera vez que escuché hablar de la perspectiva de *segunda persona* de la atribución mental fue en el año 2000, a instancias de su participación en el “Coloquio Internacional sobre temas actuales de Psicología Cognitiva y Filosofía de la Mente”, como una actividad del proyecto “Lenguaje, comprensión y cognición”, dirigido por Eduardo Rabossi y Aníbal Duarte (Duarte y Rabossi, eds., 2003). Esa fue una de las primeras ocasiones en que Gomila presentó públicamente sus ideas sobre la perspectiva de segunda persona (que cuajaron en Gomila 2001, 2002, 2003). Por entonces afirmó que Eduardo Rabossi lo puso en la pista de la obra de Martin Buber (1922) a la que califica como aquella obra que podría considerarse la primera que trató de captar lo específico de la perspectiva de segunda persona, aunque desde un planteo existencialista, como un tipo específico de relación entre *personas*. Desde aquella ocasión, he creído que dicha perspectiva no es una mera enmienda de la concepción tradicional moderna de la subjetividad, sino que sugiere la posibilidad de una genuina alternativa. En los años transcurridos desde aquel encuentro, la idea de una perspectiva de segunda persona ha dejado de sonar a atrevida excentricidad, y diversos autores, desde diversos contextos reflexivos, han utilizado esa expresión y han explorado posibilidades relacionadas.

<sup>196</sup> Como advierte Scotto (2002), en la literatura filosófica y psicológica hay algunas excepciones aisladas. Entre ellas, destaca a: J.C. Gómez (1996) y el libro editado por E. Thompson (2001). Asimismo, creo que cabe señalar la obra de Gallagher (2005, cap. 9). Desde el ámbito de la psicología, cabe subrayar diversos trabajos de la discípula de Colwyn Trevarthen, Vasudevi Reddy, la primera en utilizar la expresión “segunda persona” dentro del debate sobre la “teoría de la mente” (Reddy, 1996; 2003; Reddy y Morris, 2004) así como de Peter Hobson (Hobson, 2005). Y, como veremos en breve, recientemente se la ha incorporado al ámbito de la reflexión ética, con el libro de Stephen Darwall (2006).

Como observamos en el capítulo en el que presentamos el debate en torno a la “teoría de la mente”, a juicio de Gomila la idea de la perspectiva de segunda parece funcionar como un “ariete” frente a las dicotomías subjetivo/objetivo, o frente a las perspectivas de primera y tercera persona, entendidas como exclusivas y excluyentes. En su opinión, la cuestión de fondo es que los teóricos de la simulación y los teóricos de la teoría se mueven en los mismos términos dialécticos. Por un lado, su enfrentamiento reproduce, desde la opción estratégica por la perspectiva de primera persona, la bien conocida problemática de la concepción moderna del sujeto afectada -entre otros problemas- por el del “conocimiento de las otras mentes”, a la que deben responder, tanto la versión racionalista como la empirista, y que acaban defendiendo versiones del conocimiento por analogía, plagadas de dificultades. Por su parte, la opción estratégica por la perspectiva de tercera persona lleva a la contrapartida con la que se encuentran los materialismos objetivistas del siglo XX, los cuales resuelven el problema del “conocimiento de las otras mentes”, convirtiéndolo en un conocimiento teórico, pero al precio de despojar al autoconocimiento y a la autoconciencia de su carácter distintivo, esto es, de pasar por alto la especificidad del agente, del sujeto. Por ello, el reconocimiento de la perspectiva de segunda persona como una perspectiva genuina y distintiva, debe situarse en la línea de los recursos teóricos para superar o evitar los problemas característicos de la dialéctica sujeto/objeto de la filosofía moderna. Pues así como las filosofías subjetivistas al conceder al sujeto la posición privilegiada de fundamento ontológico y epistémico lo condenan a quedar fuera del mundo (Descartes, Kant, Wittgenstein en el *Tractatus*), las filosofías objetivistas y materialistas al concebirlo en términos objetivos excluyen los aspectos distintivos de la experiencia humana.

Esta concepción, no obstante, no exige convertir la *segunda persona* en la única existente. Al contrario, reconocer lo genuino de la perspectiva de segunda persona permite reconocer, del mismo modo, lo genuino de las de primera y tercera persona, en lugar de verlas como alternativas en conflicto. En realidad, bien puede ocurrir que atribuyamos estados mentales a los demás de formas muy distintas, y sobre bases muy diferentes, del mismo modo que también somos capaces de autoatribuirnos estados mentales de modos distintos a nosotros mismos.

Para Gomila es simple advertir que las atribuciones de segunda persona son distintas de las autoatribuciones y de las atribuciones en tercera persona, donde se adopta una posición

objetiva y distanciada de tratar de explicar o reconstruir las circunstancias de otra persona, al margen de nuestra relación con ella. No obstante, asume que puede no ser tan sencillo aceptar que constituye una perspectiva distinta y genuina, es decir, no reducible a una de las dos anteriores. En efecto, podría sostenerse que la atribución de segunda persona -igual que se ha defendido respecto de las atribuciones de tercera persona- se estructura sobre la base de las atribuciones de *primera persona*, y que luego se proyecta, por analogía- o en la terminología contemporánea, por *simulación*- a los demás. Inversamente, podría tratar de argumentarse que las atribuciones de segunda persona resultan, en realidad, de la *inferencia mediada*, desde la *perspectiva de tercera persona*, por la teoría de la mente de que se dispone. Gomila sostiene que para ver que no puede tratarse de la primera opción, es preciso considerar que, con frecuencia, las atribuciones de segunda persona conllevan la atribución de un estado distinto del que experimenta quien hace la atribución. Esto es especialmente claro en la interacción emocional: mi reconocerte o atribuirte duelo por el fallecimiento de un ser querido hace que experimente compasión y que te “acompañe en el sentimiento”, pero de ningún modo puedo experimentar ese mismo duelo, que responde a la especial relación afectiva que mantenías con el ser querido en cuestión.

En cuanto a que no se trata de una inferencia basada en una teoría, Gomila cree que hay dos tipos de consideraciones que hacen poco creíble tal eventualidad. En primer lugar, que la atribución se da implícitamente en la propia experiencia perceptiva social, no mediante una inferencia a partir de los gestos o expresiones faciales percibidos; al contrario, se trata de un caso de *percepción significativa*, de *reconocimiento perceptivo* de una instanciación de un estado psicológico - esto es, no es posible separar un aspecto del otro, igual como, en el caso del habla, no podemos separar la percepción acústica de la percepción del significado-. Ciertamente, en ambos casos, la percepción significativa está mediada por el conocimiento adquirido, pero sin que de ahí pueda deducirse que ese conocimiento es de carácter teórico (compárese, dice Gomila, con un caso claramente mediado por el conocimiento teórico, como el interpretar que alguien hace algo debido a su complejo de Edipo).

Asimismo, su *dimensión práctica* puede apreciarse en la dependencia de la atribución del contexto práctico involucrado: la atribución depende del contexto de interacción en que nos encontramos -junto con el conocimiento previo acumulado de interacciones anteriores-. La atribución no está guiada por un interés explicativo o predictivo o interpretativo, sino por la

necesidad de interacción en tiempo real, lo que lo lleva a hablar de la *dimensión reactiva*, del modo en que nuestra propia reacción a la situación depende de la atribución. En ese sentido, el conocimiento involucrado es de la misma naturaleza que el *conocimiento práctico* que guía nuestra acción en casos de “expertise” técnica o práctica en general, esto es, como un “*know how*” y no como un “*know that*” (Vega, 2001). No tiene sentido, por tanto, como pretende el simulacionismo, pensar que se trata de un funcionamiento “off line” de nuestro propio sistema psicológico: a diferencia de los procesos off-line -como pueden ser el razonamiento espacial, la imaginación hermenéutica, o la toma reflexiva de decisiones- la comprensión intencional tiene lugar claramente “on-line”, como un tipo de percepción significativa, un “ver-en” la conducta la intención que la guía o la emoción que expresa (Wollheim, 1980). Este carácter *práctico* resulta todavía más claro cuando se considera el modo en que ese conocimiento se almacena y se recupera, esto es, según los defensores de la teoría- teoría: como representaciones mentales. En realidad, el desarrollo de los debates en torno a la teoría de la mente obtuvo su impulso más consistente de esta asociación entre teoría y representación mental, y en la singularización de la creencia falsa -la capacidad de atribuir creencias falsas a otros -como criterio de haber obtenido la teoría de la mente, según vimos en el capítulo 5. Puede decirse que una de las líneas de ataque más poderosas en contra de la naturaleza teórica de la “teoría de la mente” -de la capacidad de atribución intencional- lo ha constituido la demostración de que los niños son capaces de interacciones intencionales mucho antes de ser capaces de atribuir *creencias falsas*; en realidad, desde el primer año de vida (Reddy y Morris, 2004). Derivadamente, si la capacidad de atribución intencional fuera de naturaleza teórica, debería ser sensible a la adquisición de nuevo conocimiento. Esto es lo que ocurre, en realidad, en el caso de la perspectiva de tercera persona, donde la explicación que podamos hacer de la conducta de alguien puede variar si nos aportan alguna nueva información que desconocíamos (como cualquier otro caso de inferencia a la mejor explicación); mientras que la atribución de segunda persona es más rígida, dependiente de la percepción de los elementos expresivos y conductuales en que se basa (más semejante, por tanto, al reconocimiento de patrones).

Además, la perspectiva de segunda persona puede activarse también en situaciones no canónicas, lo que avala asimismo su naturaleza distintiva genuina: en la contemplación de narraciones visuales, por ejemplo, como en el caso del cine, e incluso en el caso de la audición musical, es característico que nuestra experiencia dependa de la atribución de emociones a

patrones visuales o sonoros de los que sabemos positivamente que no son sujetos de experiencia. La capacidad del arte de generar estas experiencias constituye parte de su valor y explica por qué nos conmueve.

Por estas razones, la *segunda persona* es caracterizada - tal como anticipamos en el capítulo 6- como una perspectiva de atribución mental, diferenciada de las de primera y tercera persona. Veamos, entonces, con más detenimiento los elementos que se consideran como constitutivos de esta perspectiva de atribución mental. A tal efecto, voy a seguir, especialmente, a Scotto (2002). En este trabajo, ella describe los rasgos que identifican a la perspectiva de la segunda persona como una perspectiva *genuina, específica y básica* de la atribución mental y, sobre esa caracterización se propone defender su adecuación para dar cuenta de ciertos fenómenos básicos de atribución mental que quedan fuera del alcance y del interés de la primera y la tercera persona. Así como también su capacidad para evitar o disolver los problemas escépticos que de ellas derivan: el de las otras mentes y el del conocimiento de la experiencia subjetiva (ella habla de disolver problemas escépticos porque éstos surgen de una excesiva importancia concedida a los propósitos “epistémicos” de la atribución mental, los que no juegan un papel decisivo en la perspectiva de segunda persona).

Scotto defiende la tesis de que ni la perspectiva de primera ni la de tercera son excluyentes ni básicas, y sugiere, en cambio, que se trata de perspectivas que se adquieren gradual y tardíamente en el desarrollo psicológico, y que son complementarias entre sí y con la de segunda persona, permitiendo cada una formas diferenciadas de comprensión según los fenómenos de que se trate y los rasgos propios de quien las adopte. En segundo lugar, identifica algunos fenómenos interactivos *propios* de esta perspectiva (en especial, ciertos fenómenos emocionales, con el objeto de mostrar cómo funciona y a quiénes es correcto atribuir competencia, según ella, para la atribución intencional en estos casos). Por último, extrae algunas conclusiones generales que se siguen de la aceptación de esta perspectiva y de su importancia en la explicación de la atribución intencional ordinaria o de la *folk psychology*, en un debate signado por las opciones contradictorias de las perspectivas de primera y tercera persona.

El enfoque de la segunda persona pretende dar cuenta, como veremos, de aquellos casos de comprensión intencional que pueden atribuirse a legos e incluso a niños (así como también a otras especies animales) haciendo lugar a fenómenos que ponen de manifiesto el carácter

constitutivo de los vínculos intersubjetivos y, por lo tanto, sin depender de “nociones bien definidas de yo y otro que domina un filósofo adulto” (Hutto 2002:52).

Debe reconocerse, sin embargo, que ciertas caracterizaciones ofrecidas por los defensores de una y otra perspectiva al proporcionar reconstrucciones de las mismas en las que se destacan o priorizan fenómenos que, como veremos, es más adecuado describir desde esta nueva perspectiva, se acercan a una comprensión más satisfactoria del fenómeno complejo de la atribución mental. Tal es el caso, a juicio de Scotto, de la caracterización preferentemente *práctica y social* de la “actitud intencional” que propone Dennett (1987) y con más nitidez aún, en el caso de las anticipadoras y minuciosas reconstrucciones wittgensteinianas de los fenómenos ordinarios de la interpretación *emocional* y el papel que ella tiene en la expresión gestual, en su caracterización de los fenómenos de la *percepción visual* (los casos de “ver como”) y en general, en la comprensión de significados lingüísticos y no-lingüísticos en contextos prácticos, de interacción “cara a cara”. En igual dirección, Scotto valora las críticas de Wittgenstein a los intentos de describir a la atribución mental como un desafío escéptico acerca de las “otras mentes” o como el resultado de una inferencia analógica a partir del propio caso, tanto como de presentarla bajo el paradigma de la autoatribución transparente, en el dominio privado de la propia subjetividad.<sup>197</sup> Esto, a juicio de Scotto, significa varias cosas relativas a la perspectiva de segunda persona, a saber:

i-Se manifiesta en *conductas públicas*; el énfasis puesto en las conductas obedece a que en esta forma de vinculación intencional con el otro intervienen básicamente componentes públicos o bien componentes mentales que tienen una inequívoca y directa expresión pública, expresiones faciales o posturales, y diversas formas de conducta. Este rasgo marca una diferencia con el esfuerzo enteramente privado en el que consistiría la introspección –o la imaginación simuladora- y sus objetos de conocimiento, en el caso de la primera persona, así como también respecto de la intervención de los mecanismos inferenciales internos al intérprete de la tercera

---

<sup>197</sup> Scotto sostiene que los aportes de Wittgenstein deben destacarse especialmente, entre otros fundamentos, por haber insistido en la relación que tiene la atribución semántica con la interpretación más inmediata de las manifestaciones naturales no verbales, tales como *gestos y reacciones corporales*, ofreciendo de ese modo un *enfoque más abarcador de la intencionalidad humana* que es el que resulta necesario para una perspectiva de segunda persona.

persona, mecanismos que además intentan comprender la evidencia comportamental pública mediante la postulación de entidades mentales inobservables.<sup>198</sup>

ii- Se desarrolla como una *habilidad o destreza* (como un *know how*); en este sentido, es una *práctica*. Describir la interpretación ordinaria como básicamente procedural, registra el hecho suficientemente bien establecido -por biólogos, etólogos y psicólogos- de que ciertas habilidades para mentalizar (no todas) son no aprendidas, están universalmente extendidas - algunas son comunes a la especie y reclaman una explicación evolucionista (y se manifiestan muy temprano en el desarrollo psicológico de los niños)-. Esta visión está en las antípodas de la que proponen los teóricos de la tercera persona al asemejarla a la teorización científica, para quienes la capacidad de comprender la vida social de los demás debe ser explicada como el resultado de captar una teoría. Bogdan llama a esa concepción “la visión del espectador de la interpretación” y la define como una suma de “teoricismo” -la psicología ordinaria anticipa la teorización científica- “mentalismo” -la interpretación postula entidades mentales ocultas- y “explicativismo” -la interpretación se propone explicar y predecir comportamiento-. Frente a estos rasgos, la interpretación intencional ordinaria, en cambio, tendría según él una doble significación práctica: “opera como una práctica y tiene propósitos prácticos” (Bogdan 1997: 104-5). Opera como una práctica porque concibe y manipula relaciones sujeto-mundo y atributos y comportamientos de un sujeto en términos de estrategias efectivas de involucramiento; y tiene propósitos prácticos porque selecciona aquellas relaciones sujeto-mundo y atributos que dan lugar a estrategias efectivas. A estos propósitos Scotto se refiere como “evaluativos”. Pero, según vimos, esta perspectiva también se opone al modelo cartesiano del autoconocimiento, que da cuenta del acceso de primera persona como un acceso cognoscitivo de carácter teórico, privilegiado y fundacional, y del conocimiento de las otras mentes, como el resultado de realizar inferencias analógicas a partir del propio caso. En su versión actual, la de los teóricos simulacionistas, se afirma que la propia mente es tomada como un modelo accesible al propio sujeto sobre la base del cual puede “empatizar” o imaginar lo que ocurre en la mente de los demás.

---

<sup>198</sup> Con todo, si se piensa en Davidson como un defensor de la perspectiva de tercera persona –tal como suele hacerse- cabe señalar que en función de sus críticas a la doctrina de los objetos “ante la mente”, que abordamos en la primera y segunda partes de este trabajo, esta apreciación general acerca del compromiso de la tercera persona con la postulación de entidades mentales inobservables, no parece resultarle legítimamente aplicable.

iii- Posee, por definición, un carácter *constitutivamente social*, puesto que se realiza *con* otros (no en aislamiento o hacia otros). A juicio de Scotto, la dimensión social pone de manifiesto el carácter *básico* de las relaciones *intersubjetivas* en la comprensión recíproca<sup>199</sup>. En este sentido, al revés de las imágenes clásicas de la mente, la subjetividad individual es explicada como fruto de esta habilidad intersubjetiva<sup>200</sup>.

iv- Por su parte, la perspectiva de segunda persona describe a la atribución intencional como una forma de comprensión *recíproca*: lo que un sujeto hace tiene una respuesta sensible en el otro sujeto, y así sucesivamente, de tal modo que acciones y reacciones tienen los rasgos de la reciprocidad. Como se ve, hay una diferencia crucial entre estos rasgos y las versiones más extremas de las otras perspectivas, donde el “yo” y el “otro” están distanciados entre sí y funcionan como espectadores u observadores de sí mismos o de las conductas de otros, y se relacionan -cuando lo logran- sólo contingentemente y *a posteriori*. Desde la perspectiva de segunda persona, en cambio, *lo primario es la interacción misma*, los vínculos que mantienen a cada sujeto ligado o involucrado en relaciones *con* otros, *como un rasgo constitutivo de su hacerse personas*.

---

<sup>199</sup> En esta línea, Scotto desde sus intereses empíricos llama la atención sobre estudios recientes en psicología del desarrollo y psicología social que evidencian la importancia de lo que se denomina “comprensión social”, “cognición social” o “inteligencia social” para la conformación temprana, en los niños, de una “teoría de la mente”. Sostiene que diversos experimentos evidencian que la perspectiva de segunda persona sería la que primero adoptan los niños (en las interacciones primordiales con la madre, con otros niños, con los demás adultos) y es a partir de ella que construyen una noción de mentalidad, aplicable a otros y a sí mismos.

<sup>200</sup> Este enfoque es, para Scotto, del todo compatible con una concepción evolucionista de la mentalidad, desde la cual se explica el logro de estas habilidades y la razón de ser de los mecanismos básicos de la especie como el fruto de fuerzas selectivas directamente dependientes de los rasgos propios de la vida social de los homínidos. Sin dudas, no es éste el enfoque predominante en filosofía de la mente, y la situación es apenas diferente en la psicología, donde ha ingresado en las investigaciones actuales sobre el desarrollo ontogenético de una teoría de la mente en los niños, donde a las hipótesis individualistas se oponen los resultados surgidos de hipótesis sociales de la subjetividad (cfr. Garfield, et. al. 2001), que contiene una amplia evaluación de las dos visiones antitéticas sobre los factores que intervienen en la adquisición de una teoría de la mente en los niños: la de innatistas, modularistas y teóricos de la teoría, por una parte, y la de quienes defienden una hipótesis socio-lingüística. La tesis que se defiende es que “el aprendizaje de habilidades básicas para interactuar con otros es lo que permite que la negociación en un mundo social sea la condición necesaria para aprender a leer mentes, junto con la adquisición del lenguaje” p. 525. Estas interacciones sociales son, básicamente, de cooperación, competencia, *nurturance* y reproducción.

En esta línea, Scotto cita a las teorías socio-funcionales de las emociones, por ejemplo, que dependen de aceptar el carácter constitutivamente social de las personas, quienes enfrentarían los problemas de la supervivencia mediante interacciones sociales coordinadas (Keltner y Haidt 1999). Gomila sugiere que la imagen contraria resulta implausible al concebir la vida social “simplemente como el resultado de una decisión estratégica tomada por individuos constitutivamente auto-interesados” (2001:19-20).

v- Estos intercambios son también *inmediatos*, dependiendo su grado del carácter más o menos “primitivo” de la reacción o conducta de que se trate.<sup>201</sup> Los sujetos saben, perciben, etc., “interpretan” sin la mediación de hipótesis o inferencias, qué le está sucediendo a la otra persona y esta interpretación produce en ellos comportamientos imitativos, transformaciones emocionales. Este rasgo explicita aún más el carácter no teórico y muchas veces inconsciente de estas formas básicas de atribución mental. Ahora bien, Scotto advierte que la inmediatez de la comprensión en segunda persona hace difícil verla como un caso de atribución intencional. Ella subraya que se trata de una inmediatez fenomenológica, es decir, a nivel de la experiencia personal (y no implica ninguna tesis realista metafísica acerca de las propiedades intencionales).

