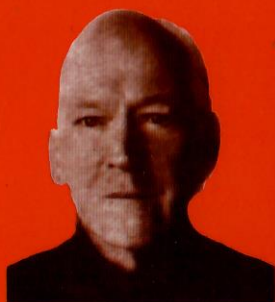


Donald Davidson



**SUBJETIVO,
INTERSUBJETIVO,
OBJETIVO**



CÁTEDRA
colección teorema

*Subjetivo,
intersubjetivo,
objetivo*

Colección Teorema

Donald Davidson

*Subjetivo,
intersubjetivo,
objetivo*

Traducción de Olga Fernández Prat

CÁTEDRA

TEOREMA

Título original de la obra:
Subjective, Intersubjective, Objective

1.ª edición, 2003

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaran, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© in this collection Donald Davidson 2001.
Subjective, Intersubjective, Objective was originally published in English in 2001. This translation is published by arrangement with Oxford University Press
© Ediciones Cátedra (Grupo Anaya, S. A.), 2003
Juan Ignacio Luca de Tena, 15. 28027 Madrid
Depósito legal: M. 34.762-2003
I.S.B.N.: 84-376-2085-6
Printed in Spain
Impreso en Lavel, S. A.
Humanes de Madrid (Madrid)

ADVERTENCIA

ESTA ES UNA COPIA PRIVADA PARA FINES
EXCLUSIVAMENTE EDUCACIONALES



QUEDA PROHIBIDA
LA VENTA, DISTRIBUCIÓN Y COMERCIALIZACIÓN

- El objeto de la biblioteca es facilitar y fomentar la educación otorgando préstamos gratuitos de libros a personas de los sectores más desposeídos de la sociedad que por motivos económicos, de situación geográfica o discapacidades físicas no tienen posibilidad para acceder a bibliotecas públicas, universitarias o gubernamentales. En consecuencia, una vez leído este libro se considera vencido el préstamo del mismo y deberá ser destruido. No hacerlo, usted, se hace responsable de los perjuicios que deriven de tal incumplimiento.
- Si usted puede financiar el libro, le recomendamos que lo compre en cualquier librería de su país.
- Este proyecto no obtiene ningún tipo de beneficio económico ni directa ni indirectamente.
- Si las leyes de su país no permiten este tipo de préstamo, absténgase de hacer uso de esta biblioteca virtual.

"Quién recibe una idea de mí, recibe instrucción sin disminuir la mía; igual que quién enciende su vela con la mía, recibe luz sin que yo quede a oscuras" ,

—Thomas Jefferson



Para otras publicaciones visite
www.lecturasinegoismo.com
Referencia: 3644

Índice

ORIGEN DE LOS ENSAYOS Y AGRADECIMIENTOS	11
INTRODUCCIÓN	15

LO SUBJETIVO

1. La autoridad de primera persona (1984)	25
2. Conocer nuestra propia mente (1987)	41
3. El mito de lo subjetivo (1988)	72
4. ¿Qué está presente en la mente? (1989)	91
5. El indeterminismo y el antirrealismo (1997)	110
6. La irreducibilidad del concepto de yo (1998)	130

LO INTERSUBJETIVO

7. Animales racionales (1982)	141
8. La segunda persona (1992)	156
9. La aparición del pensamiento (1997)	176

LO OBJETIVO

10. Una teoría coherentista de la verdad y el conocimiento (1983)	193
Reflexiones posteriores (1987)	216
11. El contenido empírico (1982)	222
12. Epistemología y verdad (1988)	245
13. La externización de la epistemología (1990)	265
14. Tres variedades de conocimiento (1991)	280

LISTA DE CONTENIDOS DE LOS VOLÚMENES DE ENSAYOS DE DONALD DAVIDSON	301
REFERENCIAS BIBLIOGRÁFICAS	305
ÍNDICE ANALÍTICO	311

A Marcia Cavell

Origen de los ensayos y agradecimientos

El Ensayo 1, «La autoridad de primera persona», fue expuesto en una conferencia sobre intencionalidad organizada por Henri Lauener y celebrada en Biel, Suiza, en 1983. Expuse versiones previas de éste en la Universidad de Illinois en el Chicago Circle, en la Universidad de Stanford y en la Universidad de Colorado. Se publicó por primera vez en *Dialectica*, 38 (1984), 101-11.

«Conocer nuestra propia mente», el Ensayo 2, fue una conferencia presidencial del *Sixtieth Annual Pacific Division Meetings of the American Philosophical Association* celebrado en Los Ángeles el 28 de marzo de 1986 y publicado en los *Proceedings and Addresses of the American Philosophical Association* (1987), 441-58. La American Philosophical Association dio su permiso para que se reimprimiera en este volumen. Estoy muy agradecido a Akeel Bilgrami y Ernie Lepore por sus críticas y consejos y a Tyler Burge por haber intentado generosamente corregir mi comprensión de sus trabajos.

El Ensayo 3, «El mito de lo subjetivo», lo leí en un congreso sobre Consciencia, Lenguaje y Arte que se celebró en Viena en 1986 y se publicó en las actas *Bewusstsein, Sprache und die Kunst*, editadas por Michael Benedikt y Rudolf Berger y editadas por S. Verlag der Österreichischen Staatsdruckerei, 1988.

El Ensayo 4, «¿Qué está presente en la mente?», lo presenté en la Second International France Veber Colloquium, celebrado en Bad Radkersburg, Austria, y Gornja Radgona y que estaba entonces en Yugoslavia, en lo que ahora es Eslovenia.

Se publicó en un libro de ensayos sobre mi obra la mayoría de los cuales se presentaron en la conferencia. Johannes Brandl y Wolfgang Gombocz fueron los editores del libro que apareció como un volumen especial de *Grazer Philosophische Studien* (vol. 36, 1989) bajo el título *The Mind of Donald Davidson* (Amsterdam, Rodopi).

«El indeterminismo y el antirrealismo», el Ensayo 5, lo presenté en una conferencia sobre realismo y antirrealismo en la Universidad de Santa Clara a comienzos de 1992. Se publicó en *Realism/Antirealism and Epistemology*, editado por C. B. Kulp (Lanham, Maryland, Rowman & Littlefield, 1997).

El Ensayo 6, «La irreducibilidad del concepto de yo», lo escribí con la ayuda y consejo de Marcia Cavell para un libro concebido en honor de la obra de Dieter Henrich. El libro, *Philosophie in synthetischer Absicht*, fue editado por Marcelo Stamm (Stuttgart, Klett-Cotta, 1998).

El Ensayo 7, «Animales racionales», lo presenté en un congreso que organizó Henri Lauener y que tuvo lugar en Biel, Suiza, en 1981. Se publicó por primera vez en *Dialectica*, 36 (1982), 317-27.

«La segunda persona», el Ensayo 8, fue una charla que ofrecí en un congreso sobre Wittgenstein en París en 1989. Se publicó bajo el título «Jusqu'où va le caractère public d'une langue?» en *Wittgenstein et la philosophie aujourd'hui*, y fue editado y traducido al francés por J. Sebestik y A. Soulez (París, Méridiens Klincksieck, 1992). En el mismo año se publicó una versión ligeramente modificada en inglés en *Midwest Studies in Philosophy*, 17. Fue editada por P. French, T. E. Uehling y H. Wettstein (Indianapolis, University of Notre Dame Press). Alrededor de cuatro páginas del mismo que se encuentran cerca del final de este artículo las tomé de «The Conditions of Thought», un artículo escrito para una sesión plenaria del World Congress of Philosophy que tuvo lugar en Brighton en 1988 y que se presentó en el mismo. Se publicó en *Le Cahier du Collège International de Philosophie* (París, Éditions Osis, 1989).

El Ensayo 9, «La aparición del pensamiento», fue presentado oralmente en un seminario sobre emergentismo celebrado en la Universidad de Frankfurt en 1993. T. Marschner

lo tradujo al alemán y se publicó con el título «Die Emergenz des Denkens» en *Die Erfindung des Universums? Neue Überlegungen zur philosophischen Kosmologie*, editado por W. G. Saltzer, P. Eisenhardt, D. Kurth y R. E. Zimmermann (Frankfurt am Main, Insel Verlag, 1997). Posteriormente se publicó en inglés bajo el título actual en *Erkenntnis*, 51 (1999), 7-17.

El Ensayo 10, «Una teoría coherentista de la verdad y el conocimiento», fue mi contribución a un coloquio organizado por Richard Rorty como parte del Stuttgart Hegel Congress de 1981. W. V. Quine y Hilary Putnam fueron los otros de los participantes en el coloquio. Nuestros artículos se publicaron dos años después en las actas de ese congreso, *Kant oder Hegel?* Después de Stuttgart nosotros cuatro intercambiamos nuestros puntos de vista sobre los mismos temas de manera más pausada en la Universidad de Heidelberg. Cuando la Pacific Division of the American Philosophical Association se reunió en marzo de 1983, Rorty presentó un artículo titulado «Pragmatism, Davidson, and Truth». En él hizo comentarios sobre lo que escribí en «Una teoría coherentista de la verdad y el conocimiento»; su artículo fue posteriormente publicado (con revisiones) en *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Repliqué a este artículo mediante «Afterthoughts, 1987», el cual fue publicado por primera vez en conjunción con una reimpression de «Una teoría coherentista de la verdad y el conocimiento» que se encontraba en *Reading Rorty*, editado por Alan Malachowski (Cambridge, Blackwell, 1990). «Reflexiones posteriores», tal y como se ha impreso aquí y dejando aparte algunos detalles bibliográficos, es una reimpression [y traducción] de «Afterthoughts, 1987».

«El contenido empírico», el Ensayo 11, fue presentado en un congreso en Viena (en la casa que Wittgenstein contribuyó a diseñar) para celebrar el centenario de los nacimientos de Moritz Schlick y Otto Neurath. Se publicó en el mismo año en el que se celebró el congreso en *Grazer Philosophische Studien*, 16-17 (1982).

El Ensayo 12, «Epistemología y verdad», lo presenté en la Fourth Panamerican Philosophy Conference celebrada en Córdoba, Argentina, en septiembre de 1987. Lo publicó la Uni-

versidad Nacional de Córdoba en 1988 en las actas de esta conferencia.

El Ensayo 13, «La externización de la epistemología», lo presenté en una versión previa en un encuentro de la SADAF [Sociedad Argentina de Análisis Filosófico] en Buenos Aires en 1989. Al año siguiente lo presenté en una conferencia en Biel, Suiza, y se publicó en 1991 en *Dialectica*, 45 2-3, 191-202.

«Tres variedades de conocimiento», el Ensayo 14, lo escribí para ser presentado en febrero de 1991 como una conferencia en memoria de A. J. Ayer, patrocinada por el Royal Institute of Philosophy. Cuando llegó la fecha de la conferencia una tormenta de nieve inutilizó el sistema de transportes en el sur de Inglaterra hasta el punto de que la conferencia tuvo que ser cancelada. Se publicó después en otras conferencias en memoria del mismo filósofo en *A. J. Ayer Memorial Essays: Royal Institute of Philosophy Supplement*, 30, editado por A. Phillips Griffiths (Cambridge University Press, 1991). Fue presentado también, con algunas modificaciones, como una conferencia Heisenberg, en una de las caballerizas reformadas del palacio de Nymphenburg de Munich y fue publicado en alemán en *Merkur* con el título en inglés correspondiente al del presente libro.

Introducción

Los ensayos de este libro se ocupan de tres clases de conocimiento proposicional y las relaciones que mantienen entre sí. Todos nosotros tenemos conocimiento sobre nuestras propias mentes, conocimiento de los contenidos de otras mentes y conocimiento del entorno que compartimos. Las subsecciones del libro se titulan Subjetivo, Intersubjetivo y Objetivo. Estas palabras señalan diferencias reales. El conocimiento de primera persona se distingue porque podemos afirmar legítimamente que tenemos una clase de autoridad única respecto de lo que creemos, queremos, pretendemos y otras actitudes. No existe esta clase de autoridad respecto al conocimiento de segunda persona y el conocimiento del resto del mundo natural, pero difieren el uno del otro en que el conocimiento sobre otras mentes es normativo de una manera en que el segundo no lo es. Sin embargo, las tres variedades de conocimiento son objetivas en el sentido de que su verdad es independiente de que se crea que son verdad. Esto es evidente en los dos últimos casos, pero también puede mantenerse esto en el caso de las creencias acerca de nuestras propias creencias y otras actitudes: tales creencias pueden estar equivocadas. Todo nuestro conocimiento es también objetivo en el sentido de que en su mayor parte podría expresarse mediante conceptos que tienen un lugar en un esquema de cosas públicamente compartido.

El Ensayo 1, «La autoridad de primera persona», se cuestiona qué es lo que explica la suposición de que el hablante está en lo cierto cuando sinceramente atribuye a su presente yo

una creencia, un deseo o una intención, mientras que esta suposición no es apropiada cuando otros hacen atribuciones similares respecto a él. Se argumenta que las «soluciones» al problema de las otras mentes que solamente reafirman la asimetría dejan el camino abierto al escéptico. Se ofrece una nueva explicación de la autoridad de primera persona que encuentra el origen de esa autoridad en un rasgo necesario de la interpretación del habla.

El Ensayo 2, «Conocer nuestra propia mente», trata de una patente dificultad acerca de la autoridad de primera persona: ¿cómo podemos reconciliar el hecho de que los contenidos de nuestras mentes estén determinados en parte por factores externos que desconocemos con la afirmación de que conocemos los contenidos de nuestras mentes sin necesitar (normalmente) evidencias ni apelar a ellas? Defiendo que la respuesta depende, entre otras cosas, de renunciar a la idea de la existencia de «objetos ante la mente» que carecen de atributos que pueden permanecer ocultos al agente.

En el Ensayo 3, «El mito de lo subjetivo», intento aclarar qué se quiere decir al negar que tenemos objetos ante la mente cuando tenemos sensaciones o pensamos. La idea de que existe una distinción fundamental que se puede trazar entre la experiencia sin interpretar y una estructura de conceptos organizada se relaciona estrechamente con la supuesta dicotomía entre lo subjetivo y lo objetivo. Estos dualismos han dominado y definido los problemas de buena parte de la filosofía moderna, no solamente los problemas que trata la epistemología, sino también los de la filosofía del lenguaje y la filosofía de la mente. En este ensayo me enfrento a la visión de la mente que se presupone en tales dualismos y con ello a los propios dualismos y a las posiciones epistemológicas y metafísicas que se basan en ellos. Ataco la concepción según la cual lo subjetivo es el fundamento del conocimiento empírico objetivo; afirmo que el conocimiento empírico no tiene un fundamento epistemológico ni tampoco lo necesita.

El Ensayo 4, «¿Qué está presente en la mente?», continúa la temática de los dos últimos ensayos con nuevos ejemplos y con argumentos nuevamente formulados.

Algunos filósofos se han preguntado si hay alguna «cuestión objetiva» (*fact of the matter*) respecto de las actitudes proposicionales. El Ensayo 5, «El indeterminismo y el antirrealismo», intenta eliminar estas dudas. En particular se discute la afirmación de que si se acepta la tesis de la indeterminación de Quine, como yo lo hago, entonces se debe abandonar la autoridad de primera persona.

El Ensayo 6, «La irreducibilidad del concepto de yo», enfatiza los rasgos de nuestras creencias acerca de nuestras actitudes presentes que se mantienen tras renunciar al mito de lo subjetivo y sus objetos mentales. Estos rasgos incluyen, por supuesto, la autoridad especial que tienen estas creencias y el papel irreductible de las oraciones indécicas. Lo que nos relaciona a nosotros y a nuestro lenguaje con el mundo que nos rodea son los pensamientos que expresan tales oraciones. También analizo brevemente el hecho de que no haya nada a lo que apelar que esté por encima de nuestros propios patrones de racionalidad, tema que trato nuevamente en el Ensayo 14.

El Ensayo 7, «Animales racionales», es uno de los intentos que he hecho para especificar algunos de los ingredientes que constituyen la racionalidad. Por racionalidad entiendo todo aquello que implica pensamiento proposicional. En un ensayo anterior («Thought and Talk», el Ensayo 11 de *Inquiries into Truth and Interpretation*) defendí que hay una dependencia mutua entre el pensamiento y el lenguaje. Muchos lectores no resultaron convencidos. Aquí lo intento de nuevo siguiendo un rumbo diferente, el mismo que después desarrollo en los siguientes ensayos. Las consideraciones que aduzco a favor de una estrecha conexión entre el lenguaje y el pensamiento no constituyen una demostración ni una prueba. En parte dependen de lo que creo que conocemos sobre unas criaturas como nosotros.

En el Ensayo 8, «La segunda persona», me explico en la idea de que el lenguaje es necesariamente social. Defiendo que para tener pensamientos y, por ello, para decir algo cuando hablamos, es necesario comprender y ser comprendidos por una segunda persona. Si Wittgenstein mantuvo que el lenguaje es necesariamente social, la tesis central de este ensa-

yo es wittgensteiniana. Sin embargo, se niega que la comunicación requiera que una persona hable como lo hacen las otras. Más bien, la objetividad que el pensamiento y el lenguaje requieren depende de las respuestas mutuas y simultáneas de dos o más criaturas a los estímulos distales comunes y a las respuestas de uno a otro. Llamo «triangulación» a esta relación a tres bandas entre dos hablantes y un mundo que comparten. Al final, la idea es tan simple como la del aprendizaje por ostensión, pero insistiendo en que la triangulación no es una cuestión, de que una persona capte un significado que está ya ahí, sino de una actuación que (cuando se detalla completamente) confiere un contenido al lenguaje. Esta tesis y sus ramificaciones para la filosofía de la mente, el lenguaje y la epistemología, reaparecen una y otra vez en mis trabajos posteriores a 1982 (Ensayo 7). Carol Rovane, Akeel Bilgrami y Marcia Cavell fueron críticos tempranos de esta idea. Sus sugerencias y dudas sinceras me ayudaron en gran medida a dar forma a mi pensamiento.

En el Ensayo 9, «La aparición del pensamiento», me pregunto cómo podemos describir la transición entre la mente prelingüística y preconceptual de un bebé a la de un niño con lenguaje, creencias y otras actitudes proposicionales. Defiendo que no tenemos vocabulario para describir los primeros estadios de tal desarrollo y que no está claro que podamos satisfacer la necesidad que sentimos de un vocabulario tal. Se sugiere que pueden distinguirse unos cuantos pasos importantes mediante la consideración del contenido de las teorías semánticas que se requieren para diversas lenguas o lenguajes.

Escribí el Ensayo 10, «Una teoría coherentista de la verdad y el conocimiento», en 1981 antes de los primeros nueve ensayos de esta colección. No hay ningún artículo de los que haya escrito que me gustaría más rehacer. Comprensiblemente ha sido objeto de muchas críticas y por esto lo reimprimo aquí sin ningún cambio. En cualquier caso lo he estado rehaciendo desde que lo escribí; los nueve ensayos que lo preceden evidencian parcialmente mis recelos posteriores y el Ensayo 14 es un testimonio añadido. También he intentado enmendar los errores en las réplicas a varias críticas en *The Philosophy of Donald Davidson* (en Library of Living Philosophers, editado por

Lewis E. Hahn, Chicago, Open Court, 1999) y en las «Replies» a Barry Stroud, John McDowell y Richard Rorty en *Crítica* 30 (1998). Lo que más me gustaría corregir es la impresión de que creo que la experiencia y la percepción no tienen ningún papel en nuestras creencias sobre el mundo. «Experiencia» y «percepción» son palabras perfectamente adecuadas para aludir a aquello, sea lo que sea que pasa en nuestras mentes cuando miramos a nuestro alrededor, olemos, tocamos, oímos y degustamos. Tenía tal ansia para que se captase la idea (la cual debería haber dicho que se remonta a Wilfrid Sellars) de que es un error postular intermediarios epistémicos entre el mundo y nuestras creencias que provoqué que muchos lectores pensasen que rechazaba todo intercambio significativo entre el mundo y la mente. De hecho mi tesis entonces como ahora es que la conexión es causal y que, en el caso de la percepción, es directa. El percibir que está nevando consiste, en las circunstancias adecuadas, en que los propios sentidos causen (del modo apropiado) la creencia de que está nevando cuando realmente la nieve está cayendo. No hay duda de que las sensaciones desempeñan un papel, pero ese papel no es el de proporcionar evidencias para las creencias.

El Ensayo 11, «El contenido empírico», ofrece un trasfondo histórico y un comentario sobre el tema del Ensayo 10. Aunque Neurath y Schlick estaban más enredados en un viejo debate de lo que ellos mismos se daban cuenta, la idea de que ellos estaban salvando la filosofía de su pasado metafísico dio un nuevo empuje y un giro lingüístico a un problema muy tratado.

«Epistemología y verdad», el Ensayo 12, analiza la relación de la epistemología con la verdad. Frecuentemente se observan dos posiciones contrapuestas: que la verdad es «radicalmente no epistémica» (en palabras de Putnam) o que se debe entender en términos de lo que nos es posible conocer (en la práctica, en la teoría o idealmente). Defiendo que ninguna de las dos alternativas es aceptable. La verdad no se puede limitar a lo que podemos o podríamos determinar como verdadero; no obstante, hay razones sólidas para conectar la verdad con la creencia verdadera de una manera u otra. Esbozo una manera de reconciliar parcialmente las dos posiciones basándome en las ideas que se exploran en el Ensayo 10.

Desde los tiempos de Descartes la mayor parte de la epistemología se ha basado en el conocimiento de primera persona. Según la idea usual debemos empezar con aquello de lo que poseemos mayor certeza: el conocimiento de nuestras propias sensaciones y pensamientos. De una manera u otra vamos progresando, si podemos, hacia el conocimiento del mundo externo objetivo. Al final podemos dar un débil paso hacia el conocimiento de las otras mentes. En el Ensayo 13, «La externización de la epistemología», defiendo una revisión total de esta concepción. Todo pensamiento proposicional, sea éste positivo o escéptico, sea sobre lo interno o lo externo, requiere la posesión del concepto de verdad objetiva, y este concepto es accesible solamente a aquellas criaturas que se comunican con otras. Por tanto, el conocimiento de tercera persona —el conocimiento de otras mentes— es esencial para cualquier otro conocimiento, pero el conocimiento de tercera persona es imposible sin el conocimiento de un mundo compartido de objetos en momentos de tiempo y en espacios compartidos. Por ello, la adquisición de conocimiento no se basa en la progresión desde lo subjetivo a lo objetivo; emerge holísticamente y es interpersonal desde el principio. Examino varias formas de externismo que encuentro insatisfactorias. Defiendo que la triangulación, que figura en muchos otros ensayos de este libro, corrige y aumenta una versión del externismo perceptual y una versión del externismo social.

«Tres variedades de conocimiento», el Ensayo 14, es el que más se aproxima a la presentación de una manera unificada de las principales ideas de este libro. Si todos los ensayos se hubiesen escrito después de que mis pensamientos se hubieran solidificado, «Tres variedades de conocimiento» debería aparecer el primero. El lector que desee leer antes una panorámica de mis ideas puede empezar por aquí.

En los ensayos de este libro he tratado de mejorar la gramática o el estilo y en ocasiones eliminar repeticiones o lo que ahora considero que son errores menores. Estoy seguro de que persisten muchos más errores y de que es probable que se haya de trabajar más en ciertas tesis. Cuando consideré la po-

sibilidad de reescribir los ensayos me di cuenta de que o tenía que reproducir mis trabajos pasados casi como estaban o empezar totalmente de nuevo y esto me llevaría años. Me consuelo con el pensamiento de que un nuevo comienzo significaría borrar lo que viene a ser una historia de mis intentos por enfrentarme al problema del triángulo que componen una persona, su sociedad y el entorno que comparten. Con ello también privaría a mis críticos de algunos de sus blancos preferidos.

De entre los críticos estoy especialmente agradecido a Richard Rorty, quien durante años ha insistido en que reuniese y publicase estos ensayos. Ernest Lepore concedió generosamente una semana de su tiempo a clasificar y poner orden a los ensayos que aparecen en este volumen y a los ensayos que aparecen en los dos volúmenes siguientes. Estoy en deuda con Ariela Lazar, que anteriormente me había dado sus consejos cualificados para organizar mis trabajos, y a Arpy Khatchirian por corregir errores ortográficos y gramaticales y pensar en muchos de mis ensayos. Peter Momtchiloff, de Oxford University Press, ha sido un editor alentador, amable y que me ha disculpado a menudo; su ayuda ha convertido un trabajo que consideraba desagradable en una tarea mucho más tolerable de lo que yo hubiera imaginado.

Cierto número de personas, invitaciones para dar cursos, universidades y otras instituciones me han proporcionado oportunidades para exponer muchas de las ideas que se formulan en estos ensayos. He disfrutado de intercambios siempre bienvenidos de mis alumnos en Berkeley y de los oyentes que asistieron a series de charlas que ofrecí en México (1992), Roma (1993), Munich (Kant Lectures, 1993), Girona (Ferrater Mora Lectures, 1994), Leuven (Francqui Lectures, 1994) y Buenos Aires (1995). Finalmente, debo mencionar mis Jean Nicod Lectures (1995) que ofrecí en Caen y París y a partir de las cuales espero desarrollar de manera más unificada y detallada los pensamientos que se encuentran dispersos en este volumen.

Estoy en deuda con todos aquellos que les añadieron algo o les quitaron algo a mis pensamientos, o bien los modificaron de otros modos. No llevé la cuenta de quienes me hicie-

ron indicaciones especialmente incisivas, así que cualquier lista que haga estará vergonzosamente incompleta. Lo que es seguro es que debo incluir en ella a Rosario Egidi, Pascal Engel, Dagfinn Føllesdal, Olbeth Hansberg, Dieter Henrich, Pierre Jacob, Carol Rovane y aquellos que asistieron a las diez conferencias y seminarios que ofrecí en Girona, de manera particular a W. V. Quine, Burton Dreben, Akeel Bilgrami, Ernest Lepore, Barry Stroud y Bruce Vermazen. Marcia Cavell no solamente asistió a muchas de mis charlas en las que ponía a prueba nuevas versiones de mis ideas, sino que durante estos años ha sido una compañera intelectual que de manera muy cuidadosa ha tratado de atemperar mis especulaciones de sillón con una perspectiva más empíricamente orientada y más psicoanalíticamente educada.

Lo subjetivo

1

La autoridad de primera persona

Cuando un hablante afirma que tiene una creencia, una esperanza, un deseo o una intención, 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 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?

La misma afirmación y la misma pregunta se pueden formular para el lenguaje o para la epistemología, pues si alguien puede hablar con una autoridad especial, el rango del propio conocimiento debe estar de alguna manera en consonancia y si el propio conocimiento muestra alguna diferencia sistemática, las afirmaciones de conocimiento deben reflejar esta diferencia. Así pues, doy por supuesto que si se puede explicar la autoridad de primera persona en el habla habremos hecho mucho, si no es que todo, de lo que necesitamos para caracterizar una teoría de los hechos epistemológicos.

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 planteo el primer problema existen dos diferencias importantes. La autoridad de primera persona es el problema más restringido puesto que lo trataré sólo en tanto que se aplique a actitudes proposicionales como la

creencia, el deseo y la intención; estar complacido, admirado, asustado u orgulloso de que algo sea el caso; o saber, recordar, notar o percibir que algo es el caso. Pero no analizaré lo que a menudo se considera que es central en el problema de las otras mentes: los dolores y otras sensaciones y el conocimiento, la memoria, la atención y la percepción tanto dirigidas a objetos como a personas, calles, ciudades, cometas y otras entidades no-proposicionales. Me parece que lo que se mantenga para las actitudes proposicionales ha de ser relevante para las sensaciones y lo restante, pero no voy a explorar aquí esta conexión.

Todas las actitudes proposicionales exhiben 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 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 cuestión de si alguien trata de cerrar la puerta girando una llave depende en parte de si esa persona quiere cerrarla y cree que al girar la llave la cerrará; y si esta creencia y este deseo han causado, de la manera apropiada, un deseo de girar la llave. 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 consideraré aquí es el caso general. Dado que casi en cada ejemplo, si no en todos, la autoridad de primera persona recae al menos parcialmente en un componente de creencia, me concentraré en el caso de la creencia.

Aunque haya autoridad de primera persona respecto a creencias y otras actitudes proposicionales, el error es posible;

esto es una consecuencia del hecho de que las actitudes son disposiciones que se manifiestan de diversas maneras y en un intervalo 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 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 preguntar a la persona que se autoatribuye algo por qué cree que tiene las creencias, los deseos o las intenciones que afirma que tiene. Wittgenstein señaló este rasgo de las autoatribuciones: «¿Cuál es el criterio 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»¹. La mayoría de los filósofos han seguido a Wittgenstein en esto y, tal y como veremos, han extendido este criterio a las actitudes proposicionales.

Este rasgo de la autoridad de primera persona, aunque pueda ser sugerente, no ayuda a explicar esa autoridad. Esto se debe parcialmente a las siguientes matizaciones: «normalmente» no hacemos autoatribuciones sobre la base de evidencias, pero algunas veces lo hacemos; «usualmente» no tiene sentido preguntarle a alguien por qué cree que tiene una cierta creencia, deseo o intención, pero a veces lo tiene. Sin embargo, incluso en casos excepcionales la autoridad de primera persona persiste; incluso cuando una autoatribución es dudosa, o es apropiado el cuestionarla, la persona que tiene la actitud habla acerca de ella con especial fuerza.

Sin embargo, la existencia de excepciones no es la razón principal por la que la autoridad de primera persona no resulte ser explicada mediante el hecho de que las autoatribuciones no se basan en evidencias; la razón principal es simple-

¹ Ludwig Wittgenstein, *Philosophical Investigations*, § 377.

mente 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 más probablemente correctas.

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 explicar el autoconocimiento sobre la base de la introspección y es fácil ver el porqué, 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.

Pocos filósofos han negado la existencia de la asimetría; Ryle es un tenaz ejemplo de la negación. En *El concepto de lo mental*, Ryle 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 que las maneras que tiene John Doe de saber algo acerca de Richard Roe». 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 del 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 muy buena posición para saber aquello que el oyente frecuentemente está en muy mala posición para saber. Los giros que toma una conversación de un hombre no sorprenden o dejan perpleja a su mujer tanto como sorprendieron y desorientaron a su prometida, ni tampoco los profesores que son colegas cercanos tienen que hacerse entender unos a otros tanto como tienen que hacerse entender a sus nuevos alumnos².

² Gilbert Ryle, *The Concept of Mind*, 156, 179.

Estoy de acuerdo con Ryle en que cualquier intento de explicar la asimetría entre las afirmaciones en 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. Cualquier explicación de ese tipo debe aceptar la asimetría, pero no puede a su vez explicarla. Sin embargo, Ryle ni acepta ni explica la asimetría; simplemente niega que exista. Puesto que creo que es obvio que existe la asimetría, pienso que es un error sostener a partir de la ausencia de una manera especial de conocer o un modo o clase de conocimiento especial, la ausencia de una autoridad especial; en vez de ello, lo que debemos hacer es buscar otro origen de la asimetría.

En cierto momento Ayer tomó una ruta similar a la de Ryle. En *The Concept of a Person* (*El concepto de persona*) enfatiza que las adscripciones de primera persona pueden ser erróneas y admite que estas adscripciones son privilegiadas³. No obstante, cuando quiere describir la autoridad de las autoadscripciones, las compara con la autoridad que le otorgamos a un testigo presencial cuando se le compara con los informes de segunda mano. Esta analogía me parece insatisfactoria por dos razones. La primera es que no nos dice por qué una persona es como un testigo presencial respecto a sus propios estados y acaecimientos 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 ve superada en determinados casos: se desacredita a un testigo presencial y las evidencias que ofrece no se toman en cuenta si es un observador que 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 ac-

³ A. J. Ayer, «Privacy».

titudes incluso cuando su afirmación se pone en cuestión o resulta refutada.

Joseph Agassi ha mantenido que conocemos los estados y acaecimientos 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)⁴.

Con la excepción del caso de Freud, hay poco en este texto que amenace la autoridad de primera persona. Freud, al extender los conceptos de intención, creencia, deseo y lo demás 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 que 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 creo que la existencia de las actitudes in-

⁴ Joseph Agassi, *Science in Flux*, 120.

conscientes amenace la importancia de la autoridad de primera persona.

A continuación voy a tratar de 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 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 sujeto⁵.

No se puede considerar que ésta 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 en preguntar cómo sabemos que la

⁵ Peter Strawson, *Individuals*, 108. (Anita Avramides ha señalado que Strawson ha contribuido más a esta problemática de lo que yo le concedo. Véase su «Davidson and the New Sceptical Problem».)

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. (Sobre la aparente ambigüedad de los predicados mentales, véase el Ensayo 2.)

Richard Rorty ha intentado ofrecer una tal explicación. Se nos pide que imaginemos que originalmente las autoadcripciones se hicieran sobre la base de la misma clase de observaciones o evidencias de comportamiento que las adcripciones 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 autoadcripciones ofrecían mejores explicaciones del comportamiento que las adcripciones de tercera persona. De ese modo, tratar las autoadcripciones 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»⁶.

Esta teoría no pretende que se la tome seriamente como un fragmento de antropología de sentido común, sino que lo que pretende es que parezca razonable tratar las autoadcripciones como algo que tiene una autoridad especial. Pero la cuestión sigue ahí: ¿cuál es la razón que ofrece Rorty para mostrar que las autoadcripciones que no se basan en evidencias tienen que ver con los mismos estados y acaecimientos que las adcripciones 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 adcripciones y la manera en que explican el comportamiento es también diferente. Lo que Rorty describe

⁶ Richard Rorty, «Incorrigibility as the Mark of the Mental», 416. La teoría de Rorty se deriva de la de Wilfred Sellars, «Empiricism and the Philosophy of Mind». Con respecto al tema que trato ahora, la teoría de Sellars no difiere de la de Rorty.

como el descubrimiento de que las autoadcripciones 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 he considerado, en realidad no han tratado el viejo problema del escepticismo respecto el conocimiento de las otras mentes, pero creo que tiene una fácil explicación. 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 autoadcripciones y las adcripciones a otros. Podemos aún preguntar por qué creemos que estos dos tipos de adcripciones conciernen a lo mismo. Ésta es una buena pregunta reconozcamos o no su tradicional estirpe escéptica.

Quizás debiera señalarse que ningún concepto aparte de aquellos que se aplican a sensaciones, actitudes proposicionales y posiciones de nuestras extremidades muestran la clase de asimetría que estamos analizando. Muchos conceptos pueden aplicarse sobre la base de múltiples criterios, pero no hay otros en los que quienes realizan la adscripción *deban*, en determinadas ocasiones, utilizar criterios distintos. 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.

El primer paso hacia una solución depende de poner en claro cuáles son las entidades a las que se aplica la autoridad de primera persona. William Alston propone el siguiente principio para caracterizar el estatuto especial de las autoatribuciones: «Cada persona está relacionada con las proposicio-

nes que le adscriben a uno mismo estados mentales presentes de un modo tal que es lógicamente imposible que esa persona crea que tal proposición es verdadera y no esté justificado en tener esta creencia, mientras que nadie más está relacionado de esta manera con tales proposiciones»⁷.

Para que esta idea sea plausible, hemos de suponer que la proposición que Jones expresa con la oración «creo que Wagner murió feliz» es la misma proposición que la proposición que Smith expresa con la oración «Jones cree que Wagner murió feliz». Es evidente que ésta es una suposición muy cuestionable. Una vez más queda sin explicar el contraste epistémico, y, en ausencia de una explicación, se plantea la cuestión de qué *razón* tiene alguien en un caso particular para creer que la proposición que Jones y Smith consideran *es* la misma. Cuando sólo se ofrece solamente una descripción de la diferencia epistémica, la conclusión natural es que esas proposiciones son distintas.

Paso, pues, a una formulación de Sidney Shoemaker que hace una mención explícita del lenguaje: «Entre los enunciados incorregibles encontramos enunciados acerca de... acaecimientos 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»⁸.

Dejaremos a un lado la condición de incorregibilidad y la sustituiremos por algo menos exigente —algo que equivalga a la autoridad de primera persona—. (Es posible que hacer esto sea razonable ya que a Shoemaker le interesan principalmente sensaciones como el dolor mientras que yo trato exclusivamente las actitudes proposicionales.) Lo que es importante 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 autori-

⁷ William Alston, «Varieties of Privileged Access», 235.

⁸ Sidney Shoemaker, *Self-Knowledge and Self-Identity*, 215-16.

dad de primera persona si la clase de esas preferencias pudiera especificarse en términos sintácticos. Desafortunadamente esto no puede hacerse. Si Shoemaker está en lo cierto, debe suponerse que un hablante que usa sinceramente un cierto tipo de oración está en lo cierto respecto a lo que dice. Pero por supuesto esto solamente se mantiene si es que el hablante sabe que está usando este tipo de oración privilegiada; si no lo está, entonces está malemployando el lenguaje. ¿Qué es lo que constituiría un mal uso en estos casos? Ante todo se diría que lo es aseverar sinceramente una oración sobre la cual no se tiene una especial autoridad para hacerlo. Quizás sea así; pero esto solamente reitera la afirmación no-informativa y no explicada de que es una convención del lenguaje tratar las autoadscripciones con un respeto especial. Visto desde el punto de vista del intérprete, ello implica que habríamos de interpretar las autoadscripciones de tal manera que las hiciésemos verdaderas, o asignar una prioridad especial a su verdad. El punto de vista del intérprete es el único que podemos adoptar, según el principio de Shoemaker, y esto implica que el principio tenga una aplicación independiente: la única razón que tenemos para decir que el hablante tiene una autoridad especial en una ocasión es que estamos preparados para tratar su oración como una autoadscripción. En otras palabras, las autoadscripciones gozan de una autoridad especial, lo cual es verdad, pero es ahí donde empezamos.

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 las proposiciones o los significados es un paso hacia una nueva dirección que promete. La razón de ello es relativamente simple. En tanto que 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 (o una oración con una determinada interpretación), estamos constreñidos a dar cuenta de las diferencias simplemente postulando diferentes tipos de fuentes de información. De manera alternativa, podemos postular distintos criterios de aplicación para los conceptos o palabras clave («cree que», «preten-

de», «desea que», etc.). Pero estos pasos, tal y como hemos visto, no hacen más que volver a plantear el problema y con ello invitan al escepticismo acerca del conocimiento de las mentes de los demás (o de nuestra propia mente). Sin embargo, si planteamos el problema en términos de las relaciones entre agentes y preferencias podemos evitar el punto muerto.

En este punto se hace necesario distinguir dos asimetrías diferentes pero que se encuentran relacionadas. Por una parte, 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. Si esto que sostenemos lo ponemos en palabras, se nos presenta la dificultad de decidir qué par de preferencias están relacionadas de manera adecuada para garantizar que esto que sostenemos tiene el «mismo contenido». 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 supuesto que estas dos asimetrías están conectadas desde el momento en que estamos inclinados a decir que la justificación que usted posee para pensar que digo algo verdadero al decir «creo que Wagner murió feliz» debe estar estrechamente relacionada con la justificación que usted posee para pensar que usted diría algo verdadero si dijera «Davidson cree que Wagner murió feliz». Pero por razones que se harán pronto evidentes debemos tratar la segunda versión de la asimetría.

La cuestión se convierte entonces 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 mismo 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 (o la palabra «cree»). Consideremos pues simplemente una preferencia mía más corta: profiero entonces la oración «Wagner murió feliz». Claramente si

usted o yo o quien sea sabe que pienso que esta oración es verdadera en la presente ocasión de proferencia y sabe lo que he pretendido decir con esta oración en esta ocasión de proferencia entonces sabe lo que creo —qué creencia he expresado.

Una vez más sería circular el tratar de explicar la asimetría básica asumiendo una asimetría en la seguridad que usted y yo tenemos de que pienso que la oración que acabo de proferir es una oración verdadera. Por descontado que debe de haber tal asimetría, pero no puede suponerse que contribuya a la explicación deseada. Sin embargo, podemos aceptar sin prejuicio que ambos sabemos, sea cual sea el origen o naturaleza de nuestro conocimiento, que en esa ocasión pienso que la oración que he proferido es verdadera. De igual manera, sería una petición de principio el explicar la asimetría básica apelando a alguna asimetría en nuestro conocimiento del hecho de que sé lo que significaba mi oración tal y como se ha proferido en esa ocasión. Así pues, vamos simplemente a aceptar que ambos sabemos esto, sea cual sea la fuente o el carácter de nuestro conocimiento.

Hasta ahora, por tanto, no hemos postulado o aceptado ninguna asimetría. Los supuestos son exactamente éstos: 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 proferencia 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. Esa presunción le es esencial a la naturaleza de la interpretación —el proceso por el cual comprendemos las proferencias de un hablante—. Este proceso no puede ser el mismo para el hablante que para sus oyentes.

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; sin embargo, 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 tanto 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.

El contraste entre el fundamento que un autoadscriptor tiene para su autoadcripción y el fundamento que tiene un intérprete para aceptar esa misma adscripción serían sólidos si asumiésemos que no hay nada que preguntarse acerca de la interpretación que un hablante hace de sus propias palabras. Pero, por supuesto, hay cosas que cuestionar, ya que lo que sus palabras significan depende en parte de las pistas para la interpretación que aquél le ha ofrecido al intérprete, o de otras evidencias que crea justificadamente que el intérprete posee. 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 preferencia de la oración “Wagner murió feliz” es verda-

dera 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 imaginando una situación en la que a dos personas que hablan lenguas no emparentadas y no conocen otras lenguas 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 poseen a su disposición. Sin embargo, esto no afecta al punto que deseo enfatizar, puesto que lo que el intérprete que he imaginado puede tratar conscientemente como evidencias son exactamente lo que condiciona a quien aprende una lengua por primera vez para llegar a ser un usuario del lenguaje. Imaginemos ahora que una de las personas del par habla mientras que la otra trata de comprenderlo. No importa si el hablante habla su lengua «nativa», ya que su situación social pasada es irrelevante. (Estoy suponiendo que el hablante no tiene interés en enseñar al oyente a habérselas con la comunidad de habla original del hablante.) 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 en qué apoyarse sino la configuración de sonidos que el hablante muestra en conjunción con otros acaecimientos (incluyendo, por supuesto, acciones posteriores por parte de ambos, el hablante y el intérprete). 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. Dado que la «lengua» que está hablando no tiene otros oyentes, la idea de que el hablante está malemployando su lengua no se

aplica. 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.

2

Conocer nuestra propia mente

No hay ningún secreto sobre la naturaleza de las evidencias que utilizamos para decidir lo que otras personas piensan: observamos sus actos, leemos sus cartas, 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 formar una imagen convincente de una mente; sabemos cómo hacerlo sin saber necesariamente cómo lo hacemos. 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?”»¹. Robert Motherwell expresó un pensamiento similar: «Diría que los mejores pintores no saben lo que piensan hasta que lo pintan.»

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. Ryle estaba

¹ Graham Wallas, *The Art of Thought*.

equivocado. Rara vez sucede que necesitemos o apelemos a las 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 equivocado sobre 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 equivocado sobre los 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 yo lo veo, no tiene una fácil explicación. 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 (o deseamos, dudamos, pretendemos) sin necesidad de recurrir a las evidencias (incluso cuando están al 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 persona, ni tampoco una afirmación en primera persona pero en otro tiempo verbal. Sin embargo, reconocer este hecho no significa explicarlo.

Desde Wittgenstein se ha convertido en rutina el intentar aliviar las preocupaciones sobre «el conocimiento de las otras mentes» señalando que un aspecto esencial de nuestro uso de ciertos predicados mentales es que los aplicamos a los demás sobre la base de evidencias conductuales, pero a nosotros mismos los aplicamos sin esta ayuda. Esta observación es verdadera y, cuando se elabora apropiadamente, debería responder a aquellos que se preguntan cómo puede ser que conozcamos

las mentes de los demás. Pero como respuesta al escéptico, la idea de Wittgenstein (si es que es de Wittgenstein) sería poco satisfactoria. En primer lugar, es una idea extraña el que afirmaciones que se hacen sin el apoyo de evidencias o de observaciones hayan de resultar favorecidas respecto de afirmaciones que poseen este tipo de apoyo. Por supuesto, si no se citan evidencias en apoyo de una afirmación, esa afirmación no puede ser impugnada cuestionando la verdad o la relevancia de esas evidencias. Pero decir esto no basta para mantener que las afirmaciones generales que carecen del apoyo de las evidencias sean más fiables que las que lo tienen. La segunda y principal dificultad es la siguiente. Se diría que normalmente lo que se consideran evidencias para la aplicación de un concepto ayuda a definir el concepto, o al menos restringe su identificación. 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 de ninguno), la conclusión obvia parece que sería que la expresión es ambigua. ¿Por qué entonces suponer que no es ambiguo un predicado como «*x* cree que Ras Dashan es la montaña más alta de Etiopía», que a veces se aplica sobre la base de evidencias conductuales y a veces no? Si es ambiguo, 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 (como deberíamos hacerlo) que el carácter necesariamente público e interpersonal del lenguaje garantiza que, de manera frecuente, aplicamos correctamente 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* piensa. La respuesta de estilo wittgensteiniano puede solucionar el problema de las otras mentes, pero crea el correspondiente 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 enfrentamos se puede especificar del siguiente modo: sé donde buscar para atribuir pensamientos a los demás. Me aplico a mí mismo 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* lo que me estoy atribuyendo a mí mismo. Pero dado que las evidencias de que me sirvo en el caso de los demás son públicas, no hay ninguna razón por la que no debiera atribuirme pensamientos a mí mismo del mismo modo en que lo hago a los demás, al modo de Graham Wallas, Robert Motherwell y Gilbert Ryle. En otras palabras, aunque no lo hago, podría tratar mis propios estados mentales de la misma manera en que trato a los de los demás. Quien busca el mismo tipo de autoridad con respecto a los pensamientos de los demás que la que aparentemente tiene al habérselas con sus propios pensamientos no puede servirse de esta estrategia. Así pues, la asimetría entre los casos sigue siendo un problema y es la autoridad de primera persona la que crea ese problema.

He propuesto una respuesta a este problema en el primer ensayo de este libro. Allí he argumentado que la atención a cómo atribuimos pensamientos y significados a los demás explicaría la autoridad de primera persona sin que ello conlleve dudas escépticas. Sin embargo, recientemente algunos de los mismos hechos acerca de la atribución de actitudes en los cuales me apoyé para defender la autoridad de primera persona se han utilizado para atacar esa autoridad; sobre la base de lo que se cree que son nuevas razones, se ha argumentado que mientras que los métodos del intérprete en tercera persona determinan lo que usualmente consideramos que son los contenidos de la mente de un agente, los contenidos así determinados pueden serle desconocidos al agente. En el presente ensayo tengo en cuenta algunos de estos argumentos y defiendo que éstos no constituyen una auténtica amenaza para la autoridad de primera persona. La explicación que ofrecí en mi ensayo anterior de la asimetría entre las atribuciones de actitudes en primera persona y en las otras personas me parece que, en todo caso, resulta reforzada por las nuevas consideraciones o por aquellas que parecen válidas.

Debo remarcar de nuevo que el problema que trato no requiere que nuestras creencias acerca de nuestros propios estados mentales presentes sean infalibles o incorregibles. Podemos cometer, y de hecho cometemos, errores acerca de lo que creemos, deseamos, aprobamos y pretendemos; también existe la posibilidad del autoengaño. Pero tales casos, aunque no sean infrecuentes, no son ni pueden ser la norma; no voy a argumentar esto ahora, pero tomémoslo como uno de los hechos a explicar.

Dejando, pues, a un lado el autoengaño y otros fenómenos anómalos o limítrofes, la cuestión es si podemos, sin irracionalidad, inconsistencia o confusión, pensar simple y llanamente que tenemos una creencia que no tenemos o pensar que no tenemos una creencia que tenemos. Algunos filósofos y psicólogos con mentalidad filosófica han mantenido opiniones que comportan o sugieren que esto podría ocurrir fácilmente, incluso que debe ocurrir todo el tiempo.

La amenaza estaba ya en la idea de Russell sobre las proposiciones que se podía saber que eran verdaderas aunque contuvieran «ingredientes» con los que la mente del sujeto no estuviera familiarizada; y a medida que fue evolucionando el estudio de las actitudes *de re*, ese peligro se hizo mayor.

Pero fue Hillary Putnam quien apagó el interruptor, quien cortó la corriente. Consideremos el argumento de Putnam de 1975 para mostrar que los significados, dicho en sus propias palabras, «simplemente no están en la cabeza»²... Putnam arguye persuasivamente que lo que las palabras significan depende de algo más que de lo que «está en la cabeza». Explica varios casos imaginarios que tienen como moraleja que los aspectos de la historia natural de cómo alguien aprende el uso de una palabra necesariamente marcan una diferencia en lo que la palabra significa. Parece que se sigue que dos personas podrían estar en estados físicos idénticos y aún así significar cosas distintas mediante las mismas palabras.

Las consecuencias son de largo alcance puesto que si la gente puede (generalmente) expresar correctamente sus pen-

² Hilary Putnam, «The Meaning of "Meaning"», 227. [Pág. 144 de la traducción al castellano reseñada en la bibliografía.]

samientos en palabras, entonces sus pensamientos —sus creencias, deseos, intenciones, esperanzas, expectativas— también en parte deben identificarse mediante acaecimientos y objetos externos a la persona. Si los significados no están en la cabeza, entonces parecería que tampoco lo están las creencias, los deseos y demás.

Dado que se puede estar un poco cansado del caso expuesto por Putnam del Doppelgänger de la Tierra Gemela, déjenme que explique mi propia historia de ciencia-ficción, si es que lo es. Mi historia elude algunas dificultades que resultan ser irrelevantes en la de Putnam, aunque lleva consigo algunos problemas nuevos³. (Volveré sobre el tema de la Tierra y la Tierra Gemela un poco más adelante.) Supongamos que cae un rayo sobre un árbol muerto en un pantano y yo estoy cerca. Mi cuerpo se ve reducido a sus elementos mientras que por coincidencia (y a partir de moléculas distintas) el árbol se convierte en una copia física de mí mismo. Mi copia, el Hombre del Pantano, se comporta exactamente como yo lo hacía antes; de acuerdo con su naturaleza, sale del pantano, encuentra a mis amigos y parece reconocerlos, devolviéndoles en inglés sus saludos según todas las apariencias. Se traslada a mi casa y parece escribir artículos sobre la interpretación radical. Nadie puede notar cuál es la diferencia.

Sin embargo, *hay* una diferencia. Mi copia no puede reconocer a mis amigos; no puede *reconocer* nada ya que, para empezar, nunca conoció nada. No puede saber el nombre de mis amigos (aunque desde luego lo parezca); no puede recordar mi casa. Por ejemplo, con la palabra «casa» no puede querer decir lo mismo que yo, ya que el sonido «casa» que produce el Hombre del Pantano no fue aprendido en un contexto que le proporcionara el significado correcto —o algún tipo de significado—. Incluso no veo cómo puede decirse que mi co-

³ No pretendo aquí ser original; Stephen Stich ha utilizado un ejemplo muy parecido en «Autonomous Psychology and the Belief-Desire Thesis», 573 y ss. Debo remarcar que no estoy sugiriendo que un objeto creado de manera accidental o artificial no pueda pensar; el Swampman (Hombre del Pantano) simplemente necesita tiempo para adquirir una historia causal que diera sentido a la afirmación de que habla, recuerda, identifica o piensa sobre cosas en el mundo. (Volveré sobre esta cuestión más adelante.)

pia quiere decir algo mediante los sonidos que produce ni que tiene pensamientos.

Putnam quizás no apoyaría esta última afirmación puesto que afirma que si dos personas (u objetos) están de manera relevante en estados físicos similares sería «absurdo» pensar que sus estados psicológicos «difieren en algo de los míos»⁴. Sin embargo, sería un error asegurar que Putnam y yo estamos en desacuerdo en este punto ya que todavía no está claro cómo se usa la expresión «estado psicológico».

Putnam mantiene que muchos filósofos se han equivocado al suponer que los estados psicológicos como creer y conocer el significado de una palabra son (I) «internos» en el sentido de que no presuponen la existencia de otro individuo más que el sujeto al cual el estado se adscribe y que son (II) los propios estados que normalmente identificamos e individualizamos como lo hacemos con las creencias y las otras actitudes proposicionales. Dado que normalmente identificamos e individualizamos los estados mentales y los significados en parte en términos de relaciones con objetos y acaecimientos distintos del sujeto, Putnam cree que (I) y (II) resultan ser incompatibles: en su opinión no hay estados que puedan satisfacer ambas condiciones.

Putnam llama a los estados psicológicos que satisfacen la condición (I) «restringidos». Concibe tales estados como solipsistas y los asocia con la visión de Descartes sobre lo mental. Puede que Putnam considere estos estados como los únicos estados psicológicos «auténticos»; en buena parte de su artículo omite el calificativo de «restringido» a pesar de que los estados psicológicos restringidos (llamados así) no corresponden a las actitudes proposicionales tal y como normalmente se las identifica. No todo el mundo se ha convencido de que se puede establecer una distinción inteligible entre estados psicológicos restringidos (o internos, cartesianos o individualistas —todos estos términos son de uso corriente—) y estados psicológicos identificados en términos de factores externos (sociales u otros), si es que hay algunos. Así, John

⁴ Hilary Putnam, «The Meaning of "Meaning"», 227. [Pág. 144 de la traducción al castellano reseñada en la bibliografía.]

Searle ha afirmado que nuestras actitudes proposicionales comunes satisfacen la condición (I), y que de este modo no son necesarios estados que satisfagan la condición (II), mientras que Tyler Burge ha negado que haya, en ningún sentido interesante, actitudes proposicionales que satisfagan la condición (I)⁵. A pesar de todo, parece que hay un acuerdo universal en que no hay estados que satisfagan ambas condiciones.

La tesis de este ensayo es que no hay ninguna razón para suponer que los estados mentales comunes no satisfacen ambas condiciones (I) y (II): creo que tales estados son «inter-nos», en el sentido de que son idénticos a los estados del cuerpo, y que son identificables sin hacer referencia a objetos o acaecimientos externos al cuerpo; al mismo tiempo son «no individualistas» en el sentido de que, en parte, pueden ser identificados mediante sus relaciones causales con acaecimientos y objetos externos al sujeto de esos estados, y generalmente lo son. Un corolario de esta tesis resultará ser que, contrariamente a lo que frecuentemente se supone, la autoridad de primera persona puede aplicarse sin contradicción a estados que se identifican regularmente mediante sus relaciones con acaecimientos y objetos externos a la persona.

Comenzaré con el corolario. ¿Por qué es natural aceptar que estados que satisfacen la condición (II) pueden no serle conocidos a la persona que está en ellos?

Debo ahora hablar sobre la Tierra Gemela de Putnam. Nos invita a imaginar dos personas que son exactas físicamente y (por ello) iguales respecto a todos los estados psicológicos «restringidos». Una de esas dos personas, la que habita en la Tierra, ha aprendido a usar la palabra «agua» mostrándosele agua, leyendo y oyendo cosas sobre ella, etc. La otra, habitante de la Tierra Gemela, ha aprendido a usar la palabra «agua» bajo condiciones que no son diferentes mediante la observación, pero la sustancia a la cual ha sido expuesta no es agua, sino una sustancia que se le parece y que podemos llamar «gagua». En estas circunstancias, Putnam afirma que el primer hablante se refiere al agua cuando usa la palabra «agua»; su ge-

⁵ John Searle, *Intentionality*; y Tyler Burge, «Individualism and Psychology».

melo se refiere a gagua cuando usa la palabra «agua». Así pues, parece que tenemos un caso en el que los estados psicológicos «restringidos» son idénticos pero los hablantes significan cosas distintas mediante la misma palabra.

¿Qué sucede con los pensamientos de estos dos hablantes? El primero se dice a sí mismo, cuando tiene delante un vaso de agua, «aquí hay un vaso de agua»; el segundo murmura para sí mismo exactamente los mismos sonidos cuando tiene delante un vaso de gagua. Cada uno dice la verdad, ya que sus palabras significan cosas distintas. Y puesto que cada uno de ellos es sincero, es natural suponer que creen cosas distintas: el primero cree que hay un vaso de agua delante suyo, el segundo cree que hay un vaso de gagua delante *suyo*. Pero ¿saben lo que creen? Si los significados de sus palabras y, con ello, las creencias que expresan al usar esas palabras están parcialmente determinados mediante factores externos que los agentes ignoran, sus creencias y significados no son restringidos en el sentido de Putnam. No hay, por tanto, ninguna base por la que ninguno de los hablantes pueda decir en qué estado se encuentra, puesto que no hay claves internas o externas que hagan ver la diferencia. Parece que deberíamos concluir que ninguno de los hablantes sabe lo que quiere decir o lo que piensa. Un cierto número de filósofos han adoptado explícitamente esta conclusión, Putnam entre ellos. Putnam declara que «abandona totalmente la idea de que si hay una diferencia en el significado... entonces *debe* haber alguna diferencia en nuestros conceptos (o en nuestro estado psicológico)». Lo que determina el significado y la extensión «en general, el hablante no lo conoce totalmente»⁶. En este contexto «estado psicológico» significa estado psicológico *restringido*, y se supone que solamente estos estados se «conocen plenamente». Jerry Fodor cree que las actitudes proposicionales comunes están (dicho con bastante aproximación) «en la cabeza», pero está de acuerdo con Putnam en que *si* las actitudes proposicionales se identificaran parcialmente mediante factores externos al agente, no estarían en la cabeza y los agentes no ne-

⁶ Hilary Putnam, «The Meaning of “Meaning”», pág. 245. [Pág. 164 de la traducción al castellano reseñada en la bibliografía.]

cesariamente las conocerían⁷. Aunque sus razones no sean las de Fodor, también John Searle mantiene que los significados están en la cabeza («no hay ningún otro lugar en el que puedan estar»), pero parece aceptar la inferencia de que si éste no fuera el caso, la autoridad de primera persona se perdería⁸. Quizá la formulación más simple de esta posición aparece en la introducción de Andrew Woodfield a un libro de ensayos sobre los objetos del pensamiento. Al referirse a la afirmación de que los contenidos de la mente se determinan frecuentemente mediante factores externos a, y quizá desconocidos para, la persona de cuya mente se trata, afirma: «Dado que la relación externa no se determina subjetivamente, el sujeto no tiene autoridad sobre ello. Una tercera persona podría muy bien estar en mejor posición que el sujeto para saber sobre qué objeto está pensando el sujeto y por ello, estar en mejor posición de saber cuál era el pensamiento»⁹. Quienes aceptan la tesis de que los contenidos de las actitudes proposicionales se identifican parcialmente en términos de factores externos parece que tienen un problema similar al problema del escéptico que piensa que podemos estar completamente equivocados sobre el mundo «externo». En el presente caso, el escepticismo común sobre los sentidos se evita suponiendo que el propio mundo más o menos determina correctamente los contenidos de los pensamientos acerca del mundo. (El hablante que piensa que es agua está probablemente en lo cierto puesto que aprendió el uso de la palabra «agua» en un entorno con agua; el hablante que piensa que es gagua está probablemente en lo cierto puesto que aprendió el uso de la palabra «agua» en un entorno con gagua.) Pero no se ha vencido al escepticismo; sólo se lo ha desplazado hacia el conocimiento de nuestras propias mentes. Nuestras creencias comunes sobre el mundo externo están (según esta concepción) dirigidas hacia el mundo, pero no sabemos lo que creemos.

⁷ Jerry Fodor, «Cognitive Science and the Twin Earth Problem», 103. Véase también su «Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology», pág. viii.

⁸ John Searle, *Intentionality*, cap. 8.

⁹ Andrew Woodfield, «Introduction», en Woodfield (ed.), *Thought and Object*, pág. viii.

Entre agua y gagua existe, por supuesto, una diferencia que puede descubrirse mediante medios normales, se descubra finalmente o no. Así pues, una persona podría averiguar lo que cree descubriendo la diferencia entre agua y gagua y averiguando lo bastante sobre sus propias relaciones con ambas para determinar de cuál de las dos trata su habla y sus creencias. La conclusión escéptica que parece que hemos alcanzado tiene que ver con el alcance de la autoridad de primera persona: éste es mucho más limitado de lo que suponíamos. La mayoría de nuestras creencias sobre el mundo son verdaderas, pero fácilmente podemos equivocarnos sobre lo que pensamos. Ésta es una imagen traspuesta del escepticismo cartesiano.

Quienes sostienen que los contenidos de nuestros pensamientos y los significados de nuestras palabras frecuentemente los fijan factores que ignoramos no se han interesado mucho en la aparente consecuencia de sus puntos de vista en la que he estado insistiendo. Por supuesto que se han dado cuenta de que, si estuvieran en lo cierto, tanto la idea cartesiana de que la única cosa sobre la que podemos tener certeza es de los contenidos de nuestras propias mentes, como la noción fregeana de los significados que se «captan» plenamente, deberían estar equivocadas. Pero, por lo que sé, no han hecho muchos intentos de resolver el aparente conflicto entre sus puntos de vista y la fuerte intuición de que existe la autoridad de primera persona.

Una de las razones para esta falta de interés puede ser que parece que algunos ven el problema como si estuviera confinado a un ámbito limitado de casos, casos en los que los conceptos o las palabras se fijan a los objetos que se discernen o a los que uno se refiere usando nombres propios, indécicos y palabras de clases naturales. Sin embargo, otros defienden que los vínculos entre el lenguaje y el pensamiento, de una parte, y las cuestiones externas, por la otra, están tan omnipresentes que no queda intacto ningún aspecto del pensamiento tal y como usualmente se lo concibe. A este respecto Daniel Dennett señala que «se debe estar muy informado sobre, e íntimamente conectado con, el mundo, con quienes lo integran y sus propiedades, para poder decir con propiedad que

tenemos creencias»¹⁰. Prosigue afirmando que los factores externos no-subjetivos que se reconoce que operan en la clase de caso que se ha discutido impregnan la identificación de *todas* las creencias. Burge también enfatiza la medida en que nuestras creencias se ven afectadas por factores externos, aunque, por razones que no explica, aparentemente no piensa que esto sea una amenaza para la autoridad de primera persona¹¹.

El tema ha adquirido un giro inquietante. En una época se recurrió al conductismo para mostrar cómo era posible que una persona conociera la mente de otra; luego se rechazó el conductismo, en parte porque no podía explicar uno de los aspectos más evidentes de los estados mentales: el hecho de que, en general, la persona que los tiene los conoce sin necesitar evidencias provenientes del comportamiento. La moda reciente, aunque no sea estrictamente conductista, vuelve a identificar parcialmente los estados mentales en términos de factores sociales u otros factores externos, restituyéndolos así al ámbito de lo que puede ser descubierto públicamente. Sin embargo, al mismo tiempo se reinstaura el problema de dar cuenta de la autoridad de primera persona.

Quienes están convencidos de la dimensión externa de los contenidos de los pensamientos tal y como comúnmente se identifican e individualan han reaccionado de diversas maneras. Una de las respuestas ha sido hacer una distinción entre, por una parte, los contenidos de la mente que están determinados subjetiva e internamente y, por otra, las creencias, deseos e intenciones comunes tal y como normalmente los atribuimos sobre la base de conexiones sociales o de otras conexiones externas. Ésta es claramente la línea que sigue el argumento de Putnam (aunque la palabra «agua» tenga diferentes significados, y se use para expresar creencias distintas cuando se la utiliza para referirse a agua y a gagua, las personas que usan la palabra para estos distintos propósitos pueden estar en «el mismo estado psicológico»). Jerry Fodor acepta la dis-

¹⁰ Daniel Dennett, «Beyond Belief», 76.

¹¹ Tyler Burge, «Other Bodies», «Individualism and the Mental», «Two Thought Experiments Reviewed», 284-93.

tinción para ciertos propósitos, pero defiende que la psicología debería adoptar la postura llamada «solipsismo metodológico», esto es, debería habérselas exclusivamente con estados internos, que son los estados psicológicos verdaderamente subjetivos que no le deben nada a sus relaciones con el mundo externo¹².

Stephen Stich hace esencialmente la misma distinción, pero extrae una moraleja más severa: donde Fodor piensa que sólo necesitamos jugar un poco con las actitudes proposicionales tal y como se conciben usualmente para aislar el elemento puramente subjetivo, Stich mantiene que los estados psicológicos tal y como pensamos sobre ellos en estos momentos pertenecen a una «psicología de sentido común» tosca y confusa que debería ser reemplazada por una «ciencia cognitiva» todavía por inventar. El subtítulo de su reciente libro es «El proceso contra la creencia» («The Case against Belief»)¹³.