Como dice Wittgenstein en relación con la interpretación de emociones:

Uno ve la emoción ...No vemos las muecas de una persona e *inferimos* que siente alegría, aflicción, aburrimiento. Describimos directamente su rostro como triste, radiante, aburrido aunque no seamos capaces de dar ninguna otra descripción de sus rasgos. La tristeza se personifica en el rostro, quisiera uno decir. Esto resulta esencial para lo que llamamos “emoción” (*Observaciones sobre la Filosofía de la Psicología*, II, #570; cfr. también *Zettel*, # 225).

vi- Asimismo, la atribución desde la segunda persona tiene un carácter *dinámico*. Este rasgo es fácil de reconocer puesto que se desarrolla en el curso de las interacciones a través del tiempo y porque es recíproca o sensible a las respuestas del otro.

vii- Se trata de casos de comprensión *situada*, es decir, que necesita ser máximamente flexible a las variaciones contextuales, a la información adicional subsiguiente y, básicamente, a la conducta variable de los demás.

viii- En cuanto a los *propósitos evaluativos* que tiene la atribución mental, según esta perspectiva, resultan de las necesidades y propósitos de la interacción misma: saber cómo actuar, corregir las propias expectativas, coordinar esfuerzos para lograr un objetivo común, persuadir a los otros de las propias intenciones, para ser aceptado o perdonado, obtener beneficios, etc. “Evaluar”, dice Scotto, aquí significa saber cómo actuar de un modo que sea más conveniente a los propios fines en función de la información contextual proporcionada por la conducta o las reacciones del otro. Estas metas de la atribución intencional no son las que han enfatizado las

---

<sup>201</sup> En este contexto, “inmediato” y “primitivo” deben tomarse como sinónimos aproximados de automático, espontáneo, no aprendido, no propositivo, no controlable. Estas son las propiedades distintivas justamente del así llamado “contagio emocional primitivo”, el fenómeno por el cual se capturan/experimentan/expresan las emociones de otra persona de un modo que es ampliamente automático y no controlable.

perspectivas de primera y tercera persona, a saber: metas *epistémicas* de un *observador desinteresado*. Estos fines *evaluativos* están esencialmente ligados con fenómenos *emocionales*<sup>202</sup>.

ix- Finalmente, Scotto señala que las interacciones en las que consiste la perspectiva de segunda persona de la atribución mental constituyen casos de *comunicación intencional*. Múltiples intercambios humanos significativos, nos dice, ocurren sin la mediación de signos lingüísticos. Decimos que son significativos porque consisten en intercambios de formas de conducta mediante las cuales un sujeto hace manifiestos sus propósitos, creencias o intenciones a otro sujeto, esperando que éste sea capaz de específicos actos de reconocimiento.<sup>203</sup>

Ahora bien, ¿cuáles son los casos de atribución intencional de segunda persona? Fenómenos como la mirada (y sus efectos), las reacciones que evidencian atención visual conjunta, los movimientos corporales coordinados, los casos de imitación y sincronización gestual, las distintas formas que adopta la identificación y el contagio emocional<sup>204</sup>, las reacciones “simpáticas” espontáneas o las “emociones empáticas”, y una gran diversidad de formas de cognición sensorio-motoras y actitudes o reacciones no verbales hacia los demás, todos los cuales cuentan como casos de atribución mental porque expresan y/o resultan de reconocer en otro agente, sentires, intenciones, necesidades, motivaciones, estados emocionales, etc.

---

<sup>202</sup> Aunque no voy a detenerme a analizar aquí la naturaleza compleja de las emociones, cabe señalar que Scotto suscribe la concepción ampliamente extendida que relaciona a las emociones con el modo como favorecen –o perturban– los fines o metas del sujeto, esto es, con sus propósitos vitales, que les reconoce contenidos intencionales (bajo una acepción que, en principio, parece que ha de ser más abarcadora que la davidsoniana) y componentes cognitivos, así como que poseen manifestaciones expresivas características. En relación con estos propósitos evaluativos, las emociones lo son tanto para el sujeto que las tiene como para quien es afectado por ellas de algún modo, pudiendo resultar apropiadas, irracionales, desproporcionadas, inoportunas, etc. A. Baier afirma que “nuestras reacciones emocionales son reacciones evaluativas” (Baier 1990:18), es decir reaccionamos con emoción de modos que nos parecen apropiados a las circunstancias en las que nos encontramos, sean estas objetivas o interpersonales. Se trata, más brevemente, de “evaluaciones espontáneas”. Nuestras interacciones emocionales son interacciones *evaluativas*. Volveré sobre este punto al abordar la relevancia de la perspectiva de segunda persona para la dimensión *evaluativa* de la normatividad de lo mental que, a mi juicio, se pone en juego en la noción de *persona*.

<sup>203</sup> A juicio de Scotto, los tres niveles de intencionalidad en que consiste el modelo comunicacional griceano pueden reconocerse también cuando no existen intercambios *verbales* sino sólo signos expresivos naturales. Y afirma que estas formas de comunicación no verbal han sido reconocidas también entre miembros de diferentes especies animales (cfr. Dennett 1987). Atribuir la misma estructura intencional a estas formas de comunicación y poder así comprender mejor la naturaleza de los intercambios entre algunos primates fortalece, a su juicio, una concepción *no lingüística* de la intencionalidad, o dicho de otro modo, desacredita las teorías que hacen depender exclusivamente de la competencia lingüística la capacidad para tener actitudes proposicionales, como es el caso de Davidson, según vimos en el cap.5.

<sup>204</sup> Cf. Hatfield, E. Cacioppo, J. y Raspón, R.L. (1992). Cabe señalar que estos fenómenos son estudiados, asimismo, por su interés para dar cuenta de los vínculos interactivos especiales que se dan en la experiencia estética o la comprensión ficcional. Al respecto, pueden verse Español (2008) y Shifres (2008).

Como lúcidamente advierte Scotto, prácticamente ninguno de estos fenómenos ha sido objeto de atención por parte de los filósofos de la mente. Sí, en cambio, son objetos de programas de investigación en neuropsicología, psicología evolutiva, del desarrollo, del comportamiento, social, entre otras. No obstante, en lo que respecta al tema de nuestro interés puede disputarse que cuenten como casos de atribución “intencional”: reaccionar más o menos inconsciente o automáticamente a las conductas de los demás no puede asimilarse a la compleja y radical operación de asignar significados y contenidos mentales a un sujeto de una cultura desconocida. La tesis de Scotto es que sin los mecanismos básicos las operaciones complejas no podrían tener lugar. En otras palabras, y bajo mi interpretación: sin la perspectiva de *segunda* persona no podría tener lugar la de *tercera*.

Ahora bien ¿en qué sentido esta noción de *segunda persona* permite echar luz sobre la peculiar normatividad de lo mental en la filosofía davidsoniana, según la elucidación que propuse? Tras habernos despojado de la dicotomía descriptivo/normativo -que en el capítulo anterior he denominado como el “cuarto dogma del empirismo”- podremos apreciar ahora los modos en que la descripción y la evaluación, el *factum* y la *norma*, se entrelazan en la noción de *persona* y, de esta manera, repensar el tratamiento davidsoniano de los problemas de los que nos hemos ocupado aquí, a saber: el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la atribución mental.

## **8.2. Ser humano, ser persona, tener una mente.**

Como hemos señalado en el capítulo 4, a nuestro juicio, Davidson está tratando de difuminar la distinción entre la mente como *espectadora* y como *participante responsable* en las prácticas sociales. En efecto, al combinar una teoría del significado con una teoría de la acción, Davidson entrelaza la distinción *mente-cuerpo* con la distinción *persona-cosa*. Mientras que la noción de *mente* se vincula con una noción de racionalidad teórica entendida como *posesión de creencias y deseos*, la noción de *persona* se vincula con una noción de racionalidad práctica que

involucra la idea de ser *uno de nosotros*, esto es, ser un miembro de nuestra comunidad, en la que estamos ligados por nuestras responsabilidades recíprocas.

Como cuestión preliminar, voy a mostrar la índole de la conexión que hay entre dichas nociones, esto es, entre *tener una mente y ser persona*. Para ello, voy a reanudar el diálogo con Daniel C. Dennett que comencé en el capítulo 4. El interés que reviste, para mí, esta interlocución es el de exhibir de qué modo alguien que comparte muchos de los presupuestos de la filosofía de la mente de Davidson, nos permite, sin embargo, continuar la discusión ahí donde Davidson la bloquea por principio. En efecto, como es sabido, Dennett (1987) adopta la perspectiva interpretacionista de la *tercera persona* como “elección táctica”, bajo la asunción de la “actitud intencional”. En una primera aproximación, dicha actitud consiste en tratar al objeto cuyo comportamiento se quiere predecir como un agente racional con estados mentales que exhiben lo que Brentano llamo “intencionalidad”. Su idea seminal es que cualquier sistema cuyo comportamiento esté bien pronosticado por esta estrategia es, en el más completo sentido, un creyente. En suma: ser un “verdadero creyente” es ser un sistema intencional, un sistema cuyo comportamiento se puede predecir en forma fiable por medio de la actitud intencional. Asimismo, según vimos, Davidson parece coincidir con la estrategia de Dennett, quien muestra cómo es posible adoptar un hablar acerca de actitudes proposicionales, concibiéndolas como relaciones de un individuo con una proposición, y aceptando que tales actitudes desempeñan un rol causal sin que por ello debamos suponer que hay algo concreto en los cerebros de los individuos (como, por ejemplo, oraciones-caso del “lenguaje del pensamiento”) que intervenga en la historia causal en la que está involucrada la actitud proposicional, ni tampoco intermediarios psicológicos o epistémicos peculiares como los “objetos ante la mente” postulados en la modernidad. En este espíritu, hemos señalado que Dennett comparte con Davidson la suscripción de la tesis de la *indeterminación de la interpretación*. Si uno niega que al atribuir una creencia uno esté identificando un objeto con el cual el sujeto está en una relación peculiar de “captación”, no resulta necesario elegir una entre muchas interpretaciones exitosas como *el* contenido de la creencia. A su vez, Dennett (1987) se compromete también con el carácter *holista de lo mental*. En este sentido sostiene que “las creencias y deseos no son atribuibles aisladamente, independientemente de otras atribuciones de creencia y deseo” (1987/1998: 63). Y, a su vez, ofrece una concepción de la racionalidad en términos de la proyección de algunos principios *normativos*. En este espíritu en “Tres clases de psicología

intencional” afirma que (1) las creencias de un sistema son aquellas que *debería* tener, dadas su capacidad perceptiva, sus necesidades epistemológicas y su biografía; (2) Los deseos de un sistema son aquellos que *debería* tener, dadas sus necesidades biológicas y los medios más factibles para satisfacerlos y (3) la conducta de un sistema consistirá en aquellos actos que *sería racional* que un agente con esas creencias y deseos ejecutara (1987/1998:55).

Sin embargo, a pesar de todos los presupuestos que Dennett comparte con la concepción davidsoniana de lo mental, su tratamiento no monolítico acerca de qué es *ser humano*, *ser persona* y *tener una mente* habilita, en mi opinión, la posibilidad de contar una historia de continuidades, allí donde Davidson abre un hiato entre lo mental y lo no mental y cancela *a priori* toda posible discusión en torno al problema de la emergencia del pensamiento y el lenguaje. Recordemos que a su juicio:

Tenemos muchos vocabularios para describir la naturaleza cuando la consideramos no mental y tenemos un vocabulario mentalista para describir el pensamiento y la acción intencional; de lo que carecemos es de una manera de describir lo que ocurre entre estas dos cosas...Esto quiere decir que quizás haya un *problema insuperable* para dar una descripción plena de la aparición del pensamiento. Doy gracias de no estar en el campo de la psicología del desarrollo (Davidson 1997c/2001: 127-128; mis itálicas).

La consecuencia dramática que se sigue, pues, del dualismo conceptual profundo que tiene lugar a partir de la inscripción de la noción de *creencia* en el espacio “normativo” de las razones es la de detener, por principio, la discusión. Esto es, se la clausura *a priori*.

Frente a este estado de la cuestión, veamos entonces qué horizonte se puede abrir con la propuesta de Dennett, una vez que nos hemos despojado de la dicotomía tajante entre una dimensión normativa propia de lo mental y una dimensión meramente descriptiva de lo no mental, y una vez que advertimos las posibilidades que ofrece la perspectiva de segunda persona de la atribución mental. Si bien es cierto que Dennett no adopta explícitamente dicha perspectiva, en su caracterización práctica y social de la “actitud intencional” da cabida a fenómenos que, a mi juicio, resulta más adecuado describir desde la perspectiva de *segunda persona* de la atribución mental. La promesa que, en mi opinión, conlleva esta perspectiva (una vez que nos

desembarazamos del “cuarto dogma”) es la de ofrecer la posibilidad de “llenar” las discontinuidades davidsonianas.

En “Condiciones de la cualidad de persona”, Dennett (1989), examina seis tesis conocidas, cada una de las cuales pretende identificar una condición necesaria de la cualidad de persona, y cada una de las cuales resulta, en su opinión, correcta bajo cierta lectura. Lo que pretende mostrar es, primero, cómo de acuerdo con su interpretación, éstas son dependientes las unas de las otras; segundo, por qué son condiciones necesarias de la cualidad *moral* de persona; y tercero, por qué es tan difícil afirmar que, en conjunto, constituyen condiciones suficientes de la cualidad moral de persona.

En primera instancia, Dennett pretende distinguir tres nociones que en el tratamiento davidsoniano resultan coextensivas: *ser humano* (adulto y normal), *ser persona* y *tener una mente*. En su opinión, lo que importa respecto de nosotros no es que pertenezcamos a la misma especie biológica, sino que ambos seamos personas. La dignidad propia no depende de nuestros orígenes. Generalmente ignoramos esto, y consideramos la humanidad como la marca decisiva de la cualidad de persona, sin duda debido a que los términos, concede Dennett, son localmente coextensivos o casi coextensivos. Hoy por hoy, los seres humanos son las únicas personas que reconocemos, y reconocemos a casi todos los seres humanos como personas. Pero argumenta: por una parte, podemos fácilmente considerar la existencia de personas biológicamente muy diferentes -quizás habitantes de otros planetas- y, por la otra, reconocemos condiciones que niegan la cualidad de persona a los seres humanos, o al menos algunos elementos muy importantes de ella. Por ejemplo, a los seres humanos de poca edad, a los seres humanos con deficiencias mentales y a los seres humanos declarados dementes por psiquiatras facultados, se les niega la cualidad de persona, o cuando menos, elementos fundamentales de ésta.

Uno podría esperar que un concepto tan importante, aplicado y denegado con tal seguridad, tuviera condiciones de atribución suficientes y necesarias, y claramente formulables; pero, de ser así, a juicio de Dennett, aún no las hemos descubierto. En última instancia, podría ser que no hubiera nada que descubrir. En último caso, dice, podríamos llegar a darnos cuenta de que el concepto de persona es incoherente y obsoleto. Skinner es uno de los que, entre otros, ha sugerido esto; pero la doctrina no prosperó, sin duda debido, en parte, a que sería difícil, incluso imposible, concebir lo que sucedería si abandonáramos el concepto de *persona*. Puede argumentarse que la idea de que podamos dejar de ver a otros y a *nosotros mismos* como

personas (si esto no significa, simplemente, que nos aniquilemos y que, por ende, dejemos de ver nada como nada) es contradictoria en sí misma.<sup>205</sup> Por lo que, dejando a un lado lo correcto o erróneo de los fundamentos de la afirmación de Skinner, resulta difícil ver cómo podría ganarle la partida a una noción tan intuitivamente invulnerable. Si, entonces, el concepto de persona es, en cierta medida, una parte imprescindible de nuestro esquema conceptual, podría hallarse aún en peores condiciones de lo que quisiéramos. Podría resultar, por ejemplo, que el concepto de persona no sea sino un nebuloso término honorífico que gustosamente nos aplicamos a nosotros mismos y a los demás, como mejor nos parece, guiados por cuestiones de costumbre y similares -así como los *chic* son sólo aquellos que pueden lograr que los consideren *chic* otros que a sí mismos se tienen por *chic* -. Dennett afirma que ser una persona es, sin duda, algo similar a esto, y si no fuera más que esto, tendríamos que reconsiderar, de ser posible, la importancia que hoy atribuimos al concepto. Suponiendo que el ser una persona sí es algo más, el investigador de condiciones necesarias y suficientes encontrará, a su juicio, aún dificultades en el caso de que haya más de un concepto de persona, y existen motivos para sospecharlo. Parece, a grandes rasgos, que se encuentran aquí *entretajadas dos nociones de persona*, que Dennett propone denominar la noción *moral* y la noción *metafísica*. A continuación, nos interpela con la siguiente batería de preguntas: ¿Acaso la noción metafísica -*grosso modo*, la noción de un agente inteligente, consciente y sensible - *coincide* con la noción moral -*grosso modo*, la noción de un agente que es responsable, que tiene tanto derechos como obligaciones-? ¿O es tan sólo que el ser una persona en el sentido metafísico es una condición necesaria pero no suficiente para ser una persona en el sentido moral? ¿Acaso ser una entidad a la que se le atribuyen estados de conciencia o autoconciencia *es lo mismo* que ser un fin en sí mismo, o es solamente una condición *previa*?

Su tesis principal es la siguiente: aún suponiendo que la noción *metafísica* (tener una mente) y la noción *moral* (ser persona como fuente de derechos y obligaciones) son diferentes, parece haber motivos suficientes para creer que la cualidad metafísica de persona es condición necesaria de la cualidad moral de persona. Ahora bien, ¿por qué no nos es posible afirmar que estas condiciones necesarias para la cualidad moral de persona son también suficientes? Simplemente porque, dice Dennett, *el concepto de persona es "inevitablemente normativo"* (1989: 35; mis itálicas). Los seres humanos u otras entidades sólo pueden aspirar a ser

---

<sup>205</sup> Remite a Dennett (1973).

aproximaciones al ideal, y no hay forma de establecer una “calificación aprobatoria” que no sea arbitraria. Aún considerando suficientes las seis condiciones que va a explorar, no podrían garantizar que una entidad real cualquiera fuese una persona, pues nunca nada podría satisfacerlas. A juicio de Dennett, pues, la noción *moral* de persona y la noción *metafísica* de persona no son conceptos separados y distintos, sino sólo *dos puntos de apoyo diferentes e inestables de un mismo continuo* (1989:36)<sup>206</sup>. Hagamos, entonces, un breve recorrido por su tratamiento de las condiciones en cuestión. Este derrotero tiene para nosotros el interés de permitarnos retomar, especialmente, el *problema de la relación entre el pensamiento y el lenguaje* y el *problema de las perspectivas de primera y tercera persona de la atribución mental* (tratados en los capítulos 5 y 6, respectivamente).

La primera tesis -y para Dennett, la más obvia- afirma que las personas son *seres racionales*. Ésta aparece, por ejemplo, en las teorías éticas de Kant y Rawls y en las teorías metafísicas de Aristóteles y Hintikka. Por su parte, según vimos en el capítulo 4 de este trabajo, tal es la posición de Davidson, paradigmáticamente en “Mental Events” (1970/1980) y “Rational Animals” (1982/2001). En efecto, él afirma que:

En la medida en que no logramos descubrir un esquema *coherente y plausible* en las actitudes y acciones de los otros, simplemente renunciamos a la posibilidad de tratarlos como *personas* (1970/1980: 222; mis itálicas).

---

<sup>206</sup> A juicio de Dennett, esta relatividad contamina la satisfacción de las condiciones de la cualidad de persona en todos los niveles. No existe condición suficiente objetivamente satisfactoria que sostenga que una entidad *realmente* tiene creencias y, en la medida que descubrimos la irracionalidad que se manifiesta bajo una interpretación intencional de una entidad, se debilitan nuestros fundamentos para atribuir creencias, particularmente cuando tenemos –lo que en principio, a su juicio, siempre podemos tener- una descripción mecanicista, no intencional de la entidad. Exactamente en la misma forma, nuestra suposición de que una entidad es una persona se tambalea justo cuando se ha hecho algo mal y cuando surge la pregunta respecto de la responsabilidad. Pues en estos casos las bases para decir que una persona es culpable –la evidencia de que obró mal, de que estaba consciente de que obraba mal y de que obró mal por su propia y libre voluntad- son, en sí mismas, bases para dudar de que estemos de hecho tratando con una *persona*. Y si se preguntara qué podría resolver nuestras dudas, en su opinión, la respuesta sería: nada. Cuando surgen semejantes dudas ni siquiera podemos decir en nuestro propio caso si somos personas.

La segunda tesis sostiene que las personas son seres a los que se les atribuyen estados de conciencia o enunciados de intencionalidad, o psicológicos o mentales.<sup>207</sup> Según vimos en el capítulo 4 de nuestro trabajo, en este punto Davidson hace una inflexión respecto de la tradición cartesiana, para proponer una acepción más acotada de lo mental, esto es: se restringe a *actitudes proposicionales*. En este sentido, los animales racionales son aquellos que tienen actitudes proposicionales como creencias, deseos e intenciones.