Está claro que quienes trazan esa distinción han asegurado con ésta que el problema de la autoridad de primera persona, tal y como lo he expuesto, no pueda ser resuelto puesto que el problema que he señalado es el de cómo explicar la asimetría entre la manera en que una persona conoce sus estados mentales presentes y la manera en que los conocen los demás. Los estados mentales en cuestión son creencias, deseos, intenciones, etc., tal como se los concibe comúnmente. Quienes han aceptado algo parecido a la distinción de Putnam ni siquiera tratan de explicar la autoridad de primera persona respecto a estos estados; si realmente hay una autoridad de primera persona va unida a estados completamente distintos. (En el caso de Stich no es obvio que vaya unida a cosa alguna.)

Creo que Putnam, Burge, Dennett, Fodor, Stich y otros hacen bien en llamar la atención hacia el hecho de que los estados mentales comunes, al menos las actitudes proposicionales, se identifican parcialmente mediante las relaciones con la sociedad y con el resto del entorno, relaciones que la persona

¹² Jerry Fodor, «Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology».

¹³ Stephen Stich, *From Folk Psychology to Cognitive Science*.

que está en estos estados puede que no conozca en algunos aspectos. En mi opinión también están en lo cierto al mantener que por esta razón (aunque no hubiera otra), los conceptos de la «psicología de sentido común» no pueden incorporarse a un sistema de leyes coherente y global como el que busca la física. Estos conceptos son parte de una teoría de sentido común para describir, interpretar y explicar el comportamiento humano que va un poco a su aire, pero que es (así lo pienso yo) indispensable. Puedo imaginar una ciencia que se ocupe de cómo piensan y actúan las personas y que esté libre de «psicología de sentido común» pero no puedo ver qué interés tendría. Sin embargo, éste no es el tema de este ensayo.

Lo que aquí me ocupa es el sorprendente descubrimiento de que al parecer no conocemos lo que pensamos —al menos de la manera que pensamos que lo hacemos—. Esto es un problema real si, como yo, usted cree que es verdad que los factores externos determinan parcialmente los contenidos de los pensamientos y también cree que en general sabemos lo que pensamos de una manera en que los demás no lo saben. El problema surge porque admitir el papel identificador e individuador de los factores externos parece conducirnos a la conclusión de que puede ser que no conozcamos nuestros propios pensamientos.

¿Pero realmente se sigue esta conclusión? La respuesta depende de la manera en que se piense que la identificación de los contenidos mentales depende de los factores externos.

Por ejemplo, la conclusión se sigue en cualquier teoría que mantenga que las actitudes proposicionales se identifican mediante objetos (tales como proposiciones, ejemplares de proposiciones o representaciones) que están en o «ante» la mente y que contienen o incorporan (como «ingredientes») objetos o acaecimientos externos al agente, puesto que es evidente que todo el mundo ignora los rasgos sin fin de cualquier objeto externo. En general se concede que tal conclusión se sigue de estos supuestos¹⁴. Sin embargo, por razones que men-

¹⁴ Véase, por ej., Gareth Evans, *The Varieties of Reference*, 45, 199, 201.

cionaré después, rechazaré los supuestos sobre los que en este caso se basa la conclusión.

Tyler Burge ha sugerido que hay otra manera en la que los factores externos participan en la determinación de los contenidos del habla y del pensamiento. Resulta que uno de sus «experimentos mentales» se me aplica a mí bastante bien. Hasta hace poco creía que la artritis consistía en una inflamación de las articulaciones causada por depósitos de calcio; no sabía que cualquier inflamación de las articulaciones, por ejemplo la gota, también se considera que es artritis. Así que cuando un doctor me dijo (falsamente como se vio luego) que tenía gota, creí que tenía gota pero no que tenía artritis. Llegados aquí Burge nos pide que imaginemos un mundo en el que yo fuese físicamente el mismo, pero en el que ocurriese que la palabra «artritis» se aplica solamente a la inflamación de las articulaciones causada por los depósitos de calcio. Entonces la oración «la gota no es una forma de artritis» sería verdadera, no falsa, y la creencia que expresé con esta oración no sería la falsa creencia según la cual la gota no es una forma de artritis sino que habría sido una creencia verdadera sobre una enfermedad que no era la artritis. Con todo, en ese mundo imaginado todos mis estados físicos, mis «experiencias cualitativas internas» y mi comportamiento y disposiciones a comportarse habrían sido los mismos que en el presente mundo. Mi *creencia* habría cambiado, pero no tendría ninguna razón para suponer que lo hubiera hecho, por lo que no se podría decir que sabría qué es lo que creo.

Burge remarca que el hecho de que este argumento depende de «la posibilidad de que alguien tenga una actitud proposicional a pesar de su dominio incompleto de alguna de las nociones que conforman su contenido... *para que* el experimento mental funcione, debemos considerar que en algún momento el sujeto cree un contenido (o tiene alguna otra actitud que se caracteriza mediante el mismo), a pesar de una comprensión incompleta o de una aplicación errónea»¹⁵.

Parece que se sigue que si Burge está en lo cierto entonces en cada ocasión en que una persona está equivocada, confusa

¹⁵ Tyler Burge, «Individualism and the Mental», 83.

o parcialmente mal informada sobre el significado público de una palabra, está equivocada, confusa o parcialmente mal informada sobre cualquiera de sus creencias que se expresan (¿o se expresarían?) utilizando esa palabra. Dado que, según Burge, la «comprensión parcial» es «común o incluso normal en el caso de gran número de expresiones de nuestros vocabularios» debe ser igualmente común o normal que estemos equivocados sobre lo que creemos (y por supuesto sobre lo que tememos, tenemos esperanzas, deseamos, dudamos, etc.).

Al parecer Burge acepta esta conclusión; al menos es así como interpreto su negación de que una «plena comprensión de un contenido es en general una condición necesaria para creer el contenido». Explícitamente rechaza «el viejo modelo según el cual una persona debe estar en contacto directo con, o aprehender inmediatamente, los contenidos de sus pensamientos... el *contenido* del pensamiento de una persona no se fija mediante lo que en ésta sucede, o mediante lo que le es accesible simplemente mediante una reflexión cuidadosa»¹⁶.

No estoy seguro de cómo entender estas afirmaciones ya que no estoy seguro de cuán seriamente tomar la expresión «contacto directo» con, y «aprehensión inmediata» de, un contenido. En cualquier caso estoy convencido de que si lo que queremos decir y pensar está determinado por los hábitos lingüísticos de aquellos que nos rodean de la manera en que Burge cree que lo están, entonces la autoridad de primera persona se ve muy seriamente comprometida. Puesto que el grado y el carácter de ese compromiso me parece incompatible con lo que sabemos sobre la clase de conocimiento que tenemos de nuestras propias mentes, he de rechazar alguna de las premisas de Burge. Estoy de acuerdo en que lo que quiero decir y pensar no se «fija» (exclusivamente) mediante lo que a mí me pasa, de manera que lo que debo rechazar es la explicación de Burge de cómo los factores sociales y otros factores externos controlan los contenidos de la mente de una persona.

Debido a diferentes razones, me inclino a rebajar la importancia de los rasgos de nuestras atribuciones de actitudes que

¹⁶ *Ibid.*, 90, 102, 104.

Burge señala. Supongamos que yo, que pienso que la palabra «artritis» se aplica a la inflamación de las articulaciones sólo si ésta está causada por depósitos de calcio, y mi amigo Arthur, que sabe más, sinceramente proferimos dirigiéndonos a Smith las palabras «Carl tiene artritis». De acuerdo con Burge, si otras cosas son más o menos igual (Arthur y yo somos ambos hablantes competentes del inglés, ambos hemos aplicado frecuentemente la palabra «artritis» a los casos genuinos de artritis, etc.), entonces nuestras palabras en esta ocasión significan la misma cosa, Arthur y yo significamos la misma cosa con nuestras palabras y expresamos la misma creencia. Mi error sobre el significado de la palabra que se encuentra en el diccionario (o sobre lo que es la artritis) no marca diferencia sobre lo que significo o pienso en esta ocasión. La evidencia que ofrece Burge para esta afirmación parece que se apoya en su convicción de que esto es lo que cualquiera (que no esté echado a perder por la filosofía) diría acerca de Arthur y yo. Dudo de que Burge esté en lo cierto acerca de esto, pero incluso si lo está no creo que ello demuestre su afirmación. La atribución común de significados y de actitudes se apoya en supuestos amplios y vagos acerca de lo que comparten o no comparten (tanto lingüísticamente como de otra manera) la persona que realiza la atribución, la persona a la que se le hace la atribución y la audiencia a la que se dirige quien realiza la atribución. Cuando alguno de estos supuestos resulta ser falso, probablemente cambiaremos las palabras que usamos para realizar el informe en cuestión, frecuentemente de manera importante. Cuando nada depende de ello, tendemos a escoger el camino más fácil: le tomamos la palabra a alguien incluso aunque ello no acabe de reflejar algún aspecto del pensamiento del hablante o del significado. Sin embargo, la razón no es que estemos obligados a ponernos legalistas acerca de ello (al menos fuera de los tribunales). Y a menudo no nos ponemos así. Si Smith (que no está echado a perder por la filosofía) le informa a otro (quizás a un doctor que a distancia intente diagnosticar sobre la base de un informe telefónico) de que Arthur y yo hemos dicho y que creemos que Carl tiene artritis, puede con ello llevar a confusión a *su* oyente. Si este problema pudiera darse, Smith, que está alerta ante esa

posibilidad, no diría simplemente «Arthur y Davidson creen que Carl tiene artritis», sino que añadiría algo como «pero Davidson piensa que la artritis la causan los depósitos de calcio». La necesidad de hacer este añadido la tomo como una muestra de que la atribución simple no era totalmente correcta; había una diferencia relevante entre los pensamientos que Arthur y yo expresamos al decir «Carl tiene artritis». Burge no tiene por qué cambiar su posición en lo más mínimo por este argumento ya que puede insistir en que el informe es literalmente correcto, pero que como cualquier informe podría llevar a conclusiones erróneas. Pienso, por otra parte, que esta réplica pasaría por alto la medida en que los contenidos de una creencia dependen necesariamente de los contenidos de otras. Los pensamientos no son átomos independientes y por ello no puede haber una regla simple, rígida para la correcta atribución de un pensamiento aislado¹⁷.

¹⁷ Burge sugiere que la razón por la que normalmente interpretamos que una persona quiere decir mediante sus palabras lo que otros de su misma comunidad lingüística quieren decir con ellas, tanto si el hablante sabe lo que los otros quieren decir como si no, es que «Cuando hay un mal uso de las palabras o un malentendido, las personas se atienen a las normas de la comunidad y los demás consideran que se atienen a ellas». También afirma que tales casos «dependen de cierta responsabilidad para con las prácticas comunes» (*ibid.*, 90). No dudo del fenómeno, pero dudo de su relación con lo que se supone que muestra: (a) Frecuentemente es razonable considerar que las personas son responsables de saber lo que significan sus palabras; en tales casos podemos tratarlas como comprometidas con posiciones con las cuales no sabían o creían que estaban comprometidas. Esto no tiene nada que ver (directamente) con lo que quiere decir con sus palabras ni con lo que creen. (b) Como buenos ciudadanos y padres queremos dar apoyo a prácticas que aumenten las posibilidades de comunicación; usando las palabras como pensamos que los demás lo hacen podemos mejorar la comunicación. Este pensamiento (esté o no justificado) puede ayudar a explicar por qué algunas personas tienden a atribuir significados y creencias de una manera legalista; esperan fomentar la conformidad. (c) Un hablante que desee ser comprendido debe tratar de que sus palabras sean interpretadas (y por ello interpretables) de ciertas maneras; se puede contribuir a esta intención usando las palabras tal y como los demás lo hacen (aunque frecuentemente éste no es el caso). De manera similar, un oyente que desee comprender a un hablante debe tratar de interpretar las palabras del hablante tal y como el hablante pretendía (sea o no «estándar» la interpretación). Estas intenciones recíprocas llegan a ser moralmente importantes en un sinfín de situaciones que no tienen una conexión necesaria con la determinación de lo que alguien tiene en mente.

Aunque rechazo la insistencia de Burge en que forzosa-mente hemos de dar a las palabras de una determinada persona el significado que tienen en su comunidad lingüística e interpretar sus actitudes proposicionales sobre la misma base, creo que hay un sentido algo distinto pero muy importante según el cual los factores sociales controlan lo que el hablante puede significar con sus palabras. Si un hablante desea ser entendido debe tratar de que sus palabras se interpreten de una determinada manera y debe tratar de proporcionar a su audiencia las claves que necesita para llegar a la interpretación deseada. Esto se mantiene tanto si el oyente tiene un uso refinado de la lengua que el hablante conoce como si es el que aprende una primera lengua. Es el requisito de que una lengua sea aprendible o sea interpretable el que proporciona el factor social irreductible y el que muestra por qué una persona no puede significar algo mediante sus palabras que no pueda ser correctamente descifrado por otra. (Parece que el propio Burge hace esta observación en un artículo posterior)¹⁸.

Querría ahora volver al ejemplo de la Tierra Gemela de Putnam, el cual no depende de la idea de que el uso lingüístico social dicta (bajo condiciones más o menos estándar) lo que los hablantes significan con sus palabras ni, por supuesto, cuáles son sus estados psicológicos (restringidos). Como ya he dicho, estoy persuadido de que Putnam está en lo cierto; lo que nuestras palabras significan se fija en parte mediante las circunstancias en las que aprendimos y usamos esas palabras. El único ejemplo de Putnam (el del agua) quizás no es suficiente para establecer esta idea ya que es posible insistir en que el término «agua» no se aplica exactamente sólo a la sustancia que tiene la misma estructura molecular que el agua, sino también a sustancias que son suficientemente parecidas al agua en su estructura como para ser inodoras, potables, algo en lo que nos podamos bañar y podamos navegar, etc. (Me doy cuenta de que esta observación, como muchas otras en este ensayo, puede mostrar que no reconozco un designador rígido cuando lo veo. Pues es cierto que no lo reconozco.)

¹⁸ Tyler Burge, «Two thought Experiments Reviewed», 289.

El tema que tratamos no depende de casos especiales como éstos ni de cómo los resolvemos o deberíamos hacerlo. El tema depende simplemente de cómo se establece la conexión básica entre las palabras y las cosas o los pensamientos y las cosas. Mantengo, al igual que Burge y Putnam si es que los interpreto correctamente, que se establece mediante las interacciones causales entre las personas y las partes y aspectos del mundo. Las disposiciones a reaccionar de manera distinta a los objetos y acaecimientos así establecidas son capitales para la interpretación correcta de los pensamientos y el habla de una persona. Si éste no fuera el caso, no tendríamos manera de descubrir lo que piensan los otros o lo que significan con sus palabras. El principio es tan simple y obvio como lo siguiente: una oración a la que alguien se siente inspirado (es decir, se ve llevado causalmente) a considerar verdadera debido a la visión de la luna y sólo a ella es apropiada para significar algo así como «Ahí está la luna»; el pensamiento expresado puede ser que la luna está ahí; el pensamiento que la visión de la luna y sólo ella inspira es el adecuado para ser el pensamiento de que la luna está ahí. Adecuado para constituir informes de segunda mano, dejando espacio para errores inteligibles, etc. No todas las palabras y oraciones están tan condicionadas a lo que tratan; podemos aprender perfectamente a usar la palabra «luna» incluso sin verla. Lo que afirmo es que todo pensamiento y lenguaje debe tener un fundamento en tales conexiones históricas directas, y estas conexiones restringen la interpretación de los pensamientos y el habla. Quizás habría de señalar que los argumentos a favor de esta afirmación no descansan en intuiciones que tienen que ver con lo que diríamos si determinados contrafácticos fueran verdaderos. No se necesita la ciencia-ficción ni ningún experimento mental¹⁹.

¹⁹ Burge ha descrito experimentos mentales que no involucran para nada el lenguaje; uno de estos experimentos le impulsa a afirmar que alguien que haya crecido en un entorno sin aluminio no podría tener «pensamientos sobre el aluminio» («Individualism and Psychology», 5). Burge no dice por qué cree esto, pero no es de ninguna manera obvio que se necesiten suposiciones contrafácticas para establecer este punto. En cualquier caso, los nuevos experi-

Estoy de acuerdo con Putnam y Burge en que (en palabras de Burge) «el contenido intencional de las actitudes proposicionales comunes ... no puede explicarse en términos de estados o procesos físicos, fenoménicos, causal-funcionales, computacionales o sintácticos que estén especificados no-intencionalmente y se definan puramente en términos del individuo aislado de su entorno físico y social»²⁰. La cuestión que sigue en pie es si este hecho es una amenaza para la autoridad de primera persona, tal y como Burge parece creer, y como Putnam y otros realmente piensan. He rechazado uno de los argumentos de Burge, el cual, si fuera correcto, supondría una amenaza así, pero todavía nos queda la posición descrita en el párrafo previo, que mantengo lo hagan o no otros, ya que creo que se requiere este grado de «externismo» para explicar cómo se aprende el lenguaje y cómo un intérprete puede identificar las palabras y las actitudes.

¿Por qué Putnam piensa que si la referencia de una palabra algunas veces se fija mediante la historia natural de cómo la palabra ha sido adquirida, el usuario de la palabra puede perder la autoridad de primera persona? Putnam afirma (correctamente en mi opinión) que dos personas pueden ser iguales en todos los aspectos físicos (químicos, fisiológicos, etc.) relevantes y aún con eso significar cosas diferentes mediante sus palabras y tener diferentes actitudes proposicionales (tal y como normalmente se identifican). Las diferencias se deben a diferencias en el entorno que en algunos aspectos los agentes pueden desconocer. ¿Por qué en estas circunstancias habríamos de suponer que estos agentes pueden no saber lo que significan y piensan? Hablar con ellos no lo mostrará fácilmente. Tal y como he remarcado, cada uno de ellos, cuando están delante de un vaso de agua o gagua, sinceramente afirman «Aquí hay un vaso de agua». Si se encuentran en sus entornos

mentos mentales parece que descansan en intuiciones bastante diferentes de las intuiciones en las que se apoya en «Individualism and the Mental»; no es claro el papel de las normas sociales en estos nuevos experimentos, y los hábitos lingüísticos de una comunidad son aparentemente irrelevantes. Llegados aquí, es posible que la posición de Burge sea cercana a la mía.

²⁰ «Two Thought Experiments Reviewed», 288.

habituales, cada uno está en lo cierto; si cada uno se traslada a la tierra del otro, entonces los dos se equivocan. Si les preguntamos a cada uno lo que quieren decir con la palabra «agua», nos ofrecen la respuesta correcta usando las mismas palabras, por supuesto. Si les preguntamos lo que creen, responderán correctamente. Estas respuestas son correctas porque, aunque verbalmente sean idénticas, deben ser interpretadas de manera distinta. ¿Y qué es lo que no saben (de la manera autorizada usual) acerca de sus propios estados? Como hemos visto, Putnam distingue los estados de que hemos estado tratando de los estados psicológicos «restringidos» que no presuponen la existencia de ningún otro individuo más que el sujeto que se encuentra en ese estado. Podemos empezar a preguntarnos por qué Putnam se interesa por los estados psicológicos restringidos. Parte de la respuesta es, por supuesto, que cree que son estos estados los que él piensa que tienen la propiedad «cartesiana» de ser conocidos de una manera especial por la persona que se encuentra en ellos. (La otra parte de la respuesta tiene que ver con la construcción de una «psicología científica»; esto no es lo que aquí nos interesa.)

Según creo el razonamiento depende de dos supuestos que generalmente no se cuestionan. Éstos son:

- (1) Si un pensamiento se identifica mediante una relación con algo externo a la cabeza, éste no está totalmente en la cabeza (no está en la cabeza).
- (2) Si un pensamiento no está totalmente en la cabeza, no puede ser «captado» por la mente de la manera que se requiere para que haya autoridad de primera persona.

Que ésta es la manera de razonar de Putnam lo sugiere su afirmación según la cual si dos cabezas son la misma, los estados psicológicos restringidos deben ser los mismos. Así pues, si suponemos que dos personas son «molécula a molécula» la misma («en el sentido en que dos corbatas pueden ser “idénticas”»); si se quiere podríamos añadir, que cada una de las dos personas «tiene los mismos pensamientos verbalizados... los mismos datos sensoriales, las mismas disposiciones, etc.»), entonces «es absurdo pensar que [un] estado psicológico difiere en algo de» el otro. Éstos son, por supuesto, estados psicoló-

gicos restringidos, no aquellos que normalmente atribuimos, los cuales no están en la cabeza²¹.

No es fácil decir exactamente de qué manera los pensamientos verbalizados, los datos sensoriales y las disposiciones pueden ser idénticos sin recurrir al ejemplo de las corbatas, de modo que vamos a volver sobre ello. La idea es la siguiente: los estados psicológicos restringidos de dos personas son idénticos cuando sus estados físicos no pueden distinguirse. No tiene sentido discutir esto puesto que es el propio Putnam quien define los estados psicológicos restringidos; lo que quiero poner en cuestión es el supuesto (1) que he señalado más arriba y que conduce a la conclusión de que las actitudes proposicionales usuales no están en la cabeza y que por ello no se les puede aplicar la autoridad de primera persona.

Debe quedar claro que del hecho de que los significados se identifiquen en parte mediante relaciones con objetos externos a la cabeza no se sigue que los significados no estén en la cabeza. Suponer esto sería tan erróneo como argumentar que puesto que el hecho de que esté quemado por el Sol presupone la existencia del Sol, mi quemadura no es una condición de mi piel. Mi piel quemada por el Sol puede ser indistinguible de la piel de algún otro que se quemó por otros medios (nuestras pieles pueden ser idénticas en el sentido «de las corbatas»); aún con todo sólo uno de nosotros se ha quemado realmente con el Sol. Esto es suficiente para mostrar que apreciar los factores externos que toman parte en nuestras maneras comunes de identificar los estados mentales no desacredita a la teoría de la identidad entre lo mental y lo físico. Andrew Woodfield parece pensar que la desacredita. Éste escribe: «Ningún estado *de re* sobre un objeto que sea externo al cerebro de una persona puede de ninguna manera ser idéntico con un estado de ese cerebro, ya que ningún estado del cerebro presupone la existencia de un objeto externo»²². Los estados y acaecimientos individuales *conceptualmente* no presuponen nada por sí mismos, pero algunas *descripciones* sí que lo

²¹ «Meaning of “Meaning”», 227. [Pág. 144 de la traducción al castellano reseñada en la bibliografía.]

²² Andrew Woodfield, en *Thought and Object*, pág. viii.

hacen. Mi abuelo paterno no me presuponía, pero si alguien puede ser descrito como mi abuelo paterno, deben existir algunas personas además de mi abuelo, incluyéndome a mí.

Puede ser que Burge cometa el mismo error en el siguiente pasaje:

ninguna aparición o caso particular de un pensamiento... puede tener un contenido diferente y ser el mismo acaecimiento-ejemplar... por tanto... el acaecimiento de un pensamiento en una persona no es *idéntico* a ningún acaecimiento que se dé en ella descrito por la fisiología, la biología, la química o la física. En efecto, digamos que *b* es un acaecimiento que se describe en términos de alguna de las ciencias físicas que se da en un sujeto mientras éste tiene el pensamiento relevante. Denotemos mediante «*b*» el mismo acaecimiento físico que se da en el sujeto en nuestra situación contrafáctica... *b* no necesita verse afectado por diferencias contrafácticas [que no cambien los contenidos del acaecimiento de pensamiento en cuestión]. Así pues... *b* [el acaecimiento físico] no es idéntico al caso particular de pensamiento que se da en el sujeto²³.

Burge no pretende haber justificado la premisa de este argumento y por tanto tampoco su conclusión. Sin embargo, mantiene que la negación de la premisa es «intuitivamente muy implausible». Añade: «las teorías materialistas de la identidad han instruido a la imaginación a concebir el contenido de un acaecimiento mental como algo que varía mientras el acaecimiento permanece fijo. Pero si tales imaginaciones son hechos posibles o solamente una fantasía filosófica es una cuestión aparte». El motivo por el que mantiene que «los experimentos mentales no-individualistas hacen implausibles» a las propias «teorías materialistas de la identidad» es que piensa que la negación de la premisa es muy improbable²⁴.

Acepto la premisa de Burge; pienso que su negación no es meramente implausible, sino que es absurda. Si dos acaecimientos mentales tienen contenidos distintos, son con segu-

²³ «Individualism and the Mental», 111.

²⁴ «Individualism and Psychology», 15 n. 7. Cfr. «Individualism and the Mental», 111.

ridad acaecimientos distintos. Lo que creo que muestran los casos imaginados por Burge y Putnam (y que el ejemplo del «hombre del pantano» muestra más directamente) es que las personas que sean similares (o «idénticas» en el sentido de las corbatas) en todos los aspectos físicos relevantes pueden diferir en lo que quieren decir (significan) o piensan, tal y como difieren en ser abuelos o en estar quemados por el Sol. Pero por supuesto que hay *algo* distinto acerca de ellos, incluso en el mundo físico; sus historias causales son distintas y son objetos físicos discretos.

Con ello concluyo que el mero hecho de que los estados mentales usuales y los acaecimientos se individuen en términos de relaciones con el mundo externo no provoca el descrédito de las teorías de la identidad físico-mental como tales. En conjunción con varios otros supuestos (plausibles), creo que el «externismo» de determinados estados y acaecimientos mentales se pueden usar para desacreditar las teorías de la identidad de tipo a tipo; sin embargo, en todo caso apoyaría a las teorías de la identidad de ejemplar a ejemplar. (No veo ninguna buena razón para denominar «materialistas» a todas las teorías de la identidad; si algunos acaecimientos mentales son acaecimientos físicos, esto no los hace ser más físicos que mentales. La identidad es una relación simétrica.)

Putnam y Woodfield se equivocan, por tanto, al afirmar que es «absurdo» pensar que dos personas no pueden ser físicamente idénticas (en el «sentido de las corbatas») y diferir en sus estados psicológicos usuales. Burge, a menos de que esté dispuesto a hacerles desempeñar un papel mucho más fuerte a los supuestos esencialistas del que les hace tener, se equivoca en pensar que ha mostrado que todas las teorías de la identidad son implausibles. Por tanto, podemos libremente mantener que las personas pueden ser idénticas en todos los aspectos físicos relevantes (idénticas en el «sentido de las corbatas») aunque difieran psicológicamente; ésta es, de hecho, la posición del «monismo anómalo», el cual he defendido en otro lugar²⁵.

²⁵ «Mental Events», en Donald Davidson, *Essays on Actions and Events*.

En este punto se ha eliminado un obstáculo frente al conocimiento no basado en evidencias de nuestras propias actitudes proposicionales usuales, ya que si nuestras creencias usuales y las otras actitudes pueden estar «en la cabeza» aunque se identifiquen como las actitudes que son parcialmente en términos de lo que no está en la cabeza, entonces la amenaza a la autoridad de primera persona no puede provenir simplemente del hecho de que los factores externos sean relevantes para la identificación de las actitudes.

No obstante, sigue habiendo una aparente dificultad. Es verdad que mi quemadura de Sol, aunque sea describible como tal solamente en relación con el Sol, es idéntica a una condición de mi piel que puede ser descrita (lo supongo) sin referencia a tales factores externos. Con todo, si como un científico formado en todas las ciencias físicas, tengo acceso solamente a mi piel y no tengo el conocimiento de la historia de esa condición, entonces por hipótesis de ninguna manera puedo decir que me he quemado con el Sol. Quizás, entonces, alguien tiene autoridad de primera persona con respecto a los contenidos de su mente solamente en cuanto esos contenidos puedan ser descritos o descubiertos sin referencia a factores externos. En la medida en que los contenidos se identifiquen en términos de factores externos, la autoridad de primera persona necesariamente desaparece. Puedo decir al examinar mi piel lo que es mi condición privada o «restringida», pero nada de lo que pueda aprender en este reino restringido me dirá que es una quemadura solar. La diferencia entre referirse a y pensar en agua y referirse a y pensar en gaga es como la diferencia entre tener una quemadura solar y que la piel de alguien esté exactamente en la misma condición pero por otra causa. La diferencia semántica se encuentra en el mundo externo, más allá del alcance del conocimiento subjetivo o sublunar. Así es como podría ser más o menos el argumento.

Esta analogía entre la visión limitada del dermatólogo y la visión como en un túnel del ojo de la mente es fundamentalmente errónea. Su atractivo depende de una imagen defectuosa de la mente, una imagen que comparten quienes atacan el carácter subjetivo de los estados psicológicos usuales con

aquellos a quien éstos atacan. Si logramos abandonar esta imagen, 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.

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. 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 ama el corazón, 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 destacada, 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, incluya el teatro, el escenario, los actores y la audiencia, pero, por otra, lo que se conoce y se registra pertenece solamente a la audiencia. 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 estos momentos no me ocupo de objetos de la mente como los datos sensoriales (a los que generalmente se rechaza en el presente) sino de sus primos enjuiciables, los presuntos objetos de las actitudes proposicionales, ya se conciben como proposiciones, ejemplares de proposiciones, representaciones o fragmentos de «mentales». La idea central que quiero atacar es que éstas son entidades que la mente pueda «abrigar», «asir», «tener ante ella» o «estar en contacto» con. (Es probable que estas metáforas sean instructivas: los *voyeurs* sólo quieren tener representaciones ante el ojo de la mente, mientras que los más agresivos las asen; un inglés puede ser que meramen-

te esté en contacto con los contenidos de la mente, mientras que los caracteres más afables los abrigan.)

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 he 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 igualmente ser captada si es que la mente conoce en qué estado se encuentra ella. 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 en este punto en el que el concepto de lo subjetivo —de un estado mental— parece que 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.

Como analogía nos encontramos con el problema del experto en quemaduras solares que mediante la inspección de la piel no puede distinguir si se trata de un caso de quemadura solar o meramente de una condición «idéntica» que tiene otra causa. Podemos resolver el problema de la quemadura solar distinguiendo entre quemadura solar y quemadura solarosa; la quemadura solarosa es como la quemadura solar excepto en que puede ser o no que el Sol la haya producido. El experto puede reconocer un caso de quemadura solarosa solamente al mirarla, pero no pasa lo mismo con una quemadura solar. Esta solución funciona porque no se requiere, a diferencia de los objetos de la mente, que las condiciones de la piel sean tales que haya alguien especial que pueda decir solamente al mirar si se da la condición en cuestión o no.

La solución en el caso de los estados mentales es diferente y más simple; se trata de deshacerse de la metáfora de los objetos ante la mente. La mayoría de nosotros abandonamos hace tiempo la idea de que las percepciones, los datos senso-

riales 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 y las dudas son *entidades* en o ante la mente, o que estar en tales estados requiera que haya los objetos mentales correspondientes.

Esto ya se ha dicho antes de diversas maneras pero por razones diferentes. Por ejemplo, los escrúpulos ontológicos no forman parte de lo que me interesa. Siempre necesitaremos una fuente infinita de objetos que nos ayuden a describir e identificar actitudes como la creencia; no estoy sugiriendo ni por un momento que las oraciones de creencia y las oraciones que atribuyen otras actitudes no sean de naturaleza relacional. Lo que estoy diciendo es que los objetos con los que relacionamos a las personas para describir sus actitudes no necesitan ser objetos *psicológicos* en ningún sentido, objetos que puedan ser captados, conocidos o abrigados por la persona cuyas actitudes se están describiendo.

También esto es algo conocido; Quine lo señala cuando sugiere que podemos usar nuestras propias oraciones para seguir la pista de los pensamientos de las personas que no conocen nuestra lengua. El interés de Quine es semántico y no dice nada en este contexto sobre los aspectos epistemológicos y psicológicos de las actitudes. Nosotros necesitamos abordar estos aspectos conjuntamente. Las oraciones sobre las actitudes son relacionales; por razones *semánticas* deben ser entonces objetos con los cuales relacionar los sujetos de las actitudes. Pero tener una actitud no es tener una entidad ante la mente; por razones muy convincentes, tanto de orden *psicológico* como *epistemológico*, deberíamos negar que haya objetos de la mente.

El origen del problema es el dogma de que tener un pensamiento es tener un objeto ante la mente. Putnam 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. Estoy de acuerdo en que no hay objetos que puedan servir a ambos propósitos. Putnam (y al-

gunos de los otros filósofos que he mencionado) piensan 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 de que un objeto *semántico* puede 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 pienso que la razón es clara. El único objeto que podría satisfacer ambos requerimientos, el de estar ante la mente y también el de que determine el contenido de un pensamiento, debería, como las ideas e impresiones de 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 mi opinión, tantas otras variantes del argumento simple y gene-

ral que he 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 comúnmente los concebimos se identifiquen en parte por su historia natural no sólo no habla en 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 acaecimientos 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 entender 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 y acaecimientos 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 tienen. 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 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 lenguaje y los determinantes externos del pensamiento y el significado se dan conjuntamente.

3

El mito de lo subjetivo

Éste es un ensayo sobre un viejo tema, la relación entre la mente humana y el resto de la naturaleza, entre lo subjetivo y lo objetivo tal como hemos llegado a concebirlos. Este dualismo, aunque a su manera demasiado evidente para ponerse en cuestión, acarrea consigo en nuestra tradición una larga y no necesariamente apropiada carga de ideas asociadas. Algunas de estas ideas se están poniendo ahora bajo el escrutinio crítico, y el resultado promete marcar un cambio capital en el pensamiento filosófico contemporáneo. El presente ensayo, aunque claramente es tendencioso, no está diseñado primariamente para convertir al escéptico; su principal objetivo es describir, desde un cierto punto de vista, una tendencia bastante reconocida en el pensamiento reciente acerca de los contenidos de la mente, y sugerir algunas de las consecuencias que creo que se siguen de la misma.

Las mentes son muchas, la naturaleza es una. Cada uno de nosotros tiene su propia posición en el mundo y por ello su propia perspectiva del mismo. Es fácil pasar de esta perogrullada a alguna noción confusa de relativismo conceptual. El primer relativismo, inocuo, es sólo el relativismo bien conocido de la posición en el espacio y el tiempo. Dado que cada uno de nosotros ocupa un volumen de espacio-tiempo, dos de nosotros no podemos estar exactamente en el mismo lugar al mismo tiempo. Las relaciones entre nuestras posiciones son inteligibles porque podemos situar a cada persona en un mundo común y único, y en un marco temporal compartido.

El relativismo conceptual parece similar, pero la analogía es difícil de llevar a cabo puesto que podemos preguntarnos cuál es el punto de referencia común o el sistema de coordenadas respecto al cual cada esquema es relativo. Sin una buena respuesta a esta pregunta la afirmación según la cual cada uno de nosotros en algún sentido habita su propio mundo pierde su inteligibilidad.

Por esta y otras razones he sostenido durante mucho tiempo que hay límites respecto a cuánto pueden diferir los sistemas de pensamiento, tanto individuales como sociales. Si por relativismo conceptual queremos dar a entender la idea de que los esquemas conceptuales y los sistemas morales, o las lenguas asociadas con ellos, pueden diferir enormemente —hasta el punto de ser mutuamente ininteligibles o inconmensurables, o por siempre fuera del alcance de resolución racional—, entonces rechazo el relativismo conceptual¹. Es evidente que en las diferentes épocas, culturas y personas, hay diferencias de ciertos tipos respecto de lo que reconocemos y por lo que luchamos; pero se trata de diferencias que con empatía y esfuerzo podemos llegar a explicar y comprender. Los problemas aparecen cuando tratamos de aceptar la idea de que debe haber diferencias más globales, ya que esto parece pedirnos (absurdamente) que adoptemos una posición externa a nuestros propios modos de pensar.

No comprendemos la idea de un esquema que nos sea totalmente extraño. Sabemos más o menos qué son los estados mentales y cómo se identifican correctamente; son precisamente aquellos estados cuyos contenidos pueden descubrirse de maneras bien conocidas. Si otras personas o criaturas se encuentran en estados que no puedan descubrirse por estos métodos, no puede ser porque nos fallen nuestros métodos, sino porque esos estados no deberían llamarse estados mentales —no son creencias, deseos, esperanzas o intenciones—. Lo asignificativa que resulta la idea de un esquema conceptual situado por siempre más allá de lo que podemos captar no se debe a nuestra incapacidad de comprender tal esquema ni a

¹ He defendido esta posición en «On the Very Idea of a Conceptual Schemes», Ensayo 13 de *Inquiries into Truth and Interpretation*.

otras limitaciones humanas; se debe simplemente a lo que entendemos por sistema de conceptos.

Muchos filósofos no están satisfechos con argumentos como éstos porque piensan que hay otra manera en que el relativismo conceptual puede hacerse inteligible, pues parece que podríamos darle sentido a un relativismo tal si encontraríamos un elemento en la mente que no se viera afectado por la interpretación conceptual. De ese modo varios esquemas podrían verse como relativos a este elemento común, y se les podría asignar el papel de organizar el mismo. Ese elemento común es, por supuesto, una u otra versión de las intuiciones kantianas, las ideas o impresiones de Hume, los datos sensoriales, las sensaciones no interpretadas o lo dado por los sentidos. Kant pensaba que solamente había un esquema posible; pero una vez que se hace explícito el dualismo del esquema y el contenido, la posibilidad de los esquemas alternativos se hace patente. La idea de que podemos optar entre diversos esquemas se encuentra de forma explícita en el obra de C. I. Lewis. «En nuestra experiencia cognitiva hay dos elementos; los datos inmediatos, como los de los sentidos, que se presentan o son dados a la mente, y una forma, construcción o interpretación que representa la actividad del pensamiento»².

Si pudiéramos concebir la función de los esquemas conceptuales de esta manera, el relativismo parecería ser una posibilidad abstracta a pesar de las dudas sobre cómo podría ser descifrado un esquema ajeno; la idea sería que distintos esquemas o lenguas constituyen diferentes maneras en las que puede ser organizado lo que se da en la experiencia. Según esto, no habría un punto de vista desde el cual pudiéramos examinar tales esquemas y quizás no habría manera en que pudiéramos compararlos o evaluarlos en general; aún con todo, en tanto que pensáramos que entendíamos la dicotomía esquema-contenido, podríamos imaginar a la corriente incontaminada de experiencias como algo que se reelabora de manera distinta por parte de mentes y culturas distintas. De

² C. I. Lewis, *Mind and the World Order*, 38. Lewis afirma que la tarea de la filosofía es «revelar los criterios categoriales que la mente aplica a lo que a ella le es dado» (pág. 36).

esta manera, puede mantenerse que el relativismo conceptual proporciona el elemento con el cual los distintos esquemas alternativos se relacionan: ese elemento es lo dado de manera no interpretada, los contenidos no categorizados de la experiencia.

Esta concepción de la mente y de su lugar en la naturaleza ha definido muchos de los problemas que la filosofía moderna pensaba que tenía que resolver. Entre estos problemas se encuentran los que conciernen al conocimiento: cómo conocemos el mundo externo, cómo conocemos las otras mentes, incluso cómo conocemos los contenidos de nuestra propia mente. Pero también deberíamos incluir el problema de la naturaleza del conocimiento moral, el análisis de la percepción y muchos temas problemáticos de la filosofía de la psicología y la teoría del significado.

Correspondiendo a este catálogo de problemas hay una larga lista de maneras en las que se ha formulado el supuesto contraste esquema-contenido. El esquema puede pensarse como una ideología, un conjunto de conceptos que se ajustan bien a la tarea de organizar la experiencia en objetos, eventos, estados y complejos de estos mismos; o bien el esquema puede ser una lengua o lenguaje³, quizás con predicados y otros elementos asociados, interpretados para servir a una ideología. Los contenidos de tal esquema pueden ser objetos de una clase especial como datos sensoriales, perceptos, las impresiones, sensaciones o apariencias; o bien los objetos pueden disolverse en modificaciones adverbiales de la experiencia: podemos ser «me parece rojamente». Los filósofos han mostrado mucho ingenio en encontrar maneras de poner en palabras los contenidos de lo dado; así tenemos esas raras oraciones, sin verbo, como «Rojo aquí ahora», y las diversas

³ A lo largo del libro traduzco el término inglés «language» por «lengua o lenguaje» allí donde me parece conveniente dejar abierta la posibilidad de que el sistema lingüístico al que se alude no es una lengua evolucionada de forma natural, sino un sistema lingüístico quizá diseñado intencionadamente, y en este sentido artificialmente (por ejemplo, un lenguaje lógico). Cuando es claro que se trata de una lengua natural particular traduzco «language» simplemente por «lengua» y, finalmente, traduzco «language» por «lenguaje» cuando se trata del lenguaje en general. [N. de la T.]

formulaciones de las oraciones protocolares sobre las que disputaron los positivistas lógicos.

No es necesario expresar el contenido en palabras y, de acuerdo con algunas concepciones, ello no es posible. La división esquema-contenido puede sobrevivir incluso en un entorno en que se rehúyan la distinción analítico-sintética, los datos sensoriales o la suposición de que puede haber pensamientos o experiencias que estén libres de teoría. Si estoy en lo cierto, éste es el tipo de entorno que proporciona W. V. Quine. De acuerdo con la «epistemología naturalizada» de Quine, desde la filosofía del conocimiento no deberíamos pedir más que una explicación de cómo, dadas las evidencias a nuestra disposición, somos capaces de formar una teoría satisfactoria del mundo. Esta explicación descansa en la mejor teoría que tenemos: la ciencia presente. Las estimulaciones de nuestros órganos sensoriales proporcionan las evidencias de las cuales dependen en último término los significados de nuestras oraciones y todo nuestro conocimiento. Son estas estimulaciones las que proporcionan a una persona sus únicos indicios acerca de «lo que pasa a su alrededor». Por supuesto que Quine no es un reduccionista: «no podemos eliminar los aderezos conceptuales oración por oración». Sin embargo, de acuerdo con Quine se debe hacer una distinción definida entre el contenido invariante y los aderezos conceptuales variables, entre «informe e invención, sustancia y estilo, indicios y conceptualización». La razón es que

podemos investigar el mundo, y el hombre como parte de él, y averiguar así qué indicios puede conseguir de lo que ocurre en torno suyo. Restando entonces esos indicios de su conjunta visión del mundo, podemos conseguir como diferencia la contribución neta del hombre. Esta diferencia indica la extensión de la soberanía conceptual del hombre, el dominio dentro del cual puede revisar la teoría salvando los datos⁴.

⁴ Este pasaje y las citas que lo preceden provienen de W. V. Quine, *Word and Object*, 5. [Pág. 19 de la traducción al castellano reseñada en la bibliografía.]

La visión del mundo y los indicios, la teoría y los datos, éstos son el esquema y el contenido del cual he estado hablando.

Lo que importa entonces no es si podemos describir los datos en un idioma neutro libre de carga teórica; lo que importa es que debería haber una fuente última de evidencia, el carácter de la cual pueda especificarse totalmente sin hacer referencia a aquello a favor de lo cual es evidencia. Así pues, los patrones de estimulación, como los datos sensoriales, pueden identificarse y describirse sin referencia a «lo que ocurre en torno nuestro». Si nuestro conocimiento del mundo deriva enteramente de evidencias de este tipo, entonces no solamente puede que nuestros sentidos nos engañen a veces, sino que es posible que estemos engañados de manera sistemática y general.

Es fácil recordar lo que provoca este punto de vista: se piensa que es necesario aislar del mundo externo las fuentes últimas de evidencia a fin de garantizar la autoridad de la evidencia para el sujeto. Ya que no podemos estar seguros sobre cómo es el mundo externo a la mente, lo subjetivo puede mantener su virtud —su castidad, su certeza para nosotros— solamente siendo protegido de la contaminación del mundo. Un problema familiar es, por supuesto, que la desconexión crea una brecha que ningún razonamiento o construcción puede salvar de manera plausible. Cuando se ha escogido este punto de partida, no hay manera de decir de qué son evidencia las evidencias, o eso es lo que parece. Entonces nos amenazan el idealismo, formas reduccionistas de empirismo y el escepticismo.

La historia es conocida; pero déjenme continuar a mi modo, sin tomar aliento, todavía por un capítulo. Si las evidencias últimas a favor de nuestros esquemas y teorías, el material básico sobre el cual se fundamentan, son subjetivas de la manera que he descrito, entonces también lo será todo aquello que se base directamente en ellas: nuestras creencias, deseos, intenciones y lo que queremos decir o significar mediante nuestras palabras. Aunque todo esto constituya la proge- nie de nuestra «concepción del mundo» —ciertamente, tomado en conjunto, constituye nuestra concepción del mundo—, sin embargo también conserva la independencia cartesiana res-

pecto de aquello sobre lo que versa que tenían las evidencias en que se basaba: como las sensaciones, todo ello podría ser tal como es, y el mundo ser muy diferente. Nuestras creencias pretenden representar algo objetivo, pero el carácter de su subjetividad no nos deja dar el primer paso para determinar si corresponden a lo que pretende representar.

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 objetivo y subjetivo. La razón de ello es que ambos dualismos tienen un origen común: un concepto de mente con sus estados privados y objetos.

He llegado al punto al que me estaba dirigiendo puesto que me parece que el cambio más prometedor e interesante que ocurre en la filosofía actualmente es que estos dualismos se están poniendo en cuestión de nuevas maneras o que se están redefiniendo radicalmente. Se da una buena ocasión para que sean abandonados al menos en su forma presente. Estamos asistiendo a la aparición de una visión revisada de la relación entre mente y mundo.

La acción se centra en el concepto de subjetividad, lo que está «en la mente». Empecemos con lo que sabemos o captamos cuando conocemos el significado de una palabra u oración. Es un lugar común en la tradición empirista que aprendemos nuestras primeras palabras (que al principio tienen la función de oraciones) —palabras como «manzana», «hombre», «perro», «agua»— por medio del condicionamiento de los sonidos o la conducta verbal a los fragmentos de materia adecuados que son del dominio público. El condicionamiento funciona mejor con objetos que interesan al que aprende y es difícil que tanto el profesor como el alumno los pasen por alto. Sin embargo, aquí está la cuestión: esto no es solamente una historia sobre cómo aprendemos a usar las palabras, sino que debe ser también una parte esencial de una explicación adecuada de a lo que se refieren las palabras y de lo que significan.

No es necesario decir que la historia completa no puede ser así de simple. Por otro lado, es difícil creer que esta clase de

interacción directa entre los usuarios del lenguaje y los acaecimientos y objetos públicos no sea una parte básica de la historia completa, la parte que, directa o indirectamente, determina en gran medida cómo las palabras del que aprende se relacionan con las cosas. Una consecuencia de ello es que los detalles de los mecanismos que constituyen las cadenas causales de hablante a hablante, y de objeto del que se habla al hablante y al que aprende el lenguaje, no pueden ellos mismos contribuir al significado y a la referencia. La captación de los significados se determina mediante los elementos terminales del proceso de condicionamiento, y se pone a prueba solamente mediante el producto final: el uso de las palabras que concuerdan con los objetos y situaciones apropiados. Esto se ve mejor al señalar que dos hablantes que «quieren decir lo mismo» por medio de una expresión no tienen por qué tener en común más que sus disposiciones a la conducta verbal apropiada; las redes neuronales pueden ser muy diferentes. Esta cuestión puede explicarse al revés: dos hablantes pueden ser iguales en los aspectos físicos «relevantes» y, sin embargo, significar cosas muy distintas con las mismas palabras debido a diferencias en las situaciones externas en las que se aprendieron esas palabras. Por ello, en tanto que se piense en lo subjetivo o mental como algo que sobreviene a las características físicas de una persona y nada más, los significados no pueden ser puramente subjetivos o mentales. Tal y como Hilary Putnam dijo (de forma engañosa), «los significados no están en la cabeza»⁵. La cuestión es que la interpretación correcta de lo que un hablante significa no se determina solamente por lo que está en su cabeza; depende también de la historia natural de lo que está en su cabeza. El argumento de Putnam depende de complicados experimentos mentales que algunos filósofos han encontrado poco convincentes. No obstante, se puede explicar mejor apelando directamente a hechos obvios sobre el aprendizaje del lenguaje y a hechos sobre cómo interpretamos las palabras y las lenguas con las que no esta-

⁵ Hilary Putnam, «The Meaning of "Meaning"», 227. [Pág. 144 de la traducción al castellano reseñada en la bibliografía.]

mos familiarizados⁶. Los hechos relevantes ya se han mencionado más arriba; en los casos más simples y más básicos, el significado de las palabras y de las oraciones se deriva de los objetos y las circunstancias en presencia de las cuales aquéllas se aprendieron. Si mediante el proceso de aprendizaje se ha condicionado a alguien a considerar verdadera una oración en presencia de fuegos, la misma será verdadera (usualmente) cuando haya un fuego presente; cuando se ha condicionado a alguien a considerar aplicable una determinada palabra en la presencia de serpientes se referirá a serpientes. Por supuesto que muchas palabras y oraciones no se aprenden de esta manera; pero son las que así se aprenden las que anclan el lenguaje al mundo.

Si los significados de las oraciones son proposiciones, y las proposiciones son los objetos de actitudes como la creencia, la intención y el deseo, entonces lo que se ha dicho sobre los significados debe mantenerse respecto de todas las actitudes proposicionales. Esta afirmación puede hacerse sin recurrir a proposiciones y otros supuestos objetos de las actitudes, ya que del hecho de que los hablantes en general sean capaces de expresar sus pensamientos en lenguaje, se sigue que en la medida en que la subjetividad del significado sea puesta en duda, en esa misma medida se pone también en duda la del pensamiento en general.

Lo que de estas consideraciones se desprende para la teoría del conocimiento es revolucionario. Si las palabras y los pensamientos tratan necesariamente, en los casos más básicos, acerca de las clases de objetos y acaecimientos que comúnmente los causan, no hay lugar para las dudas cartesianas sobre la existencia independiente de tales objetos y acaecimientos. Puede haber dudas, por supuesto. Pero no necesita haber nada de lo que estemos indubitablemente en lo cierto para que sea seguro que la mayoría de las veces estamos en lo cierto sobre la naturaleza del mundo. Algunas veces el escepticismo parece que se apoya en una simple falacia, la falacia de concluir del hecho de que no hay nada acerca de lo cual no

⁶ Esto lo defiendo en el Ensayo 1.

pudiéramos estar equivocados que podríamos estar equivocados acerca de todo. La segunda posibilidad se descarta si aceptamos que los significados de nuestras oraciones más simples se dan mediante las situaciones que por regla general nos llevan causalmente a considerarlas verdaderas o falsas, ya que considerar que una oración que entendemos es verdadera o falsa es tener una creencia. Continuando en esta línea, vemos que el escepticismo general acerca de lo que nos suministran los sentidos no puede ni siquiera ser formulado puesto que los sentidos y lo que éstos nos suministran no tienen ningún papel teórico central en la explicación de la creencia, el significado y el conocimiento si los contenidos de la mente dependen de las relaciones causales, sean éstas las que sean, entre las actitudes y el mundo. Esto no significa, por supuesto, negar la importancia del papel causal de los sentidos en el conocimiento y en la adquisición del lenguaje.

Es un accidente empírico que nuestros oídos, ojos, papilas gustativas y órganos táctiles y olfativos desempeñen el papel de intermediario en la formación de las creencias sobre el mundo. Las conexiones causales entre el pensamiento y los objetos y acaecimientos del mundo podrían haberse establecido de maneras completamente diferentes sin que ello produjera una diferencia filosófica significativa para los contenidos o la veracidad de las creencias perceptuales. Lo que es verdad es que ciertas creencias determinadas por las experiencias sensoriales frecuentemente son verídicas y por tanto a menudo nos proporcionan buenas razones para más creencias.

Si todo esto es verdad, la epistemología (como algo aparte, quizá, del estudio de la percepción, que ahora se concibe como algo sólo lejanamente relacionado con la epistemología) no tiene necesidad de «objetos de la mente» puramente privados o subjetivos, ya sea como datos sensoriales o experiencias no interpretados, por una parte, o como proposiciones totalmente interpretadas, por la otra. El contenido y el esquema, tal y como se hizo notar en la cita de C. I. Lewis, se dan conjuntamente; podemos eliminarlos a los dos juntos. Una vez que damos este paso, no hay objetos con respecto a los cuales surja el problema de la representación. Las creencias son verdaderas o falsas, pero no representan nada. Es bue-

no librarse de las representaciones y con ellas de la teoría de la verdad como correspondencia, puesto que es pensar que hay representaciones lo que engendra indicios de relativismo. (Es totalmente inocuo afirmar que las creencias y pensamientos verdaderos lo son debido a cómo es el mundo: ambos «representan» correctamente el mundo.)

Hay una gran cantidad de problemas difíciles acerca de la sensación y la percepción, pero estos problemas no son, como he dicho, fundamentales para la epistemología. La pregunta de qué es lo que experimentamos directamente en la sensación y cómo ello se relaciona con los juicios perceptuales, aunque sea tan difícil de responder como lo fue siempre, no puede suponerse ya que sea la cuestión central para la teoría del conocimiento. La razón ya la he expuesto: aunque la sensación tiene un papel crucial en el proceso causal que conecta las creencias con el mundo, es un error pensar que desempeña un papel epistemológico en la determinación de los contenidos de tales creencias. Al aceptar esta conclusión abandonamos el dogma clave del empirismo tradicional, lo que he llamado el tercer dogma del empirismo. Pero esto era de esperar; el empirismo es la concepción según la cual lo subjetivo («la experiencia») es el fundamento del conocimiento empírico objetivo. Lo que yo estoy afirmando es que el conocimiento empírico no tiene fundamento epistemológico, y no lo necesita⁷. Hay otro conocido problema que resulta transformado cuando caemos en la cuenta de que las creencias, los deseos y otras de las llamadas actitudes proposicionales no son subjetivas de la manera en que pensábamos que lo eran. El problema es cómo alguien conoce la mente de otro. Quizá sea obvio que si es correcta la explicación que he esbozado de nuestra comprensión del lenguaje y de su conexión con los contenidos del pensamiento, la accesibilidad a las mentes de los demás queda asegurada desde el principio. El escepticismo sobre la posibilidad de conocer otras mentes queda excluido. Pero reconocer esto no es responder a la pregunta sobre cuáles son las condiciones conceptuales que ponemos a

⁷ Véase el Ensayo 10.

las pautas de pensamiento que hacen posible que un intérprete pase del comportamiento observado al conocimiento de las actitudes intencionales de otra persona. Sin embargo, el hecho de que la naturaleza del lenguaje y el pensamiento sea tal que resultan interpretables garantiza que esa pregunta tenga una respuesta⁸.

No se debería suponer que si dejamos de sentirnos intimidados o seducidos por las dicotomías esquema-contenido y subjetivo-objetivo, todos los problemas de la epistemología se evaporarán. Pero los problemas que parecen sobresalientes cambiarán. No volverá a ser un reto responder al escéptico global, la búsqueda de fundamentos epistemológicos en la experiencia preconceptual se verá como algo sin sentido y el relativismo conceptual perderá su atractivo. Aún con ello seguirá habiendo preguntas de un interés igual o mayor o serán generadas por el nuevo punto de vista. La eliminación de lo subjetivo tal y como anteriormente se lo concebía nos deja sin fundamentos para el conocimiento y nos evita la necesidad de los mismos, pero entonces aparecen nuevos problemas que tienen que ver con la naturaleza del error, puesto que es difícil identificar y explicar el error si no se constriñe de alguna manera el holismo que viene ligado a un enfoque no fundamentalista. No es problemático que el conocimiento del mundo o de otras mentes sea posible, pero sigue planteándose la cuestión de cómo obtenemos tal conocimiento y de las condiciones que las creencias deben satisfacer para constituir conocimiento. Estas cuestiones no pertenecen tanto a la epistemología tradicional como a la naturaleza de la racionalidad. Son cuestiones que, como las cuestiones epistemológicas a las que sustituyen, no tienen una respuesta final; pero, al contrario que las cuestiones a las que reemplazan, merece la pena contestarlas.

En la actualidad está bastante extendido el conocimiento de muchas de las observaciones que he venido haciendo, pero no se ha apreciado plenamente el alcance de la revolución que suponen nuestras maneras de pensar en la filosofía.

⁸ Véase el Ensayo 1.

Parte de la razón de ello puede encontrarse en ciertas cosas que se han entendido mal relativas a la naturaleza del nuevo antisubjetivismo (como podemos intentar llamarlo). A continuación señalo tres de estas cosas que se han entendido mal.

1. Se ha aceptado la dependencia de los significados de factores externos a la cabeza mediante ejemplos más que mediante argumentos generales. Existe por lo tanto una fuerte tendencia a suponer que esa dependencia se limita a las clases de expresiones que se encuentran en los ejemplos: nombres propios, palabras de clases naturales como «agua» y «oro» y los indécicos⁹. Pero en realidad encontramos ese fenómeno en todas partes puesto que es inseparable del carácter social del lenguaje. No se trata de un problema local que haya de ser resuelto mediante un truco semántico, sino de un hecho perfectamente general acerca de la naturaleza del pensamiento y el habla.
2. Si los estados mentales como la creencia, el deseo, la intención y el significado no sobrevienen solamente a los estados físicos del agente, entonces, según se ha argumentado, las teorías que identifican los estados y acaecimientos mentales con los estados y acaecimientos físicos del cuerpo deber estar equivocadas. Esto lo sugiere la afirmación de Putnam de que «los significados no están en la cabeza» y Tyler Burge y Andrew Woodfield lo afirman explícitamente¹⁰. El argumento supone que si un estado o acaecimiento se identifica (quizá necesariamente si es un estado o acaecimiento mental) por referencia a las cosas fuera del cuerpo, entonces el estado o acaecimiento mismo debe ser parcialmente externo al cuerpo, o al menos no ser idéntico con ningún acaecimiento que se dé en el cuerpo. Esto es simplemente un error: se podría igualmente defender que un pedazo de

⁹ Tyler Burge, «Individualism and the Mental»; y Hilary Putnam, «The Meaning of "Meaning"».

¹⁰ Tyler Burge, «Individualism and the Mental», 111; y Andrew Woodfield, «Introduction», en Woodfield (ed.), *Thought and Object*, pág. viii.

piel quemada por el Sol no se localiza en el cuerpo de la persona que se ha quemado con el Sol (puesto que el estado de la piel ha sido identificado por referencia al Sol). De manera similar los estados mentales se caracterizan en parte por sus relaciones con los acontecimientos y objetos externos a la persona, pero esto no muestra que los estados mentales sean estados de algo más que la persona, ni que no sean idénticos a estados físicos.

3. Un tercer malentendido está estrechamente relacionado con el segundo. Se piensa que si la determinación correcta de los pensamientos de un agente depende, al menos en cierto grado, de la historia causal de esos pensamientos y el agente puede desconocer esa historia, entonces el agente no puede saber lo que piensa (y, *mutatis mutandis*, lo que quiere significar, pretender, etc.). Por ello, el nuevo antisubjetivismo se ve como una amenaza a la autoridad de primera persona —al hecho de que las personas por regla general saben sin recurrir a inferencias a partir de las evidencias lo que ellas mismas piensan, quieren y pretenden, y por tanto lo saben de un modo en que los otros no lo saben—. La reacción ha sido recurrir a maniobras diseñadas para aislar de nuevo los estados mentales de sus determinantes externos.

No se necesitan estas maniobras puesto que la autoridad de primera persona no está en peligro. Puedo no saber la diferencia entre una equidna y un puerco espín¹¹; como resultado de ello quizá denomine puerco espín a todas las equidnas con las que me encuentre. Sin embargo, dado el entorno en el que aprendí la palabra «puerco espín», mi palabra «puerco espín» se refiere a los puerco espín y no a las equidnas; esto es a lo que pienso que se refiere, y lo que creo que veo ante mí cuando sinceramente afirmo «esto es un puerco espín». Mi ignorancia de las circunstancias que determinan lo que quiero de-

¹¹ Ambos animales son muy parecidos al ser de similar tamaño y tener el cuerpo cubierto de púas. No obstante las equidnas, que viven en Australia, Tasmania y Nueva Guinea, son animales peculiares puesto que son ovíparos a pesar de ser mamíferos. [N. de la T.]

cir y pienso no tienden a mostrar que no sepa lo que quiero decir y pienso. Suponer otra cosa es mostrar cuán fuertemente estamos unidos a la idea de estados mentales subjetivos que podrían ser tal y como son independientemente del resto del mundo y de su historia.

Otra reacción a la imaginaria amenaza a nuestras vidas internas reales es conceder que las creencias y otros estados mentales, tal y como normalmente los identificamos, no son verdaderamente subjetivos, pero al mismo tiempo mantener que existen estados mentales internos (o «restringidos») similares a ellos que sí lo son. La idea podría ser, por ejemplo, que puesto que nada en mi estado interno o en mi comportamiento distingue entre puerco espín y equidna, lo que realmente creo cuando veo una equidna (o un puerco espín) es que lo que está ante mí es un animal con ciertas características generales —características que son de hecho compartidas por los puerco espín y las equidnas—. El problema es que ya que mi palabra «puerco espín» se refiere solamente a los puerco espín, aparentemente no sé lo que significo o quiero decir cuando digo «eso es un puerco espín». Esta solución tan poco atractiva es innecesaria puesto que se basa en una confusión acerca de lo que es «interno». Del hecho de que no haya ninguna diferencia física presente entre mi estado actual y el estado en el que estaría si hubiese querido significar «equidna o puerco espín» o «animal con tales y tales propiedades» en lugar de «puerco espín» y hubiese creído lo que entonces había querido decir, no se sigue que no haya ninguna diferencia psicológica. (Puede no haber ninguna diferencia física relevante entre la piel de alguien que está quemado por el Sol y la piel de alguien quemado por una lámpara solar, pero existe una diferencia, ya que un estado fue causado por el Sol y el otro no. En este respecto los estados psicológicos son como las quemaduras solares.) Así pues, nada impide decir que puedo saber lo que significo o quiero decir cuando uso la palabra «puerco espín» y lo que creo cuando tengo pensamientos sobre un puerco espín, aunque no pueda distinguir una equidna de un puerco espín. La diferencia psicológica, que es sólo la diferencia entre significar y creer que hay un puerco espín ante mí, y significar y creer que hay una criatura con ciertas

características comunes a los puerco espín y las equidnas, es exactamente la diferencia que se necesita para asegurar que sé lo que significo y lo que pienso. Todo lo que Putnam y otros han mostrado es que esa diferencia no tiene que estar reflejada en el estado físico del cerebro.

Inventar un nuevo conjunto de verdaderos estados psicológicos «internos» o «restringidos» no es, por tanto, una manera de devolverle la autoridad de primera persona a lo mental; es más bien al contrario. Sin embargo, sigue en pie la afirmación de que una ciencia sistemática de la psicología requiere estados del agente que pueden ser identificados sin referencia a su historia o a otras conexiones con el mundo externo. Se dice que de otra manera no se daría cuenta del hecho de que yo, que puedo referirme solamente a puerco espín con la palabra «puerco espín», no pueda explicar la diferencia entre un puerco espín y una equidna más de lo que podría hacerlo si yo (sin cambios físicos) quisiera en cambio significar «puerco espín o equidna».

Las perspectivas que pueda tener una psicología científica no son directamente relevantes para el tema de este ensayo, y por tanto podemos dejar de lado la cuestión de si hay estados internos de los agentes que pudieran explicar su comportamiento mejor que las creencias y deseos corrientes. Pero sí que es relevante considerar si hay estados mentales que tienen más derecho a ser denominados subjetivos que las actitudes proposicionales tal y como usualmente se conciben e identifican.

Se han dicho dos cosas al respecto. La más modesta (que se encuentra en la obra de Jerry Fodor, por ejemplo) es que podríamos considerar como verdaderos estados internos o solipistas a estados seleccionados de entre las actitudes usuales y las modificaciones de éstas. Los pensamientos sobre los puerco espín y equidnas serían eliminados, ya que los contenidos de tales pensamientos se identifican por relaciones con el mundo externo; pero serían admisibles los pensamientos acerca de animales que satisficieran determinados criterios generales (por ejemplo, los mismos que usamos al decidir si algo es un puerco espín)¹².

¹² Jerry Fodor, «Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology».

Tales estados internos, si existieran, serían calificados como subjetivos siguiendo casi todas las normas: se podrían identificar y clasificar sin referencia a objetos y acontecimientos externos; se podría acudir a ellos para utilizarlos como fundamentos del conocimiento empírico; y de manera plausible se les podría atribuir autoridad de primera persona. Sin embargo, parece claro que tales estados no existen, al menos si éstos pueden ser expresados con palabras. Los «rasgos generales» o «criterios» que usamos para identificar puerco espín son los de tener cuatro pies, una nariz, ojos y púas. Pero es evidente que los significados de las palabras que se refieren a esos rasgos y los contenidos de los conceptos que esas palabras expresan dependen tanto de la historia natural de cómo las palabras y los conceptos se adquirieron como sucede con los términos «puerco espín» y «equidna». No hay palabras, o conceptos ligados a las palabras, que no se comprendan e interpreten, directa o indirectamente, en términos de las relaciones causales entre las personas y el mundo (y por supuesto, de las relaciones entre palabras y otras palabras, conceptos y otros conceptos).