A mi juicio, este tratamiento de lo mental en conjunción con la dicotomía descriptivo/normativo de la que me ocupé en el capítulo anterior, signa dramáticamente su posición en torno a los problemas de los que aquí nos ocupamos. En mi opinión, su interés seminal por la explicación de la acción racional (que se hizo patente en “Acciones, razones y causas”) apartó su mirada de la consideración de estados mentales del tipo que, precisamente, pretende recoger la perspectiva de *segunda persona* de la atribución mental, a saber: especialmente fenómenos emocionales, tales como los referidos en la primera parte de este capítulo. A mi juicio, al excluirlos del “mapa” de lo mental<sup>208</sup> en virtud de su concepción de la intencionalidad estrictamente proposicional o lingüística, Davidson perdió la posibilidad de trazar continuidades entre lo no conceptual y lo conceptual y entre las diferentes perspectivas de la atribución mental y, con ello, de dotar de genuina “carnadura” al carácter constitutivo de los vínculos *intersubjetivos*. Volveré, en breve, sobre este punto.

La tercera tesis sostiene que el que una cosa sea considerada persona, depende en cierta forma de la actitud que se adopta hacia ella. Esta tesis sugiere que no es una vez establecido el hecho objetivo de que algo es una persona cuando empezamos a tratarla de un modo particular, sino que este modo particular de tratarla forma parte, de alguna manera y hasta cierto punto, del

---

<sup>207</sup> En este punto Dennett remite a Strawson, quien identifica el concepto de persona como “el concepto de un tipo de entidad tal, que *tanto* los enunciados que le atribuyen estados de conciencia, *como* los enunciados que le atribuyen características corpóreas le son aplicables” (Cf. Strawson, 1959). Dennett sugiere que tal vez se trate de una caracterización demasiado amplia; abarca a todas las criaturas conscientes y activas. Se puede también discutir, dice Dennett, que los estados de conciencia son sólo un subconjunto adecuado de estados psicológicos o estados caracterizados intencionalmente, pero le parece claro que Strawson pretende aquí lanzar sus redes lo suficiente como para incluir los estados psicológicos en general.

<sup>208</sup> Aquí cabe hacer la aclaración siguiente. Si bien es cierto que Davidson hace algunas (escasísimas) referencias a las emociones (véase por ejemplo Davidson 1976b/1980 y /1982b/2004), las asimila al paradigma de las actitudes proposicionales y, en consecuencia, las considera asequibles sólo desde la perspectiva de tercera persona de la atribución mental, sin tomar en cuenta los aspectos específicos que se ponen en juego con ellas (a los que nos referimos en el primer apartado de este capítulo al aludir al tipo de fenómenos mentales que se recogen desde la perspectiva de *segunda persona*).

hecho de ser persona.<sup>209</sup> Según vimos en los capítulos 2 y 3 de nuestro trabajo, la elección davidsoniana de la perspectiva de la *tercera persona* como aquella desde la cual se ha de investigar el lenguaje, el pensamiento y su relación, está motivada por la idea de que el lenguaje es esencial y centralmente un fenómeno social y debe ser comprendido desde una dimensión pública. En este sentido, dijimos, Davidson sigue a Quine, quien revolucionó nuestra comprensión de la comunicación al tomar seriamente el hecho de que no puede haber nada más para el significado que lo que una persona adecuadamente equipada puede aprender y observar. En otros términos: aquello que liga la empresa interpretativa de Davidson con el proyecto quineano es que, en virtud de que el lenguaje es por su naturaleza un instrumento para la comunicación, el significado debe ser accesible desde un punto de vista público: los datos

---

<sup>209</sup> En su opinión, algunas variaciones de esta tesis han sido expresadas, precisamente, por Strawson (1974; de cuya propuesta me ocuparé –especialmente– en el próximo apartado), Amélie Rorty (1962), Putnam (1964), Sellars (1966), Thomas Nagel (1972) y por Dennett mismo. En efecto, con anterioridad, Dennett (1971) hizo referencia a estas tres primeras tesis: *racionalidad*, *intencionalidad* y *actitud*, para definir no a las personas, sino a una más amplia clase de lo que llamó *sistemas intencionales* y dado que pretende partir de esa noción, presenta entonces un breve resumen. Un sistema intencional es un sistema cuyo comportamiento puede ser –al menos en ocasiones– explicado y predicho recurriendo a atribuciones de *creencias* y *deseos* (y otros rasgos caracterizados por la intencionalidad –que llama “intenciones” queriendo referirse a esperanzas, temores, intenciones, percepciones, expectativas, etc.) al sistema. Puede haber en cada caso otras formas de predecir y explicar el comportamiento de un sistema intencional–por ejemplo, formas mecanicistas o físicas–pero la actitud intencional puede ser la más sencilla o la más efectiva o, en cualquier caso, una actitud cuya adopción resulta exitosa y que basta para el objetivo que es ser un sistema intencional. Así definidos, es obvio, dice Dennett, que los sistemas intencionales no son todos personas. Atribuimos creencias y deseos a los perros y peces y, por lo tanto, predecimos su comportamiento e incluso podemos usar este procedimiento para predecir el comportamiento de algunas máquinas. Es, por ejemplo, la estrategia que conviene adoptar, de hecho la única conveniente, ante una computadora que juega bien al ajedrez. La computadora es un sistema intencional en estas circunstancias, no porque tenga algún rasgo intrínseco particular, como tampoco porque tenga realmente creencias y deseos, sino tan sólo porque ella sucumbe ante cierta actitud adoptada con respecto a ella, a saber, la actitud intencional, la actitud que procede atribuyendo a la computadora enunciados intencionales bajo las limitantes usuales, la actitud que procede considerando a la computadora como un razonador racional práctico. El mismo Dennett advierte lo “blanda” que es esta definición de sistema intencional y cuán grande puede ser la correspondiente clase de los sistemas intencionales. Si, por ejemplo, predigo que una planta particular –digamos una azalea en una maceta– va a crecer dando la vuelta a una esquina, hacia arriba y en dirección a la luz, porque “busca” la luz, “quiere” salir de la sombra en que ahora se encuentra, y “cree” y “espera” que habrá luz al doblar la esquina, habremos adoptado la actitud intencional con respecto a la planta y dentro de límites muy estrechos ésta funciona. Dado que funciona, algunas plantas son sistemas intencionales de muy bajo nivel. Esta manera de hablar no es sólo pintoresca y no es en absoluto “supersticiosa”; es simplemente una forma eficaz de dar sentido, controlar, predecir y explicar el comportamiento de las plantas de un modo que sutilmente elude nuestra ignorancia de los mecanismos de control –de nivel más básico–. La abstención total del uso del lenguaje intencional puede volverse casi tan heroica y engorrosa y carente de sentido, dice, como el tabú paralelo de los conductistas estrictos al hablar de ratas y pichones. Y aunque las expresiones intencionales referentes, por ejemplo, a las actividades de la planta tengan un valor heurístico casi imperceptible, a Dennett le parece más inteligente dar por sentado que la planta es un sistema intencional soso e insignificante, que tratar de trazar una línea por arriba de la cual las interpretaciones intencionales sean “objetivamente verdaderas”.

relevantes se restringen, pues, a lo disponible para el intérprete radical, a lo asequible desde la perspectiva de la *tercera persona*. En este sentido, leemos:

Pues no podemos empezar a descifrar lo que dice una persona si no comprendemos sus actitudes en torno a sus oraciones, tales como sostener, desear, o querer que sean verdaderas. Empezando por estas actitudes, debemos construir una teoría de lo que la persona quiere decir, dando así, simultáneamente, contenido a sus actitudes y a sus palabras. En nuestra necesidad de hallarla comprensible, trataremos de construir una teoría *que la haga consistente, creyente de verdades y amante de lo bueno* (1970/1980:222; mis itálicas).

Para Dennett es obvio que ser un sistema intencional no es condición suficiente para ser persona, pero sin duda es condición necesaria. Nada hacia lo cual no podamos adoptar satisfactoriamente una actitud intencional, con sus presupuestos de racionalidad, podrá considerarse persona. ¿Podemos entonces definir a las personas como una subclase de los sistemas intencionales? A primera vista podría parecer provechoso suponer que las personas son sólo la subclase de los sistemas intencionales que *realmente* tienen creencias, deseos y demás, y de los que no sólo *suponemos* que los tienen por mor de una predicción expedita. Pero, según vimos en nuestro capítulo 4 (a instancias de lo elusiva que resulta la posición de Davidson respecto de la dicotomía realismo/antirrealismo), parece que todos los intentos por definir aquello de lo que se puede decir que *realmente* tiene una creencia -de manera que ni un perro, ni un árbol, ni una computadora pudieran resultar idóneos- terminan poniendo, a juicio de Dennett condiciones a la creencia genuina, condiciones que (1) son demasiado estrictas para nuestras intuiciones, y (2) aluden a distintas condiciones de la cualidad de persona que se encuentran en los últimos lugares de su lista.

Considera, entonces, que se puede lograr un mayor avance si considera la cuarta tesis, la de *reciprocidad*, para ver qué tipo de definición podría recibir en términos de sistemas intencionales. Esta tesis dice que el objeto con respecto al cual se adopta esta postura intencional debe ser capaz de *reciprocarse* de alguna manera. Dennett considera que algunas versiones muy variadas de esto han sido expresadas o insinuadas por Rawls, MacKay, Grice y Strawson. La tesis sugiere que una persona debe ser capaz de “reciprocarse” la actitud, lo que sugiere que un sistema intencional que, a su vez, adoptara una actitud intencional hacia otros objetos pasaría la

prueba. Dennett define un “sistema intencional de segundo orden” como aquel al cual le atribuimos no sólo creencias, deseos y demás intenciones simples, sino creencias, deseos y demás intenciones respecto de creencias, deseos y demás intenciones. Así, un sistema intencional *S* será un sistema intencional de segundo orden si, entre las atribuciones que le hacemos, se encuentran las del tipo *S cree que T desea que p*, *S espera que T tema que p*, y casos reflexivos como *S cree que S desea que p*<sup>210</sup>.

Ahora bien, ¿son acaso los seres humanos, hasta donde sabemos, los únicos sistemas intencionales de segundo orden? Para Dennett ésta es una pregunta *empírica*. En este punto es donde se evidencia claramente, a mi juicio, la diferencia en el talante de la indagación de los dos filósofos que pusimos en interlocución: ahí donde Dennett hace explícito su interés por la dimensión empírica de la cuestión, Davidson hace manifiesto que se restringe a estudiar rasgos de los conceptos. Y estos compromisos dispares tienen, desde luego, consecuencias. Dice, Dennett, atribuimos creencias, por ejemplo, a los perros, los gatos, los leones, los pájaros y los delfines y a partir de ahí predecimos con frecuencia su comportamiento, pero es difícil imaginar un caso en el que la conducta del animal fuese tan sofisticada que necesitémos atribuirle intenciones de segundo orden para poder predecir o explicar su comportamiento. Por supuesto, que si alguna versión del fisicalismo mecanicista es verdadera –como Dennett cree que es– no *necesitaremos* nunca atribuir intención alguna a nada, pero suponiendo que por razones heurísticas y pragmáticas tuviésemos que atribuir intenciones a los animales ¿sentiríamos acaso el impulso pragmático de atribuirles intenciones de segundo orden?<sup>211</sup> Por un lado, Dennett admite por ejemplo, que Fido *quiere* su cena y *cree* que su amo le dará la cena si gime frente a él. Pero por otro lado, y contrariamente al espíritu de Davidson (según vimos en el capítulo 5, a propósito del caso de la sorpresa, desde el cual Davidson va a sostener la premisa que identificamos como (B2) en su argumento a favor de la dependencia del pensamiento respecto

---

<sup>210</sup> Algunos pensarán que los casos reflexivos convierten automáticamente a todos los sistemas intencionales en sistemas de segundo orden e incluso en sistemas de orden *n*, sobre la base de que creer que *p* implica creer que se cree que *p*, y así sucesivamente; pero éste, dice Dennett, es un error fundamental. La repetición de creencias y demás intenciones no es nunca redundante, y así, aún cuando algunas repeticiones son normales nunca son triviales o automáticas.

<sup>211</sup> Con frecuencia los psicólogos han recurrido a un principio llamado el canon de parsimonia de Lloyd Morgan, que puede considerarse como un caso especial de la navaja de Occam; el principio es que se debe atribuir a un organismo tan poca inteligencia o conciencia o racionalidad o pensamiento, como baste para dar cuenta de su comportamiento. Este principio puede interpretarse –y ha sido interpretado– nada menos que como un conductismo radical; pero esto le parece un error a Dennett y dice que podemos interpretarlo como el principio que requerimos al adoptar la actitud intencional con respecto a un objeto, para atribuirle creencias, deseos y demás estados más simples, menos sofisticados y de menor orden, que darán cuenta de su comportamiento.

del lenguaje, a saber: que para tener una creencia es necesario tener el *concepto* de creencia), Dennett dice que no necesitamos atribuirle a Fido la *creencia* posterior de que su gemido produce en su amo la creencia de que él, Fido, *quiere* su cena. De manera similar, mi *expectativa* al depositar una moneda en la máquina de dulces no depende de la creencia posterior de que al insertar la moneda, se induce en la máquina la *creencia* de que yo *quiero* un dulce. Esto es, que aún si el gemido de Fido se parece mucho a una verdadera interacción de segundo orden-en la que Fido trata a su amo como a un sistema intencional- si suponemos que para Fido su amo es tan sólo una máquina proveedora de alimento que se activa mediante su gemido, obtendremos una atribución predictiva igualmente buena, más modesta, pero que sigue siendo intencional.

¿Son entonces los perros, los chimpancés y demás animales “superiores” incapaces de elevarse al nivel de los sistemas intencionales de segundo orden? Y de ser así ¿por qué? Dennett dice que solía pensar que la respuesta era afirmativa, pensando que la razón era la falta de lenguaje en los animales no humanos, y que el lenguaje era necesario para representar las intenciones de segundo orden. Lo atraía la hipótesis de que los animales no pueden, por ejemplo, tener creencias de segundo orden, creencias acerca de creencias, por la misma razón por la que no pueden tener creencias sobre el día viernes o sobre poesía. Hasta aquí, su planteo era sumamente consonante con el de premisa que en el capítulo 5 identificamos como (B3) en el argumento de Davidson a favor de la dependencia del pensamiento respecto del lenguaje, a saber: para tener el concepto de creencia es necesario tener un lenguaje.

Pero, dice Dennett (1989), si bien es cierto que algunas creencias no pueden adquirirse sin lenguaje, es falso que todas las creencias de segundo orden se encuentren entre ellas y es falso también que los no humanos no puedan ser sistemas intencionales de segundo orden.<sup>212</sup> A su juicio, cuando *x* trata de provocar una conducta en *y*, conducta que resulta inapropiada para el entorno y las necesidades *reales* de *y*, pero apropiada para el entorno y las necesidades que *y* *percibe* o *crea*, nos vemos obligados a atribuir intenciones de segundo orden a *x*. Una vez establecido de esta manera, este aspecto parece familiar, y frecuentemente es utilizado por los críticos del conductismo: se puede ser conductista en la explicación y control de la conducta de

---

<sup>212</sup> Ahí donde Davidson (1982) considera el caso del perro “Arturo” para respaldar la idea de que “para tener cualquier actitud proposicional es necesario tener el concepto de creencia, tener una creencia sobre alguna creencia” y afirmar que para ello es necesario tener lenguaje, Dennett considera, en cambio, el ejemplo de la perra de Peter Ashley. En este caso, dice, parece que tenemos que atribuirle a la perra (un animal no lingüístico) la *intención* de que su amo *crea* que ella *quiere* salir –una intención no sólo de segundo, sino de tercer orden-. La clave de este ejemplo es que la creencia que la perra pretende causar en su amo es falsa. Propone que pensemos en algunos casos de engaño, en los que el animal creyendo que *p* trata de hacer que otro sistema intencional crea *no-p*

animales de laboratorio sólo en tanto se pueda confiar en que no hay una grave dislocación entre el entorno real del experimento y el entorno percibido por los animales. Una táctica para poner en aprietos a los conductistas en el laboratorio consiste en efectuar experimentos que engañen a los sujetos; si el engaño tiene éxito, su conducta será predecible a partir de sus *creencias* falsas respecto del entorno, no a partir del entorno real. Ahora bien, un sistema intencional de primer orden es conductista; no atribuye intención alguna a nada. Así pues, si deseamos tener evidencia efectiva de que un sistema dado *S* no es conductista -sino que es un sistema intencional de segundo orden-, esto sólo ocurrirá en los casos en que las teorías conductistas no se adecuen a los datos, sólo en aquellos casos en que el conductismo no explicaría el éxito del sistema *S* cuando manipula la conducta de otro sistema (por *default* del abordaje conductista). A su juicio, entonces, esto sugiere que el ejemplo de Ashley no es, después de todo, tan convincente; en el sentido de que puede ser rechazado suponiendo que su perra es una conductista menor. No es preciso que ella crea que al arañar la puerta llevará a Ashley a creer que ella quiere salir; basta con que ella crea, como buena conductista, que ha condicionado a Ashley para que vaya a la puerta cuando ella la araña. Así, ella aplica el estímulo usual, obtiene la respuesta usual, y eso es todo. El caso de Ashley se derrumba si ésta es la forma normal como su perra logra que se le abra la puerta, como probablemente lo es, pues entonces una hipótesis más modesta será que la perra cree que su amo está condicionado para ir a la puerta cuando ella la araña. Si la perra hubiera hecho algo novedoso (como correr hacia la ventana, mirar hacia afuera y gruñir sospechosamente), entonces tendríamos, dice Dennett, que reconocer que el levantarse de la silla no fue simplemente una respuesta condicionada de Ashley, y no podría ser “considerada” como tal por su perra, pero entonces, tal virtuosidad en un perro resultaría altamente improbable.<sup>213</sup>

---

<sup>213</sup> ¿Cuál es la diferencia entre el caso improbable y los muy conocidos casos en los que un pájaro que tiene su nido al ras del suelo finge que se le ha roto un ala, para alejar al depredador del nido? El efecto logrado es una novedad, en el sentido de que el pájaro con toda seguridad no ha condicionado repetidamente a los depredadores de los alrededores mediante este estímulo; por ello nos vemos obligados a explicar la artimaña como un pequeño y genuino engaño, mediante el cual el pájaro *busca* provocar una *creencia* falsa en el depredador. Llevados a aceptar esta interpretación de la conducta, nos podría impresionar enormemente la ingeniosidad del pájaro, de no ser por el hecho de que sabemos que dicho comportamiento es “meramente instintivo”. Afirmar que es instintivo es afirmar que todas las especies de pájaros lo hacen; lo hacen aún cuando las circunstancias no son totalmente apropiadas; lo hacen cuando existen mejores motivos para permanecer en el nido; el patrón de comportamiento es rígido, un tropismo de menor relevancia y, presumiblemente los controles fueron transmitidos genéticamente, no aprendidos ni inventados. No obstante, dice Dennett, debemos tener cuidado de no llevar muy lejos esta actitud minimizadora; no es que el pájaro realice el truco “irreflexivamente”, pues si bien es claro que en ningún momento cruzó por su mente un argumento o esquema, un hombre puede hacer algo igualmente astuto y con genuina inteligencia, novedad y adecuación, sin por ello valerse tampoco de “pensamientos conscientes”. El *pensar los pensamientos*, como quiera que esto se caracterice, no es lo que hace realmente inteligente a una conducta inteligente... Cuando nos

Dice Dennett: yo sería el último en negar o ignorar la enorme diferencia que hay entre una conducta instintiva o tropista, y la conducta más versátil e inteligente de los humanos y de otros; pero en lo que quiere insistir aquí es en que, si uno está preparado para adoptar sin escrúpulos una actitud intencional como herramienta para predecir y explicar la conducta, el pájaro es entonces un sistema intencional de segundo orden, como cualquier hombre.<sup>214</sup>

Ahora bien, llegados a este punto cabe preguntarnos, entonces, por la peculiaridad del *reciprocarse* que se pone en juego en el caso humano. A mi juicio, aquí la perspectiva de segunda persona puede permitirnos arrojar luz. Recordemos, en primera instancia, que Gomila (2002) concibe a la segunda persona como una perspectiva de atribución psicológica. Pero aquí tenemos que ser cautos con el término "atribución", puesto que no tiene por qué involucrar el empleo de conceptos por parte de los individuos que hacen la atribución. En efecto, según indicamos en la primera parte de este capítulo, se trata de mecanismos de reactividad emocional, en los cuales un individuo reacciona ante el estado psicológico del otro, cambiando sus propios estados psicológicos, sin que ello suponga que los dos sujetos en juego están en posesión del concepto psicológico relevante para describir el estado psicológico del otro. Por ejemplo, la madre se asusta y por un mecanismo de "contagio emocional" el bebé se asusta. Esto no supone que el bebé deba por ello poseer el concepto de miedo que un humano adulto normal tiene. Según indicamos, pues, hay *reciprocidad* en tanto cada uno atribuye estados psicológicos al otro y en el mismo acto hay modificación de los propios contenidos mentales. Desde luego, soy plenamente consciente de que, en lo que respecta al tema de nuestro interés, puede disputarse que el ejemplo en cuestión cuente como caso de atribución "intencional", esto es: reaccionar más o menos inconsciente o automáticamente a las conductas de los demás no puede asimilarse a la compleja y radical operación de asignar significados y contenidos mentales a un sujeto de una cultura desconocida. En este sentido, cabe aclarar que la consideración que propugno de la perspectiva de segunda persona no pretende, en absoluto, socavar el tipo de atribución intencional sofisticada que tiene lugar entre humanos adultos normales, sino simplemente permitirnos advertir que si se la concibe de un modo más abarcador (esto es, tal que no restrinja

---

preguntamos por qué los pájaros evolucionaron con este tropismo, lo explicamos subrayando la utilidad de poseer un medio para *engañar* a los depredadores, o de provocar en éstos creencias erróneas; lo que debe explicarse es el origen de las intenciones de segundo orden del pájaro.