Llegados a este punto, podemos imaginar una propuesta al efecto de que existen criterios fenoménicos inexpresables a los cuales los criterios públicamente expresables se puedan reducir; y aquí es de esperar que los recuerdos de los fracasos pasados de tales fantasías reduccionistas sirvan para eliminar el pensamiento de que tal propuesta se podría llevar a buen término. Pero incluso dejando a un lado nostálgicos murmullos acerca de reducciones fenomenalistas, es instructivo hallar que el esfuerzo de hacer científica a la psicología se convierte en una búsqueda de estados proposicionales internos que pueden ser detectados e identificados aparte de sus relaciones con el resto del mundo, en gran medida igual que los filósofos anteriores buscaban algo «dado en la experiencia» que no contuviera ningún indicio necesario acerca de lo que sucedía ahí fuera. La motivación es similar en los dos casos: se piensa que un paso sólido, ya sea para el conocimiento ya sea para la psicología científica, requiere algo interno en el sentido de ser no-relacional.

La segunda y más revolucionaria afirmación fue que los estados mentales que se necesitaban para hacer una psicología

científica, aunque tuvieran de manera aproximada un carácter proposicional, no guardaban ninguna relación directa con las creencias, deseos e intenciones comunes¹³. De hecho estos estados se estipulan como aquellos que explican el comportamiento y por lo tanto son internos o subjetivos solamente en el sentido de caracterizarse exclusivamente mediante lo que físicamente se encuentra debajo de la piel. Pero no hay razón para suponer que las personas puedan decir sin observación cuándo se encuentran en tales estados y por tanto no hay razón alguna para llamarlos subjetivos.

En resumen, he hecho cinco observaciones interrelacionadas acerca de los «contenidos de la mente».

Primera, los estados de la mente como las dudas, deseos, creencias y esperanzas se identifican en parte por el contexto social e histórico en que se adquieren; a este respecto son como otros estados que se identifican por sus causas, como el sufrir de ceguera provocada por la nieve o el favismo (una enfermedad causada por el contacto con las habas).

Segunda, esto no muestra que los estados mentales no sean estados físicos de una persona; cómo describimos e identificamos los acaecimientos y los estados no tiene nada que ver directamente con lo que sean estos estados y acaecimientos.

Tercera, el hecho de que los estados de la mente, incluyendo lo que un hablante significa, se identifiquen por sus relaciones causales con objetos o acontecimientos externos es esencial para que sea posible la comunicación y hace que una mente sea en principio accesible a otra; pero este aspecto público e interactivo de la mente no tiende a disminuir la importancia de la autoridad de primera persona.

Cuarta, la idea de que hay una división básica entre la experiencia no-interpretada y un esquema conceptual organizador es un profundo error nacido de una imagen esencialmente incoherente de la mente como una espectadora pasiva pero crí-

¹³ Stephen Stich ha fomentado esta idea en *From Folk Psychology to Cognitive Science*.

tica de un espectáculo interno. Una explicación adecuada del conocimiento no apela a tales intermediarios epistemológicos como los datos sensoriales, los qualia o las sensaciones físicas. Como resultado, el escepticismo global acerca de los sentidos no es una posición que podamos sostener de manera coherente.

Finalmente, he argumentado contra la postulación de «objetos del pensamiento», tanto si éstos se conciben según el modelo de los datos sensoriales como teniendo carácter proposicional. Hay muchos estados mentales, pero su descripción no requiere que haya entidades fantasmagóricas que la mente contemple. Prescindir de tales entidades es eliminar más que resolver un número de fastidiosos problemas, puesto que no podemos preguntar cómo esos objetos pueden representar el mundo si no hay tales objetos, ni podemos quedarnos perplejos con la cuestión de cómo la mente puede estar en contacto directo con ellos.

¿Qué queda entonces del concepto de subjetividad? En lo que se me alcanza, hay dos rasgos de lo subjetivo tal y como se conciben clásicamente que siguen en su sitio. Los pensamientos son privados en el obvio pero importante sentido en que una propiedad puede ser privada, esto es, perteneciente a una persona. Y el conocimiento de los pensamientos es asimétrico en el sentido en que una persona que tiene un pensamiento por regla general sabe que lo tiene de una manera en que los otros no lo saben. Pero esto es todo lo que hay de subjetivo. Lejos de constituir una reserva tan aislada que sea un problema el que pueda producir conocimiento del mundo externo o ser conocido por los demás, el pensamiento es necesariamente parte de un mundo público común. No sólo pueden los demás saber frecuentemente lo que pensamos al notar las dependencias causales que proporcionan a nuestros pensamientos su contenido, sino que la propia posibilidad del pensamiento exige normas compartidas de verdad y objetividad.

4

¿Qué está presente en la mente?

Hay un sentido en el que cualquier cosa acerca de la cual pensemos está ante la mente mientras estamos pensando en ella. Pero hay otro sentido, bien conocido por los filósofos, según el cual solamente se puede decir que están ante la mente algunas cosas. Algunas de ellas están ante la mente cuando pensamos en cualquier cosa; con algunas otras sucede que no tenemos que pensar en ellas para que estén ante la mente. Ambas son cosas que se supone que están tan directamente ante la mente que es imposible identificarlas erróneamente, o solamente podemos identificarlas mal si no sabemos lo que pensamos; en esto difieren de los objetos físicos ordinarios o incluso de cualquier otro objeto, los cuales con facilidad se pueden identificar erróneamente. No podemos confundir estas entidades con otras simplemente porque son estos objetos los que fijan los contenidos de nuestros pensamientos. Quiero hablar sobre cosas de esta segunda clase.

Algunas de estas cosas que pertenecen a esta segunda clase son objetos, o así se les llama, de los deseos, creencias, intenciones, preocupaciones y esperanzas; son las proposiciones con las cuales tenemos las diversas actitudes, los pensamientos (tal y como Frege los denominó) que nuestras oraciones expresan. Debemos también incluir los constituyentes de las proposiciones, cosas como las propiedades y las relaciones.

Los objetos que he mencionado tienen una relación especial con la mente: solamente podemos conocerlos a través de la mente. Puesto que son abstractos, no se puede acceder a

ellos mediante los sentidos. No tienen poder causal, y por ello no pueden actuar sobre nuestras mentes, cerebros o nosotros mismos, y no puede ser que nuestras mentes, cerebros o nosotros mismos, actuemos sobre ellos.

Esta situación suscita la cuestión de qué clase de relaciones se piensa que nosotros o nuestras mentes tienen con esos objetos cuando decimos que están ante la mente. Tenemos muchas palabras para expresar nuestras relaciones con proposiciones: las captamos cuando comprendemos una oración, las consideramos, las rechazamos, deseamos que sean verdaderas, esperamos que no lo sean o pretendemos que sean verdaderas. Pero ¿qué clase de relaciones son éstas? Suenan a relaciones psicológicas, como si hubiera alguna clase de transacción mental entre nosotros y esas entidades. Pero ¿qué clase de transacción se puede dar con un objeto abstracto?

Se supone entonces que los objetos proposicionales de la mente y sus constituyentes tienen estas dos propiedades: *identifican* un pensamiento fijando su contenido; y *constituyen* un aspecto esencial de la psicología del pensamiento al ser captados o conocidos de otro modo por la persona que tiene el pensamiento. Es un problema cómo se pueden reconciliar estas dos propiedades. Mi tesis es que no se pueden reconciliar.

Aquí está la principal dificultad. Doy por supuesto que en la mayoría de casos conocemos lo que pensamos, aunque haya desviaciones de lo que sea un autoconocimiento total. (Uso «pensar» cubriendo todas las actitudes proposicionales.) Pero si un pensamiento se constituye en el pensamiento que es por medio del conocimiento que tiene la mente del objeto que lo identifica, entonces alguien sabe qué pensamiento está pensando solamente si conoce qué objeto tiene en mente. Sin embargo, parece que la idea de saber qué objeto uno tiene en mente no posee un significado claro. El problema radica en que la ignorancia de incluso una sola propiedad de un objeto puede, bajo las circunstancias apropiadas, ser considerada como un caso en que no se conoce qué objeto es. Ésta es la razón por la que los filósofos que han querido fundamentar el conocimiento en la identificación infalible de objetos han recurrido a objetos que, como las impresiones e ideas que propuso Hume, «son lo que parecen y parecen lo que

son», esto es, tienen todas y solamente aquellas propiedades que pensamos que tiene. Pero ¡ay!, tales objetos no existen. Todo objeto tiene una infinidad de propiedades lógicamente independientes, incluso objetos como los números de los cuales especificamos todas sus propiedades «esenciales»¹.

La discusión reciente sobre las creencias *de re* pone de relieve esta idea. Si estamos de acuerdo con Russell en que una persona no puede formar un juicio acerca de un objeto a menos que conozca qué objeto es, o (dicho de la manera que Russell prefería) que sea un objeto con el que la persona esté en contacto directo (*acquainted*), y en que en el caso de las proposiciones esto requiere que esta relación especial se dé entre quien juzga y cada parte de la proposición juzgada, entonces se plantea el siguiente problema acerca de las actitu-

¹ En *The interpretation of Frege's Philosophy*, 50, Michael Dummett describe la siguiente característica de los sentidos fregeanos: «Un sentido no puede tener ninguna característica no discernible mediante reflexión sobre, o deducción a partir de, lo que implica expresarlo o captarlo. Solamente pertenece al sentido de una expresión lo que es relevante para la determinación del valor de verdad de la oración en la que aparece; si hay algunos rasgos de su contribución a las condiciones de verdad de determinadas oraciones que no logremos captar, entonces no captamos plenamente su sentido, mientras que por otro lado, cualquier aspecto de su significado que no tenga que ver con las condiciones de verdad de las oraciones que lo contienen no es parte de su sentido. Por tanto, no puede ser que el sentido tenga toda clase de rasgos que no podamos detectar... Un pensamiento es transparente en el sentido de que, si lo captamos, al hacerlo sabemos de él todo lo que hay que saber acerca de cómo es en sí mismo.» Dummett limita los rasgos de los sentidos que no podemos dejar de detectar a sus «propiedades internas», pero no está claro cuál es el principio en que debe basarse la distinción de estas propiedades de otras que no caiga en una petición de principio.

Steve Yablo pregunta (en una comunicación privada) por qué no podríamos prestar atención a solamente algunas de las propiedades de los objetos, quizá las propiedades «esenciales». No, no si nos apegamos a la idea de que los contenidos del pensamiento se individualizan mediante el objeto, puesto que si dos objetos tienen las mismas propiedades «esenciales», pero difieren en alguna otra propiedad, son objetos distintos, y por ello los pensamientos deben diferir.

De este modo quizá no es el *objeto* lo que está ante la mente sino solamente algún *aspecto* de él. Podríamos equivocarnos respecto al objeto, pero no importaría. Pero entonces el aspecto es el objeto del pensamiento, no el objeto del cuál es aspecto; y reaparecen las mismas dificultades.

des: hay una receta para hacer pan de maíz que Juana cree que es fácil. La verdad de esta atribución no solamente requiere que haya una receta para hacer pan de maíz, sino que también parece requerir que Juana sepa de qué receta se trata —o que de alguna manera esté en contacto con ella—. Cuando Quine enfatizó por primera vez la distinción entre las oraciones de creencia *de re* y *de dicto* en términos modernos, se inclinó a considerar la referencia *de re* como una isla de claridad en el opaco mar de la intensionalidad². Desarrollos posteriores llevaron a un cambio de opinión. En «Intensions Revisited» escribió: «La noción de conocer o creer quién o qué es alguien o algo depende finalmente del contexto»³. Lo que condujo al cambio de opinión fue la dificultad en explicar la relación entre una persona y un objeto que justificaría la afirmación de que la persona sabe de qué objeto se trata. Se han hecho algunos intentos para clarificar la relación: Føllesdal afirmó que sólo los nombres «genuinos» podían caber en ella; Kripke habló de «designadores rígidos»; David Kaplan llamó «vividros» a los nombres elegidos⁴. Gareth Evans estudió el problema en profundidad; pensó, como otros, que la única relación psicológica que podría ser considerada como algo que proporciona el tipo requerido de «identificación fundamental» de un objeto era la identificación demostrativa. Solamente en tal caso se podría decir que el objeto formaba parte de la proposición que uno estaba considerando. Siguiendo a Russell, Evans concluyó que cuando una persona piensa que está considerando un pensamiento singular pero está usando un nombre no-referencial, entonces no existe proposición alguna que ella contemple y por ello no tiene un pensamiento. Si usa una oración que contenga un nombre no-referencial, no expresa ningún pensamiento⁵. Si, como yo, tiene usted problemas para confiar en los criterios para los casos genuinos de «identificación fundamental», apreciará por qué Rus-

² W. V. Quine, «Quantifiers and Propositional Attitudes».

³ W. V. Quine, «Intensions Revisited», 121.

⁴ Dagfinn Føllesdal, «Knowledge, Identity, and Existence»; Saul Kripke, «Naming and Necessity»; David Kaplan, «Quantifying In».

⁵ Gareth Evans, *The Varieties of Reference*.

sell limitó tales casos a situaciones en que la mente está directamente en contacto con sus objetos, algo que pensaba que solamente era posible con los datos sensoriales (y quizá con uno mismo).

Lo que subyace a algunos de estos intentos de caracterizar la relación especial entre la mente y sus objetos es, por supuesto, el impulso cartesiano de identificar una clase de conocimiento que está garantizado contra el error. Si esta búsqueda se combina con el supuesto de que el conocimiento consiste en que la mente esté en contacto psicológico con un objeto, entonces se habrán de encontrar objetos acerca de los que el error sea imposible —esto es, objetos que deben ser lo que parecen y parecer lo que son.

Simplemente tales objetos no existen. Ni incluso las apariencias son todo lo que pensamos que son. Ni los «aspectos» de los datos sensoriales pueden, si realmente son objetos, estar protegidos de una clase u otra de identificación errónea. La razón por la que Quine está en lo cierto en pensar que no podemos identificar «la» relación que constituye el conocimiento de qué objeto sea un objeto dado es simplemente que *cualquier* propiedad de un objeto puede, bajo condiciones adecuadas, ser considerado el identificador relevante.

Me he detenido brevemente en el problema tal y como se ha estudiado en relación con los nombres propios porque es ahí donde los filósofos parecen haber llegado más cerca de apreciar la naturaleza de la dificultad. Pero si mi diagnóstico es correcto, el problema realmente no tiene una relación especial con los nombres propios; es un problema perfectamente general acerca de los objetos de la mente. Si la mente puede pensar solamente teniendo la relación adecuada con algún objeto que ella pueda distinguir con certeza de todos los demás, entonces el pensamiento es imposible. Si una mente puede conocer lo que piensa solamente identificando de manera impecable los objetos ante ella, entonces debemos conceder que frecuentemente no sabemos lo que pensamos.

Si conllevan tantos problemas, ¿por qué suponemos que *hay* objetos proposicionales del pensamiento? Bien, por una parte ciertamente hablamos como si hubiera tales entidades: pensamos en pensamientos profundos, compartimos pensa-

mientos, descartamos o aceptamos creencias, abrigamos, consideramos, reflexionamos sobre, contemplamos ideas y proposiciones —podríamos alargar mucho más la lista—. Éste es el tipo de comentarios que con razón se nos ha enseñado a ver con sospechas ontológicas. Pero es mucho más difícil tomar a la ligera el problema en el que nos encontramos cuando intentamos dar una explicación sistemática de lo que significan determinadas oraciones —aquellas que usamos para atribuir pensamientos a los sujetos—, puesto que es difícil ver cómo oraciones usuales como «Pablo cree que el diamante Koh-i-noor es una de las joyas de la corona» puedan ser analizadas si no es relacionando a Pablo con alguna entidad identificada por la frase «que el diamante Koh-i-noor es una de las joyas de la corona», o quizá mediante la oración «el diamante Koh-i-noor es una de las joyas de la corona». Se han hecho muchos intentos de evitar hablar de «creencias» como de un término relacional, esto es, como relacionando dos objetos, alguien que cree y algo distinto. Pero por lo que sé ninguno de estos intentos ha tenido éxito.

Una de las propuestas es tratar el resto de la oración después de «cree» como un adverbio complejo. La lectura sería algo así como: «Pablo cree a la manera que-el-diamante-Koh-i-noor-es-una-de-las-joyas-de-la-corona». Pero esta idea es insatisfactoria puesto que nadie sabe cómo derivar el significado de un modificador adverbial como este de los significados de las palabras constituyentes y, sin embargo, es intuitivamente obvio que comprendemos las oraciones que siguen al «que» en las oraciones de creencia porque comprendemos las palabras constituyentes. Si los significados de tales oraciones contenidas (que proporcionan los «contenidos de las actitudes proposicionales») no se construyen a partir de los significados de sus partes, deberían ser aprendidas independientemente —como si fueran palabras nuevas, a veces muy largas—. Parece claro que ésta es una idea equivocada y probablemente imposible ya que cualquier oración declarativa puede figurar como una oración-contenido, y las hay de manera ilimitada y por ello presumiblemente no podrían ser aprendidas.

Se han hecho con mucha frecuencia otras propuestas en la misma línea, pero nadie ha mostrado cómo implementar tal

propuesta incorporándola a una teoría semántica desarrollada⁶. El principio que rige aquí parece claro: dejando a un lado las oraciones, se requiere la ontología cuando reconocemos una categoría gramatical a la cual debemos asignar una infinidad de expresiones. Debe haber una infinidad de objetos que pueden ser referidos o indéxicamente o por medio de descripciones; éstos son los dos mecanismos disponibles para permitir que con un vocabulario finito se pueda dar cuenta de una potencial infinidad de objetos. Los objetos nos permiten manejar adjetivos; los acaecimientos hacen lo mismo para al menos algunos adverbios. Los números hacen lo mismo cuando medimos⁷.

No hay entonces una alternativa plausible a considerar las oraciones de creencia como relacionales y por lo tanto no hay alternativa alguna a considerar la oración que da el contenido («El diamante Koh-i-noor es una de las joyas de la corona» más, quizá, la palabra «que» cuando está presente) como un término singular que, por su referencia a una entidad apropiada, especifica la creencia que es relevante⁸.

Aparentemente tenemos un dilema. Por una parte, nos encontramos con el hecho de que tener una creencia u otra actitud proposicional es estar en relación con un objeto de una clase u otra; por otra parte, nos encontramos con el hecho de que parece que no hay una explicación satisfactoria de la relación psicológica con el objeto apropiado en que debe encontrarse una persona para tener esa actitud. La dificultad en ofrecer tal explicación depende de la idea de que ya que una per-

⁶ Para una discusión de diversos intentos de eliminar el compromiso ontológico con «objetos del pensamiento» en el análisis de las oraciones de creencia y otras oraciones que atribuyen actitudes proposicionales, véase el Ensayo 7 en *Inquiries into Truth and Interpretation*.

⁷ Véase W. V. Quine, «Events and Reification».

⁸ Stephen Schiffer enfatiza correctamente que la posibilidad de la «semántica composicional» (la idea de que los significados de las expresiones complejas deben verse como una función de los significados de las expresiones de las cuales están compuestas) depende de manera crucial del análisis relacional de las oraciones que se usan para atribuir actitudes. Por razones que yo no comparto, Schiffer no tiene esperanzas de encontrar un análisis satisfactorio y por ello renuncia al proyecto de una semántica composicional. Stephen Schiffer, *Remnants of Meaning*.

sona conoce generalmente lo que piensa, debe estar directamente en contacto con, o ser capaz de una manera especial de identificar o individuar el objeto u objetos que definen (dan el contenido de) su pensamiento.

Si nos liberamos de concepciones previas, creo que es fácil ver dónde nos hemos equivocado. Del hecho de que un sujeto conozca lo que piensa y de que lo que éste piensa pueda ser fijado relacionándolo a él con un cierto objeto, no se sigue que el sujeto esté en contacto con el objeto o incluso que sepa alguna cosa sobre el mismo. He argumentado que alguien que atribuya un pensamiento a otro debe relacionar esa persona con algún objeto, y por tanto quien realiza la atribución debe, por supuesto, identificar un objeto apropiado, ya sea señalándolo ya sea describiéndolo. Sin embargo, no hay razón alguna por la que quien realiza la atribución deba estar en una relación especial con el objeto identificador; todo lo que tiene que hacer es referirse a él de la manera como se refiere a todo lo demás.

Especificamos el estado subjetivo de un sujeto relacionándolo con un objeto, pero no hay razón para afirmar que el propio objeto tenga un carácter subjetivo que sea «conocido» por el sujeto o que esté «ante la mente» del sujeto. Esta consecuencia ya estaba implícita en algunos análisis de las oraciones acerca de actitudes proposicionales, por ejemplo en la propuesta —considerada en un momento dado por Carnap y analizada por Quine— de que las oraciones de creencias se consideren como algo que relaciona al sujeto de la creencia con una oración de quien realiza la atribución. De este modo Quine, que mantiene que un gato puede tener una creencia, señala que no hay ninguna razón para suponer que un gato está en contacto con la oración «aquí hay comida» sólo porque se pueda decir correctamente que el gato cree que ahí hay comida. Para aquellos que dudan de que los gatos tengan actitudes proposicionales, lo mismo puede decirse remarcando que podemos identificar una creencia de Sebastián diciendo que cree que Nápoles está al norte de San Francisco, aunque no sepa ni una palabra de castellano. Lo que quiero señalar aquí no es que las oraciones de creencia relacionen a los sujetos de la creencia con oraciones, sino que esta conocida pro-

puesta supone que los objetos que se usan para identificar una creencia puede que no estén al alcance del sujeto que tiene la creencia. Una vez que aceptamos esta posibilidad, nos vemos libres de separar la necesidad semántica de objetos especificadores-de-contenido de la idea de que debe haber algún otro tipo de objetos con los que alguien que tiene una actitud está en contacto psíquico.

Veamos una analogía: consideremos el peso. Algunas cosas pesan más que otras; algunas cosas no pesan nada; de vez en cuando dos cosas pesan lo mismo. Una cosa puede pesar el doble que otra. Es sobre estas relaciones entre objetos sobre las que queremos informar cuando les asignamos pesos. Introducir una unidad de medida no altera la situación. Así, una libra (moneda) en el tiempo de Guillermo I pesaba 12 onzas; 20 peniques pesaban una onza; y un penique pesaba lo mismo que 32 granos de trigo sacados del centro de la espiga. Así pues, una libra pesaba 7.680 granos de trigo. Todo lo que deseamos decir sobre cuánto pesan las cosas puede ponerse en términos de estas comparaciones: por ejemplo, el diamante Koh-i-noor en su condición actual pesa casi lo mismo que $\frac{1}{3}$ de 15 peniques o que 490 granos de trigo.

Todas estas comparaciones pueden ser aburridas y en cualquier caso hacemos perspicuas las comparaciones relevantes usando números. Así, lo que conviene es establecer una manera de representar las relaciones entre objetos con respecto al peso usando directamente números. De esta manera, decimos que el diamante Koh-i-noor pesa 109 quilates o 345 gramos. Pero hablar de esta manera no requiere que incluyamos los quilates o los gramos en nuestra ontología. Los únicos objetos que necesitamos son los números y las cosas que pesan. Decir que el peso en quilates del diamante Koh-i-noor es de 109 no nos obliga a considerar los *pesos* como objetos: solamente es asignar el número 109 al diamante como una manera de relacionarlo con otros objetos en la escala de los quilates.

Visto de esta manera, hablar de cuánto pesan las cosas es relacional: se relacionan los objetos con los números, y por tanto, unos con otros. Pero nadie supone que los números sean en ningún sentido intrínsecos a los objetos que tienen

peso, o que de alguna manera formen «parte» de ellos. Lo que es básico son ciertas *relaciones* entre objetos: nosotros seguimos convenientemente la pista a estas relaciones asignando números a los objetos, y recordando cómo las relaciones entre los objetos se reflejan en los números.

Un aspecto importante de la medición numérica es que típicamente sólo algunas de las propiedades de los números son relevantes para el uso de informar sobre los pesos relativos. Así pues, es relevante que sean cuales sean los números que usamos para seguir la pista a los pesos éstos preserven las proporciones: si una cosa pesa el doble de gramos que otra, debe pesar también el doble de libras. Por otro lado, es irrelevante lo grande que sea un número de manera absoluta: medir en libras, granos de trigo y gramos produce diferentes números pero los mismos pesos *relativos* cuando se los compara con otros objetos⁹.

La analogía con las creencias es la siguiente. Así como al medir el peso necesitamos una colección de entidades que tengan una estructura en la que podamos reflejar las relaciones entre los objetos que tienen peso, en la atribución de estados de creencia (y otras actitudes proposicionales) necesitamos una colección de entidades que se relacionen de maneras que nos permitan seguirles la pista a las propiedades relevantes de diversos estados psicológicos y a las relaciones pertinentes entre ellos.

Al pensar y hablar de los pesos de los objetos físicos no necesitamos suponer que hay cosas tales como los pesos que los objetos los tengan. De manera similar, al pensar y hablar de creencias de las personas no necesitamos suponer que haya entidades tales como las creencias. Y tampoco tenemos que inventar objetos que sirvan como «objetos de la creencia» o como lo que está ante la mente o en el cerebro. Tal invención

⁹ Paul Churchland en *Scientific Realism and the Plasticity of Mind*, 105, ha comparado el atribuir creencias con la medición numérica. Pero comete el error de suponer que podemos interpretar la expresión «345 gramos» que aparece en «El diamante Koh-i-noor pesa 345 gramos» como un adverbio de «pesa», y de esta manera librarnos del problema ontológico (él diría lo mismo de las oraciones que siguen a «cree»). Pero como hemos visto, esta propuesta no puede apoyarse en una semántica seria.

es innecesaria porque las entidades que mencionamos para ayudar a especificar un estado de la mente no tienen que desempeñar ningún papel *psicológico* o epistemológico para la persona que se encuentra en ese estado, del mismo modo en que los números no tienen ningún papel físico. Como consecuencia, no hay ninguna razón para concluir del hecho de que el sujeto que piensa no conozca las entidades que usa para seguirles la pista a sus pensamientos, que *él* no pueda saber lo que piensa.

Lo que estoy proponiendo acerca de la naturaleza de las actitudes proposicionales se aplica directamente al problema que ha preocupado a un determinado número de filósofos en años recientes. Hay convincentes argumentos para mostrar que la determinación correcta de los contenidos de las creencias (y los significados y otras actitudes proposicionales) depende en parte de las conexiones causales entre el sujeto y los acontecimientos y objetos del mundo que el sujeto puede desconocer. Un ejemplo ya familiar es el caso de la Tierra Gemela de Putnam¹⁰. Se requiere que imaginemos que hay una tierra gemela de nuestra propia Tierra que es en todos los aspectos inmediatamente discernibles como nuestra propia Tierra. En la Getierra, mi réplica es molécula a molécula idéntica a mí (esto es, similar en los «aspectos relevantes») al haber estado expuesta a los mismos condicionamientos que yo y tener exactamente las mismas disposiciones lingüísticas que yo tengo. Con todo, uno de nosotros cree que es agua lo que ve ante él (yo) donde el otro cree que es gagua. La explicación es que mientras que hay agua en la Tierra, hay gagua en la Getierra aunque nadie haya detectado todavía la diferencia. Putnam argumenta que no hay diferencia interna o psicológica entre yo y mi gemelo. Ninguno de nosotros tiene razón alguna para decir que cree una cosa en lugar de la otra. En consecuencia (continúa Putnam), ninguno de nosotros sabe lo que cree. (No sé si creo que esto es agua o gagua.) Así pues, puede haber y quizás siempre haya factores no subjetivos —factores desconocidos para el sujeto— que deciden cuál es el «objeto

¹⁰ Hilary Putnam, «The Meaning of "Meaning"».

del pensamiento». Si la identidad del «objeto del pensamiento» depende parcialmente de factores que la persona que tiene el pensamiento ignora, ¿no se sigue que la persona puede no saber lo que piensa?

La respuesta es que no se sigue. Se seguiría si el objeto usado para identificar mi pensamiento fuera algo que tuviera que poder discriminar para saber lo que pienso. Pero éste es precisamente el presupuesto que he recomendado encarecidamente que deberíamos abandonar. Lo que veo ante mí creo que es agua; no hay ningún riesgo de que piense que es gagua, puesto que no sé lo que es el gagua. Si estoy en la Tierra, también creo que pienso que estoy viendo agua, y en esto estoy en lo cierto. Si, sin que me diera cuenta, se me transfiriera a la Getierra creería que el guaga era agua, lo que sería un error. En ambos casos sabría lo que creo. Por supuesto que mi gemelo en la Getierra se referiría a gagua con la palabra «agua». Él se refiere a gagua al usar la palabra «agua» porque piensa que es gagua. Esto es también lo que piensa que piensa. Así pues, ninguno de nosotros estaría equivocado sobre lo que pensaba. La posibilidad de error, o de no distinguir el propio estado mental, debida solamente a los elementos externos que ayudan a determinar ese estado mental, es inteligible solamente bajo el supuesto de que tener un pensamiento requiere una relación psicológica especial con el objeto que se usa para identificar ese estado mental¹¹.

La consecuencia de este ejercicio puede sorprendernos. Consiste en que los estados subjetivos no sobrevienen al estado del cerebro o del sistema nervioso: dos personas pueden estar en estados físicos similares¹² y ser aún disimilares en sus estados psicológicos. Por supuesto, esto no significa que los estados mentales no sobrevengan a los estados físicos, puesto que debe haber una diferencia física *en algún sitio* si es que los estados psicológicos difieren. Pero la diferencia física interesante puede no estar en la persona; del mismo modo en que

¹¹ Para una discusión más amplia de esta cuestión véase el Ensayo 2.

¹² Por supuesto, no estados físicos idénticos, como frecuentemente se ha dicho de manera errónea, puesto que los estados físicos no pueden ser, en un sentido relevante, idénticos a menos que se den en el mismo objeto.

la diferencia entre agua y gagua puede estar (según estamos suponiendo) en otro lugar.

Muchos filósofos, confrontados con este resultado, han decidido que las creencias y otras de las llamadas actitudes proposicionales no son, tal y como usualmente las identificamos, tan subjetivas como se pensaba que lo eran. Así, Jerry Fodor, al recomendar el «solipsismo metodológico» a los psicólogos, piensa que está recomendando que el psicólogo trate con estados *verdaderamente* subjetivos. Estados cuyas identidades se determinan solamente por lo que está en la cabeza. David Kaplan y Daniel Dennett han sugerido cosas similares¹³. La idea de Stephen Stich sigue la misma línea pero de manera más portentosa, puesto que piensa que debemos abandonar completamente la psicología de sentido común si queremos tener una ciencia seria del comportamiento humano¹⁴. Estoy de acuerdo en que los conceptos de creencia, deseo, intención y otros similares no son adecuados para una ciencia como la ciencia física y estoy de acuerdo en que una de las razones es que los estados mentales se identifican parcialmente sobre la base de sus causas y efectos. Pero dado que lo mismo vale para las acciones humanas tal y como normalmente las identificamos, parece que no hay ninguna posibilidad de tener explicaciones más «científicas» en este área (suponiendo que lo que queremos decir con explicación «científica» es la clase de explicación que esperamos de la física).

Pero sea lo que sea que pensemos sobre el futuro científico de la psicología de sentido común, no hay ninguna buena razón para afirmar que las creencias y las otras actitudes proposicionales no son verdaderos estados subjetivos. No hemos descubierto ninguna razón para decir que los sujetos generalmente no saben lo que piensan, o que no haya siempre una presunción a favor de que estén en lo cierto cuando están en desacuerdo con otros sobre lo que hay en sus mentes. Y ésta parece ser una prueba de subjetividad tan buena como cualquier otra que podamos tener.

¹³ Jerry Fodor, «Methodological solipsism considered as a Research Strategy in Cognitive Psychology»; David Kaplan, «Quantifying In»; Daniel Dennett, «Beyond Belief».

¹⁴ Stephen Stich, *From Folk Psychology to Cognitive Science*.

Hay dos cuestiones importantes que no he abordado y a las que dedicaré algunos comentarios. En primer lugar, no he dicho nada acerca de cuáles son los objetos a los cuales nos referimos cuando queremos especificar lo que alguien cree. Y, en segundo lugar, he argumentado que el descubrimiento de que lo que se cree está parcialmente fijado por hechos que el sujeto puede ignorar no muestra que éste no sepa lo que cree. Pero esto es sólo decir algo negativamente; ni siquiera indica *por qué* hay una presunción —usualmente correcta— de que lo sabe. Estas dos cuestiones están estrechamente relacionadas.

Entonces, primero, ¿cuáles son los objetos que nombramos o describimos para caracterizar los estados mentales? Bien, ¿qué es lo que *decimos*? Decimos cosas como «Pablo cree que el diamante Koh-i-noor es una de las joyas de la corona». Las palabras «cree que el diamante Koh-i-noor es una de las joyas de la corona» caracterizan un aspecto del estado mental de Pablo. La palabra de relación es «cree» y lo que sigue nombra al objeto, no del pensamiento —según he insistido— sino al objeto que indica de alguna manera regular el estado mental de Pablo. En el pasado sugerí que interpretáramos la palabra «que» en tales oraciones como un demostrativo que señala o se refiere a la siguiente preferencia del hablante que está haciendo la atribución. Lo que sigue da la idea: Pablo cree lo que yo creería si yo asertara sinceramente lo que digo a continuación:

El diamante Koh-i-noor es una de las joyas de la corona.

Dado que no estoy intentando hacer observaciones semánticas sutiles, podemos tomar como «objeto» tanto mi preferencia real como la oración de la cual es una preferencia (relativizada a un tiempo y a un hablante). Si se desea, incluso se puede tomar como objeto una proposición. Puesto que las preferencias, las oraciones y las proposiciones están relacionadas estrechamente, lo más probable es que si una de estas cosas sirve, las otras también sirvan. Pero las preferencias tienen ciertas ventajas iniciales puesto que no son abstractas y están ligadas a un hablante, a un tiempo y a un contexto. Así pues, supondré que nos hemos decidido por las preferencias —las

preferencias mismas que se producen al atribuir actitudes— como objetos que sirven para individuar e identificar los distintos estados mentales.

Tales objetos sirven muy bien a su papel de caracterizar los estados mentales. Ciertamente, hay tantas preferencias distintas a nuestra disposición (o preferencias potenciales) como estados mentales que seamos capaces de distinguir al atribuir las; de hecho hay más. Incluso las dificultades especiales que tienen que ver con las expresiones demostrativas o indécicas no son insolubles si deseamos la idea de que estamos intentando identificar objetos con los cuales quien tiene actitudes está en misterioso contacto psíquico.

Las preferencias se relacionan las unas con las otras de forma muy parecida a como lo hacen las creencias: mediante relaciones de consecuencia y de apoyo evidencial. Las preferencias, como las creencias, son verdaderas o falsas. Aparte de las complicaciones debidas a los elementos indécicos, identificamos una creencia profiriendo una oración que tiene las mismas condiciones de verdad que la creencia para identificar la cual se utiliza aquella preferencia. Tampoco esto es nada sorprendente puesto que frecuentemente expresamos nuestras creencias profiriendo oraciones con las condiciones de verdad de la creencia que estamos expresando.

Llegados a este punto, puede usted pensar que me estoy acercando peligrosamente a la recuperación de la misma teoría que provocó todos los problemas al comienzo ya que, ¿por qué no decir que, puesto que las preferencias han determinado los significados, y son los significados los que se emparejan con los estados de creencia, los objetos que estamos nombrando cuando proferimos las oraciones que dan el contenido en el contexto de la atribución de creencias son precisamente los significados de estas oraciones, es decir, son proposiciones? Esto quedaría así: cuando un francés atribuyera el mismo estado mental a Pablo que yo, ambos estaríamos nombrando el mismo objeto; esto no sería el caso con la teoría que he considerado anteriormente, puesto que la oración relevante de la persona francesa no era la mía.

En el contexto presente solamente tengo objeciones marginales a dar este paso, puesto que no es el propio paso lo que

provoca el problema. (Puede producir otra clase de problemas como que los significados y las proposiciones sean el tipo de cosa con trampa que resulta ser.) El problema que hemos tratado proviene de la identificación del objeto que se usa para caracterizar un estado mental con un objeto que la mente «conoce» o con el que está «en contacto directo», un «objeto del pensamiento». Si evitamos esta identificación, podemos evitar los problemas que he estado explorando. Pero si evitamos esta identificación tampoco ganamos nada con el paso de las preferencias a las oraciones y a los significados o proposiciones. De modo que está bien atenerse a las preferencias para ayudar a tener bien presente mi observación principal. No habrá entonces ningún riesgo de suponer que en general el sujeto de la creencia está en contacto directo con los objetos que se usan para caracterizar sus estados mentales. No nos debería preocupar que el francés y yo usemos preferencias distintas para caracterizar el mismo estado mental de Pablo; ello es como medir el peso en quilates u onzas: diferentes conjuntos de números hacen desempeñar el mismo cometido.

Esta última idea dirige nuestra atención hacia un tema mayor. Cuando usamos números para registrar las relaciones entre pesos, longitudes o temperaturas no mostramos tendencia alguna a responder al hecho de que diferentes conjuntos de números sirvan tan bien como otros para registrar todo lo que es empíricamente relevante quejándonos de que los pesos, longitudes o temperaturas no son «reales». Sabemos que no hay contradicción entre decir que la temperatura del aire es de 32 grados Fahrenheit y decir que es de 0 grados centígrados; no hay nada en este «relativismo» que muestre que las propiedades que se han medido no sean reales, aunque curiosamente se ha llegado a esta conclusión en varias ocasiones. Por ejemplo, John Searle encuentra incomprensible que al mismo pensamiento (o preferencia) de una persona se le puedan dar dos interpretaciones completamente distintas¹⁵. Sin embargo, a la luz de las consideraciones que se han ofrecido, esto meramente equivale al reconocimiento de que más de

¹⁵ John Searle, «Indeterminacy and the First Person».

un solo conjunto de las preferencias de una persona podría captar por igual los contenidos de los pensamientos o el habla de alguna otra persona. Del mismo modo en que los números pueden captar todas las relaciones empíricamente significativas entre los pesos o las temperaturas de infinidad de maneras distintas, las preferencias de una persona pueden captar todos los rasgos significativos de los pensamientos y el habla de otra persona de maneras distintas. Este hecho no pone en peligro la realidad de las actitudes o de los significados de los que se informa de esas múltiples formas. Jerry Fodor es otro de los filósofos que piensa que el holismo o la indeterminación de la traducción asociada a él suponen una amenaza para el realismo con respecto a las actitudes proposicionales¹⁶. Éste es el mismo error. La indeterminación de la traducción significa que conjuntos distintos de preferencias (u oraciones, como dice Quine) pueden interpretar igualmente bien el lenguaje de un hablante (o sus pensamientos); esto no quiere decir que los estados mentales del hablante o el sujeto de los pensamientos así captados sean de alguna manera vagos o irreales. El holismo mantiene que los contenidos del habla y el pensamiento dependen de las relaciones entre significados y entre pensamientos. Pero de nuevo no hay nada en esta afirmación que suponga una amenaza para la realidad de los estados que se relacionan de estas maneras. La amenaza para la realidad de los pensamientos y los significados que Searle y Fodor piensan que detectan es de hecho una amenaza bastante distinta, una amenaza para el supuesto de que las entidades que se utilizan para identificar los pensamientos y los significados se «captan» de alguna manera mediante la mente, y por tanto, si esas entidades son distintas, los propios pensamientos deben ser distintos. Es como si la «diferencia» entre que una cosa mida un metro y que mida 100 centímetros con una barra de medir fuera una diferencia de la propia barra de medir.

Voy a abordar ahora la última cuestión; dado que una concepción correcta de la manera en que los «objetos del pensa-

¹⁶ Jerry Fodor, *Psychosemantics*.

miento» determinan la identidad de los diversos estados mentales no supone una amenaza para la autoridad de primera persona, ¿qué es lo que explica tal autoridad?

Una de las dificultades la creó la gramática y las falsas inferencias provenientes de ella; ellas llevaron a la idea de un objeto interno que se conoce de manera privilegiada. Este obstáculo fue eliminado al mostrar que no había razón alguna para suponer que existan tales objetos. Una segunda dificultad provenía de la convicción de que lo verdaderamente subjetivo —aquello de lo cual la persona tiene conocimiento privilegiado— puede no deber ninguna de sus cualidades subjetivas a las conexiones con el mundo externo. He argumentado que aunque un intérprete deba examinar las relaciones entre la mente que está interpretando y su entorno si es que quiere acertar, ello no prejuzga el autoconocimiento del sujeto cognoscente.

Habiendo eliminado los obstáculos, queda poco por decir acerca de cómo sabemos qué es lo que pensamos. En los casos interesantes y originales que causan perplejidad no *hay manera* de saberlo, puesto que no hay evidencias que se puedan buscar, ni ningún objeto interno que se pueda escrutar, ni hipótesis que compitan para ser sopesadas. Esta cuestión queda clara al considerar la siguiente situación: supongamos que digo, «creo que el diamante Koh-i-noor es una joya de la corona». O simplemente no lo supongamos puesto que lo digo. Y supongamos, como es el caso, que sé lo que significan las palabras que he acabado de proferir y que estoy haciendo una afirmación sincera. Finalmente, supongamos que ambos, usted y yo, estamos de acuerdo en estos puntos, es decir, que he proferido estas palabras y que al hacerlo estaba profiriendo sinceramente palabras que comprendía. De estas suposiciones se sigue que sé lo que creo, pero no se sigue que usted sepa lo que yo creo. La razón es sencilla: usted puede no saber lo que yo quiero decir. Su conocimiento de lo que mis palabras significan se ha basado en evidencias y en inferencias; usted probablemente supone que está en lo cierto, y probablemente lo está. Sin embargo, es una hipótesis. Por supuesto puede ser que yo tampoco sepa lo que quiero decir con esas palabras. Pero existe la *presunción* de que lo sé, ya que no tiene sentido

suponer que *de manera general* estoy equivocado sobre lo que mis palabras significan; la presunción de que de manera general no estoy equivocado sobre lo quiero decir le es esencial al hecho de que poseo un lenguaje, es decir, al hecho de que se me pueda interpretar. Recurriendo a una idea familiar aunque frecuentemente mal entendida: al enunciar las condiciones de verdad de mi preferencia de la oración «El diamante Koh-i-noor es una joya de la corona» no puedo hacer nada mejor que decir que es verdadera si y sólo si el diamante Koh-i-noor es una joya de la corona. Si digo esto, profiero una tautología, pero si usted da las condiciones de verdad de mi preferencia con las mismas palabras, usted estará haciendo una afirmación empírica, aunque probablemente verdadera¹⁷.

De acuerdo con Dummett, Brentano «rechaza admitir que un acto mental... tenga un objeto interno distinto de su objeto externo, es decir, una representación mental... mediante la cual el objeto externo se presente a la mente»¹⁸. Dummett señala que esto le deja a Brentano con el problema de los pensamientos (o pensamientos aparentes) acerca de objetos que no existen; un problema, remarca Dummett, que Brentano «no resolvió satisfactoriamente». Pero el problema se resuelve fácilmente si desechamos la idea de que hay objetos internos o representaciones mentales en el sentido que se requiere. No hay necesidad de suponer que si no existen tales objetos internos, solamente quedan los objetos externos para que nos ayuden a identificar los diversos estados mentales. El hecho simplemente es que tenemos los recursos necesarios para identificar los estados mentales, incluso si se da el caso de que tales estados mentales se dirigen, como nos gusta decir, a objetos que no existen, puesto que podemos hacer esto sin suponer que haya objetos del tipo que sea ante la mente.

¹⁷ Este párrafo resume la conclusión del Ensayo 1 de este volumen.

¹⁸ Michael Dummett, *The Interpretation of Frege's Philosophy*, 57. Dummett nos remite a Franz Brentano, *Psychology from an Empirical Standpoint*, 79.

5

El indeterminismo y el antirrealismo

El antirrealismo es 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. De este modo, debería considerarse a Parménides como un antirrealista temprano y extremo, ya que afirmó que el Uno no cambiante, homogéneo y globular era todo lo que era real sobre la base de que era el único objeto de conocimiento posible. Debe considerarse también, por supuesto, que Platón era un antirrealista, ya que mantuvo que el mundo físico, y todo lo que hay en él, es menos que real porque no puede ser conocido. La mayoría de los ismos reductivos deberían considerarse como formas de antirrealismo: el idealismo, el pragmatismo, el empirismo, el materialismo, el conductismo, el verificacionismo. Cada uno de ellos trata de recortar la realidad para que encaje en su epistemología. Cada una de estas posiciones ofrece consuelos: los objetos comunes como las mesas o los servicios de mesa son reales, se nos dice, pero existen solamente en la mente, los objetos físicos no son nada más que posibilidades permanentes de sensación; los estados mentales no son nada más que patrones de comportamiento; los fenómenos intencionales no son nada más que acaecimientos físicos y objetos, etc. Se nos permite la terminología de nuestra vieja ontología en tanto que estemos de acuerdo en aceptar solamente lo que se puede cons-

truir a partir de las entidades o experiencias que podamos conocer con certeza. Pero estos regalos al escepticismo no deberían engañarnos; los antirrealismos siguen siendo filosofías de uvas verdes. Su lema es: si no puedes alcanzar las uvas (en un sentido apropiado), no es que estén verdes, es que, para empezar, nunca existieron. Por razones que mencionaré a continuación, algunos antirrealismos se explican mejor en términos de limitaciones epistémicas del concepto de verdad. Así pues, puede sostenerse que cuando nuestras facultades epistémicas son deficientes con respecto a la determinación de la verdad o la falsedad de alguna oración, debería considerarse que la oración no tiene valor veritativo, o deberíamos emplear algún sentido reducido de verdad. El resultado es el mismo: lo real o lo verdadero se recorta a la medida de una forma de conocimiento que uno aprueba.

No veo al realismo como la alternativa al antirrealismo. Podemos rechazar una u otra versión de antirrealismo sobre la base de que los argumentos a su favor fallan sin que tengamos que aceptar una vaga posición llamada realismo. Una caracterización común de realismo es la siguiente: hay algo en o acerca del mundo que hace que nuestras preferencias o aseveraciones o pensamientos sean verdaderos cuando lo son, tengamos o no el poder de determinar su verdad. La mayoría de las críticas a esta formulación se apoyan comprensiblemente en la dificultad de explicar qué significan en este contexto «poder» y «determinar». Pero no creo que necesitemos adentrarnos en esas profundidades ya que no creo que comprendamos la afirmación previa de que hay algo en o acerca del mundo que hace que nuestros pensamientos o aseveraciones sean verdaderos cuando lo son. El hecho es que nadie ha sido nunca capaz de decir de una manera no trivial qué clase de «cosa» es lo que hace verdadera a una oración (u a otro portador de verdad). Algunas oraciones, preferencias o creencias son verdaderas, y hay muchas cosas en el mundo, pero no explica nada decir que las cosas «hacen» verdaderos a los vehículos de verdad. El resultado es que los realistas mantienen el concepto de verdad, pero no pueden explicar cómo la realidad lo explica. El fallo de las teorías de la correspondencia en sustanciar el concepto de verdad demuestra, al mismo tiem-

po, la vaciedad de las caracterizaciones del realismo en términos de correspondencia.

Aquí discutiré esas versiones del antirrealismo que suscitan dudas sobre la realidad de los estados y acaecimientos mentales a la luz de la indeterminación de la traducción o la interpretación. Me ocuparé igualmente de aquellos que argumentan que *si* aceptamos la indeterminación de la interpretación, entonces deberemos tener dudas acerca del estatuto de las actitudes proposicionales —o, al menos, del estatuto de las atribuciones de actitudes proposicionales—. Dado que acepto la tesis de la indeterminación de la interpretación, me sentiría ofendido si averiguara que implica alguna clase de antirrealismo. Pero, tal y como intentaré explicar, no creo que lo implique. Pienso que ni el indeterminismo muestra que las actitudes proposicionales no sean plenamente reales (sea lo que sea lo que esto signifique) 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.

Son tres las consideraciones que parece que impiden aceptar esta tesis. La primera es una forma de cientifismo. Las actitudes proposicionales no parece que sean apropiadas para incorporarse a una concepción científica y unificada del mundo. Así, Quine ha dicho que «el lenguaje esencialmente dramático de las actitudes proposicionales» no tiene lugar en una ciencia seria¹, o que nuestro vocabulario mentalista está «testarudamente en desacuerdo con los patrones científicos». Creo que Quine está en lo cierto en mantener que el vocabulario mentalista no puede reducirse o incorporarse al vocabulario de la física, o a cualquier otra de las llamadas ciencias «duras». Una razón de esta irreducibilidad proviene del hecho de que cualquier explicación correcta de las creencias y otras actitudes de un agente debe reconocer el elemento normativo inherente en los contenidos cognitivos. Una actitud tiene de

¹ *World and Object*, 219. [Pág. 228 de la traducción al castellano reseñada en la bibliografía.]

terminadas relaciones lógicas con otras; estas relaciones, aunque distorsionadas y trastornadas por los límites de nuestras capacidades, sirven, sin embargo, para ubicar, y de este modo identificar, los contenidos de nuestros pensamientos. Cuando tratamos el mundo como algo sin mente, tal y como pasa con las ciencias naturales, nada se corresponde con esta dimensión de lo mental. Una propiedad que menos frecuentemente se ha señalado 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 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 encajarse en un sistema cerrado de leyes. Hay otra consideración (más debatida) que milita contra la reducción nomológica o definicional de los conceptos mentales a los de la física, a saber, el hecho de que las actitudes proposicionales y otros acaecimientos y estados relacionados se identifican en parte en términos de sus relaciones causales o de otro tipo con acaecimientos ajenos en tiempo y en lugar al agente que ellos caracterizan. Así pues, se sostiene, correctamente en mi opinión, 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. También pienso que 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.

Estas observaciones no cumplen otra función que la de recordar por qué muchos filósofos, incluido yo mismo, estamos convencidos de que la psicología, en tanto que incluya conceptos como acción intencional, creencia, percepción y las actitudes afectivas, no puede formar parte de la física o de cualquier otra ciencia «natural». La cuestión es si la irreducibilidad definicional y nomológica de las actitudes proposicionales es una razón para dudar del estatuto objetivo de ellas.

Quine ha dicho que está de acuerdo con mi versión del monismo de ejemplares, que sostiene que cada objeto o acaecimiento que esté identificado en términos mentales es idéntico a un objeto o acaecimiento 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. Esta posición, que yo llamo monismo anómalo, se acerca a reconciliar la mayor parte de lo que Quine ha afirmado acerca de las actitudes proposicionales, aunque no todo ello. 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. El monismo anómalo no sugiere que los acaecimientos y estados mentales sean meramente proyectados por el que atribuye algo a un agente; por el contrario, mantiene que los acaecimientos mentales son tan reales como los acaecimientos físicos, al ser idénticos a ellos, y que las atribuciones de estados son también tan objetivas. La descripción de Quine de las atribuciones de actitudes como retratos dramáticos no implica que no haya nada que retratar.

El hecho de que el vocabulario mental no encaje de manera que pueda ser incluido en ciencias como la física o la fisiología no puede por sí mismo tomarse como una crítica a la realidad de los estados, acaecimientos y objetos para describir

los cuales se utiliza tal vocabulario. Una física perfeccionada debe comprender cada objeto y acaecimiento, 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.

Todavía no he abordado directamente la cuestión de la indeterminación de la traducción o la interpretación. Si se mantiene la tesis de la indeterminación, ¿no implica esto que las actitudes proposicionales no llegan a ser reales? Puede pensarse que Quine insinúa esto cuando dice: «La tesis de Brentano de la irreducibilidad de las locuciones intencionales es inseparable de la tesis de la indeterminación de la traducción», ya que «aceptar literalmente el uso intencional ... es postular que las relaciones de traducción son algo objetivamente válido, pero pese a ello indeterminado en principio respecto de la totalidad de las disposiciones verbales»². ¿No significa esto decir que aceptar las locuciones mentales implica postular algo para lo cual no hay evidencia empírica y, por tanto, algo para lo cual no tenemos razón para creer que exista?

De cualquier modo, sea o no sea el caso que las ideas de Quine impliquen que haya algo no real en las actitudes proposicionales, es una idea compartida por un cierto número de filósofos que el indeterminismo socava la realidad de los estados mentales. Fodor y otros han pensado que esto constituye una reducción al absurdo del indeterminismo. Daniel Dennett, por otro lado, acepta el indeterminismo, pero está de acuerdo en que ello resta realidad a los estados mentales. Yo discuto esta inferencia. No veo la manera de dejar a un lado el indeterminismo, pero creo que deja intacta la realidad de lo mental.

En primer lugar voy a defender los pasajes de Quine que acabo de citar en los que se parece afirmar que el «uso intencional» no tiene contenido empírico. Hay una confusión que puede en parte ser debida al propio Quine pero que es fácil de resolver. Si por «uso intencional» Quine quiere decir hablar de significados y proposiciones como cosas hipostasiadas

² *Ibíd.*, 221. [Pág. 230 de la traducción al castellano reseñada en la bibliografía.]

por los filósofos, es verdad que él las rechaza como cosas que no tienen contenido empírico o explicativo. Pero no se sigue que las atribuciones comunes de actitudes, incluyendo las interpretaciones del habla, sean vacías. Incluso podemos preguntarnos: ¿cómo podía ser de otra manera si el habla sobre lo mental es, como Quine acepta, algo inevitable?

Creo que hay dos cosas que ante todo nos tientan a suponer que no puede haber ninguna indeterminación en lo que queremos decir mediante nuestras palabras o en lo que pensamos. Una de ellas es la autoridad de primera persona, el hecho de que se presume que sabemos lo que pensamos de una manera en la que nadie más puede hacerlo. La otra la engendra la semántica de las oraciones que se usan para atribuir actitudes. Voy a abordar primero esta segunda cuestión y por simplicidad me atenderé a las oraciones de creencia. Es obvio que potencialmente existe un número infinito de creencias que se podrían asignar a alguien y está claro que potencialmente existe un número infinito de oraciones a nuestro alcance, puesto que podemos crear una oración bien formada poniendo cualquiera de una infinidad de oraciones después de las palabras «Inés cree que...». Pero los únicos mecanismos semánticos viables que están disponibles para acomodar esta situación implican tomar «cree» (o «cree que») como un verbo de relación, y tratar lo que sigue como un término singular o una descripción. Los términos singulares no pueden ser todos nombres propios no estructurados; no hay los suficientes. Así, si lo que sigue a «cree» o «cree que» es un término singular, debe haber ahí un mecanismo demostrativo; o bien debemos tener descripciones. Pero en cualquier caso, debe haber almacenadas infinitas *entidades* para escoger. Si bien una vez que se haya escogido la entidad apropiada, como sujetos que atribuimos actitudes habremos asignado un contenido a la creencia.

Quizá 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 por supuesto, sobre las otras actitudes), y también dada la suposición natu-

ral de 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 especificaría el mismo estado de creencia igualmente bien, ¿qué entidad deberíamos decir que capta el sujeto, o cómo podemos decir que el sujeto sabe lo que cree?

Este acertijo puede resolverse desechando la idea de tratar los objetos gramáticos en oraciones de creencia como términos que nombran objetos psicológicos reales, objetos que quien tiene una creencia, conoce, considera o capta. El único *objeto* que se requiere para la existencia de una creencia es la de un sujeto que crea algo. Tener una creencia no es como tener un gato favorito, es estar en un estado, y estar en un estado no requiere que haya una entidad llamada estado en el que uno está. Lo único que se necesita para la verdad de una atribución de una actitud es que el predicado que se emplee sea verdad respecto de la persona que tiene esa actitud.

Se le debe atribuir a Dummett el mérito de caracterizar el antirrealismo en términos de la verdad y no de la referencia, en términos de oraciones y no de nombres o descripciones. Es así puesto que la cuestión real acerca de las oraciones que atribuyen estados mentales no es ontológica; las creencias no son entidades, ni tienen que ser objetos los «objetos de las creencias». La cuestión real es si es o no el caso que las atribuciones de actitudes son objetivamente verdaderas o falsas³.

La atribución de actitudes es análoga en muchos aspectos a la medición con diversas magnitudes. Podemos asignar números para captar las tallas, pesos y velocidades de los objetos, si los objetos exhiben un patrón del tipo apropiado. No suponemos que los objetos tengan entidades empíricas que se llamen pesos o tallas o velocidades. Como señaló Carnap hace mucho tiempo, no deberíamos pensar en la oración «esta caja pesa 8 libras» como si identificara dos entidades lla-

³ Por supuesto, está la cuestión de si existen los significados y las proposiciones, pero esto lo entiendo como un asunto de decisión. Si existen, son entidades abstractas y no necesitan más que ser definidas. Una vez definidas, todavía se ha de mostrar si tienen algún papel explicativo o descriptivo.

madas «el peso de esta caja» y «8 libras», sino como que equi-para el peso de la caja en libras con el número 8. Así pues, la ontología que se requiere consiste en objetos que tienen pesos, y en números. Los números no son parte de los objetos con peso; no pertenecen al mundo empírico sino a nosotros que somos los que los necesitamos para captar ciertas relaciones entre objetos. De la misma manera, las entidades a las cuales relacionamos a los sujetos cuando les atribuimos creencias y otras actitudes proposicionales, no están en los sujetos —no están en sus mentes o ante sus mentes—. Tener una creencia es solamente ejemplificar una propiedad, el que un cierto predicado se le aplique a uno con verdad; pero para tener suficientes predicados para todas las creencias que queramos distinguir, deberemos construir los predicados usando verbos de relación y rellenando uno de los lugares de la relación con una referencia a un objeto que se haya sacado de algún almacén potencialmente infinito. Un predicado así es « x cree que la nieve es blanca».

Una de las virtudes de esta analogía es que aclara por qué diferentes asignaciones de objetos pueden captar toda la información relevante acerca de una situación sin comprometer la verdad o la «realidad» de la situación. Nadie piensa que el hecho de que podamos registrar el peso ya sea en libras o en kilogramos muestre 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 kilos. Supóngase que el hecho es que C pesa el doble que D; entonces el número asignado para medir el peso de C debe ser el doble que el número asignado para medir el peso de D, tanto si los pesos se dan en libras o como en kilos. 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. Los números pueden hacer esto de un sinfín de maneras.

Otra de las virtudes de esta analogía es que aclara por qué Quine está justificado a insistir, contra Chomsky y otros, en

que la indeterminación de la traducción es distinta de la infra-determinación de las teorías ante todas las evidencias posibles. Ciertamente es un problema el de si puede darse la infra-determinación en el sentido de Quine, puesto que ésta requiere que haya teorías empíricamente equivalentes pero incompatibles⁴. 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.

¿Cuáles son los objetos que podemos usar para captar las actitudes? Claramente deben constituir aproximadamente un campo tan complejamente configurado como lo son las propias actitudes individuales. Las características más sobresalientes de las actitudes individuales 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 que la nieve es blanca es verdadera si y sólo si la nieve es blanca)⁵.

Nuestras *oraciones* son entidades que tienen este tipo de propiedades que se requieren, y no está claro que cualquier otro conjunto de entidades sirviera igualmente bien (excepto las preferencias)⁶. En cualquier caso, hay un sentido en el que nuestras atribuciones normales de actitudes utilizan oraciones de quien realiza la atribución para especificar los contenidos de la actitud atribuida. Con ello solamente quiero decir que debemos usar una oración nuestra después de palabras como «Juana cree que...» para identificar la creencia de Juana. Debemos preguntarnos ahora qué propiedades de esta oración son relevantes para la identificación de la creencia, igual que nos preguntamos qué propiedades de los números son esenciales

⁴ Quine ha cambiado de opinión sobre esta cuestión en más de una ocasión. Véase W. V. Quine, *Pursuit of Truth*, 97 y ss.

⁵ No se debería pensar que por hablar de las relaciones entre las actitudes y el mundo esté aceptando una teoría de la correspondencia. Una teoría de la verdad de la clase que tengo en mente depende de establecer una relación entre ciertas palabras y objetos (la relación de «satisfacción» de Tarski), pero no utiliza para nada objetos con los que pudieran corresponderse oraciones.

⁶ Véase el Ensayo 4.

para su papel de registrar pesos. Es evidente que la propiedad de tener un valor veritativo no es suficiente. La sintaxis sería más que suficiente si no fuera por el hecho de que la misma oración puede usarse para querer decir cosas totalmente distintas, dependiendo de la interpretación (la misma oración puede pertenecer a lenguas o idiolectos diferentes, o puede ser ambigua). A causa de esta dificultad, he propuesto que tomemos la *preferencia* (o inscripción) real proporcionada por quien hace la atribución como el objeto al cual se está refiriendo para dar el contenido de una actitud. Así pues, he propuesto una explicación paratáctica de las oraciones que atribuyen actitudes, según la cual el «que» de «Juana cree que...» debe interpretarse como un demostrativo que se refiere a la preferencia que viene a continuación. La preferencia tiene una sintaxis única y puede interpretarse que resuelve la cuestión de cuál es la lengua o idiolecto apropiado, que ha eliminado ambigüedades y que ha proporcionado los parámetros que se necesitan para fijar las referencias indécicas. Como equivalente aproximado de la explicación paratáctica, podríamos pensar en el contenido de la oración como encerrado entre comillas de cita, donde se da por entendida o se hace explícita la lengua y otras características del hablante y el contexto. Esta segunda idea capta mejor la intuición según la cual las expresiones que atribuyen actitudes son oraciones únicas, pero es más complicada.

No debería sorprendernos que lo que podemos decir y entender de las actitudes proposicionales de los demás debería ser lo que podemos captar al casar nuestras propias oraciones (o preferencias; no me molestaré en distinguirlas) con esas actitudes. Esto no significa, por supuesto, que otro no pueda pensar cosas que yo no pueda; pero entonces eso habrá de ser algo cuyo contenido yo tampoco puedo expresar. Tampoco querría excluir pensamientos proposicionales que yo tengo pero que no puedo expresar. Para mi argumento es suficiente que cuando pueda expresar un pensamiento, ya sea propio o de otro, haya de recurrir al mecanismo básico de representarlo en el tejido de nuestras oraciones. Nuestras oraciones proporcionan la única medida de lo mental.

¿Qué conclusiones se siguen para la indeterminación? Todo el mundo estará dispuesto a mostrarse de acuerdo en que algunas veces oraciones bastantes distintas del mismo informante pueden utilizarse para atribuir exactamente el mismo pensamiento, de manera que algún grado de indeterminación se da automáticamente con la redundancia del lenguaje. Pero el tipo de indeterminaciones de las que habla Quine, la mayoría de las cuales acepto, son otra cuestión, puesto que sugieren que oraciones que nadie diría que ni siquiera son aproximadamente sinónimas pueden sin embargo ser usadas para especificar el mismo pensamiento.

Voy a explicar brevemente en qué Quine y yo diferimos y en qué estamos de acuerdo respecto a estos temas. Los dos tomamos las actitudes de una persona hacia las diversas preferencias y las circunstancias que causan esas actitudes como las evidencias básicas para la interpretación. Quine sostiene que son ciertos patrones de asentimiento y disentimiento los que proporcionan la interpretación de las funciones oracionales veritativo-funcionales, y así lo hago yo también; sin embargo, al contrario que Quine, sigo usando el mismo método para identificar los mecanismos de la cuantificación y la referencia cruzada. Así pues, al revés que Quine, yo no veo que esté indeterminada la estructura interna de las oraciones más simples. Ambos aceptamos que el ámbito de los fenómenos observados que nos induce a asentir y disentir a ciertas oraciones nos permite conectar esas oraciones con acaecimientos y objetos del mundo. Por lo que respecta a las oraciones que están menos directamente vinculadas con lo fácilmente observable, creo que podemos hacerlo mejor de lo que Quine cree, y ello porque mantengo que podemos usar nuestra capacidad para detectar grados de apoyo evidencial entre las oraciones como una clave para la interpretación correcta. Quedan todavía dos clases importantes de indeterminación en las cuales estamos de acuerdo: la indeterminación debida a lo que Quine denomina inescrutabilidad de la referencia y la indeterminación que resulta de desdibujar la distinción entre lo analítico y lo sintético.

La tesis de la inescrutabilidad de la referencia afirma que no hay forma de decir qué manera de conectar palabras con

cosas es la correcta; si una manera funciona, entonces habrá infinidad de otras maneras que también lo harán. Desde un punto de vista técnico esto significa que para la relación de satisfacción estándar (la satisfacción es una forma refinada de referencia) podemos sustituir una infinidad de otras relaciones sin que se alteren las condiciones veritativas de cualquier oración o las relaciones lógicas entre oraciones. Dado que todas las evidencias para interpretar el lenguaje deben provenir del nivel oracional (puesto que solamente las oraciones tienen un uso en la comunicación), el resultado es que no puede haber ninguna evidencia de que una de las relaciones de satisfacción (o referencia) sea la correcta.