<sup>214</sup> Así pues, la *reciprocidad*, siempre que se entienda por ésta simplemente la capacidad de los sistemas intencionales para exhibir intenciones de orden superior, si bien depende de las tres primeras condiciones, es independiente, según veremos, de la quinta y la sexta.

la intencionalidad a direccionalidad a contenidos proposicionales) es posible, al menos en principio, trazar una historia de continuidades entre los estadios no conceptuales y los conceptuales en el sentido cabal davidsoniano. La idea que está detrás de este planteo es que sin los mecanismos básicos las operaciones complejas no podrían tener lugar. En otras palabras: sin la perspectiva de *segunda* persona no podría tener lugar la de *tercera*, esto es, sin transitar una dimensión participativa o de involucramiento con otro próximo, no podríamos llegar a ser intérpretes distantes de la conducta de un tercero.

En la caracterización de la perspectiva de *segunda persona* que he tomado de la reflexión contemporánea en filosofía de la mente, no parece haber necesariamente un mundo objetivo compartido. Más que un “triángulo”, la situación paradigmática respecto de la perspectiva de *segunda persona* parece ser la de una díada. Sin embargo, dado el hecho de que algunas emociones son intencionales, el mundo entra a través de nuestras emociones; a este rasgo Gomila lo denomina el “carácter triádico de la perspectiva de segunda persona”<sup>215</sup> y es lo que, en mi opinión, promete la posibilidad de trazar la conexión conceptual con la comunicación lingüística y la triangulación en sentido davidsoniano.<sup>216</sup>

Desde la psicología del desarrollo, se ha puesto el énfasis en ciertos hechos que constituyen un conjunto de estadios intermedios entre la situación de la llamada “intersubjetividad primaria” – que atañe a los primeros meses de vida (3 a 8 meses), donde los bebés no exhiben conductas que nos autoricen a atribuirles concepto psicológico alguno- y la situación de un nene que pasa exitosamente el test de falsa creencia, considerada –según vimos en el capítulo 5- la prueba<sup>217</sup> de que posee nuestro concepto humano adulto normal de *creencia*. Formulados de manera sumamente esquemática, los estadios que trazan las continuidades entre dichas situaciones son los siguientes<sup>218</sup>:

0. Desde el nacimiento el bebé es capaz de hacer imitaciones de las expresiones del rostro del adulto (Meltzoff y Moore 1977, 1989, 1994) y también es capaz de reaccionar ante la emoción del adulto (Riviere y Sotillo 1999/2003).

---

<sup>215</sup> Carácter que parece recoger una acepción de la intencionalidad que se aparta del sentido davidsoniano más estrecho que la restringe a la direccionalidad hacia contenidos proposicionales.

<sup>216</sup> Desde luego, los detalles de tal conexión exceden mi propósito en el presente trabajo, en el que simplemente estoy intentando habilitar un espacio para comenzar a discutirlos.

<sup>217</sup> Cabe aclarar que el test en cuestión ha sido blanco de múltiples críticas, cuya consideración, no obstante, trasciende completamente los límites del presente trabajo.

<sup>218</sup> Sigo los lineamientos generales de la reconstrucción que ofrece Pérez (inédito)

1. Durante el transcurso del primer año de vida aparecen una serie de capacidades atencionales y perceptuales que suelen ser citadas como precursores de la teoría de la mente, como el mecanismo de detección de la mirada (seguir la mirada del otro: 6 meses), el mecanismo de alternancia de la mirada (mirar a un objeto y a otra persona alternadamente: 9 meses), el mecanismo de atención conjunta (en este estadio es donde se suele ubicar el “primer triángulo”: bebé-madre-mundo: 9 meses; instancia que en la terminología de Trevarthen (1980) recoge la denominada “intersubjetividad secundaria”, esto es, la comunicación adulto-bebé gira, no en torno a la relación entre ambos, sino en torno a objetos y situaciones del mundo compartido)<sup>219</sup>.
2. Hacia los doce meses aparecen los protoimperativos y protodeclarativos: Los primeros pretenden influir en la acción del otro; los segundos suponen la capacidad de comunicar al otro algo sobre algo que está ahí afuera de los dos. Se suele considerar que ambos ponen en juego lo que marca la diferencia entre monos y humanos.
3. A los 18 meses aparece el juego de ficción, que es considerado por muchos como el primer precursor de la teoría de la mente porque allí aparece el desacoplamiento que está detrás de una mente representacional, esto es: algo (banana) que está en lugar de otra cosa (teléfono) que está ausente.
4. A los tres años hay lo que se suele denominar una "psicología de percepción y deseos". Durante este período los niños adquieren un conjunto de conceptos para actividades mentales representacionales tales como pensar, soñar, imaginar y simular y muestran maestría en el uso de los términos para referirse a esos conceptos entre el segundo y el cuarto año de vida (Olson, Astington y Harris, 1988, Wellman, Phillips y Rodríguez 2000, Wellman 2004). Según Harris (1992) aquí es cuando el nene logra entender que el otro puede diferir de uno en cuanto a lo que conoce del mundo, o lo que ve del mundo (una forma de conocer), o diferencias respecto de los deseos, o sea que puede haber diferencias respecto de la *referencia* de los estados mentales.
5. Al pasar el test de falsa creencia, hacia los 4 años y medio, se dice que se ha alcanzado el desarrollo completo de la "psicología de creencias y deseos", que es la que pervive aún en la

---

<sup>219</sup> Trevarthen (1980) diferencia entre intersubjetividad primaria e intersubjetividad secundaria, con la finalidad de distinguir entre las respuestas que los bebés están en condiciones de dar ante la figura de apego, antes de la adquisición del lenguaje. La intersubjetividad primaria hace referencia a la capacidad de compartir los estados afectivos que surgen en el seno de la diada figura de apego-bebé. Por su parte, la intersubjetividad secundaria hace referencia a la capacidad de compartir estados mentales relacionados con los objetos, es decir, estados mentales relacionados con cosas externas a la diada figura de apego-bebé.

edad adulta. Se manifiesta, entonces, la capacidad metarrepresentacional; se puede decir que el niño adquirió completamente los conceptos psicológicos comunes a todos los humanos adultos normales, indispensables para la comprensión de las acciones y mentes ajenas.

La quinta tesis dennettiana afirma que las personas deben ser capaces de *comunicación verbal*<sup>220</sup>. Esta condición deniega fácilmente a los animales no humanos la plena cualidad de persona y la responsabilidad moral concomitante, y parece al menos implícita en todas las teorías de contrato social de la ética. Es también una tesis que ha sido sostenida enfáticamente o presupuesta por muchos escritores de la filosofía de la mente –desde luego, Davidson (1975 y 1982) y el propio Dennett entre ellos- y en la cual, sostiene, no se ha sometido a discusión la dimensión *moral* de la cualidad de persona (tal es el sentido que intentaré alumbrar en el apartado próximo, también a partir de la perspectiva de segunda persona).<sup>221</sup>

---

<sup>220</sup> Finalmente, la sexta tesis sostiene que las personas se diferencian de otras entidades por ser conscientes de una manera particular; nosotros somos conscientes de una forma en que ninguna otra especie lo es. En ocasiones, ésta se conoce como *autoconciencia*, de uno u otro tipo. A su juicio, tres son los filósofos que afirman –de muy diversas maneras- que un tipo especial de consciencia es una condición previa para un ser un agente moral, a saber: Anscombe en *Intention*, Sartre en *La Trascendencia del ego*, y Harry Frankfurt en “Freedom of the will and the concept of a person”. Por razones de pertinencia y extensión, no me ocuparé aquí de tematizar esta tesis, sino que me limitaré a señalar la afinidad de Davidson con la posición de Anscombe. En efecto, *si ha de hacerse responsable de una acción* (una fracción de mi conducta descrita de una manera particular) debí estar *consciente* de esa acción descrita de esa manera particular. Porque sólo estando consciente de la acción puedo decir lo que intentaba hacer, y participar desde una posición privilegiada en *el juego de preguntas y respuestas para dar los motivos de mis actos*. Y lo importante de poder participar en este juego es que sólo a los que son capaces de participar en el juego de dar explicaciones de les puede o no hacer responsables del curso de sus actos. Y si somos incapaces de “oír razones” en algún asunto, entonces no se nos puede hacer responsables de dicho asunto. Las capacidades de comunicación verbal y de conciencia de nuestros propios actos son, pues, esenciales para quien esté dispuesto a oír argumentos y abierto a la persuasión, y semejante persuasión, semejante ajuste recíproco de intereses logrado mediante la utilización mutua de la racionalidad, es un rasgo de la modalidad óptima de la interacción personal.

<sup>221</sup> Para esta quinta condición, la capacidad de comunicación verbal, recurre a la teoría del significado de Grice. Como es sabido, Grice trata de definir lo que él llama “significado no natural”, lo que pretende quien se expresa al expresar algo, en términos de las *intenciones* del que se expresa. Su definición inicial es la siguiente (los escritos clave son “Meaning” y “Utterer’s Meaning and Intentions”. Su formulación inicial, desarrollada en el primer artículo, está sujeta a una serie de revisiones en el segundo artículo, del cual se extrae la formulación siguiente):

“U pretendió decir algo expresando x” es verdad si, para algún oyente A, U expresó x con la intención de que:

- (1) A produjera una respuesta particular r.
- (2) A pensara (reconociera) que U tenía la intención de (1).
- (3) A cumpliera (1) sobre la base de su cumplimento de (2)

Obsérvese que la intención (2) atribuye a U no sólo una intención de segundo, sino de tercer orden: U debe *tratar* de que A reconozca que U *trata* que A produzca r.

Del análisis del significado no natural de Grice se desprenden dos puntos de gran importancia para Dennett. Primero, dado que el significado no natural, pretender significar algo al decir algo, debe ser un rasgo de toda comunicación verbal verdadera, y dado que depende de las intenciones de tercer orden por parte del que se expresa, Dennett sostiene que la condición 5 se apoya en la condición 4, y no a la inversa. Segundo, Grice nos muestra que no bastan las simples intenciones de segundo orden para producir reciprocidad genuina; para ello se requieren intenciones de tercer orden. Debe haber, dice, un “encuentro real”, un “reconocimiento mutuo” entre el que se expresa y el oyente para que el que se expresa logre significar algo; pero los encuentros pueden darse en ausencia de significado no natural (la perra de Ashley)<sup>222</sup> y las artimañas que dependen de intenciones de tercer orden no necesitan involucrar encuentros (por ejemplo, A puede tratar de que B crea que C desea que p). Así, las intenciones de tercer orden son una condición necesaria pero no suficiente para la ocurrencia de casos de significado no natural, esto es, casos de comunicación verbal. El aspecto que me interesa destacar de este punto señalado por Dennett, es que abre la posibilidad de contar una historia de continuidad acerca del *problema de la relación del pensamiento, el lenguaje y su emergencia* allí donde, desde el talante estrictamente conceptual del planteo de Davidson, tal posibilidad se cancela *a priori*. Podría argumentarse que aquello que se pone en juego en el “reconocimiento mutuo” o el “encuentro real” es la perspectiva de *segunda persona*.

En mi opinión, parece innegable aceptar que las primeras interacciones sociales a las que están sometidos los bebés de la especie humana poseen características peculiares y excluyen el uso, por parte del bebé, de conceptos psicológicos humanos adultos normales, o al menos de algunos conceptos sofisticados como "creencia". Pero cabe señalar aquí que no ocurre lo mismo con los adultos cuando interactúan con el bebé. Esta asimetría es básica. Esto hace que ninguna interacción humana esté completamente “ayuna” de conceptos, o, en otros términos: ninguna interacción humana real es completamente describible en términos meramente físicos, de la manera que, según vimos, pretende Davidson describir el triángulo inicial en el proceso de

---

No importa en absoluto que Grice se haya visto obligado por una serie de contraejemplos a ir de esta definición inicial hacia versiones mucho más complejas, ya que todas ellas reproducen la intención de tercer orden de (2).

<sup>222</sup> En este sentido, la reciprocidad bajo la acepción de Dennett parece altamente compatible con el tipo de interacción que se pone en juego en la perspectiva de segunda persona de la atribución mental, según vimos en la caracterización ofrecida en la primera parte de este capítulo.

aprendizaje del lenguaje.<sup>223</sup> En este sentido, no habría un hiato entre una instancia meramente descriptivo/causal de reconocimiento automático de patrones desprovista de toda normatividad y una instancia normativa de articulación conceptual plena. Si la interacción en cuestión está “atravesada” desde el principio por los conceptos mentales y su concomitante normatividad (toda vez que el adulto humano trae al chico al mundo de la cultura), no hay por qué concebir la emergencia del lenguaje y el pensamiento como un paso de “prestidigitación”, sino que se habilita la posibilidad de encontrar continuidades.

Como señalamos en el capítulo 5, Davidson admite tres aplicaciones diferentes de la idea de triangulación: primero, como una descripción de las condiciones necesarias (completamente no-intencionales) para el pensamiento; segundo, como una descripción de la situación de aprendizaje en la que la criatura aprende su primer lenguaje (la situación en la que pensamiento y lenguaje “emergen”); y tercero, la situación en la que se determinan los contenidos de nuestras creencias perceptuales. El punto que deseo enfatizar es que no hay interacción humana que sea meramente descriptivo-causal. Recordemos que- como señalamos en el capítulo 3- Davidson (1987b/2001) caracteriza su singular tesis externista, la triangulación, a la vera de las posiciones de Putnam y Burge. Frente a la pregunta acerca de cómo puede una persona determinar lo que está en otra mente, sostiene que una parte de la respuesta depende de que, en los casos más simples, los eventos y objetos que causan una creencia determinan también los contenidos de la misma. Así, la creencia que es causada distintivamente y en condiciones normales por la presencia de algo amarillo o de la propia madre, es la creencia de que algo amarillo o la propia madre están presentes. La idea, dijimos, no es, desde luego, que la naturaleza garantiza que nuestros juicios más llanos sean siempre correctos sino que la *historia causal* de tales juicios representa un importante rasgo constitutivo de sus contenidos, punto que lo emparenta con el externismo de Putnam. A su vez, su peculiar propuesta externista, aunque expresamente distinta de la de Burge, pretende ubicar el *rol de la sociedad dentro del nexo causal* que incluye la interacción entre las personas y el resto de la naturaleza. La identificación de los objetos del pensamiento reposa así en una *base social*: sin una criatura para observar a la otra, la

---

<sup>223</sup> Pérez (inédito) advierte que esto marca una diferencia entre una posible historia ontogenética y una filogenética del origen de la psicología *folk*, en el sentido de que en esta última debe de haber habido en algún momento un triángulo puramente físico. No obstante, la consideración filogenética excede ampliamente los límites del presente trabajo, en el que –deliberadamente– sólo me atengo (muy acotadamente) a la dimensión ontogenética. Aquí simplemente pretendo habilitar un horizonte conceptual para empezar a explorarla y, en consecuencia, no persigo en absoluto su explicitación acabada.

triangulación que ubica a los objetos relevantes en un espacio público no podría tener lugar. En este sentido, subrayamos que Davidson afirma que no podemos resolver la cuestión de los contenidos de los estados mentales desde el punto de vista de una criatura única. De este modo, su externismo, lejos de postular “objetos ante la mente” ubica a los pensamientos en un mapa público.

Ahora bien, esta mediación intersubjetiva de la causalidad ilustra, a mi juicio, el modo en que la dimensión descriptiva está *constitutivamente atravesada* por la normatividad que resulta, así, endémica. Recordemos que desde la perspectiva interpretativa, la normatividad es siempre normatividad *atribuida*. De este modo, a partir de interactuar con un adulto intérprete que no puede sino proyectar las restricciones estructurales del juego interpretativo, el niño que con él triangula, se vincula con los objetos y eventos del mundo mediante lazos *causales* que son *inescindibles* de la *dimensión normativa*. Y esto es así, específicamente, porque, según vimos, en el esquema davidsoniano hay un fuerte nexo entre los conceptos de *segunda persona*, *error* y *normatividad*. Recordémoslo brevemente.

¿De dónde sacamos la idea de que podemos estar equivocados, de que las cosas pueden que no sean como pensamos que son? Para Davidson es Wittgenstein quien nos dio la pista: no tendríamos el concepto de hacer las cosas correcta o incorrectamente si no fuera por nuestras interacciones con los otros. Lo que introduce la posibilidad de error es el fallo ocasional en lo que se espera que el otro haga, en las expectativas insatisfechas, en las reacciones que no se correlacionan. Desde luego, como señalamos, la idea no es que el consenso defina el concepto de verdad, sino que crea el espacio para su aplicación. Y de este modo resulta, pues, que el pensamiento y el lenguaje son necesariamente sociales. Una criatura no puede tener pensamientos ni lenguaje sin poseer el concepto de verdad intersubjetiva, y este es un concepto que no se puede tener sin compartir y saber que uno comparte con alguien más un mundo y una manera de pensar acerca del mismo.

De este modo, una vez que prescindimos del “corsé” que introduce la dicotomía descriptivo/normativo, habilitamos un horizonte para poder contar desde el ámbito de la indagación conceptual <sup>224</sup> cómo es que pasamos de un triángulo en el cual hay dos seres

---

<sup>224</sup> Cabe advertir que la perspectiva de segunda persona de la atribución mental resulta pasible tanto de un abordaje conceptual, como de un abordaje empírico, en el que los esfuerzos se cifran, especialmente, en clarificar el mecanismo psicológico subyacente. El modo en el que yo la abordo pretende, no obstante, ser más afin al espíritu estrictamente conceptual de la reflexión davidsoniana.

humanos y un mundo conectados causalmente entre sí, al triángulo humano adulto normal, donde hay dos humanos y un mundo externo, humanos que nos atribuimos los unos a los otros estados intencionales acerca de un mundo compartido, que somos capaces de dar razones de nuestras acciones y creencias, y que podemos entablar un intercambio comunicativo. Esto es, una vez que prescindimos de una escisión tajante entre una dimensión meramente descriptiva relativa al contenido no conceptual que se pondría en juego en el intercambio emocional (recogido desde la perspectiva de segunda persona) y una dimensión especialmente normativa relativa al contenido conceptual de nuestras actitudes proposicionales (recogido desde la perspectiva de tercera persona), se habilita un espacio conceptual para trazar continuidades y dar una descripción de la emergencia del pensamiento y el lenguaje, ahí donde Davidson lo cancela por principio.

Un esfuerzo en un espíritu similar al que asumo aquí -aunque valiéndose de otros recursos conceptuales- puede encontrarse en Quintanilla (2008). Allí, él considera que la interpretación, si es tal, sólo puede ser *participativa* y que no hay tal cosa como la interpretación *no comprometida* (ya que la atribución de estados mentales en el proceso de triangulación involucra siempre la interacción con otros, tomando como referencia los objetos que asumimos compartir con ellos<sup>225</sup>) y esboza una teoría del contenido que entrelaza la dimensión conceptual o cognitiva (propia de la tercera persona) con la dimensión emocional (propia de la segunda). Charles Taylor (1985) piensa que al producir oraciones-T el intérprete puede dar sentido a oraciones que pretenden describir el mundo, pero no a oraciones que expresan experiencias personales y complejas. Quintanilla, en cambio, sostiene que se podría contra argumentar que las oraciones-T pretenden explicar el contenido cognitivo y no emocional, pero por dos razones, dice, esa sería una mala respuesta. En primer lugar, porque una teoría de la interpretación debe poder explicar qué es comprender a una *persona* y no sólo sus oraciones con contenido cognitivo. En segundo lugar, porque en la mayor parte de los casos es difícil, si no imposible,

---

<sup>225</sup>El rodeo de Quintanilla es vía la idea de simulación. Él desea subrayar un punto que Davidson no menciona: la interpretación requiere también del fenómeno de la simulación, esto es, la habilidad de imaginar ser el otro en condiciones contrafácticas para atribuirle los estados mentales que uno cree que tendría si estuviera pasando por lo que uno cree que está pasando el otro. A su juicio, cuando el intérprete simula ser el otro, no sólo está ampliando su perspectiva para imaginar la del otro sino que está generando un espacio compartido con el otro, que puede ser visto como la creación de una forma de vida o como una fusión de horizontes, según la metáfora que uno quiera emplear. Sin embargo, en su opinión, para que la comprensión sea posible no es necesario sostener que los interlocutores comparten una forma de vida; lo que es necesario es que los intérpretes puedan simular ser el otro bajo condiciones contrafácticas y asumiendo una realidad compartida; de esta manera comenzarán, dice, a crear un espacio compartido.

separar en una preferencia verbal el contenido cognitivo del emocional. Separar lo cognitivo de lo expresivo, sostiene, no es posible ni necesario. Quintanilla concluye que al interpretar a otro uno *no es un observador no comprometido* sino un *participante* que asume compartir con sus interlocutores significados, estados mentales y el mundo<sup>226</sup>. Al adscribir tales atribuciones al agente participamos de su perspectiva y compartimos algo de su espacio personal. Si el hablante hace lo mismo con uno, estamos creando ese espacio compartido que asumimos tener.