Aquí va un ejemplo. Supongamos que la relación de satisfacción *s* asocia la palabra «Roma» con Roma, y el predicado «es una ciudad de Italia» con ciudades de Italia. Entonces la definición de verdad mostrará que la oración «Roma es una ciudad de Italia» es verdadera si y sólo si Roma es una ciudad de Italia. Consideremos ahora otra relación de satisfacción *s*» que asocie la palabra «Roma» con un área de 100 millas al sur de Roma, y el predicado «es una ciudad de Italia» con áreas de 100 millas al sur de las ciudades de Italia. La definición de verdad dirá ahora que la oración «Roma es una ciudad de Italia» es verdadera si y sólo si el área que está 100 millas al sur de Roma es un área que está 100 millas al sur de una ciudad de Italia. Las condiciones de verdad son claramente equivalentes. La tesis de la inescrutabilidad de la referencia sostiene que no puede haber evidencia alguna de que *s* sea de alguna manera mejor que *s*» para interpretar la oración «Roma es una ciudad de Italia». No hay manera de decir «acerca» de qué es una oración o qué es lo que alguien está pensando.

¿Cómo puede ocurrir esto? Con seguridad Roma y un área que esté 100 millas al sur de ella no son la misma entidad. Esto es verdad. Sin embargo, todas las diferencias que convencionalmente pensamos que son diferencias en la referencia de los nombres o de las extensiones de los predicados serán preservadas por cualquier relación correcta de satisfacción. El hecho de que no haya diferencia empírica entre las interpretaciones ofrecidas por *s* y por *s*» no implica que la persona que está siendo interpretada no pueda distinguir entre

Roma y un área que esté 100 millas al sur; pero sí implica que no hay manera de decir a cuál de ellas se refiere su palabra «Roma». Una interpretación correcta capta una configuración compleja y ubica a las oraciones y actitudes particulares en la misma. Pero las fichas que usamos para representar esta configuración, a saber, nuestras oraciones, pueden representarla de más de una manera.

Si las dos interpretaciones de la preferencia «Roma es una ciudad de Italia» que hace alguien son igualmente correctas, ¿por qué sentimos que una sola es la correcta? Creo que la respuesta completa es bastante compleja, pero la respuesta simple es que aceptamos el método estándar, que en este caso es el más simple, de la traducción palabra a palabra. Por supuesto que cuando el manual de traducción homofónica es uno de los manuales disponibles, lo usamos; y al interpretar una lengua para la cual no haya un manual homofónico disponible usaremos el método disponible convencional que sea en su mayor parte más corto. Esto no significa que las otras interpretaciones sean erróneas; *s*» ofrece una interpretación tan correcta de «Roma es una ciudad de Italia» como *s*.

A alguien que objete, «Pero “Roma” no *significa* “un área 100 millas al sur de Roma”», la respuesta correcta es: las palabras individuales carecen de significado. Tienen un *papel* en la determinación de las condiciones de verdad de las oraciones, y este papel tanto lo capta *s*» como *s*.

Con todo, esta respuesta aún deja algo fuera: la necesidad de atenerse a algún método de interpretación al asignar contenidos al habla o a las actitudes de un agente. La oración «Roma es una ciudad de Italia» resultará refutada si la interpretamos como verdadera si y sólo si Roma está en un área que se encuentra a 100 millas al sur de una ciudad de Italia; y usar un método de interpretación para una oración y otro método para otra destruirá las relaciones de implicación entre oraciones. Del mismo modo en que debemos indicar si los números que estamos usando para medir la temperatura dan la temperatura en la escala Fahrenheit o en la centígrada, deberemos indicar qué método de interpretación estamos usando (por ejemplo, *s* o *s*»). Una de las cosas convenientes de atenerse al mismo método es que podemos hacer que la indica-

ción del método quede implícita. (De un modo parecido, si digo «En esta habitación debemos estar a 70» se entenderá de manera natural que estoy usando la escala Fahrenheit.)

Hay otra objeción que mucha gente piensa que es fatal para la tesis de la indeterminación. La idea es que la indeterminación socava la autoridad de primera persona. Estas personas pretenden decir, «Quizá el criterio empírico de Quine no pueda excluir el uso de métodos alternativos de interpretación de los demás, pero *yo sé* que con mi palabra “Roma” quiero decir Roma y no el área 100 millas al sur de Roma». Esta observación es correcta, pero no está en conflicto con el indeterminismo. La razón correcta es que el hablante no identifica Roma y un área 100 millas al sur de Roma. Cuando una persona está haciendo tales autoatribuciones, sabemos que las palabras en el metalenguaje tienen los mismos papeles que esas palabras tienen en el lenguaje objeto, puesto que el lenguaje objeto y el metalenguaje son uno y lo mismo. Las interpretaciones de primera persona están ligadas necesariamente al manual de traducción homofónico (lo que equivale a decir que la traducción o la interpretación no tienen aquí lugar).

Del hecho de que una persona esté restringida a una única manera de interpretarse a sí misma (si se le puede llamar a esto interpretación; sería mejor decir que, dejando a un lado casos patológicos, nuestra manera de interpretar a los demás no se aplica a nosotros mismos) no debería concluirse que por lo tanto sus palabras tienen una referencia única. Lo único que el agente ha descubierto en nuestro ejemplo es que sus expresiones «Roma» y «el área 100 millas al sur de Roma» no pueden sustituirse una por la otra en su lengua, y esto no fija la referencia de ninguna de esas expresiones. Por eso mismo, el hecho de que la referencia de sus palabras no esté fijada no muestra que la persona no sepa lo quiere decir o lo que piensa. Las diferentes maneras de interpretar su habla o de representar sus pensamientos no marcan diferencia alguna en los contenidos de sus actitudes, de modo que no hay nada que deje de saber si no sabe qué manera de representar sus pensamientos es la correcta, ya que no hay una sola manera correcta de hacerlo; no hay nada más que pueda saber.

No todos los casos de indeterminación dependen de la posibilidad de alterar sistemáticamente la relación de satisfacción de modos que no afecten a las condiciones de verdad de las oraciones; puede haber indeterminaciones que afecten los valores veritativos. Estas indeterminaciones se dan si uno acepta la afirmación de Quine (como yo lo hago) de que no hay ninguna manera fundamentada de aclarar la distinción entre lo analítico y lo sintético. Es fácil dar ejemplos. Encuentro que muy frecuentemente estamos en desacuerdo con otras personas sobre si llamar verde o azul al color de algún objeto. El desacuerdo es consistente: hay un grupo bastante definido de casos en que yo digo verde y los demás dicen azul. Podemos dar cuenta de esta diferencia de dos maneras: puede ser que yo (o muchas otras personas) estemos equivocados sobre el color de ciertos objetos, o puede ser que yo no use las palabras «azul» y «verde» de la misma manera que los demás. Puede que no haya manera de decidir entre estas dos explicaciones; haciendo ajustes compensatorios en una de las interpretaciones de mis oraciones y creencias se pueden acomodar ambas historias. Sin embargo, en una de las explicaciones ciertos de mis pronunciamientos sobre los colores son falsos, mientras que en la otra son verdaderos. Y en ambas explicaciones, sé lo que pienso: pienso que las cosas que llamo verdes son verdes y que las cosas que llamo azules son azules; ninguna de las explicaciones sobre lo que quiero decir y pienso amenaza el conocimiento de primera persona.

Mientras que las críticas que he estado comentando sugieren que si se acepta el indeterminismo se está tratando a los significados y a los estados mentales como no totalmente reales, Daniel Dennett se ha quejado de manera interesante de que, según mi concepción de los mismo, los estados intencionales son *demasiado* reales⁷. Dennett ve claramente por qué mi concepción, que compara las teorías rivales de la interpretación con diferentes maneras de representar la longitud o la temperatura mediante números, no puede poner en cuestión la realidad de lo que se representa. Pero piensa que dos siste-

⁷ Daniel Dennett, «Real Patterns».

mas diferentes de atribuir creencias a un individuo pueden diferir de manera sustancial, incluso hasta el punto de ofrecer diferentes predicciones del comportamiento, y, sin embargo, nada podría establecer que sea uno de esos sistemas y no el otro el que describe las creencias reales de esa persona.

Pienso que el caso de la predicción es una pista falsa. Ningún sistema de atribución de actitudes, no importa cuán completo sea, produce una predicción de las acciones sin una teoría, y es ciertamente posible en psicología diferir respecto a las teorías predictivas. Está claro que no habrán predicciones plausibles sin una descripción cuantitativa que especifique los grados de creencia (probabilidades subjetivas) y la fuerza relativa de los deseos. Pero incluso suponiendo que pudiéramos dar una descripción completa de todas las actitudes y de su fuerza para un individuo dado en un cierto momento de tiempo, no hay razón para suponer que existan leyes estrictas que predigan lo que un individuo hará a continuación, ni buenas razones para suponer que tales leyes no existan. Entonces, ¿a qué tipo de evidencias puede apelar Dennett para mostrar que sistemas de atribuir actitudes igualmente justificados son incompatibles? El hecho de que de diferentes sistemas resulten diferentes predicciones no prueba nada, ya que se puede usar el mismo sistema para apoyar predicciones diferentes.

La idea de Dennett es que lo que es real es el comportamiento y que los estados intencionales son pautas en este comportamiento. Estas pautas no se definen en términos de comportamiento, sino que son percibidas por un observador cuando éste adopta la «postura intencional». El valor de estas pautas es que reducen una vasta, y para nosotros indescriptiblemente compleja, situación física a algo que podemos captar, algo sobre la base de lo cual podemos hacer predicciones aproximadas. Las pautas son reales en algún sentido (no pleno), aunque sean abstractas, y personas distintas pueden percibir pautas distintas en el mismo terreno de comportamiento. Algunas de estas pautas pueden ser mejores a la hora de predecir o entender algunos fenómenos; otras pautas van mejor para predecir o entender otros fenómenos. Todas estas pautas son «reales», pero cuando son diferentes no tiene sentido decir cuál de ellas representa las actitudes reales del sujeto en cuestión.

No es fácil hacer un juicio sobre esta propuesta, y todavía es más difícil captar cómo se relaciona con la cuestión del realismo. Ciertamente, no *toda* pauta que percibimos en el comportamiento son pautas de actitudes proposicionales. ¿Pero lo es alguna de ellas? Dennett nos dice que las pautas son abstracciones como los centros de gravedad. Pero entonces, ¿pueden ser creencias y deseos? Nos gusta pensar que las creencias y los deseos son estados de un cuerpo físico que pueden tener consecuencias causales; según supongo 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 mis creencias e intenciones. ¿Percibimos pautas? Me parece que no: lo que percibimos es algo que *tiene* una determinada pauta, y con suerte (y la postura apropiada) podemos percibir *que* tiene esa pauta. Así pues, la cuestión no es si las pautas son reales. Al no ser un nominalista, creo que las pautas, al igual que las formas y los números, son tan reales como puedan ser. Pero no veo cómo las actitudes proposicionales de una persona pueden *ser* pautas. Si nos preguntamos qué es lo que exhibe la pauta, podemos decir que es la persona, o 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. Si las personas o su comportamiento realmente exhiben las pautas que Dennett dice que muestran, y que alguien tenga una actitud proposicional consiste en que esa persona exhiba esa pauta, entonces no hay ninguna cosa acerca de las actitudes que sea ni ontológica ni epistemológicamente secundario.

Me parece que Dennett ha confundido dos cosas. La primera es si las actitudes son entidades, y sobre eso pienso que no, a menos que usted piense que los estados son entidades. Si no se piensa esto se debería simplemente hablar de personas que tienen actitudes, lo que significaría que determinadas cosas que predicamos sobre ellas son verdad. La segunda cosa es si hay una respuesta correcta a la pregunta de si alguien tiene una determinada actitud. Esto lo interpreto no como una cuestión acerca de la vaguedad o de casos dudosos, sino 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 no creo que haya dado ninguna razón para que aceptemos esta respuesta.

¿Hay fundamentos objetivos para escoger entre hipótesis rivales? Especialmente en este caso tenemos que preguntar qué es lo que hace «objetivos» a los fundamentos. La fuente última (que no el fundamento) de la objetividad es, en mi opinión, la intersubjetividad. Si no nos comunicáramos con los demás, no habría nada en lo que basar la *idea* de estar equivocado, o por consiguiente, la de estar en lo cierto, tanto respecto a lo que decimos como a lo que pensamos. Según mi concepción, la posibilidad del pensamiento tanto como la de la comunicación depende del hecho de que dos o más criaturas estén respondiendo, más o menos simultáneamente, a *inputs* del mundo que comparten e *inputs* provinientes de cada uno de ellos. Por ejemplo, podemos decir que alguien responde de «la misma manera» a los lobos. Pero, por supuesto, aquí «misma» significa «similar». Los fundamentos para afirmar que a una persona le parece un lobo similar a otro es el hecho de que la persona responde de manera similar respecto a los lobos. Esto provoca la siguiente cuestión: ¿qué hace que las reacciones sean similares? La única respuesta es que a alguien más le parezca que tanto los lobos como las reacciones de la primera persona son similares. Esto, por supuesto, solamente aplaza la cuestión otra vez. Sin embargo, es este 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 el que proporciona las condiciones necesarias para que tenga aplicación el concepto de verdad. Tal y como Wittgenstein afirmó, sin una segunda persona no hay base para juzgar que una reacción es errónea o, por lo tanto, que es correcta⁸.

Esto me conduce a mi último punto. La analogía que he propuesto entre la medición en las ciencias físicas y la asignación de contenidos a las palabras y a los pensamientos de los

⁸ Sobre este tema y el de los párrafos de la conclusión, véase el Ensayo 14.

demás es imperfecta en un aspecto esencial. En el caso de la medición común usamos los números para captar los hechos que nos interesan. En el caso de las actitudes proposicionales usamos nuestras oraciones o preferencias. Pero se da la siguiente diferencia: podemos especificar mutuamente las propiedades de los números. Los números, como los objetos a los cuales los aplicamos, están a medio camino, por decirlo así, entre nosotros y los demás. Esto es lo que significa decir que son objetivos o que son objetos. No puede ser así con nuestras oraciones. Usted y yo no podemos llegar a estar de acuerdo en la interpretación de nuestras oraciones como un preliminar para usarlas para interpretar a otras personas, puesto que el proceso de llegar a un acuerdo tal implica una interpretación del mismo tipo que aquella para la cual creíamos que nos preparaba. No tiene sentido preguntar por una norma de interpretación común puesto que la mutua interpretación proporciona la única norma que tenemos.

No deberíamos desesperar por el hecho de que no podamos proporcionar una norma mediante la cual juzgar la norma, una prueba para ver si el metro patrón actualmente tiene un metro de largo. Nuestra conclusión debería ser más bien la siguiente: si nuestros juicios acerca de las actitudes proposicionales de los demás no son objetivos, los juicios no lo son, y el concepto de objetividad no se aplica aquí.

6

La irreducibilidad del concepto de yo

Dieter Henrich me preguntaba frecuentemente, a su manera amable e inquisitiva, si yo pensaba que la idea del yo era un concepto irreducible o primitivo. Yo respondía que así lo creía, pero por lo que puedo recordar, nunca dije por qué. Sigo teniendo dudas sobre lo que requiere una respuesta adecuada, pero aquí trataré de enumerar algunas de las consideraciones que pueden ser relevantes para ofrecer una respuesta más completa.

Un concepto es irreducible sólo relativamente a recursos específicos. Cuando se trata de los conceptos más amplios e imponentes que importan a los filósofos, como los de bien, verdad, creencia, conocimiento, objeto físico, causa y acaecimiento, pienso que un concepto es irreducible si no puede definirse en términos que sean tan generales como el concepto a reducir, que sean al menos tan claros y que no conduzcan a un círculo. Con respecto a los conceptos que he enumerado, pienso que la búsqueda de una definición o análisis de este tipo está condenada al fracaso. La cuestión es: ¿el concepto de yo es otra de las piezas conceptuales esenciales, y por ello insustituible, que construyen nuestro pensamiento y nuestro lenguaje?

Hubiera sido mejor haber confesado desde buen principio que no veo modo de abordar directamente esta cuestión, puesto que la expresión «el yo» no desempeña ningún papel claro en el lenguaje cotidiano. Los filósofos lo introducen cuando quieren discutir temas como la autoconsciencia o la cuestión

de qué es lo que une las diferentes experiencias de una persona. Así pues, necesariamente abordaré el tema de manera indirecta.

Thomas Nagel nos pide (en «The Objective Self») que imaginemos una descripción del universo que explique cuándo y dónde ocurre todo relativamente a algún marco de referencia espacio-temporal dado objetivamente; tal descripción nombra a cada persona con sus estados mentales junto con las posiciones y propiedades de todos los demás objetos. Esa descripción incluye una historia completa del propio Nagel, y todo ello desde el punto de vista de una tercera persona. Nagel afirma que una cosa que esta descripción no contiene es la información de que *él* es una de las personas descritas. Es evidente que puesto que puede no saber que su nombre es Nagel, la única manera que tiene para ubicarse a él mismo con certeza en ese mundo es usando la palabra «yo». De este modo, el pronombre personal le permite expresar el conocimiento que no puede expresarse de otra manera. Está claro que hay algo irreducible e insustituible que está incluido en el uso de los pronombres de primera persona —en todos los pronombres de primera persona—. De hecho, todas las oraciones que contienen indéxicos dependen de quién las profiere para su interpretación y para su valor veritativo. Pero si las profiero yo, sé sin necesidad de observación que soy yo quien las he proferido. De esta manera, me relaciono a mí mismo con los lugares, objetos, momentos de tiempo y con otras personas mediante mi uso de «allí» («aquí», «detrás de mí»), «eso» («esto»), «ahora» («mañana»), todos los verbos con desinencia temporal), «tú». No hay sustituto para esta manera de ubicarme a mí mismo en el mundo público.

Una función menos obvia pero igualmente importante que tienen los indéxicos es la de proporcionar un primer y esencial paso hacia el dominio del pensamiento y el lenguaje, ya que lo primero que aprendemos es a asociar lo que resultan ser oraciones de una sola palabra («Mamá», «No», «Perro», «Azul») con situaciones, acaecimientos, objetos y sus características. Los niños aprenden pronto el poder mágico de hacer sonidos que los adultos encuentran apropiados y que por tanto recompensan. Esto son sólo los prolegómenos a un habla

y un pensamiento hecho y derecho, puesto que al comienzo los niños no se percatan de la distinción entre lo que se cree y lo que ocurre, lo que se pregunta o pide y lo que se responde o hace. Pero aunque sean solamente prolegómenos, estas relaciones primitivas entre dos personas en presencia de estímulos provenientes de un mundo compartido contienen el núcleo del aprendizaje ostensivo, y solamente en el contexto de tales interacciones llegamos a captar los contenidos proposicionales de las creencias, los deseos, las intenciones y el habla.

No estoy tratando de establecer aquí la verdad de estas afirmaciones; las he argumentado en otros lugares. Sin embargo, si estoy en lo cierto, esta concepción del origen y la naturaleza del pensamiento racional apunta a otro sentido en el que la perspectiva sobre el mundo de cada persona difiere irreduciblemente de la de cualquier otra. La fuerza y el significado pleno de esta afirmación se va a ver aquí. Pero lo siguiente debería estar claro: para tener algún pensamiento debemos percatarnos del triángulo básico que forman dos personas y un mundo común. Si puedo pensar, sé que hay otras personas con mentes como la mía propia y que se encuentran en un tiempo y espacio poblado con objetos y acaecimientos muchos de los cuales les son conocidos a otros (a través de las ostensiones que nos hicieron asequibles tales pensamientos). En particular, yo, como cada una de las otras criaturas racionales, tengo tres clases de conocimiento: conocimiento del mundo objetivo (sin múltiples ostensiones exitosas no tendría pensamientos), conocimiento de las mentes de los demás y conocimiento de los contenidos de mi propia mente. Ninguna de estas tres clases de conocimiento se reduce a ninguna de las otras dos, o a cualquier combinación de las otras dos. Ciertamente del hecho de que tenga una cualquiera de estas clases de conocimiento *se sigue* que tengo las otras dos puesto que el triángulo básico es una condición del pensamiento pero ninguna es conceptual o temporalmente anterior a las demás.

Si mi conocimiento fuera un conocimiento exclusivamente de los contenidos mentales, quizá de los de mi propia mente y los de los demás, no habría manera en que pudiera cons-

truir mi conocimiento del mundo en que vivimos, puesto que ello requiere las conexiones causales con el mundo que proporciona la percepción. Por supuesto que si conociera los contenidos de mi propia mente o de la mente de otro, conocería el mundo público y sabría que lo tengo, pero esto no significa que pudiera *construir* tal conocimiento a partir del conocimiento de los contenidos mentales.

El conocimiento que tengo de otra mente no es reducible a mi conocimiento de los contenidos de mi propia mente y del mundo natural, aunque mi conocimiento de las mentes de los demás dependa de mi conocimiento de los movimientos de sus cuerpos mediante la percepción. Pero suponer que los estados mentales pueden *definirse* en términos extraídos exclusivamente de las ciencias naturales (como algunos conductistas en un tiempo mantuvieron) es suponer que lo intencional se puede reducir a lo extensional, y con seguridad tal reducción no es posible.

Como se sabe, parte de los sueños de cartesianos y empiristas es que el conocimiento de mi propia mente es especial y básico para todo mi conocimiento. Al menos lo siguiente es correcto: tal conocimiento es básico en el sentido de que sin él no sabría nada (aunque el autoconocimiento no sea suficiente para el resto), y especial por el hecho de ser irreduciblemente diferente de las otras clases de conocimiento.

Lo que une a las tres variedades de conocimiento y hace que su interdependencia sea el fundamento de todo pensamiento se ilustra mejor describiendo el proceso de ostensión. Considérese el caso de dos personas que ya poseen su propia lengua. Una pregunta señalando: «¿Qué es eso?», y la otra responde: «Un cormorán.» Puede que se necesiten más preguntas y más acciones de señalar antes de que el que aprende lo capte, pero es asombroso lo frecuente que es que hábitos de generalización compartidos hagan que ello se logre en un solo intento. Aquí tenemos en funcionamiento las tres formas maduras de conocimiento. Tanto el que aprende como el profesor saben qué hay en la mente del otro, y un elemento vital de esta comprensión depende de estímulos perceptuales compartidos. Sin embargo, es más importante la manera en que este conjunto triangular de relaciones revela las bases de la ob-

jetividad del pensamiento, su capacidad para adherirse a algo que se encuentre en el mundo público. En nuestro escenario imaginado, 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. Sin embargo, al comienzo de la ostensión, el que aprende no tiene sino un estímulo sin conceptos ni palabras que pueda conectarse con un concepto y una palabra. Estamos suponiendo 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 de fallo. 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 requiere compañía. Así pues, el engranar las respuestas verbales (y de otro tipo) con situaciones, acaecimientos y objetos a través de la ayuda y las percepciones de los demás desempeña un papel clave tanto para la adquisición de una primera lengua como para el aprendizaje de una segunda lengua en ausencia de un intérprete o de un diccionario bilingüe.

Es evidente que aprender una primera lengua o una segunda son empresas muy distintas. Lo primero significa entrar en el dominio del pensamiento por primera vez, mientras que lo segundo significa que alguien que ya está en el reino del pensamiento entra en el pensamiento de otra persona. Sin embargo, ambas empresas dependen de mecanismos parecidos y de indicios similares. Además, el contraste se debilita cuando nos damos cuenta de que en el caso de un niño ya iniciado se mezclan las dos formas de aprendizaje, puesto que al captar la idea de un mundo objetivo, los niños están aprendiendo simultáneamente a comunicarse con los demás, lo que requiere adentrarse en los pensamientos e intenciones de los demás.

Dar sentido al comportamiento verbal y no verbal de los demás no es una empresa limitada a los casos especiales de

aprender una primera lengua u otras, o de añadir nuevas palabras a la lengua propia. O como podríamos decir si tratáramos la noción de lengua de manera más estricta de lo que se trata en el habla común: todas estas cosas tienen que ver con aprender una nueva lengua, ya que cualquier adición o alteración en nuestros recursos verbales hace que nuestra lengua sea diferente y nueva. Si miramos el detalle, no hay dos idiolectos idénticos, ni hay una lengua de una persona que sea susceptible de permanecer inalterada durante mucho tiempo. Como intérpretes estamos constantemente determinando (frecuentemente de manera automática) lo que otros quieren decir a través de lo que dicen. Por el hecho de la interdependencia del pensamiento y el habla y por otras razones obvias estamos también constantemente decidiendo qué es lo que piensan, pretenden o quieren los demás. Muchas veces damos por sentado que para los propósitos inmediatos ya entendemos suficientemente bien cómo son las cosas, y la mayoría de las veces, cuando se necesitan pequeños ajustes, éstos se nos ocurren tan fácilmente que no nos percatamos de ningún esfuerzo mental. Sólo ocasionalmente nos damos cuenta de que nos estamos enfrentando con palabras viejas pero usadas de maneras que son nuevas para nosotros, o de que algún elemento de la gramática no puede tratarse como algo conocido. En tales casos nos hacemos conscientes de que se están poniendo a prueba nuestras habilidades interpretativas. Es entonces cuando la contribución única del punto de vista de primera persona —el «yo»— se hace más manifiesta.

Un sujeto conceptualmente desarrollado tiene dos recursos interpretativos básicos a su disposición para intentar comprender las preferencias y acciones de los demás: la suposición de que se da la suficiente racionalidad para hacer esas acciones inteligibles, y el conocimiento de cómo la percepción produce los contenidos de las creencias. En el caso del habla, esto es fácil de ilustrar, y esta lección se aplica también a las actitudes proposicionales. Las oraciones, o mejor las actitudes que expresan, deben su contenido, es decir, su significado, a dos cosas: sus relaciones con otras oraciones o actitudes, y sus relaciones directas o indirectas con el mundo a través de la percepción. Por ello a un intérprete le es imposible desaten-

der completamente las relaciones lógicas entre las oraciones o actitudes de un hablante. No es una cuestión de que los sujetos se esfuercen por ser consistentes; se trata de que su habla y su comportamiento tienen el significado que tienen en virtud de cómo se ajustan uno a otro. A menos de que haya una coherencia suficiente es imposible asignar contenidos proporcionales a su habla, sus creencias, sus deseos o sus intenciones. Así, la suposición del intérprete de un grado de racionalidad por parte de aquellos que desea entender no es más que una condición para poder entenderlos.

El papel fundamental de la ostensión en el aprendizaje y la interpretación del habla garantiza que un intérprete no pueda equivocarse en suponer de manera general que las preferencias de un hablante provocadas de manera fiable por rasgos evidentes del mundo observable son verdaderas y tratan de esos rasgos. Es de esperar que haya errores por parte del hablante o el intérprete, pero éstos no pueden ser la regla, ya que los errores toman su contenido de un fondo de pensamientos verídicos y aseveraciones sinceras. La diferencia crucial entre las creencias y supuestos predominantes, que son en su mayor parte banales, trillados y certeros, y las desviaciones ocasionales es la siguiente: los errores, las confusiones y las irracionalidades tienen explicaciones particulares; es de esperar que se acierte excepto en los casos difíciles.

He dicho poco acerca del conocimiento de los contenidos de nuestras propias mentes. Como todo conocimiento, no puede existir asilado de sus comienzos sociales; el concepto de yo como una entidad independiente depende de darse cuenta de la existencia de los demás, un darse cuenta que se forma con la comunicación. Pero el vocabulario de las actitudes se aplica igualmente a uno mismo y a los demás, y los contenidos de las actitudes que así se atribuyen se expresa igualmente en conceptos que son del dominio público. Mi pensamiento de que Shakespeare era una mujer es mío, pero todo el mundo puede pensar lo que yo pienso. ¿Qué es entonces lo que hay de tan especial en mi conocimiento de mi propia mente?

No solamente sé lo que pienso. También sé una infinidad de cosas que puedo expresar y que sé que algún otro podría

creer, dudar o cuestionarse; esta lista es en un sentido tan amplia como la lista de cosas expresables en mi lengua, en conceptos que domino. Éstos son los contenidos proposicionales a los que me refiero cuando atribuyo actitudes a mí mismo o a los demás. Para interpretar las preferencias o acciones de aquellos que me rodean debo usar este rico repertorio de recursos conceptuales. Cuando hablo de «usar» estos recursos no quiero decir que generalmente ello requiera reflexión o esfuerzo. Se da sin que lo notemos, de manera automática. Incluso cuando los demás tienen lapsus, cometen omisiones, cometen errores gramaticales o cambian accidentalmente un nombre por otro, yo los corrijo en silencio y a menudo inconscientemente. La mayoría de las veces sus palabras llegan a mi oído y las entiendo sin que ningún proceso mental consciente haga de intermediario. Pero mi comprensión prueba que ha tenido lugar un proceso por más que esté inarticulado o que se encuentre lejos de la investigación introspectiva, un proceso que merece ser llamado, si no interpretación, algo parecido. Es la interpretación en la que el razonamiento consciente y la apelación explícita a las evidencias y la inducción se han reducido a cero.

La interpretación en este sentido suavizado puede describirse en términos no de los procesos observados, sino de la transición desde los inputs al resultado, puesto que esta transición es después de todo la misma que la transición llevada a cabo por la interpretación consciente. En este caso los elementos son los mismos que anteriormente: la búsqueda de coherencia basada en el supuesto de la racionalidad y la percepción de indicios externos que están al alcance del hablante o el agente. En este momento no me interesan los detalles del proceso, sino un hecho tan obvio que se nos podría escapar: que las normas de racionalidad y de realidad de las que dependo para entender a los demás son las mías propias, y que no hay dónde apelar más allá de ellas. Esto no comporta negar que las normas que tengo no existirían si no hubiera una historia de comunicación y de experiencia. Tampoco comporta suponer que no pueda reflexionar sobre mi propio razonamiento y consultar con los demás para una mayor claridad, acierto e información. Pero en tanto que busque infor-

mación directamente mediante experimentación y observación, no puedo hacer nada mejor o no puedo hacer más que emplear mis propios recursos. Si me pregunto si las normas de racionalidad que empleo al tratar de entender a los demás son correctas puedo, por supuesto, preguntar a Sebastián si soy tan objetivo o razonable como debería ser en mi explicación de los pensamientos y acciones de Basilio. Pero mi comprensión de la respuesta de Sebastián constituirá un ejercicio más de mis propias normas y métodos. Hay otra indicación evidente de la singularidad irreducible de mi contacto directo con los contenidos de mi propia mente, y es que tal conocimiento es único en tanto que, dejando a un lado casos inusuales, no está apoyado en la observación, las evidencias o las razones. Esto es debido, al menos en parte, al hecho de que aquí la interpretación no se aplica. Apoyarse en mi reserva de pensamientos potenciales no provocará nada más que tautologías cuando se apliquen a esos mismos pensamientos. La autoconsciencia de esa clase puede dirigir la atención hacia el interior y promover autocríticas, pero puede llevar a la acción sólo indirectamente.

No creo que por el hecho de que el último tribunal de apelación sea personal, mis juicios sean arbitrarios o subjetivos, puesto que están formados mediante un nexo social que asegura la objetividad de mis creencias, cuando no su corrección. La intersubjetividad es la raíz de la objetividad, no porque sea necesariamente verdad aquello en lo que las personas están de acuerdo, sino porque la intersubjetividad depende de la interacción con el mundo. Aunque no podríamos haber estado en situación de comparar anotaciones sin una interacción previa, son las anotaciones privadas las que al final se comparan. Es aquí donde cada persona, cada mente o yo, se muestra a sí misma como parte de una comunidad de yos libres. No habría pensamiento alguno si los individuos no tuvieran el indispensable y al final ineludiblemente creativo papel de árbitros últimos.

Lo intersubjetivo

7

Animales racionales

Ni un bebé de una semana ni un caracol son criaturas racionales. Si el bebé vive lo suficiente, probablemente llegará a ser racional, mientras que esto no es verdad con respecto al caracol. Si queremos, podríamos decir que los bebés son criaturas racionales desde el comienzo porque probablemente llegarán a ser racionales si es que sobreviven, o porque pertenecen a una especie que tiene esa capacidad. No importa como lo digamos, sigue habiendo esa diferencia, por lo que respecta a la racionalidad, entre los bebés y los caracoles, por una parte, y las personas adultas normales, por la otra.