En mi opinión, a la luz de esta reinterpretación del externismo mediante la inclusión de la perspectiva de *segunda persona*, se exhibe pues que lo normativo y lo descriptivo están entrelazados y que aquello que nos hace especiales, la normatividad de lo mental, consiste en *nuestra* inescapabilidad de la triangulación, en la que resulta esencial no sólo nuestro vínculo con el mundo sino nuestra interacción con los otros. En este sentido, la adopción de la perspectiva de *segunda persona* de la atribución mental permite clarificar en qué sentido nuestra peculiaridad reside en que “somos una *comunidad* de mentes” (Davidson, 1991/2001:218; mis itálicas).

Ahora bien, hemos señalado que, a nuestro juicio, al proponer una teoría unificada del pensamiento, el significado y la acción, Davidson entrelaza la distinción *mente-cuerpo* con la distinción *persona-cosa*, esto es, está tratando de desdibujar la distinción entre la noción de *mente* (vinculada con una noción de racionalidad teórica entendida como *posesión de creencias y deseos*) y la noción de *persona* (vinculada con una noción de racionalidad práctica que involucra la idea de ser *uno de nosotros*, esto es, ser un miembro de nuestra comunidad, en la que estamos ligados por nuestras responsabilidades recíprocas).

En la elucidación de la *normatividad de lo mental* que hemos propuesto, su peculiaridad reside en el *valor* de tratarnos mutuamente como *personas*. El rodeo por la propuesta de Dennett nos ha permitido advertir de qué modo la *perspectiva de la segunda persona* de la atribución mental posibilita empezar a explorar una historia de continuidades entre nuestras interacciones primigenias y la adquisición de los conceptos psicológicos comunes a todos los humanos adultos normales, indispensables para la comprensión de las acciones y mentes ajenas (lo que para Dennett se pone en juego en *tener una mente*). En este sentido, a partir de despojarnos de la

---

<sup>226</sup> El matiz que yo introduciría en esta afirmación es que el intérprete *no es un mero* observador no comprometido, sino *también* un participante. Mi interés no radica en reducir la perspectiva de tercera persona a la de segunda, sino en conservar ambas dentro de la empresa interpretativa dotándolas de un carácter igualmente genuino. Asimismo, insisto en subrayar que aquí no pretendo, en absoluto, proporcionar una teoría del contenido, sino simplemente habilitar un espacio conceptual para empezar a pensar el tránsito desde lo no conceptual a lo conceptual.

dicotomía descriptivo/normativo, ahí donde Davidson detiene *a priori* la discusión acerca del *problema de la emergencia del pensamiento y el lenguaje* intentamos proseguirla y revisar su posición en torno al *problema de las perspectivas de la atribución mental*. Dado el carácter unificado de la propuesta davidsoniana, veamos ahora si la perspectiva de *segunda persona* también nos permite empezar a avizorar una historia de continuidad entre nuestras reacciones emocionales primigenias y nuestra constitución en *personas morales*. Ello nos remitirá, finalmente, al otro problema del que nos hemos ocupado aquí: el problema mente-cuerpo y la peculiar respuesta davidsoniana: el *monismo anómalo*.

En el capítulo 4, hemos señalado que en “Mental Events” Davidson afirma que la libertad implica la anomalía de lo mental (1970/1980:207). En este sentido, advertimos que el monismo anómalo puede ser interpretado no sólo como un principio de la filosofía de la mente, sino como un principio ético-político. En efecto, cabe recordar que allí, fiel a su estilo de redacción, Davidson comienza por presentarnos un problema que opera como “acicate” para lo que sigue. El problema en cuestión, dijimos, posee una inequívoca resonancia kantiana (tal como lo evidencian las citas de la *Fundamentación de la Metafísica de las Costumbres* con las que da inicio y cierre al artículo) y consiste en la cuestión siguiente. Dado que los eventos mentales se resisten a ser capturados en la red nomológica de la teoría física ¿cómo puede reconciliarse este hecho con el papel causal que poseen en el mundo físico? Así, advertimos que Davidson parte del supuesto de que el anomalismo y la dependencia causal de los eventos mentales son hechos innegables y su desafío consiste en explicar, frente a las aparentes dificultades, cómo es ésto posible. Reconciliar la libertad con el determinismo causal, nos dice, ha de ser un caso especial de este problema.

Como dijimos más arriba, para ponderar la manera en que la perspectiva de segunda persona permite que empecemos a clarificar el peculiar valor de tratarnos como personas -en el sentido de *tener una mente*- nuestro interlocutor fue D. Dennett. Con la convicción persistente de que la reflexión filosófica supone necesariamente el diálogo, propongo ahora que nuestro interlocutor sea P. F. Strawson. Mi invitación es a que dialoguemos con él para clarificar la manera en que dicha perspectiva permite que comencemos a esclarecer el singular valor de tratarnos como *personas morales*.

### 8.3. La relevancia moral de la perspectiva de segunda persona.

Lepore y Ludwig (2005) abren su introducción con una notable reflexión de P.F. Strawson acerca de Davidson. Allí dan cuenta de que aquél, tomando prestada la siguiente frase de Virginia Woolf, describió a Davidson como un hombre “de una inteligencia pura sangre, que conduce su mente al galope a través del campo en busca de una idea” (Strawson 1984). Para Lepore y Ludwig, esta imagen captura mejor que cualquier descripción prosaica la excitación propia de la obra de Davidson, que como señalamos en la primera parte de nuestro trabajo, ha atravesado el “territorio filosófico” casi por completo, dejando una profunda influencia en el desarrollo ulterior de la filosofía analítica.

Bajo la estela sugerida por Lepore y Ludwig, voy a ensayar un posible diálogo de Davidson con Strawson. Lo que motiva mi elección es que, partiendo del mismo problema, pero asumiendo una concepción mucho más abarcadora de lo mental que, según veremos, atiende a fenómenos propios de la perspectiva de *segunda persona* de la atribución mental, Strawson ofrece una respuesta que permite eludir el altísimo costo de la contrapartida de Davidson, esto es, se trata de una respuesta que permite sortear el dualismo conceptual profundo entre lo mental y lo no mental (y sus correspondientes consecuencias). Para ello, tendremos que adentrarnos en “territorios” poco transitados de la filosofía de Strawson.

En efecto, *Libertad y resentimiento, y otros ensayos*, forma parte del grupo de obras de Strawson menos conocidas.<sup>227</sup> En “Libertad y resentimiento”, el problema que funciona como

---

<sup>227</sup> Strawson, P. (1974). Como señala Juan José Acero (1995), a este mismo grupo posiblemente pertenecen también *Subject and Predicate in Logic and Grammar*, que retoma algunos de los temas centrales de *Individuals* y de sus *Ensayos lógico-lingüísticos* (y que extiende los logros de Strawson sobre la lógica de nuestro sistema conceptual) y *Skepticism and Naturalism: Some Varieties*, un libro tan breve como esclarecedor sobre algunos problemas centrales de la filosofía del lenguaje y de la mente, la epistemología, la metafísica, la ética. El Strawson más conocido es el que pone orden en las grandes y pequeñas cuestiones de nuestro lenguaje cotidiano y del sistema conceptual que subyace a él, en los problemas del significado y de la verdad de nuestras representaciones del mundo y de los lazos que atan ese lenguaje y ese sistema a las cosas, a sus propiedades y relaciones. Para quienes tengan esa imagen de Strawson, y echen de menos en la obra de uno de los filósofos más notables de la segunda mitad del siglo XX, una reflexión que vaya más allá de los límites de la lógica y la metafísica, y se adentre en otras esferas de la filosofía como la ética, la filosofía moral y la estética, *Libertad y resentimiento, y otros ensayos*, significa un aporte de gran interés. Se trata de incursiones en los terrenos de la filosofía menos transitados por Strawson; y todavía, más importante, ese interés se ve acrecentado por el hecho de que dichos ensayos contienen elementos que permiten completar sustancialmente nuestra imagen de Strawson como filósofo analítico. De “Libertad y resentimiento” y de “Moralidad social e ideal individual” se desprende que el análisis del lenguaje es la mejor ruta para adentrarse no sólo en la composición y estructura de nuestro equipamiento conceptual, sino en la composición y estructura de lo

“acicate” para lo que viene luego, en el fondo es el mismo que, según vimos en la segunda parte de este trabajo, animó a Davidson en “Mental Events”, a saber: cómo conciliar la libertad con la necesidad natural. En el abordaje de Strawson, la cuestión se da bajo el siguiente “ropaje”. Por un lado, nos presenta la posición del “pesimista” a propósito de las implicaciones de la tesis del determinismo. Si el determinismo es verdadero, los conceptos de *obligación y responsabilidad moral* realmente no tienen aplicación alguna y las *prácticas de condena y aprobación* morales son injustificadas. Por su parte, el “optimista” sostiene que estos conceptos tienen su razón de ser, aún si la tesis del determinismo es verdadera. A Strawson lo mueve un deseo de reconciliación entre ambas posiciones, con lo cual, dice, es probable que su planteo le parezca desacertado a ambos contendientes. La estrategia que va a asumir es la siguiente. Va a trasladar la discusión a un campo vecino al asunto entre pesimistas y optimistas, con la esperanza de que, puesto que este ámbito está menos poblado de polemistas, “la cuestión podría ser más fácilmente resuelta aquí. Y si se soluciona en él, podría resolverse con más facilidad en ese otro terreno atestado de controversistas” (Strawson 1974/1995:41). Al respecto, insiste en que lo que va a decir consiste, en buena parte, en lugares comunes. La índole de su propuesta será tal que, nadie que la haya comprendido, afirma, podría insistir en una demanda ulterior de justificación de nuestras prácticas morales. Según veremos, el objetivo de estos lugares comunes es tener presente qué es estar efectivamente *involucrado en relaciones interpersonales ordinarias*.

El ámbito en el que inscribe la discusión es, entonces, el de “las actitudes y reacciones no indiferentes de la gente directamente involucrada en los intercambios entre sí” (Strawson 1974/1995:41). Esto es, el ámbito de las actitudes y reacciones de quienes se sienten ofendidos y de quienes se sienten beneficiados; de cosas tales como “la gratitud, el resentimiento, el perdón, el amor y los sentimientos heridos” (1974/1995:41). El lugar común central que va a subrayar es, pues, el de la grandísima importancia que concedemos a las actitudes e intenciones de otros seres humanos hacia nosotros. Lo que Strawson quiere contrastar es la gama de actitudes de *involucramiento* en una relación humana, frente a la actitud *no participativa, objetiva*, hacia otro ser humano. El corazón de su planteo consistirá en que, siendo humanos, no podemos abandonar

---

que, genéricamente hablando, podríamos denominar nuestro equipamiento mental: un marco de actitudes y sentimientos con que reaccionamos ante los actos de los demás, por ser nosotros parte implicada en los efectos que esos actos tengan. Si, como se suele apuntar, en las reflexiones de Strawson acerca de nuestro sistema conceptual es perceptible la figura de Kant, en aquellas en las que se ve implicado el resto de nuestro equipamiento mental ha dejado una impronta inequívoca el pensamiento de Hume.

de modo absoluto y permanente nuestra actitud *participativa*. Nuestro compromiso con actitudes interpersonales ordinarias es parte de la estructura general de la vida humana y, en este sentido, no es una opción, no es algo susceptible de ser revisado. Ahora bien, para apreciar la relevancia de su propuesta, veamos con más detalle en qué consiste la actitud *participativa*.

En primer lugar, Strawson considera las “actitudes y sentimientos reactivos *personales*” (Strawson 1974/1995:42). Se trata de nuestras reacciones ante la calidad de la voluntad de otros hacia *nosotros*. Los ejemplos que ofrece son: la gratitud, el resentimiento, el perdón, el amor y el sentirse herido. Este primer grupo descansa sobre una expectativa y una exigencia de manifestación de cierto grado de buena voluntad o consideración de otros seres humanos hacia nosotros mismos (o, por lo menos, en la expectativa y en la exigencia de la ausencia de manifestación de una mala voluntad activa o de un desinterés indiferente). Por su parte, estas actitudes tienen sus correlatos del lado del agente, esto es, así como hay actitudes asociadas con exigencias de otros hacia uno mismo, también hay actitudes *autorreactivas* asociadas con exigencias de uno mismo hacia otros. Aquí menciona fenómenos tales como sentirse avergonzado, culpable o arrepentido.

En segundo lugar, pasa a discutir los análogos *impersonales o vicarios*, de las reacciones personales. Se trata de reacciones *desinteresadas o generalizadas*, por lo que reciben la calificación de “*morales*” (1974/1995:53). Son nuestras reacciones ante la calidad de la voluntad de otros, ya no hacia nosotros, sino *hacia un tercero*. El ejemplo paradigmático que proporciona es el de la indignación moral, como un caso en el que sentimos resentimiento en nombre de otro.

Estos tres tipos de actitud están conectados no “sólo lógicamente”, sino “humanamente” (1974/1995:55). Tienen raíces en nuestra naturaleza humana y en nuestra pertenencia a comunidades humanas.

Siguiendo el análisis del aporte de Strawson que propone Acero (1995)<sup>228</sup>, cabe hacer, al menos, dos observaciones. La primera es que el sistema de actitudes y sentimientos reactivos no es algo simplemente *contiguo* al sistema conceptual. El análisis del lenguaje hace posible revelar los ingredientes de que consta nuestro sistema conceptual y las posiciones en la estructura conceptual que estos ingredientes ocupan. Pero los conceptos y las palabras no sólo guardan relaciones con conceptos y palabras, respectivamente. Unos y otras no son simplemente

---

<sup>228</sup> En Strawson (1974/1995), Introducción.

herramientas con las que clasificamos objetos y propiedades, acciones y eventos, y con las que establecemos y describimos los vínculos que se dan entre ellos y los patrones que emergen de esas relaciones. Los conceptos frecuentemente están teñidos de los sentimientos, de las emociones, que esos objetos y acciones ocasionan en nosotros. Por lo tanto, la estructura *conceptual* y la estructura de *sentimientos* y de actitudes se hallan profundamente *entrelazadas*. La actitud objetiva hacia el mundo natural y social se mezcla íntimamente con el sistema de actitudes emotivas que adoptamos ante los efectos que tienen las cosas y las acciones de los demás en nosotros. El vínculo no es, por tanto, ocasional.

Para la comprensión de la realidad moral ha de tenerse en cuenta que la suspensión de las actitudes reactivas, del resentimiento o la gratitud, sólo acontece en circunstancias que son más bien la excepción que la norma, esto es: cuando el daño que se nos causó no pudo evitarse o cuando, por ejemplo, la persona que lo infligió no puede ser considerada moralmente responsable. Ahora bien, lo que no debe olvidarse es que, al adoptar una *actitud no participativa* hacia una persona, se convierte para nosotros en algo que exige seguir una táctica adecuada, que ha de ser objeto de un cierto tipo de tratamiento o que da pie a nuestra curiosidad intelectual. Y todo ello lo percibimos como una forma de anomalía. Una percepción así tiene su origen en que, como seres humanos que somos, no podemos renunciar a la *actitud participativa* por mucho tiempo. Nuestra implicación en el sistema de actitudes reactivas y nuestro compromiso con él no es, por consiguiente, algo a lo que nos veamos ocasionalmente sujetos. Es una implicación y un compromiso naturales. Una actitud hacia el mundo que fuese sostenidamente objetiva sería, por el contrario, una actitud profundamente no natural:

El compromiso humano de participación en las relaciones interpersonales ordinarias resulta, creo, demasiado abarcador y sus raíces son demasiado profundas como para que nos tomemos en serio el pensamiento de que una convicción teórica general pueda cambiar tanto nuestro mundo que no haya en él más relaciones interpersonales conforme las entendemos corrientemente. (Strawson 1974/1995:50)

La segunda propiedad que destaca Acero, hace extensiva al sistema de actitudes emotivas y participativas el carácter internalista de la justificación del sistema conceptual. No podemos elegir, más que dentro de ciertos márgenes estrechos, ni qué sistema conceptual vamos a emplear

ni cómo vamos a modificarlo. Y tampoco podemos elegir, más que dentro de márgenes muy reducidos también, el sistema de sentimientos y actitudes reactivas con que afrontamos nuestra realidad social y moral. Puesto que no está en nuestras manos, salvo en una medida poco significativa, elegir cambiar de marco, toda estimación, modificación o crítica será *interna* al marco mismo en que nos hallamos situados.

Se sigue de esto que no está en nuestras manos elegir cambiar de marco de actitudes y sentimientos -una “elección divina”-, y se sigue también, en virtud del estrecho nexo entre la estructura emotiva y la estructura conceptual, que no podemos imaginar cómo sería un marco diferente.

La existencia de esta red de sentimientos y actitudes que se entrelaza con nuestro sistema conceptual, el hecho mismo de nuestro compromiso natural con dicho equipamiento conceptual y actitudinal y el que la evaluación moral sea siempre algo interno a él, son tres importantes novedades de la concepción del análisis filosófico strawsoniano.

En “Libertad y resentimiento” Strawson cifra, entonces, en esa red de sentimientos la clave de la solución que ofrece al problema de si la verdad de la doctrina del determinismo implica que los conceptos de condena y aprobación moral carecen de contenido y que las prácticas de la condena y la aprobación moral no tienen razón de ser. Strawson reconoce no saber muy bien qué afirma esa doctrina, pero rechaza la relación de implicación. Simplemente, no es verdad que si la doctrina fuese cierta -es decir, si estuviésemos sin saberlo sometidos al reinado universal de la causalidad natural-, de ello se seguiría que la única actitud que cabría legítimamente adoptar hacia los demás sería una “actitud de objetividad omniabarcadora”. Pues esa imagen de nosotros mismos choca frontalmente con algo que no está en nuestras manos dejar a un lado: el compromiso que tenemos con nuestras actitudes y sentimientos morales. Por lo tanto, el filósofo pesimista, el que acepta la implicación en cuestión, se equivoca al suponer que podemos adoptar de forma permanente una actitud puramente objetiva hacia las acciones propias y ajenas y que el fundamento de la condena y aprobación moral habría de buscarse, por consiguiente, en alguna suerte de principio metafísico, quizás en alguna propiedad de nuestros actos. Y también se equivoca el filósofo optimista, el que rechaza la anterior implicación por pensar que el fundamento moral de prácticas como la del castigo y la aprobación descansa en sus consecuencias sociales. Pues también éste malinterpreta, a juicio de Strawson, el citado marco de sentimientos y actitudes.

He sostenido que “Mental Events” y “Libertad y Resentimiento” están animados por un problema afín, esto es: cómo conciliar la libertad con la necesidad natural. No obstante, a la luz de la exposición que acabamos de considerar, creo que su principio de solución descansa en concepciones de lo mental sumamente diferentes. Davidson se ciñe al entramado estrictamente conceptual de las actitudes proposicionales y queda atado a un dualismo conceptual profundo que es heredero del hiato entre la normatividad peculiar de lo mental y el carácter meramente descriptivo de los conceptos no mentales. Según vimos, la consecuencia que dicho hiato trae aparejada es que se detiene *a priori* la discusión acerca del pasaje de lo no conceptual a lo conceptual, acerca de la emergencia del pensamiento y el lenguaje. Strawson, en cambio, elabora su respuesta en términos de un entramado mental en que lo conceptual y lo emocional están profundamente entretnejidos. En el primer caso, según vimos, la inviabilidad de la reducción está dada por el carácter no opcional del *principio normativo de la caridad*. En el segundo, por el carácter no opcional de la *actitud participativa*, que se pone en juego en el marco de las actitudes y sentimientos que se exhiben en nuestras relaciones interpersonales cotidianas.<sup>229</sup>

Ahora bien, ¿de qué modo este rodeo conceptual en torno al aporte de Strawson puede sernos de ayuda para articular una genuina perspectiva de *segunda persona* que, a su vez, permita arrojar luz sobre la *normatividad* propia de lo mental, en los términos en que he propuesto elucidarla? A mi juicio, en virtud de la *relevancia moral* que posee la perspectiva de segunda persona de la atribución mental. Veamos, pues, el tipo de conexión que hay entre las nociones de *segunda persona*, *actitudes reactivas*, y la *relevancia moral* que se pone en juego en la dimensión evaluativa de la normatividad de lo mental, esto es, en el valor de tratarnos mutuamente con *personas*. Esto es, exploremos de qué modo la perspectiva de *segunda persona*

---

<sup>229</sup> En este sentido, cabe advertir que pese al talante anticartesiano de sendos tratamientos de lo mental, los costos de sus respectivas soluciones son radicalmente diferentes. En efecto, por un lado, creo que hay una explícita afinidad entre ambas posiciones a la luz de la manera en que he propuesto interpretar el aporte de Davidson, esto es, siguiendo a Ramberg, a partir de la íntima conexión entre tener una *mente* (ser interpretable) y ser una *persona*. Así, en “Yo, mente y cuerpo”, Strawson sostiene que la diferencia entre el cartesiano y el anticartesiano es una diferencia de perspectiva sobre la relación entre el concepto de persona y el concepto de mente. El anticartesiano propugna que el concepto de mente tiene un estatus secundario o dependiente respecto del concepto de persona, que se revela como el concepto fundamental. En contraste, el filósofo cartesiano, al que Strawson se opone, no puede admitir esa dependencia. Sin embargo, allí donde la solución davidsoniana abre un hiato profundo entre lo mental y lo no mental, la salida strawsoniana revela una concepción de lo mental que hace lugar a dimensiones de lo mental que permiten trazar continuidades con el mundo natural.

permite esclarecer la dimensión *evaluativa* de la normatividad que inserta nuestra práctica de tratarnos mutuamente como *personas* en el seno de la reflexión ética.

Si bien Strawson no emplea la expresión “perspectiva de segunda persona”, la manera en que elabora su solución a la polémica entre optimistas y pesimistas, constituye un aporte fundamental en la formulación de este punto de vista y lleva a reconocer su relevancia moral.

Lo interesante en nuestro contexto es que Strawson ilustra la comprensión implícita de la moralidad a través de las *actitudes reactivas*. Según vimos, dichas actitudes son contempladas por él como esenciales a las prácticas humanas de considerar a las personas como moralmente responsables de sus actos. No obstante, cabe señalar que en tales actitudes no se trata de atribuir una emoción al otro con quien se interactúa, sino en *reaccionar emocionalmente* a su conducta. Strawson no tematiza este aspecto previo de la interacción, pero parece obvio que este nivel de interacción cara a cara, expresivo y espontáneo entre los agentes (esto es, la perspectiva de *segunda persona* en la caracterización que ofrecimos) es una *precondición* para que la actitud reactiva *personal* tenga sentido, en la medida en que la reacción depende de reconocer una emoción en el otro.

A su vez, hemos visto que a partir de generalizar las actitudes reactivas sobre la expectativa de buena voluntad por parte de otros, no simplemente hacia uno mismo, sino hacia todos los hombres, tienen lugar las reacciones *impersonales* que son, a juicio de Strawson, las propiamente “morales”. De este modo, la perspectiva de *segunda persona* se revela como un *presupuesto de la moralidad* y aparece como un recurso promisorio para contar una historia conceptual acerca de cómo llegamos a constituirnos y reconocernos como *personas*, como fuente de obligaciones morales. En este mismo espíritu, Gomila (2008) sostiene que las “actitudes reactivas” implicadas en las emociones morales presuponen constitutivamente la perspectiva de segunda persona y, en esa medida, constituyen la base última de la obligación moral. Dicho en otros términos, la base de la moralidad radica en nuestra capacidad de interactuar con los demás desde la perspectiva de segunda persona. A su juicio, las prácticas inmorales pero, por desgracia, persistentes en la historia moral siglo XX -como la tortura, la desaparición o el genocidio- requieren suspender, bloquear o inhibir de alguna forma, esa relación humana básica.

Por su parte, Stephen Darwall (2006) ha explorado la relevancia moral de la perspectiva de segunda persona, aunque no entendida desde la vertiente psicológica transitada por Gomila. En efecto, Darwall no está interesado en el modo en que se produce la interacción, sino en la

implicación que conlleva la emoción reactiva de lo que él denomina una “razón de segunda persona”, de una reivindicación moral que presupone la relevancia moral del otro sujeto de la interacción. Es decir, la segunda persona es la perspectiva que lleva a reconocer al otro como sujeto y, por tanto, como fuente de obligaciones morales. Según Darwall, pues las actitudes reactivas son siempre implícitamente de segunda persona, es decir, presuponen la interacción intersubjetiva entre los agentes implicados, o más propiamente, su carácter de agentes responsables, que deben rendir cuenta de sus actos. Según dijimos, las actitudes reactivas *personales* se experimentan como derivadas de la interacción con otro agente, mientras que las *impersonales* se sienten en tanto que miembros de la comunidad moral. Lo importante es que estas reacciones tienen, para Darwall, condiciones de corrección, lo que significa que quienes las experimentan cuentan con las razones del tipo correcto –razones morales– que las justifican, y que son razones de segunda persona, razones relativas a la interacción entre agentes. Así, la conducta de alguien puede generar indignación en otros, en la medida en que para éstos hay razones para reprocharle su conducta, por la que debe rendir cuentas –bien, desde el plano personal, bien desde el plano impersonal, al margen de la implicación directa en la situación de quienes la sienten-. Tanto en un caso como en otro, no obstante, Darwall habla de razones de segunda persona: razones que regulan las relaciones entre sujetos, de respeto, de obligaciones morales recíprocas.

No ha sido mi objetivo aquí entrar en profundidad en la discusión de la interesante propuesta de Darwall, en particular sobre si ésta es ciertamente la única o la principal base de obligación moral, o si puede reinterpretarse en estos términos el imperativo categórico kantiano, sino tan sólo subrayar la *relevancia moral de la perspectiva de segunda persona*.

#### 8.4. La normatividad de lo mental y el rol de la segunda persona: conclusiones.

En este trabajo mi propuesta ha sido ofrecer un *correctivo* y un *complemento* al fascinante entramado conceptual de Davidson, de modo que podamos repensar su controvertido tratamiento de los problemas de los que nos hemos ocupado: esto es, el problema mente-cuerpo, el problema de la relación entre el pensamiento y el lenguaje y el problema de las perspectivas de primera y tercera persona de la atribución mental. El correctivo ha consistido en despojarnos de la dicotomía descriptivo/normativo que subyace a su filosofía de la mente (esto es, lo que presenté como el “cuarto dogma del empirismo”); el complemento reside en la adopción de la perspectiva de segunda persona de la atribución mental.

La motivación que estuvo por detrás del correctivo ha sido la de tornar más “davidsoniano” a Davidson. Esto es, volverlo más solidario con su propia inscripción en la tradición crítica de las dicotomías, en vistas de su adhesión a las críticas quineanas al reduccionismo neopositivista y a la dicotomía analítico-sintético y de su propia denuncia del “tercer dogma” del empirismo (la dicotomía esquema-contenido). Al despojar a su proyecto de lo que podría pensarse como una suerte de “cuarto dogma” del empirismo, se nos abrió un nuevo horizonte conceptual en el que puse en juego la perspectiva de *segunda persona* de la atribución mental.

He señalado que prácticamente ninguno de los fenómenos que son recogidos desde la perspectiva de segunda persona de la atribución mental ha sido objeto de atención por parte de los filósofos de la mente. Especialmente, para el caso que nos ocupa, advertí que el paradigma de la explicación de la acción racional que signó el esfuerzo seminal de Davidson (1963), encorsetó su concepción de lo mental de manera que lo llevó a excluir *a priori* de su dominio el tratamiento de fenómenos del tipo aludido. Esta exclusión, en conjunción con su adhesión a una dicotomía entre una dimensión normativa y una dimensión descriptiva, determinó dramáticamente, en mi opinión, su tratamiento de los tres problemas de los que nos ocupamos aquí. En efecto, en el seno del *problema de las perspectivas de primera y tercera persona de la atribución mental*, esta “ceguera” respecto de ciertos fenómenos mentales que no son asequibles desde la perspectiva distante de la tercera persona, del intérprete como observador desinteresado, lo llevó a no poder articular la perspectiva de “segunda persona”, que sólo llegó a bautizar sin explorar en toda su

hondura. En este sentido, no terminó de precisar su propia pretensión de que el intérprete no sea un mero observador “ocioso” sino un “jugador esencial” en el juego de la mutua comprensión. Como resultas de ello, allí donde logró ofrecer una caracterización precisa de los conceptos de objetividad y de la subjetividad, no logró lo propio acerca de la noción de intersubjetividad. Respecto del *problema mente-cuerpo*, su restringida concepción de lo mental lo llevó a abrazar un dualismo conceptual profundo entre lo mental y lo no mental que, según vimos, abre una brecha imposible de zanjar entre el mundo de la libertad y el mundo de la necesidad natural. Quedamos, así, inexorablemente escindidos. Como consecuencia de dicho dualismo, la discusión acerca de cómo pasamos de lo no conceptual a lo conceptual, de nuestras interacciones primigenias a la adquisición de conceptos psicológicos del tipo que posee el humano adulto normal, queda detenida por principio. De esta suerte, el *problema acerca de la relación entre el pensamiento y el lenguaje y su emergencia* se tornó en un problema “insuperable”, en el sentido de que por razones conceptuales nunca podremos dar una descripción plena de su aparición.

Frente a este estado de la cuestión, he argumentado que si se prescinde de la dicotomía descriptivo/normativo y se debilita la concepción “teoricista” estándar de la interpretación radical al subrayar las habilidades que pone en juego, es posible complementarla con una perspectiva de la atribución mental signada por su dimensión práctica y reactiva, esto es, una perspectiva que involucra un conocimiento de la misma naturaleza práctica que aquel que guía nuestra acción en caso de “expertise” técnica o práctica, esto es, que conlleva un saber-cómo. Así, al despojarnos del “cuarto dogma del empirismo” hemos podido dar cabida a la consideración de fenómenos básicos de la atribución intencional que quedan fuera del alcance y del interés de las perspectivas de primera y tercera persona de la atribución mental. Esto es, hemos podido contemplar los fenómenos interactivos propios de la perspectiva de segunda persona, especialmente, los fenómenos emocionales.

En efecto, según vimos en la primera sección de este capítulo, se trata de fenómenos en los que se “entretajan” aspectos descriptivo-causales y aspectos evaluativos. Esto es, en ellos se ponen en juego el reconocimiento automático e inconsciente de patrones y, a la vez, propósitos evaluativos. Según la perspectiva de segunda persona de la atribución mental, los propósitos evaluativos que tiene la atribución mental resultan de las necesidades y propósitos de la interacción misma: saber cómo actuar, corregir las propias expectativas, coordinar esfuerzos para lograr un objetivo común, persuadir a los otros de las propias intenciones, para ser aceptado o

perdonado, obtener beneficios, etc. “Evaluar”, hemos dicho, aquí significa saber cómo actuar de un modo que sea más conveniente a los propios fines en función de las reacciones del otro. En relación con estos propósitos evaluativos, hemos señalado que las emociones lo son tanto para el sujeto que las tiene como para quien es afectado por ellas de algún modo, pudiendo resultar apropiadas, desproporcionadas, inoportunas, etc. En este sentido, hemos enfatizado que nuestras reacciones emocionales son reacciones evaluativas, es decir, reaccionamos con emoción de modos que nos parecen apropiados a las circunstancias en las que nos encontramos. Se trata, más brevemente, de “evaluaciones espontáneas”. Nuestras interacciones emocionales son, así, interacciones evaluativas.

Ahora bien, en el seno de mi propuesta de elucidación de la *normatividad de lo mental*, aquel rasgo que a juicio de Davidson porta la clave de lo humano, he argumentado que para comprenderla en toda su profundidad no basta que la capturemos su dimensión *deóntica* sino que tenemos que internarnos en su dimensión *evaluativa*. De este modo, defendí que la *normatividad de lo mental* debe concebirse en términos del principio de racionalidad como *regla constitutiva* de la práctica por la que nos tratamos mutuamente como *personas*, tal es el valor que se pone en juego. Asimismo, enfatiqué que con su teoría unificada del pensamiento, el significado y la acción, Davidson intenta difuminar la distinción entre *tener una mente* (vinculada con una concepción de la racionalidad teórica) y ser *persona* (vinculada con una concepción de la racionalidad práctica). En este sentido, hemos visto que la perspectiva de la *segunda persona* tiene un *doble rol*: esto es, a mi juicio, permite esclarecer tanto cómo es que llegamos a adquirir conceptos psicológicos comunes a todos los humanos adultos normales, indispensables para la comprensión de las acciones y mentes ajenas, así como permite arrojar luz acerca de cómo llegamos a constituirnos en personas en un sentido moral. Para precisarlo, establecí un diálogo con las propuestas de Dennett y Strawson, respectivamente. El interés que, en este contexto, revistieron para mí sus posiciones fue el de exhibir puntos de partida consonantes con los de Davidson pero exentos de sus dificultades, en la medida en que asumen una concepción más abarcadora de lo mental. En efecto, estos autores permiten dar cabida a fenómenos emocionales que Davidson excluye *a priori* del dominio de lo mental. De este modo, ahí donde él detiene por principio la discusión, me ha resultado posible proseguirla a partir de considerar la perspectiva de *segunda persona* de la atribución mental para ofrecer una articulación de aquel rasgo clave que asume haber dejado “crudo”, “vago” o “demasiado rígido”: la *normatividad de lo mental*.

El diálogo con Dennett, nos permitió retomar el *problema de la relación entre el pensamiento y el lenguaje*, así como el de su *emergencia*, y frente al paso de “prestidigitación” que se sigue del planteo davidsoniano, habilitar un espacio conceptual para trazar una historia de continuidades. En efecto, la consideración de fenómenos emocionales en los que están imbricados aspectos descriptivo/causales y aspectos normativos torna plausible contemplar estadios intermedios entre la situación de la llamada “intersubjetividad primaria” donde los bebés no exhiben conductas que nos autoricen a atribuirles concepto psicológico alguno y la situación en que llegan a poseer nuestro concepto humano adulto normal de *creencia*. Asimismo, señalamos que ninguna interacción humana real es completamente describable en términos meramente físicos, de la manera que, según vimos, pretende Davidson describir el triángulo inicial en el proceso de aprendizaje del lenguaje. En este sentido, sostuvimos que no habría un hiato entre una instancia meramente descriptivo/causal de reconocimiento automático de patrones desprovista de toda normatividad y una instancia normativa de articulación conceptual plena. Si la interacción en cuestión está “atravesada” desde el principio por los conceptos mentales y su concomitante normatividad no hay por qué concebir la emergencia del lenguaje y el pensamiento como un salto “mágico”, sino que se habilita la posibilidad de encontrar continuidades.

A su vez, el diálogo con Dennett nos permitió reanudar el tratamiento del *problema de las perspectivas de primera y tercera persona de la atribución mental*. Según nuestro diagnóstico, Davidson ofrece articulaciones precisas de las nociones de objetividad y subjetividad, esto es, de las perspectivas de tercera y primera persona, pero no logra lo propio respecto de la noción de intersubjetividad, en la medida en que sólo “bautiza” la “segunda persona” sin caracterizarla como una perspectiva genuina y específica. Frente a ello, la consideración de fenómenos propios de la perspectiva de segunda persona de la atribución mental tal como se la concibe desde la filosofía de la mente contemporánea, nos permitió advertir que *lo primario es la interacción misma*, esto es, que los vínculos que mantienen a cada sujeto ligado o involucrado en relaciones *con* otros resultan un rasgo constitutivo de su hacerse *personas*. De este modo, logra dotarse de “carnadura” a la idea de la intersubjetividad como una dimensión distinta respecto de la subjetividad y la objetividad. En otras palabras: la dimensión social propia de la perspectiva de segunda persona pone de manifiesto el carácter básico de las relaciones *intersubjetivas* en la comprensión recíproca. En este sentido, al revés de las imágenes

clásicas de la mente, la subjetividad individual es explicada como fruto de esta habilidad intersubjetiva. De esta suerte, no he pretendido defender que la perspectiva de segunda persona sea la única existente. Por el contrario, simplemente he querido destacar que las perspectivas de primera y tercera persona de la atribución mental no son ni básicas, ni excluyentes, sino que resultan complementarias entre sí y con la perspectiva de segunda persona. Reconocer lo genuino de la perspectiva de segunda persona permite, a su vez, advertir lo específico de las perspectivas de primera y tercera persona, en lugar de verlas como alternativas en conflicto.

Finalmente, el diálogo con Strawson nos permitió retomar el *problema mente-cuerpo* desde la interpretación que propusimos del anomalismo de lo mental. Esto es, no como un mero principio de la filosofía de la mente sino como un principio ético-político. De este modo, nos vimos conducidos al seno de la reflexión ética y vimos que, pese a que el problema que opera como “acicate” es el mismo tanto en Strawson como en Davidson, el primero logra eludir el compromiso con un dualismo conceptual profundo que detiene *a priori* la discusión. Alegué que la clave de la salida strawsoniana reside en su concepción más abarcadora de lo mental. Si atendemos al modo en que en su sistema se entrelazan lo conceptual y lo emocional, es posible arrojar luz sobre un punto que en Davidson resulta, a mi juicio, sumamente vago: aquel en que los valores están “enraizados en las cosas”. En especial, el modo en que el valor de tratarnos como personas es, a la vez, *factum* y *norma*, precondition y resultado del proceso de interpretación.

En efecto, a mi juicio, aquello que se pone en juego en el modo de vida compartido que para Davidson es precondition de la comprensión, son fines *evaluativos* ligados esencialmente a las *emociones* por las que nos vamos constituyendo recíprocamente en *personas*. Antes de compartir una visión del mundo, las ideas no existen. Recordemos que el pensamiento depende de una configuración triangular de relaciones que involucran al menos dos individuos y una serie de eventos compartidos. Lo que es dado al individuo, en última instancia, es la triangulación comunicativa. En este sentido, la intersubjetividad está a la base del conocimiento. Podría sostenerse entonces que “¿el consenso está a la base de la objetividad?”. Esta es una de las preguntas con las que Giovanna Borradori (1994) interpela a Davidson. Frente a ella Davidson responde:

El consenso (*consensus*) es una manera equivocada de formularlo, porque hablar de consenso lo hace sonar como si cada uno de nosotros tuviera sus propias ideas y llegáramos a un acuerdo (*agreement*), mientras que yo

estoy diciendo que no tenemos ninguna idea hasta que compartimos una visión del mundo (*we share a picture of the world*)<sup>230</sup>.

En este sentido, la perspectiva de *segunda persona* parece operar como una suerte de “puente” entre dicho modo de vida compartido y la comprensión a la vera de la cual nos consideramos cabalmente como “creyentes de verdades y amantes de lo bueno” (1970/1980:222). En otras palabras: el *participante* se revela como una condición constitutiva del *intérprete*. En efecto, a mi juicio, si no se atiende a la dimensión participativa que es posibilitada por la perspectiva de segunda persona, no logra articularse de manera precisa el tipo de *interacción* peculiar que subyace a la *interpretación* y que torna inteligibles a sus condiciones de posibilidad: la adopción metodológica de la perspectiva de tercera persona (objetividad) y la presunción a favor de la verdad del hablante (subjetividad). Tal es la manera en que he entendido que la noción de *persona* hace manifiestos los modos en que descripción y evaluación están entrelazados: siendo precondition y a la vez ideal regulativo de la interpretación, la noción de *persona* exhibe la manera en que –desafiando al “cuarto dogma”- *factum* y norma resultan “imbricados”. Creo que esta complementación que he propuesto nos permite advertir pues, en palabras del propio Davidson, la manera en que el valor de tratarnos como *personas* está “enraizado en las cosas” (*value are rooted to things*; 1995c/2004:52).

Recordemos que al decir de Davidson: si comprendemos las palabras del hablante, “existe un terreno común, un modo de vida compartido” (1984c/2004:37). La perspectiva de *segunda persona* permite, a mi juicio, arrojar luz sobre la índole de ese “compartir”, que en la propuesta de Davidson queda inexplorado. En este sentido, con ella se abre la posibilidad de precisar el tipo de “interacción” del grupo (*interacting group*) que a juicio de Davidson está a la base de la comprensión. Tal es la manera en la que he propuesto clarificar su pretensión de que el intérprete no sea un mero espectador ocioso sino un “jugador esencial” (1995/2004:49) de este “juego” de la mutua comprensión, que no podemos elegir dejar de jugar sin dejar, a la vez, de ser quienes somos. En este sentido, frente a los propósitos fundamentalmente epistémicos de

---

<sup>230</sup> Borradori señala que esta reflexión introduce una clara originalidad del planteo de Davidson frente a la tradición pragmatista, desde Charles S. Peirce a Clarence I. Lewis. En efecto, ahí donde ellos entendían el proceso comunicativo en términos de alcanzar el consenso, para Davidson, la noción de consenso resulta secundaria. Hablar de consenso presupone que las ideas existen previo a él, y que cuando se confrontan con otras ideas, dan lugar al acuerdo; mientras que para Davidson, antes de compartir una visión del mundo las ideas no existen. En este sentido, el acuerdo no es un logro de la comprensión, sino un presupuesto de la misma.

las perspectivas clásicas de la atribución mental, la perspectiva de segunda persona permite advertir propósitos evaluativos que apuntan en una dirección diferente: la dimensión práctica que se pone en juego en nuestro tratarnos mutuamente como personas.<sup>231</sup>

Tras este derrotero conceptual, sin duda, cabe la pregunta acerca de cuál es la ventaja de mi propuesta (si es que reviste alguna). Por un lado, creo que frente al carácter inarticulado en que deja Davidson dos nociones que parecen centrales a su planteo, a saber, la *normatividad de lo mental* y la *segunda persona*, he ofrecido una elaboración que ha intentado exhibir las conexiones que, a mi juicio, hay entre ambas. En efecto, he pretendido mostrar que una elucidación de la segunda, permite arrojar luz sobre la primera. Por otro lado, mi propuesta ha intentado restaurar un elemento –a mi juicio- vital de nuestra autoimagen, que en la historia de la filosofía (salvo honrosas excepciones) ha sido dejado de lado: nuestras emociones. Para posibilitar su regreso como parte de la reflexión acerca de qué nos hace ser quienes somos, he sostenido la necesidad de despojarnos de la dicotomía descriptivo/normativo. La ventaja que, en mi opinión comporta este movimiento es la de permitir que continuemos la discusión que, en Davidson, queda bloqueada *a priori*.

En suma, a lo largo de este trabajo he intentado abrir un horizonte conceptual que nos permita ir más allá de la propuesta de Davidson. Tarea que, sin embargo, habría sido imposible sin él.<sup>232</sup>

---

<sup>231</sup> En este punto, creo que es interesante señalar la posibilidad que se abre para dialogar con otras tradiciones filosóficas para las cuales la afirmación del otro ha de proceder, ya no desde una dimensión epistemológico-semántica, sino desde una dimensión eminentemente ética. Al respecto, puede verse por ejemplo el tratamiento que Cragolini (2007:69) ofrece de las obras de Nietzsche y Derrida, en términos de “la sustitución de la ontología de la identidad por una ética de la responsabilidad”.

<sup>232</sup> Pierre Bourdieu diferenciaba, a propósito de su interpretación de la obra de Foucault, entre dos lecturas posibles: una lectura de lector y otra de *auctor*. Según la primera, se trata simplemente de comentar una obra; de acuerdo con la segunda, se trata de acceder a su sentido más hondo, no siempre explícito, de un modo que puede requerir incluso ir más allá de los textos. Es en esta segunda vertiente en la que he pretendido inscribir, humildemente, la tarea asumida aquí.

## BIBLIOGRAFÍA

- ACERO, J.J. (1995) "Introducción" en Strawson, P.F. (1974/1995), *Libertad y resentimiento y otros ensayos*, Barcelona, Paidós, pp.9-34.
- AGASSI, J. (1975) *Science in Flux*, Dordrecht, Reidel.
- ALLEN, C. y BEKOFF, M. (1997) *Species of Mind: The Philosophy and Biology of Cognitive Ethology*, Cambridge, MA, MIT Press.
- AMORETTI, M.C. y VASSALLO, N. (2008) *Knowledge, Language, and Interpretation. On the Philosophy of Donald Davidson*, Ontos Verlag, Frankfurt.
- ANDREWS K (2002) "Interpreting autism: a critique of Davidson on thought and language", *Philosophical Psychology*, vol. 15, no.3.
- ANSCOMBE, G.E.M. (1957) *Intention*, Blackwell, Oxford.
- ANTONY, L. (1994) "The inadequacy of anomalous monism as a realist theory of mind" en Gerhard Preyer, F. Siebelt & A. Ulfig (eds.), *Language, Mind, and Epistemology: On Donald Davidson's Philosophy*. Dordrecht: Kluwer.
- ARMSTRONG, D. (1981) *The Nature of Mind and other essays*, Ithaca, NY, Cornell University Press.
- ASTINGTON, J., HARRIS, P. & OLSON, D. (1988) (eds.) *Developing Theories of Mind*, Cambridge, U.K.: Cambridge University Press,
- ATKINSON, R.L. et. al (1990) *Introduction to Psychology*, San Diego, Harcourt Brace Jovanovich.
- AVRAMIDES, A. (1999) "Davidson and the New Sceptical Problem" en Zeglen U. (ed.) *Donald Davidson: Truth, Meaning and Knowledge*, London, Routledge.
- AYER, A. J. (1963) "Privacy" en *The Concept of a Person and other essays*, New York, Macmillan.
- BAIER, A. (1990) "What Emotions are About" en Tomberlin, J., (ed) *Action Theory and Philosophy of Mind*, California, Ridgeview Publ. Co.
- BAKER, G.P. y HACKER, P.M.S. (1985) *Wittgenstein, Rules, Grammar, and Necessity*, vol.2 of *An Analytical Commentary on the "Philosophical Investigations"*, Oxford, Basil Blackwell.
- BAR-ON, D. y RISJORD, M. (1992) "Is There such a Thing as Language?", *Canadian Journal of Philosophy*, 22: 163-90.
- BARON-COHEN, S. LESLIE, A. & FRITH, U. (1985) "Does the autistic child have a "theory of mind"?", *Cognition*, 21, 37-46.
- BICKERTON, D. (1990) *Language and Species*, Chicago, University of Chicago Press.
- \_\_\_\_\_ (1995) *Language and Human Behaviour*, Seattle, University of Washington Press, London, UCL Press.
- BILGRAMI, A. (1998) "Self-Knowledge and Resentment" en C. Wright, B. Smith & C. Macdonald (eds.) *Knowing Our Own Minds*, Oxford, Oxford University Press, pp. 207-241.
- BLOCK, N. (1986) "Advertisement for a semantics for psychology", *Midwest Studies in Philosophy* 10:615-78

- \_\_\_\_\_ (1998) "Mental and semantic holism" en Craig, E. (ed.) *Routledge Encyclopedia of Philosophy*. Routledge.
- BOGDAN, R. (1997) *Interpreting Minds*, A Bradford Book, M.I.T.
- BORRADORI, G. (1994) *The American Philosopher: Conversations with Quine, Davidson, Putnam, Nozick, Danto, Rorty, Cavell, MacIntyre, Kuhn*, Chicago, The University of Chicago Press.
- BRANDOM, R. (1994) *Making it Explicit*, Cambridge, Harvard U.P.
- \_\_\_\_\_ (2000) (ed.) *Rorty and His Critics, with and Introduction by Robert Brandom*, Oxford, Blackwell Publishers.
- \_\_\_\_\_ (2002) *Tales of the Mighty Dead*. Cambridge, MA: Harvard University Press.
- BUBER, M. (1922) *Yo y Tú*, versión castellana: Traducción de Horacio Crespo. Ed. Nueva Visión, Buenos Aires, 1969.
- BURG, F. (2007) *Davidson and Spinoza: mind, matter and morality*, Ashgate Publishing Company, Burlington.
- BURGE, T. (1979) "Individualism and the Mental", *Midwest Studies in Philosophy* 4: 73-121
- BURGE, T. (1988) "Individualism and Self-Knowledge", *Journal of Philosophy*, 85: 649-63.
- BYRNE, R. & WHITEN, A. (1988) (eds.) *Machiavellian Intelligence*, Oxford, Oxford University Press.
- CAORSI, C. (2001) *De una teoría del lenguaje a una teoría de la acción intencional*, Montevideo, Asociación Cultural Factótum Ed.
- CAORSI, C. & SILVA FILHO, W. (2008) (eds.) *Razones e interpretaciones. La filosofía después de Donald Davidson*, Buenos Aires, Ed. Del Signo.
- CAREY, S. (1985) *Conceptual Change in Childhood*, Cambridge, Mass., MIT Press.
- CARPENTER, A. (2002) "Davidson's transcendental Arguments", en Malpas, J. (2002) (ed.), cap.12.
- CARRUTHERS, P. (1996) *Language, Thought and Consciousness: An Essay in Philosophical Psychology*, Cambridge, Cambridge University Press.
- CARRUTHERS, P. Y BOUCHER, J. (1998) (ed.) "Introduction: opening up options" en *Language and Thought Interdisciplinary Themes*, Cambridge, Cambridge University Press.
- CAVELL, M. (1996) *The Psychoanalytic Mind. From Freud to Philosophy*, Harvard University Press
- CHOMSKY, N. (1957) *Syntactic Structures*, The Hague, Mouton.
- \_\_\_\_\_ (1980) *Rules and Representations*, New York, Columbia University Press.
- \_\_\_\_\_ (1988) *Language and Problems of Knowledge*, Cambridge, Mass., MIT Press.
- \_\_\_\_\_ (1995) "Language and Nature", *Mind*, 104, 1-61.
- CRAGNOLINI, M. (2007) *Derrida, un pensador del resto*, Buenos Aires, La Cebra Ed.
- CUTROFELLO, A. (1999) "The Transcendental Pretensions of the Principle of Charity" en Hahn, L.E. (1999) (ed.), pp. 333-341.
- DANCY, J. (2000), *Normativity*, Oxford, Blackwell Publishers.
- DAVIDSON, D. (1980) *Essays on Actions and Events*, Oxford, Oxford University Press.
- \_\_\_\_\_ (1984) *Inquiries into Truth and Interpretation*, Oxford, Oxford University Press.
- \_\_\_\_\_ (2001) *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press.

- \_\_\_\_\_ (2004) *Problems of Rationality*, Oxford, Oxford University Press.
- \_\_\_\_\_ (2005) *Truth, Language and History*, Oxford, Oxford University Press.
- \_\_\_\_\_ (1963/1980) "Actions, reasons and causes" en Davidson (1980), pp. 3-20.
- \_\_\_\_\_ (1965/1984) "Theories of Meaning and Learnable Languages" en Davidson (1984), pp. 3-16.
- \_\_\_\_\_ (1967/1984) "Truth and Meaning" en Davidson (1984), pp. 17-42.
- \_\_\_\_\_ (1967b/1980) "Causal Relations" en Davidson (1980), pp. 149-162.
- \_\_\_\_\_ (1970/1980) "Mental Events" en Davidson (1980), pp. 207-228.
- \_\_\_\_\_ (1973/1984) "Radical Interpretation" en Davidson (1984), pp. 125-140.
- \_\_\_\_\_ (1974/1984) "On the very idea of a conceptual scheme" en Davidson (1984), pp. 183-198.
- \_\_\_\_\_ (1974b/1984) "Belief and the Basis of Meaning" en Davidson (1984), pp. 141-154.
- \_\_\_\_\_ (1974c/1980) "Psychology as philosophy" en Davidson (1980), pp. 229-244.
- \_\_\_\_\_ (1975/1984) "Thought and Talk" en Davidson (1984), pp. 155-170.
- \_\_\_\_\_ (1976/1984) "Reply to Foster" en Davidson (1984), pp. 171-180.
- \_\_\_\_\_ (1976b/1980) "Hume's cognitive theory of pride" en Davidson (1980), pp. 277-290.
- \_\_\_\_\_ (1977/1984) "Reality without reference" en Davidson (1984), pp. 215-226.
- \_\_\_\_\_ (1977b/1984) "The Method of Truth in Metaphysics" en Davidson (1984), pp. 199-214.
- \_\_\_\_\_ (1980b/2004) "A Unified Theory of Thought, Meaning and Action" en Davidson (2004), pp. 151-166.
- \_\_\_\_\_ (1982/2001) "Rational Animals" en Davidson (2001), pp. 95-106.
- \_\_\_\_\_ (1982b/2004) "Paradoxes of Irrationality" en Davidson (2004), pp. 169-188.
- \_\_\_\_\_ (1983/2001) "A Coherence Theory of Truth and Knowledge" en Davidson (2001), pp. 137-158.
- \_\_\_\_\_ (1984b/1984) "Communication and Convention" en Davidson (1984), pp. 265-280.
- \_\_\_\_\_ (1984c/2004) "Expressing evaluations" en Davidson (2004), pp. 19-38.
- \_\_\_\_\_ (1984d/2001) "First Person Authority" en Davidson (2001), pp. 3-14.
- \_\_\_\_\_ (1986/2005) "A nice derangement of epitaphs" en Davidson (2005), pp. 89-108.
- \_\_\_\_\_ (1986b/2004) "Deception and Division" en Davidson (2004), pp. 199-212.
- \_\_\_\_\_ (1987/2004) "Problems in the explanation of action" en Davidson (2004), pp. 101-116.
- \_\_\_\_\_ (1987b/2001) "Knowing one's own mind" en Davidson (2001), pp. 15-38.
- \_\_\_\_\_ (1988/2001) "The Myth of the Subjective" en Davidson (2001), pp. 39-52.
- \_\_\_\_\_ (1989/2001) "What is present to the mind?" en Davidson (2001), pp. 53-68.
- \_\_\_\_\_ (1989b/2005) "James Joyce and Humpty Dumpty" en Davidson (2005), pp. 143-158.
- \_\_\_\_\_ (1989c) 'The Conditions of Thought', en J. Brandl and W.L. Gombocz (eds), *The Mind of Donald Davidson*, Amsterdam, Rodopi, 1989.
- \_\_\_\_\_ (1991/2001) "Three varieties of knowledge" en Davidson (2001), pp. 205-220.
- \_\_\_\_\_ (1992) "Las condiciones del pensamiento" en Davidson, D. (1992), *Mente, mundo y acción*, Barcelona, Paidós.

- \_\_\_\_\_ (1992b/2001) "The second person" en Davidson (2001), pp.107-122.
- \_\_\_\_\_ (1993/2005), "Locating Literary Language" en Davidson (2005) pp. 167-182.
- \_\_\_\_\_ (1993b/2005) "Spinoza's causal theory of the affects" en Davidson (2005), pp.295-314.
- \_\_\_\_\_ (1993c/2005) "Thinking causes" en Davidson (2005), pp. 185-200.
- \_\_\_\_\_ (1993d/2005) "Method and Metaphysics" en Davidson (2005), pp.39-46.
- \_\_\_\_\_ (1994/2005) "The social aspect of language" en Davidson (2005), pp.109-126.
- \_\_\_\_\_ (1994b/2005) "Dialectic and Dialogue" en Davidson (2005), pp. 251-260.
- \_\_\_\_\_ (1994c/2005) "What is Quine's view of truth?" en Davidson (2005), pp.81-86.
- \_\_\_\_\_ (1995/2004) "Could there be a science of rationality?" en Davidson (2004), pp.117-134.
- \_\_\_\_\_ (1995b/2004) "The problem of objectivity" en Davidson (2004), pp. 3-18.
- \_\_\_\_\_ (1995c/2004) "The objectivity of values" en Davidson (2004), pp.39-58.
- \_\_\_\_\_ (1995d/2005) "Laws and cause" en Davidson (2005), pp.201-220.
- \_\_\_\_\_ (1996/2005) "The folly of trying to define truth" en Davidson (2005), pp.19-38.
- \_\_\_\_\_ (1997/2005) "Seeing through language" en Davidson (2005), pp.127-141.
- \_\_\_\_\_ (1997b/2001), "Indeterminism and antirealism" en Davidson (2001), pp. 69-84.
- \_\_\_\_\_ (1997c/2001) "The emergence of thought" en Davidson (2001), pp.123-134.
- \_\_\_\_\_ (1997d/2005) "Gadamer and Plato's *Philebus*" en Davidson (2005), pp.261-276.
- \_\_\_\_\_ (1997e/2005) "Truth Rehabilitated" en Davidson (2005), pp.3-18.
- \_\_\_\_\_ (1998/2005) "Appendix: Replies to Rorty, Stroud, McDowell and Pereda" en Davidson (2005), pp.316-327.
- \_\_\_\_\_ (1998b/2001) "The irreducibility of the concept of the self" en Davidson (2001), pp. 85-91.
- \_\_\_\_\_ (1999) "Reply to Martin" en Hahn (1999) (ed.), pp. 357-360.
- \_\_\_\_\_ (2001b) "Comments on Karlovy Papers" en Kotatko, P., Pagin, P. & Segal, G. (eds.), *Interpreting Davidson*, CSLI, Stanford, pp. 285-308.
- \_\_\_\_\_ (2001c), "Externalisms" en Kotatko, P., Pagin, P. & Segal, G. (eds.), *Interpreting Davidson*, CSLI, Stanford, pp.1-16.
- Nicod Lectures V
- DAVIDSON, D. y GLÜER, K. (1995) "Relations and Transitions. An Interview with Donald Davidson", *Dialectica*, 49:75-86.
- DARWALL, S. (2006) *The Second-Person Standpoint: Morality, Respect, and Accountability*, Harvard University Press.
- DASCAL, M. (2007) "Dichotomies and types of debate" en Eemeren & Garssen (eds.) *Controversy and Confrontation*, University of Amsterdam, Benjamins Company.
- DENNETT, D. (1991) *Consciousness Explained*, New York, Little Brown and Co. (UK: Allen Lane: Penguin)
- DENNETT, D. (1971) "Intentional Systems", *Journal of Philosophy*, Londres, Routledge and Kegan Paul.
- \_\_\_\_\_ (1973) "Mechanism and Responsibility" en T Honderich (ed.) *Essays on Freedom of Action*.
- \_\_\_\_\_ (1987) *The Intentional Stance*, MIT Press, Cambridge, Mass
- \_\_\_\_\_ (1989) "Condiciones de la cualidad de persona", *Cuadernos de Crítica*, nro. 45, pp. 5-36.

- DESCARTES, R. (1641) *Meditaciones Metafísicas*, en E. de Olazo y T. Zwanck (comps.) *Obras Escogidas*, Buenos Aires, Editorial Charcas, 1980.
- DEVITT, M. (1991), "Naturalistic Representation" *British Journal for the Philosophy of Science* 42 (3).
- DRAY, W. (1957), *Laws and Explanation in History*, London, Oxford University Press.
- DRESTKE, F. (1994), "Introspection", *Proceedings of the Aristotelian Society*, 94, pp. 263-278.
- DUARTE, A. y RABOSI, E. (2003) (eds.) *Psicología cognitiva y filosofía de la mente: pensamiento, representación y conciencia*, Buenos Aires, Alianza.
- DUMMETT, M. (1978) "Can Analytic Philosophy Be Systematic and Ought it to Be?" en *Truth and Other Enigmas*, Cambridge, Mass. Harvard University Press. 437-458
- \_\_\_\_\_ (1981) *Frege: Philosophy of Language*, London, Duckworth.
- \_\_\_\_\_ (1986) "A Nice Derangement of Epitaphs": Some Comments on Davidson and Hacking, en E. Lepore (ed.). *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Cambridge, Mass., Blackwell.
- \_\_\_\_\_ (1991) *Frege and Other Philosophers*, Oxford, Oxford University Press.
- ENGEL, P. (1999) "The Norms of the Mental" en Hahn (1999) (ed.), pp. 447-459.
- \_\_\_\_\_ (2001) "Is Truth a Norm?" en Kotatko, P; Pagin P. & Segal, G. (eds.), *Interpreting Davidson*, CSLI, Stanford. pp. 37-52.
- ESPAÑOL, S. (2004) *Cómo hacer cosas sin palabras*, Madrid, A. Machado Libros
- \_\_\_\_\_ (2008) "La entrada al mundo a través de las artes temporales" en *Estudios de Psicología*, vol. 29 (1), pp.81-101.
- EVANS, G. (1982) *The Varieties of Reference*, Oxford, Oxford University Press.
- EVNINE, S. (1991) *Donald Davidson*, California, Stanford University Press.
- FLAVELL, J. (1985) *Cognitive Development*, New York, Prentice Hall, second edition.
- FODOR, J. (1974) "Special Sciences", *Synthese* 28:97-115.
- \_\_\_\_\_ (1975) *The Language of Thought*, New York, Crowell. UK: Hassocks, Harvester Press.
- \_\_\_\_\_ (1978) "Propositional attitudes", reimpresso en Fodor, J. (1981) *RePresentations*, Hassocks, Harvester Press.
- \_\_\_\_\_ (1983) *The Modularity of Mind*, Cambridge, Mass., MIT Press.
- \_\_\_\_\_ (1991) "You Can Fool Some of The People All of The Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations", *Mind* 100 (397):19-34.
- \_\_\_\_\_ (1998) *Concepts: Where Cognitive Science Went Wrong*, Oxford, Oxford University Press.
- FODOR, J. y LEPORE, E. (1992) *Holism: a shopper's guide*, Cambridge, Mass., Basil Blackwell.
- GALLAGHER, S. (2005) *How the body shapes the mind*, Oxford, Oxford University Press.
- GARCÍA SUÁREZ, A. y VALDÉS VILLANUEVA, L. (1982) "Significado, verdad y comprensión", *Teorema*, vol. xii/1-2, 1982, Valencia, España.
- GARDNER, H. (1983) *Frames of Mind: the theory of multiple intelligences*, London, Heinemann.
- GARFIELD, J.L. et. al. (2001) "Social Cognition, Language Acquisition and the Development of the Theory of Mind", *Mind and Language*, 16, 5:494-541.
- GLOCK, H.J. (2003), *Quine and Davidson on Language, Thought and Reality*, Cambridge, Cambridge University Press.

- GLÜER, K. (2001) "Dreams and Nightmares: Conventions, Norms and Meaning in Davidson's Philosophy of Language" en Kotatko, P., Pagin, P. & Segal, G. (2001), pp. 53-74.
- GLÜER, K. (2006) "The status of Charity I: conceptual truth or a posteriori necessity?", *Journal of Philosophical Studies*, 14 (3):337 – 359.
- GOLDFARB, W. (1985) "Kripke on Wittgenstein Rules", *Journal of Philosophy*, 82: 471-88.
- GOLDIE, P. (2000) *The Emotions. A Philosophical Exploration*, New York, Oxford, Clarendon Press, Oxford, U. Press.
- GOLDMAN, A. (1989) "Interpretation Psychologized", *Mind and Language*, vol. 4, pp. 161-185.
- \_\_\_\_\_ (1992) "Empathy, mind and morals", *Proceedings and Addresses of the American Philosophical Association*, vol. 66, pp. 17-41.
- GÓMEZ, J.C. (1996) "Second Person Intentional Relations and the Evolution of Social Understanding", *Behavioral and Brain Sciences*, 19 (1).
- GOMILA B., A. (2001) "La perspectiva de segunda persona: mecanismos mentales de la intersubjetividad", *Contrastes* 6: 65-86.
- \_\_\_\_\_ (2002) "La perspectiva de segunda persona de la atribución mental: mecanismos mentales de la intersubjetividad", *Azafea*, vol. IV, pp. 123-138.
- \_\_\_\_\_ (2003) "La perspectiva de segunda persona" en Rabossi, E.& Duarte, A. (eds.) *Psicología cognitiva y Filosofía de la Mente*, pp.195-218, Buenos Aires, Alianza Editorial.
- \_\_\_\_\_ (2008) "La relevancia moral de la perspectiva de segunda persona" en *Cuestiones Filosóficas, Ensayos en honor de Eduardo Rabossi*, Buenos Aires, Catálogos, pp. 155-173.
- GOPNIK, A. y ASTINGTON, J.W. (1988) "Children's Understanding of Representational Change and its Relation to the Understanding of False Belief and Appearance-reality Distinction", *Child Development*, 59:26-37.
- GORDON, R. (1986) "Folk Psychology as Simulation", *Mind and Language* 1:158-71.
- \_\_\_\_\_ (1992) "The Simulation Theory: Objections and Misconceptions", *Mind and Language*: 7 (1-2):11-34.
- GRICE, H. P. (1957) "Meaning", *Philosophical Review*, 66, pp.377-88, reimpresso en Grice 1989, pp.213-23.
- \_\_\_\_\_ (1969) "Utterer's meaning and intention", *Philosophical Review*, 78.
- \_\_\_\_\_ (1989) *Studies in the Way of Words*, Cambridge, Mass., Harvard University Press.
- HABERMAS, J. y PUTNAM, H. (2008) *Normas y valores*. Introducción, traducción y notas de Jesús Vega Encabo y Francisco Javier Gil Martín, Ed. Trotta, Madrid.
- HABERMAS, J. y RAWLS, J. (1998) "Reconciliación mediante el uso público de la razón" en Habermas, J y Rawls, J. *Debate sobre el liberalismo político*, Paidós, Barcelona.
- HABERMAS, J. (2000), *Aclaraciones a la ética del discurso*, Trotta, Madrid.
- HACKING, I. (1986) "The Parody of Conversation" en E.Lepore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Cambridge, Mass., Blackwell.
- HAHN, L.E. (1999) (ed.) *The Philosophy of Donald Davidson*, Library of Living Philosophers, vol. 27, Chicago, Open Court.
- HAMPSHIRE, S. (1959) *Thought and action*, Chatto and Windus, London.
- HARE, R.M. (1981) *Moral Thinking*, Oxford: Clarendon Press.
- HART, H.L.A. y HONORÉ, A.M. (1959), *Causation in the Law*, Clarendon Press, Oxford.

- HATFIELD, E. CACIOPPO, J. Y RASPÓN, R.L. (1992) "Primitive Emotional Contagion", en Clark, M. (ed.) *Emotion and Social Behavior*, California, Sage Pub., pp.151-177.
- HERNÁNDEZ IGLESIAS, M. (1990), *La semántica de Donald Davidson. Una introducción crítica*. Visor, Madrid.
- HOBSON, R.P. (1995) , *El autismo y el desarrollo de la mente*, Madrid, Alianza.
- HONDERICH, T. (1973) (ed.) *Essays on Freedom of Action*, London and Boston, Routledge & Kegan Paul.
- HOOKEYWAY, C. (1988) *Quine: Language, Experience and Reality*. Stanford: University Press
- HUTTO, D. (2002) "The World is not Enough: Shared Emotions and Other Minds", *Understanding Emotions: Mind and Morals*, Goldie, P. (ed.), Aldershot, Hants: Ashgate Pub. Ltd.
- JANKE, T. (2008), "A Freewheeling Defense of Kant's Resolution of the Third Antinomy". *Kritike*, vol. 2, nro.1 (junio 2008), pp. 110-122.
- JEANNEROD, M. (2001) *The Cognitive Neuroscience of Action*, Oxford, B. Blackwell.
- JEFFREY, R. (1985) "Animal Interpretation" en Lepore (ed.) *Actions and Events: Perspectives on Philosophy of Donald Davidson*, pp.482-7. Oxford, Blackwell.
- \_\_\_\_\_ (1990) en *The Logic of Decision*, 2<sup>nd</sup>. Ed., University of Chicago Press.
- KELTNER, D. y HAIDT, J. (1999) "Social Functions of Emotions at Four Levels of Analysis", *Cognition and Emotion*, 13:5:505-521.
- KENNY, A. (1963) *Action, Emotion and Will*, Routledge and Kegan Paul, London.
- KIM, J. (1985) "Psychophysical Laws", reimpresso en Kim (1993) *Supervenience and Mind*, Cambridge University Press, Cambridge.
- \_\_\_\_\_ (1989), "The Myth of Nonreductive Materialism", *Proceedings and Addresses of the American Philosophical Association*, LXIII, 1989, pp. 31-47.
- \_\_\_\_\_ (1993) *Supervenience and Mind*, Cambridge University Press, Cambridge.
- \_\_\_\_\_ (2000) *Mind in a Physical World*, Cambridge, Mass., MIT Press,
- KORSGAARD, Ch. (1996) *Creating the kingdom of ends*, New York, Cambridge University Press.
- KOSSLYN, S. (1994) *Image and brain: the resolution of imagery debate*: Cambridge, MA: MIT Press.
- KOTATKO, P., PAGIN, P. & SEGAL, G. (2001) (eds.) *Interpreting Davidson*, Stanford, CSLI Publications.
- KRIPKE, S. (1982) *Wittgenstein on Rules and Private Language*, Oxford, Oxford University Press.
- LEIBNIZ, G. (1704) *New Essays on Human Understanding*, Remnant & Bennett (1996) (ed.) Cambridge, Cambridge University Press
- LEPORE, E. & MCLAUGHLIN, B. (1985) *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford, Basil Blackwell.
- LEPORE, E. (1986) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford, Blackwell.
- LEPORE, E y LUDWIG, K. (2003) "Truth and meaning" en Ludwig, K. (2003) (ed.), pp.35-63.
- LEPORE, E. (2004), "An Interview with Donald Davidson" en Davidson, D. (2004), *Problems of Rationality*, Oxford University Press, pp. 231-266.
- LEPORE, E. y LUDWIG, K. (2005) *Meaning, Truth, Language, and Reality*, Oxford, OUP.
- \_\_\_\_\_ (2006) (eds.) *The Essential Davidson/Donald Davidson; with an introduction by Ernie Lepore and Kirk Ludwig*. Oxford, Oxford University Press.

- LEVELT, W. (1989) *Speaking: From Intention to Articulation*, Cambridge, Mass., MIT Press.
- LEVI, I. (1999) "Representing Preferences. Donald Davidson on Rational Choice" en Hahn (1999) (ed.), pp. 531-570.
- LEWIS, D. (1969) *Convention*, Oxford, Blackwell.
- \_\_\_\_\_ (1983) "Language and languages" en *Philosophical Papers*, vol. 1, pp.163-88, Oxford, Oxford University Press.
- LOCKE, J. (1690) *An Essay Concerning Human Understanding*, edited with an introduction of P. Niddicht (1975), Oxford, Clarendon Press.
- LUDWIG, K. (ed.) (2003) *Donald Davidson*, Cambridge, Cambridge University Press.
- LURIA, A. (1959) "The directive function of speech in development and dissolution, *Word*, 15, 341-52.
- LYCAN, W. (1987) *Consciousness*, Cambridge, MA, MIT Press.
- MACKIE, J. (1978) *Ethics: Inventing Right and Wrong*, Viking Press.
- MALPAS, J. (1992) *Donald Davidson and the Mirror of Meaning*, Cambridge, Cambridge University Press.
- \_\_\_\_\_ (2002) (ed.) *From Kant to Davidson: Philosophy and the Idea of the Transcendental*, Londres, Routledge.
- \_\_\_\_\_ (2002b). "Gadamer, Davidson, and the Ground of Understanding", *Gadamer's Century: Essays in Honor of Hans-Georg Gadamer*, MIT Press, Malpas, Jeff, Arnsward, Ulrich and Kertscher, Jens (2002) (ed), Massachusetts, pp. 195-215.
- \_\_\_\_\_ (2007) "The nature of interpretative charity", *Dialectica*, vol 42, issue 1, pp.17-36.
- MARTIN, B. (1999): "Interpretation and Responsibility: Excavating Davidson's Ethical Theory" en Hahn (1999) (ed.) pp. 345-356.
- MCDOWELL, J. (1994) *Mind and World*, Cambridge, Mass., Harvard University Press.
- \_\_\_\_\_ (1998) "The constitutive ideal of rationality: Davidson and Sellars", *Critica*, vol.30, nro. 88, pp. 29-48.
- MELDEN A.I. (1961) *Free Action*, Routledge and Kegan Paul.
- MELTZOFF, A.& MOORE, M. (1977) "Imitation of facial and manual gestures". *Science*, 198, pp. 75-80.
- \_\_\_\_\_ (1989) "Imitation in newborn infants: exploring the range of gestures imitated, and the underlying mechanisms", *Developmental Psychology*, 25. pp. 954-962.
- (1994) ) "Imitation, memory and the representation of persons". *Infants Behavior and Development*, 17. pp. 83-99.
- MORAN, R. (1997) "Self-Knowledge" an Empirical Problem? Renegotiating the Space of Philosophical Explanation", *The Journal of Philosophy*, 5, pp.141-161.
- MORETTI, A. (1996) *Concepciones tarskianas de la verdad*, Oficina de publicaciones del C.B.C., UBA, Bs. As.
- \_\_\_\_\_ (2008) *Interpretar y referir (Ejercicios de análisis filosófico)*, Bs. As., Grama Editorial.
- MOYA, C. (1990), *The Philosophy of Action; An Introduction*. Polity Press, Cambridge, U.K.
- \_\_\_\_\_ (1992) "Introducción a la filosofía de Davidson: mente, mundo y acción", en Davidson (1992), pp.9-45.
- NAGEL T (1972) "War and Massacre", *Philosophy and Public Affairs*, Vol. 1, No. 2, pp. 123-144.
- \_\_\_\_\_ (1986) *The View From Nowhere*. New York: Oxford University Press

- ORENSTEIN, A. (2002) *W.V. Quine. Philosophy Now*, Princeton University Press.
- ORLANDO, E. (1999) "Caridad e interpretación" en *Manuscrito*, Universidad de Campinas, San Pablo, vol. XXII, N° 1 (abril), pp.85-115.
- \_\_\_\_\_ (1999b) "Interpretación radical y holismo semántico" en *Cuadernos de Filosofía* (Instituto de Filosofía, Facultad de Filosofía y Letras, UBA), N° 46 (primavera de 1999), pp.9-25.
- \_\_\_\_\_ (2001) "Sobre la estrategia de la semántica realista", *Diálogos*, vol. 36, nro. 77, pp.47-68.
- PAGIN, P. (2006) "Meaning Holism" en E. Lepore y B. Smith (eds), *Handbook of Philosophy of Language*, Oxford University Press, Oxford, pp. 213-32.
- PÉREZ, D. (1997), "Notas sobre la analogía entre los sistemas de medida de los objetos físicos y los estados mentales de los seres humanos", *Diánoia*, nro. 43, pp.31-46.
- \_\_\_\_\_ (1999), *La mente como eslabón causal*, Buenos Aires, Catálogos.
- \_\_\_\_\_ (2005), "Is thought without language possible?", *Principia* 9, pp. 177-191.
- \_\_\_\_\_ (inédito) "Las bases no conceptuales de la psicología folk".
- PERNER, J. (1987) *et. al.* "Three years-Olds' Difficulty with False belief: The case for Conceptual Deficit", *British Journal of Developmental Psychology*, 5: 125-37.
- PERRY, J. (1994) "Fodor and Lepore on Holism", *Philosophical Studies*, 73, pp. 123-138.
- PIAGET, J. (1932) *The Language and Thought of the Child*, London, Routledge and Kegan Paul.
- PIETROSKI (1994) "A defense of derangement" *Canadian Journal of Philosophy* 24:95-118.
- PINKER, S. (1994) *The Language Instinct*, London, Penguin Books, US, New York, William Morrow.
- PREMACK, D. y WOODRUFF, G. (1978), "Does the chimpanzee have a theory of mind?", *Behavioral and Brain Sciences* 1(4), 515-526.
- PREYER, G. SIEBELT, F. y ULFIG, A. (1994) (eds.) *Language, Mind and Epistemology*, Kluwer, Dordrecht.
- PUTNAM, H. (1964) "Robots: Machines or Artificially Created Life?", *Journal of Philosophy*, 61:668-691.
- \_\_\_\_\_ (1975) "The meaning of 'meaning'". In K. Gunderson (Ed.), *Language, Mind and Knowledge*, Vol. VII of *Minnesota Studies in the Philosophy of Science*, 131-193. University of Minnesota Press.
- \_\_\_\_\_ (1981/1988) *Razón, verdad e historia*, traducción de J. Esteban, Madrid, Tecnos, 1988, pp. 11-12; publicación original: H. Putnam, *Reason, Truth, and History*, Cambridge, Cambridge University Press, 1981.
- \_\_\_\_\_ (2002/2004) *The collapse of the fact/value dichotomy and other essays*, Cambridge, MA, Harvard University Press. (2004) versión castellana: *El desplome de la dicotomía hecho/valor y otros ensayos*, Barcelona, Paidós. (2004)
- \_\_\_\_\_ (2004b) *Ethics without ontology*, Cambridge, Mass.: Harvard University Press.
- QUINE, W. V. (1951) "Two Dogmas of Empiricism", *The Philosophical Review* 60 (1951): 20-43. Reimpreso en W.V.O. Quine, *From a Logical Point of View* (1953) Harvard University Press.
- \_\_\_\_\_ (1960) *Word and Object*, Cambridge: MIT Press.
- QUINTANILLA, P. (2008) "Taylor contra Davidson: intérpretes participantes o no comprometidos" en Caorsi & Silva-Filho (eds.) pp. 145-160.

- RABOSSA, E. (2000) "La filosofía de sentido común y la teoría de la teoría. Algunas reflexiones críticas", *Endoxa*, 12, pp. 643-655.
- RAMBERG, B. (1989) *Donald Davidson's Philosophy of Language: An Introduction*, Cambridge, Mass., Blackwell.
- \_\_\_\_\_ (1999) "The significance of charity" en Hahn (1999) (ed.) 601-618.
- \_\_\_\_\_ (2000) "Post-Ontological Philosophy of Mind: Rorty versus Davidson" en Brandom R. (ed.) *Rorty and his critics*, Oxford, Blackwell Publishers.
- \_\_\_\_\_ (2003) "Illuminating Language: Interpretation and understanding in Gadamer and Davidson" en Prado, C. (2003) *A House Divided: Comparing Analytic and Continental Philosophy*, Humanity Books.
- RAMSEY F. P. (1931), "Truth and Probability", en R.B. Braithwaite (ed.) *The Foundations of Mathematics and other Logical Essays*, pp.156-98, London, Routledge, Kegan Paul.
- RAWLING P. (2003) "Radical Interpretation" en Ludwig K. (2003) (ed.) *Donald Davidson*, pp. 85-112.
- RIVIÈRE, A. Y SOTILLO, M. (1999/2003) "Comunicación, suspensión y semiosis humana: los orígenes de la práctica y de la comprensión interpersonales" reimpresso en M. Belichón, M. et al (2003) *Angel Rivière. Obras escogidas*, vol. 1. Madrid, Ed. Panamericana.
- RORTY, A. (1962) "Slaves and Machines", *Analysis*, Volume 22, No. 5, No. 89, Abril, pp. 1 18-120.
- RORTY, R. (1970) "Incorrigibility as the Mark of the Mental", *Journal of Philosophy*, 67.
- \_\_\_\_\_ (1986) "Pragmatism, Davidson and Truth" en Lepore. E. (1986) pp. 333-355.
- \_\_\_\_\_ (1989) *Contingency, Irony, and Solidarity*, Cambridge: Cambridge University Press
- \_\_\_\_\_ (1997) "Davidson entre Wittgenstein y Tarski", *Crítica Revista Hispanoamericana de Filosofía*, vol XXX, nro. 88, 49-71.
- \_\_\_\_\_ (1999) "Davidson's Mental-Physical Distinction" en Hahn (1999) (ed.), pp. 575-594
- ROVANE, C. (1999) "Rationality and Identity" en Hahn (1999) (ed.), pp.463-479.
- RUSSELL, B. (1921) *The Analysis of Mind*, London: George Allen and Unwin.
- RYLE, G. (1949) *The Concept of Mind*, Barnes and Noble, New York.
- SCHNEIDER, H. (1979) "Explanation and Understanding in the Theory of Language" en Bäuerle et. al. (comp.), *Semantics from Different Points of View*, Springer, Berlin, 1979.
- SCOTTO, C. (2002) "Interacción y atribución mental: la perspectiva de la segunda persona", *Análisis Filosófico*, vol. XXII, nro. 2, noviembre 2002, pp. 135-151
- SEARLE, J. (1983) *Intentionality*, Cambridge, Cambridge University Press.
- \_\_\_\_\_ (1969), *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.
- SELLARS, W. (1956/1963) "Empiricism and the Philosophy of Mind" en *Minnesota Studies in the philosophy of science, vol 1*, Feigl & Scriven (eds.) reimpresso con introducción de Rorty, R., study guide de Brandom, R., Cambridge, Mass., Harvard University Press, 1997.
- \_\_\_\_\_ (1966) "Fatalism and Determinism" en K. Lehrer (ed.) *Freedom and Determinism*, Nueva York, Random House, 1966.
- SHIFRES, F. (2008) "Música, transmodalidad e intersubjetividad" en *Estudios de Psicología*, vol. 29 (1), pp.7-30.
- SHOEMAKER (1963), *Self-Knowledge and Self-Identity*, London: Oxford University Press.
- SMART J.J.C. (1987) 'Davidson's Minimal Materialism' en *Essays on Davidson: Actions and Events*, Vermazen and Hintikka (eds) Oxford: Clarendon Press. pp. 173-182.

- SOKOLOV, A. (1972) *Inner Speech and Thought*, New York, Plenum Press.
- SPERBER, D. et.al (1995). *Causal Cognition*, Oxford, Oxford University Press.
- STRAWSON, P.F. (1959) *Individuals*, London, Methuen.
- \_\_\_\_\_ (1974/1995), *Libertad y resentimiento y otros ensayos*, Barcelona, Paidós.
- \_\_\_\_\_ (1983) *Ensayos Lógico-lingüísticos*, Tecnos, Madrid
- \_\_\_\_\_ (1984) "Hot Pursuit: Inquiries into Truth and Interpretation", *London Review of Books*: 17.
- TAYLOR, Ch. (1985) *Human Agency and Language: Philosophical Papers I*, Cambridge, Cambridge University Press.
- THOMPSON, E. (2001), *Between Ourselves: Second-person issues in the study of consciousness*, Londres, Imprint Academic.
- TREVARTHEN, C. (1980) The Foundations of Intersubjectivity: Development of Interpersonal and Cooperative Understanding in Infants. In D. Olson (Ed.). *The Social Foundations of language and Thought. Essays in Honor of Jerome S. Bruner*. New York:W. W. Norton & Co.
- VEGA, J. (2001) "Reglas, medios, habilidades. Debates en torno al análisis de "S sabe cómo hacer X", *Crítica*, 33:3-40.
- VERMAZEN, B. y HINTIKKA, M. (1987) (eds.) *Essays on Davidson: Actions and Events*, Oxford: Clarendon Press.
- VIGOTSKY, L. (1934) *Thought and Language*, reimpresso en (1986), Cambridge, Mass., MIT Press.
- VON WRIGHT (1963), *Norm and action, A logical Enquiry*, London, Routledge & Kegan Paul.
- WALKER, J. (1983) *Animal Thought*, London, Routledge.
- WALSH, V. (1987) "Philosophy and Economics", en J. Eatwell, M. Milgate y P. Newman (comps.) *The New Palgrave: A Dictionary of Economics*, vol. 3, Londres, Macmillan, 1987, pp.861-869.
- WELLMAN, H., PHILLIPS, A. y RODRIGUEZ, T. (2000) "Young children understanding of perception, desire and emotion", *Child Development*, vol. 71. no. 4, pp. 895-912.
- WHORF, B.L. (1956) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*, Cambridge, Mass., MIT Press, US, New Jersey, Wiley.
- WIMMER, H. y PERNER, J. (1983), "Beliefs about beliefs: Representation and constraining of wrong beliefs in young children's understanding of deception", *Cognition*, 13, pp.103-28.
- TOMASELLO, M. CALL, J. Y HARE, B. (2003) "Chimpanzees understand psychological states- the question is which ones and to what extent" *Trends in Cognitive Science*, 7, 153-156.
- WISPÉ, L. (1986) "The Distinction Between Sympathy and Empathy: to Call Forth a Concept, a Word is Needed", *Journal of Personality and Social Psychology*, 20, 2: 314-321.
- WITHEN, A. y BYRNE, R.W. (1988) "Tactical deception in primates", *Behavioral and Brain Sciences*, 11: 233-273.
- WITTGENSTEIN, L. (1921) *Tractatus Logico-Philosophicus*, London, Routledge.
- \_\_\_\_\_ (1953) *Philosophical Investigations*, Basil Blackwell, Oxford.
- \_\_\_\_\_ (1967) *Zettel*, Basil Blackwell, Oxford.
- \_\_\_\_\_ (1980) *Observaciones sobre la Filosofía de la Psicología*, México, UNAM, 1985, VOL. 2.
- WOLLHEIM, R. (1980) "Seeing-as, seeing-in, and Pictorial Representation", Ensayo suplementario V de *Art and its objects*, 2da. ed., Cambridge University Press.
- ZEGLEN, U. (1999) (ed.) *Donald Davidson: Truth, Meaning and Knowledge*, London, Routledge.