La diferencia consiste en tener actitudes proposicionales como las creencias, los deseos, las intenciones y la vergüenza. Esto provoca la pregunta de cómo decir cuando una criatura tiene actitudes proposicionales; estamos de acuerdo en que los caracoles no las tienen, pero ¿qué pasa con los perros o los chimpancés? La cuestión no es totalmente empírica puesto que subsiste la cuestión filosófica de qué evidencias son relevantes para decidir si una criatura tiene actitudes proposicionales.

Algunos animales piensan y razonan; consideran, someten a prueba, rechazan y aceptan hipótesis; actúan por razones, algunas veces después de deliberar imaginando consecuencias y sopesando probabilidades; tienen deseos, esperanzas y odios, algunas veces por buenas razones. También cometen errores de cálculo, actúan en contra de su mejor juicio, o aceptan doctrinas basadas en evidencias inadecuadas. Cada uno de estos logros, actividades, acciones o errores es suficien-

te para mostrar que tal animal es un animal racional, puesto que ser un animal racional es solamente tener actitudes proposicionales, no importa cuán confusas, contradictorias, absurdas, injustificadas o erróneas puedan ser éstas. Propongo que ésta es la respuesta.

La cuestión es: ¿qué animales son racionales? Por supuesto que no pretendo dar nombres, aunque sea nombres de especies u otros grupos. No intentaré decidir si los delfines, los monos, los embriones humanos o los políticos son racionales, ni tampoco si todo lo que impide que los ordenadores sean racionales está en su génesis. Mi pregunta es qué es lo que hace que un animal (u otra cosa, si uno quiere) sea racional.

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, los deseos y, si es que 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.

Hacer la distinción de manera tan marcada y hacer que dependa del lenguaje provoca la acusación de antropocentrismo. La queja es justa, pero no debería dirigírseme a mí. Yo solamente describo una característica de ciertos conceptos. Después de todo, no es sorprendente que nuestro lenguaje humano sea rico en recursos para distinguir a los hombres y las mujeres de otras criaturas, del mismo modo en que según se dice los Inuit tienen un vocabulario conveniente para captar distintas variedades de nieve (se dice ahora que esto es un mito). Nos confabulamos con nuestro lenguaje para que tanto éste como nosotros parezcamos especiales.

He prometido no discutir la cuestión de si son racionales determinadas especies particulares, pero será imposible evitar la apariencia de hablar de las proezas y habilidades de las bestias porque gran parte de la discusión sobre la naturaleza del pensamiento se ha centrado tradicionalmente en los poderes mentales de los animales no humanos. Considero que este enfoque es solamente una manera colorista (y a veces cargada emocionalmente) de pensar acerca de la naturaleza del pensamiento¹.

Norman Malcolm explica la siguiente historia para mostrar que los perros piensan:

Supongamos que nuestro perro está persiguiendo al gato de nuestro vecino. El gato corre directamente hacia el roble, pero de repente se desvía en el último momento y desaparece en un arce que está cerca. El perro no ha visto esta maniobra y al llegar al roble, se levanta sobre sus patas posteriores, da zarpazos en el tronco como si intentara escalar y ladra excitadamente a las ramas de arriba. Quienes observamos todo este episodio desde una ventana decimos: «Piensa que el gato ha subido a ese roble»².

(Malcolm añadía que diríamos que el perro estaba ladrándole al árbol equivocado.) Malcolm afirma que en esas circunstancias alguien que le atribuyera esa creencia al perro casi con seguridad estaría en lo cierto; tal persona tendría exactamente la clase de evidencia que se necesita para justificar tal atribución.

¹ Mis notas me muestran que antes de que escribiera este ensayo di no menos de diez charlas desde Valdosta, Georgia, a Auckland, con el título «Why Animals can't Think» [Por qué los animales no pueden pensar]. El título era tendencioso, ya que lo que argumentaba (como aquí) es que solamente las criaturas con lenguaje pueden pensar. Sin embargo, lo que ocurre es que creo que solamente los hombres y las mujeres tienen lenguaje, o algo lo bastante parecido a un lenguaje como para justificar la atribución de pensamientos proposicionales. Acerca de la cuestión moral de cómo deberíamos tratar a las criaturas sin habla, no veo ninguna razón para ser menos bondadoso con aquellos que no tienen pensamientos o lenguaje que con aquellos que lo tienen; más bien pasa lo contrario.

² Norman Malcolm, «Thoughtless Brutes», 13.

Voy a ofrecer un argumento preliminar pensado para poner en duda la afirmación de Malcolm. Está claro que la evidencia para afirmar que el perro tiene una «creencia» depende de tomar la creencia como un determinante de la acción y la respuesta emocional. Se nos pide que inferamos de lo que vemos que el perro quiere atrapar al gato, que corre hacia donde lo hace porque tiene este deseo y una creencia acerca de dónde ha ido el gato, y que al no haber sido capaz de seguir al gato hacia arriba, está descargando su frustración ladrando, escarbando el suelo, etc. No es necesario que los detalles sean correctos. La cuestión es hasta aquí obvia: si estamos justificados en inferir creencias, también estamos justificados en inferir intenciones y deseos (y quizá mucho más).

¿Pero qué pasa con la supuesta creencia del perro de que el gato subió a ese roble? Ocurre que el roble es el árbol más viejo que está a la vista. ¿El perro cree que el gato subió al árbol más viejo que está a la vista? ¿O que el gato subió al mismo árbol que la última vez que lo persiguió? Es difícil dar sentido a estas preguntas. Pero entonces no parece posible distinguir entre cosas bastante distintas que podría decirse que el perro cree.

Una manera de reconocer que estamos atribuyendo una actitud proposicional es notar que las oraciones que usamos para hacer la atribución pueden pasar de ser verdaderas a ser falsas si, en las palabras que captan el objeto de la actitud, sustituimos alguna expresión referencial por otra expresión que se refiera a la misma cosa. La creencia de que el gato subió a ese roble no es la misma creencia que la de que el gato subió al árbol más viejo que está a la vista. Si usáramos palabras como «cree», «piensa» y «pretende» prescindiendo del rasgo de la opacidad semántica, se suscitaría la cuestión de si estamos usando esas palabras para atribuir actitudes proposicionales, ya que desde hace tiempo se ha reconocido que la opacidad semántica distingue el hablar acerca de actitudes proposicionales del hablar de otras cosas.

Alguien podría pensar que la posición que ocupa la expresión «ese roble» dentro de la oración «El perro piensa que el gato ha subido a ese roble» es transparente (dicho en la termi-

nología de Quine). La manera correcta de exponer la creencia del perro —continúa la propuesta— es «El perro piensa, con respecto a ese roble, que el gato ha subido a él» o «Ése es el roble al que el perro piensa que ha subido el gato». Pero tales construcciones, aunque liberan a quien hace la atribución de la necesidad de producir una descripción del objeto que aceptase el sujeto de la creencia, implican sin embargo que hay alguna descripción así; la descripción *de re* identifica un objeto que el sujeto de la creencia podría identificar de alguna manera. Utilizando una expresión popular aunque engañosa, el perro debe creer, bajo alguna descripción del árbol, que el gato ha subido a él. ¿Pero cuál es la clase de descripción que se acomoda al perro? Por ejemplo, ¿el perro puede creer de un objeto que es un árbol? Esto parecería imposible a menos que supongamos que el perro tiene muchas creencias generales acerca de los árboles: que son cosas que crecen, que necesitan suelo y agua, que tienen hojas o que se queman. No hay una lista fija de cosas que deba creer alguien que tenga el concepto de árbol, pero sin muchas creencias generales no habría ninguna razón para identificar una creencia como una creencia acerca de un árbol, y mucho menos acerca de un roble. Consideraciones similares se aplican al supuesto pensamiento del perro acerca del gato.

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, empezamos por donde empezamos, 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 primera parece inestable.

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. Para creer que el gato subió al

roble debemos tener muchas creencias verdaderas sobre los gatos y los árboles, sobre este gato y este árbol, sobre el lugar, apariencia y hábitos de los árboles y de los gatos, etc.; pero lo mismo ocurre si me pregunto si el gato subió al roble, si me temo que lo hizo, si espero que lo hiciera, si deseo que lo haya hecho o pretendo que lo haga. 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.

Como he señalado anteriormente, no puede haber una lista fija de creencias de la cual dependa ningún pensamiento particular. Sin embargo, se necesitan muchas creencias verdaderas. Algunas de las creencias del tipo que se necesita son generales, pero plausiblemente empíricas, como, por ejemplo, que los gatos pueden arañar o subir a los árboles. Otras son particulares, como, por ejemplo, que el gato que se ha visto correr hace un momento está todavía en los alrededores. Algunas son lógicas. Los pensamientos, como las proposiciones, tienen relaciones lógicas. Dado que la identidad de un pensamiento no puede separarse de su lugar en la red lógica formada por otros pensamientos, no se puede reubicar en la red sin que se convierta en un pensamiento distinto. Por esto es imposible que en las creencias se dé una incoherencia radical. Tener una actitud proposicional en particular es tener una lógica en gran parte correcta, en el sentido de tener una pauta de creencias que son lógicamente coherentes. Ésta es una de las razones por las que tener actitudes proposicionales significa ser una criatura racional. Esta idea se extiende a la acción intencional. La acción intencional es la acción que puede ser explicada en términos de creencias y deseos cuyos contenidos proposicionales racionalizan la acción. De manera parecida, una emoción como la de estar satisfecho de haber dejado de fumar debe ser una emoción que sea racional a la luz de las creencias y valores que uno tenga.

No es necesario decir que esto no significa negar la existencia de creencias, acciones y emociones irracionales. Uno puede tener razones para realizar una acción que uno tiene mejores razones para evitar. Una creencia puede ser razonable a la luz de algunas de las creencias de uno, pero no de su totali-

dad; y así sucesivamente. La idea es que la posibilidad de irracionalidad depende de un alto grado de racionalidad. La irracionalidad no es la mera falta de razón sino una enfermedad o perturbación de la razón.

Estoy suponiendo que un observador que esté en circunstancias favorables puede decir qué creencias, deseos e intenciones tiene un agente. De hecho, he apelado a esta suposición cuando he señalado que si una criatura no puede hablar, no está claro que se pueda mantener la intencionalidad en las descripciones de sus supuestas creencias y demás actitudes. De manera similar, me preguntaba si en ausencia de lenguaje podría haber un fundamento adecuado para atribuir las creencias generales que se necesitan para dar sentido a un pensamiento. Sin defender la suposición de que podemos conocer las mentes de los demás, voy a distinguir esta suposición de otras suposiciones más fuertes. Afirmar meramente que en condiciones favorables un observador puede decir lo que algún otro piensa no significa adoptar el verificacionismo, ni siquiera con respecto a los pensamientos. La razón es que la suposición de la observabilidad no implica que sea posible enunciar explícitamente cuáles son las evidencias necesarias o suficientes para determinar la presencia de un pensamiento particular; no se está sugiriendo que, de alguna manera, pensar se pueda reducir definicionalmente a alguna otra cosa. La suposición de la observabilidad no implica tampoco que la única manera de determinar la existencia de un pensamiento sea mediante la observación. Al contrario, está claro que normalmente las personas saben lo que creen, quieren o pretenden sin necesidad de observación o de evidencias.

La suposición de la observabilidad tampoco conlleva el conductismo. Un observador que no presencie nada más que la conducta puede descubrir actitudes proposicionales sin que ello implique que las actitudes sean de alguna manera reducibles a la conducta. Existen lazos conceptuales entre las actitudes y la conducta que, con suficiente información acerca de la conducta actual y potencial, son suficientes para permitir hacer inferencias correctas de actitudes.

Por todo lo que se ha dicho sobre la dependencia de las creencias respecto de otras creencias, y de otras actitudes pro-

posicionales respecto de las creencias, está claro que se debe observar una pauta de comportamiento muy compleja para que se justifique la atribución de un solo pensamiento. O, dicho más exactamente, debe haber una buena razón para creer que hay una pauta compleja de comportamiento. Y a menos que haya realmente tal pauta compleja de comportamiento, no hay pensamiento.

Creo que solamente hay tal pauta si el agente tiene un lenguaje. Si esto es correcto, entonces Malcolm estaba justificado en atribuir pensamiento a su perro solamente si creía, basado en buenas evidencias, que su perro tenía un lenguaje.

El punto de vista según el cual el pensamiento —las creencias, los deseos, las intenciones y demás— requiere lenguaje es controvertido, pero ciertamente no es nuevo. Se ha de distinguir la versión de la tesis que propugno de varias versiones relacionadas con ella. Por ejemplo, no creo que se pueda reducir el pensar a la actividad lingüística. No encuentro plausibilidad alguna en la idea de que los pensamientos puedan identificarse nomológicamente o ser correlacionados con fenómenos caracterizados en términos físicos o neurológicos. Tampoco veo ninguna razón para mantener que no podemos pensar aquello sobre lo que no podemos hablar. 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.

Como he señalado anteriormente, esto se ha afirmado de manera frecuente; pero ¿sobre qué base? Dada la popularidad de la idea que abarca desde los racionalistas a los pragmatistas americanos e incluso a los filósofos analíticos contemporáneos, hay una remarcable escasez de argumentos. Hasta aquí he señalado la dudosa aplicabilidad de la prueba de la intensionalidad cuando se trata de animales sin habla, y la necesidad de que haya una rica provisión de creencias generales (y verdaderas) para que exista pensamiento. Estas consideraciones apuntan al lenguaje, pero no equivalen a una demostración

de que el lenguaje sea necesario para el pensamiento. De hecho, lo que estas consideraciones sugieren es sólo que probablemente no puede haber mucho pensamiento sin lenguaje.

Contra la dependencia del pensamiento respecto del lenguaje se da la evidente observación de que tenemos éxito en explicar, y algunas veces en predecir, el comportamiento de animales sin lenguaje atribuyéndoles creencias, deseos e intenciones. Este método funciona igualmente para perros y ranas como para personas. Y podríamos añadir que no hay un marco alternativo práctico y general para explicar el comportamiento animal. ¿No conllevan estos hechos una *justificación* de la aplicación de ese método?³

No hay duda de que comportan esa justificación, pero todavía quedaría un sentido claro en que sería equivocado concluir que esos animales sin habla (=incapaces de interpretar o mantener comunicación lingüística) tienen actitudes proposicionales. Para ver esto solamente es necesario reflexionar en que podría ser muy bien el caso que alguien no tenga mejor alternativa o no tenga otra alternativa que no sea la de explicar los movimientos de un misil que siga el calor mediante la suposición de que el misil quería destruir un avión y creía que podía hacerlo moviéndose de la manera en que se observaba que se movía. Podría ser que este observador no informado estuviera justificado en atribuir un deseo y una creencia al misil; pero estaría equivocado. Por ejemplo, yo estoy mejor informado, no porque sepa cómo están diseñados los misiles, sino porque sé que se mueven como lo hacen porque han sido diseñados y construidos por personas que tenían el deseo y las creencias que mi amigo ignorante asignaba al misil. Mi explicación, aunque todavía sea teleológica y dependiente de la existencia de actitudes proposicionales, es una explicación mejor porque no atribuye al misil la potencialidad de un rico abanico de comportamientos del tipo que debe tener una criatura pensante.

El caso de una criatura sin lenguaje difiere del caso del misil en dos aspectos: muchos animales son mucho más seme-

³ Ésta es la posición que acentúa Jonathan Bennett en *Linguistic Behavior*.

jantes a los humanos por lo que se refiere al abanico de comportamientos que tienen de lo que lo son los misiles, y frecuentemente no tenemos una manera mejor de explicar su comportamiento que asignándoles actitudes proposicionales. Por ello para defender mi posición necesitamos una caracterización de qué es lo que ofrece el lenguaje que sea necesario para el pensamiento, puesto que si existe una condición necesaria de este tipo, podemos continuar explicando el comportamiento de las criaturas sin lenguaje atribuyéndoles actitudes proposicionales, pero al mismo tiempo reconociendo que tales criaturas realmente no tienen actitudes proposicionales. Estaremos obligados a reconocer que estamos aplicando un modelo de explicación que es mucho más fuerte de lo que requiere el comportamiento que se observa y que ese comportamiento observado no es lo suficientemente sutil para dar un sentido pleno a tal explicación.

En lo que queda de ensayo enunciaré la condición que se requiere para el pensamiento y que creo que solamente el lenguaje puede proporcionar, y expondré consideraciones a favor de mi posición. Aunque presento estas consideraciones en forma de argumento, quedará claro que hay varios puntos en mi razonamiento que pueden ponerse en cuestión.

El «argumento» tiene dos pasos. Creo que he mostrado que todas las actitudes proposicionales requieren un trasfondo de creencias, así que me concentraré en las condiciones para las creencias. Sin creencias no hay otras actitudes proposicionales y así pues no hay racionalidad tal y como la he caracterizado.

En primer lugar argumentaré que para tener una creencia es necesario tener el concepto de creencia.

En segundo lugar argumentaré que para tener el concepto de creencia se debe tener lenguaje.

Norman Malcolm, en el artículo que he mencionado más arriba, hace una distinción similar a la que yo quiero hacer entre tener una creencia y tener el concepto de creencia, pero su terminología difiere de la mía. He estado usando la palabra «pensamiento» para abarcar todas las actitudes proposiciona-

les. Sin embargo, Malcolm restringe la aplicación del término «pensamiento» a un nivel más alto del pensar. Según su posición, el perro puede creer que el gato ha subido al roble, pero no puede tener el pensamiento de que el gato ha subido al roble. Malcolm sostiene que se requiere lenguaje para lo último pero no para lo primero. Malcolm hace la distinción diciendo que una criatura meramente piensa (cree) que p si se percata de que p , pero que tiene el pensamiento de que p si se percata de que se percata de que p . Esta idea se acerca a la distinción que hago entre creer que p y creer que uno cree que p . La segunda es una creencia acerca de una creencia, y por ello requiere el concepto de creencia. Haciendo una tosca comparación: Malcolm mantiene que el lenguaje establece una línea divisoria entre criaturas que meramente piensan y criaturas que tienen el concepto de pensamiento; yo mantengo que para pensar se debe tener el concepto de pensamiento y que por ello se necesita el lenguaje para ambas cosas.

Donald Weiss se opone a Malcolm: Weiss piensa que es posible hacer atribuciones sensatas de ciertos actos de percatarse a criaturas sin habla⁴. Puesto que pienso que su ejemplo puede despertar simpatía en otros, voy a parafrasearlo y a continuación lo citaré extensamente. Aquí va la historia. Arturo no es un perro, sino que, digámoslo así, es un superperro de otro planeta. Arturo llega a la Tierra sin compañía y aquí es cuando sale del cascarón. No ha tenido contacto con, o conocimiento de, otras criaturas —se le observa a través de espejos de una sola dirección—. Carece de lenguaje. De acuerdo con Weiss, nos llegamos a convencer de que tiene una inteligencia reflexiva cuando observamos la siguiente escena:

Un día Arturo se encuentra con un metal brillante, lo pone en el fuego y trata de aplanarlo con un martillo, pero descubre que aparentemente no es más maleable de lo que lo era cuando estaba frío. Vuelve a intentarlo otra vez más lenta y metódicamente, pero de nuevo obtiene el mismo resultado. La regularidad en la que Arturo creía —susurramos entre nosotros— no es enteramente universal. Arturo ha descubierto una ejemplificación que no se ajusta a la regla general.

⁴ Donald Weiss, «Professor Malcolm on Animal Intelligence».

Arturo empieza a caminar agitadamente por todo el espacio en que vive. De repente se sienta; de repente se levanta otra vez; va y viene de un lado a otro. Se sienta de nuevo, pero esta vez se queda sentado. Pasan quince minutos sin que cambie de postura; los ojos de Arturo miran al frente. Entonces, de repente, se pone de pie e inmediatamente procede a apilar una gran cantidad de madera en el fuego... Entonces sumerge su metal recientemente descubierto en el fuego, y después de un tiempo, lo retira. Trata otra vez de aplanarlo con el martillo y esta vez tiene éxito. Aparentemente satisfecho con ello... se pone a pararse la comida de una manera relajada⁵.

Weiss afirma que en ese momento tenemos una fuerte evidencia de que Arturo ha reflexionado sobre sus propias creencias; lo que le impresiona especialmente es el hecho de que Arturo en respuesta a su estado de aturdimiento, se siente con los ojos bien abiertos, completamente inmóvil, y que luego de repente pase a realizar los actos que constituyen la solución a su problema⁶.

No tendré en cuenta el vocabulario que Weiss usa para describir los movimientos de Arturo, el cual comporta una petición de principio, puesto que pienso que Weiss está en el buen camino: es esencial que podamos describir a Arturo como sorprendido. Lo que pienso que está claro es que si está sorprendido, tiene pensamientos producto de una reflexión y, por supuesto, creencias.

Esto no significa afirmar que todo pensamiento sea autoconsciente, o que siempre que pensamos que p debamos percatarnos de que p , o creer que creemos que p , o pensar que pensamos que p . Mi afirmación es más bien la siguiente: para tener cualquier actitud proposicional es necesario tener el concepto de creencia, tener una creencia sobre alguna creencia. Pero ¿qué se requiere para tener el concepto de creencia? Aquí es cuando busco ayuda en el fenómeno de la sorpresa porque pienso que la sorpresa requiere el concepto de creencia.

Supongamos que creo que tengo una moneda en el bolsillo. Vacío el bolsillo y no encuentro ninguna moneda. Me

⁵ *Ibíd.*, 91-2.

⁶ *Ibíd.*

sorprendo. Está suficientemente claro que no podría sorprenderme (aunque quizá podría estar alarmado) si para empezar no tuviera creencias. Quizá está igualmente claro que tener una creencia, al menos de la clase que he tomado como ejemplo, supone la posibilidad de sorpresa. Si creo que tengo una moneda en el bolsillo podría ocurrir algo que haría que cambiase de opinión. Pero la sorpresa implica dar un paso más allá. No es suficiente con que primero crea que tengo una moneda en el bolsillo y que después de haberlo vaciado ya no tenga esa creencia. La sorpresa requiere que me percate del contraste entre lo que creía y lo que paso a creer después. Sin embargo, ese percatarse es una creencia sobre una creencia: si me sorprendo, entonces entre otras cosas paso a creer que mi primera creencia era falsa. No necesito admitir que todo caso de sorpresa implique la creencia según la cual una creencia previa era falsa (aunque me inclino a pensar que esto es así). Lo que quiero afirmar es que uno no puede tener un conjunto general de creencias del tipo necesario para tener cualquier creencia sin ser sujeto de sorpresas que implican creencias sobre la corrección de las propias creencias. La sorpresa sobre algunas cosas es una condición necesaria y suficiente para el pensamiento en general. Esto concluye la primera parte de mi «argumento».

Buena parte del sentido que tiene el concepto de creencia es que se trata del concepto de un estado de un organismo que puede ser verdadero o falso, correcto o incorrecto. Por tanto, tener el concepto de creencia es tener el concepto de verdad objetiva. Si creo que tengo una moneda en el bolsillo puedo estar en lo cierto o estar equivocado; estoy en lo cierto solamente si tengo una moneda en el bolsillo. Si me sorprende al descubrir que no tengo una moneda en el bolsillo, entonces paso a creer que mi primera creencia no se correspondía con el estado de mis finanzas. Tengo la idea de una realidad objetiva que es independiente de mi creencia.

Una criatura puede interactuar con el mundo de maneras complejas sin que abrigue proposiciones. Puede discriminar colores, sabores, sonidos y formas. Puede aprender, esto es, cambiar su conducta de manera que se preserve su vida o aumente lo que come. Puede «generalizar» en el sentido de reaccionar a nuevos estímulos tal y como ha llegado a reaccio-

nar a estímulos previos. Con todo, haya tenido el éxito que haya tenido de acuerdo con mis normas, nada de esto muestra que la criatura domine el contraste entre lo que se cree y lo que es el caso, tal y como se requiere para las creencias.

¿Qué es lo que *mostraría* el dominio de este contraste? Claramente, la comunicación lingüística es suficiente. Para comprender el habla de otra persona debo ser capaz de pensar las mismas cosas que ella; debo compartir su mundo. No tengo que estar de acuerdo con ella en todo, pero para poder estar en desacuerdo debo abrigar las mismas proposiciones, que traten de lo mismo, y tener el mismo concepto de verdad. La comunicación depende de que cada uno de quienes se comunican tengan, y piensen correctamente que el otro tiene, el concepto de un mundo compartido, un mundo intersubjetivo. Pero el concepto de un mundo intersubjetivo es el concepto de un mundo objetivo, un mundo sobre el cual todos los que se comunican pueden tener creencias.

Propongo entonces que el concepto de verdad intersubjetiva es suficiente como base para la creencia y por tanto para los pensamientos en general. Y quizá es suficientemente plausible que tener el concepto de verdad intersubjetiva depende de la comunicación en un sentido plenamente lingüístico. Sin embargo, para completar el «argumento», necesito mostrar que la *única* manera en que se podría tener el contraste creencia-verdad es a través de tener el concepto de verdad intersubjetiva. Debo confesar que no sé cómo mostrarlo, pero tampoco tengo idea de qué otra manera se podría llegar al concepto de verdad objetiva. En lugar de un argumento para ese primer paso, ofrezco la siguiente analogía.

Si estuviese clavado en la tierra no tendría manera de determinar la distancia que me separa de muchos objetos. Solamente sabría que se encuentran en una posición en la que se podría trazar una línea entre ellos y yo. Podría interaccionar de manera exitosa con los objetos, pero no habría manera de dar contenido a la cuestión de dónde están. Al no estar clavado puedo triangular. Nuestro sentido de la objetividad es consecuencia de otra clase de triangulación, una que requiere dos criaturas. Cada una de ellas interacciona con un objeto, pero lo que da a cada una el concepto de cómo son las cosas obje-

tivamente es la línea de base que se forma entre las criaturas mediante el lenguaje. Sólo el hecho de que compartan un concepto de verdad da sentido a la afirmación de que tienen creencias, de que son capaces de asignar un lugar a los objetos en el mundo público.

La conclusión de estas consideraciones es que la racionalidad es un rasgo social. Solamente la tienen quienes son capaces de comunicarse.

8

La segunda persona

... significar es como dirigirse hacia alguien.

(Wittgenstein, *Philosophical Investigations*, § 457)

¿Cuántos hablantes competentes de una lengua debe haber para que se pueda decir que alguien habla o entiende una lengua? Dado que esto es una cuestión gobernada por la tortuosa carrera de la evolución, no tengo ni idea de cuál es la respuesta; quizá se necesite una muchedumbre. No obstante, como filósofos podemos formular la pregunta con un cariz más teórico. En este ensayo me centraré en el papel —el papel por principio— de la segunda persona. Debería quizá añadir que mi tema no es la segunda persona gramatical, el «tú» o «vosotros», el «you» o «thou»; diré cosas sobre auténticas segundas personas, no sobre las palabras que usamos para dirigirnos a ellas.

Una lengua o lenguaje¹ se puede ver como un objeto abstracto complejo, definido mediante una lista de expresiones finita (las palabras), reglas para construir concatenaciones de expresiones con significado (las oraciones) y una interpretación semántica de expresiones con significado basada en los rasgos semánticos de las palabras individuales. En este ensayo no trataré los detalles de cómo podrían describirse o definirse tales objetos.

¹ Véase la nota 3 del Ensayo 3. [N. de la T.]

Concebidos de esta manera, una lengua o lenguaje es abstracto en el sentido obvio de que es inobservable, inmutable y de que sus componentes tampoco son observables ni se alteran. Si así lo queremos, las expresiones pueden verse como figuras o formas acústicas o bidimensionales que de vez en cuando podrían conformar preferencias o inscripciones reales, pero las propias expresiones seguirían siendo abstractas y su existencia sería independiente de la ejemplificación². Las funciones que interpretan algunas expresiones al hacer que se correspondan con objetos o clases de objetos son también abstractas, por supuesto. Los únicos particulares concretos que forman parte de la caracterización de una lengua o lenguaje son algunos de los objetos con los que se corresponden algunas expresiones (por ejemplo, por medio de la relación de nombrar, o dando un paso más hacia la generalidad, por medio de la relación de satisfacción de Tarski).

No hay por qué maravillarse del carácter abstracto del lenguaje. El concepto de lenguaje es del mismo tipo de conceptos que los de nombre, predicado, oración, referencia, significado, y depende de ellos. Todos éstos son conceptos teóricos y se aplican a cosas que son objetos abstractos. No los necesitamos para usar o aprender una lengua o lenguaje; obviamente no son accesibles cuando estamos aprendiendo una primera lengua. Esos conceptos los queremos cuando hablamos del comportamiento verbal. Los filósofos, los psicólogos y los lingüistas necesitan estos términos teóricos si quieren describir, teorizar sobre, y explicar las actividades verbales. El resto de nosotros también tenemos ocasión de hablar sobre el habla o de escribir sobre el hecho de escribir; así que estos conceptos teóricos tienen su lugar en las «teorías» informales e imprecis-

² O, tomando entidades igualmente abstractas, las expresiones pueden concebirse como clases de preferencias o inscripciones. Pero si tomamos las expresiones como clases de preferencias o inscripciones, todas las expresiones no proferidas o no escritas, y por tanto todas las expresiones de las lenguas o lenguajes no utilizados, serán idénticas. En algunos casos hay maneras tortuosas de sortear esto. (Véase W. V. Quine y Nelson Goodman, «Steps toward a Constructive Nominalism».) Para nuestros propósitos será mejor tomar las expresiones como figuras o formas, esto es, propiedades que pueden tener las preferencias y las inscripciones.

sas sobre el lenguaje que todos tenemos. De hecho, todos hablamos tan libremente sobre el lenguaje, o las lenguas, que tendemos a olvidar que cosas tales no existen en el mundo; solamente hay personas con sus distintas producciones escritas y acústicas. Esta idea, tan obvia en ella misma, es sin embargo fácil de olvidar, y ello tiene consecuencias que no son universalmente reconocidas.

Una de las características del concepto de lenguaje tal como lo he descrito es, por tanto, que debe haber una infinidad de «lenguas» o «lenguajes» que nadie ha hablado ni hablará nunca. Decir que alguien habla una lengua particular, pongamos por caso el francés, es decir que sus preferencias y escritos registrables son ejemplares de expresiones del francés. Para ser el ejemplar de una oración del francés, una preferencia o inscripción debe ejemplificar una oración del francés, esto es, debe tener una de las formas definidas para ser una oración del francés; y la preferencia debe tener los rasgos semánticos que la definición del francés asigna a la forma. (Otras lenguas o lenguajes, probablemente no habladas, asignarán otros rasgos semánticos a esas mismas formas). La existencia del francés no depende de que nadie lo hable del mismo modo que la existencia de formas no depende de que haya objetos con esas formas³.

Se sigue de ello que no hay nada sobre la existencia de una lengua o lenguaje particular que le dote de más interés que el que puede tener cualquier objeto abstracto; como lógicos podemos estudiarlo como un ejemplo entre incontables otros de una configuración formal. Tal investigación da resultados de varias clases, pero no tiene relación con lo que nos preocupa normalmente, con la comprensión del habla de los demás o con aprender a hacernos comprender por los demás. Nuestro interés práctico, por contraste con nuestro interés puramente teórico, en los fenómenos lingüísticos es el siguiente: queremos entender las preferencias reales de los demás y queremos que nuestras preferencias sean entendidas. ¿Qué tiene que ver el lenguaje con este interés?

³ Este concepto de lengua o lenguaje es esencialmente el de David Lewis, «Languages and Language».

La respuesta es que solamente empleando conceptos como palabra y oración podemos dar una descripción sistemática de los aspectos lingüísticos de la conducta y las aptitudes lingüísticas. Por ejemplo, no podríamos decir lo que hemos aprendido cuando aprendemos que «demain» significa mañana en francés si no pudiéramos hablar de palabras —aquellas misteriosas formas abstractas (acústicas) que comparten las preferencias de la palabra «demain»—. Así pues, una preferencia de la palabra «demain» se refiere al día después del día de la preferencia. Si no pudiéramos referirnos a palabras y oraciones, no habría una manera fácil en que pudiéramos especificar qué preferencias constituyen preferencias de oraciones, y por ello constituyen preferencias inteligibles.

Por tanto, la principal función del concepto de lenguaje y de sus conceptos concomitantes, como los de predicado, oración y referencia, es capacitarnos para dar una descripción coherente del comportamiento de los hablantes y de lo que saben los hablantes y sus intérpretes que les permite comunicarse. No estoy diciendo que los hablantes y quienes les entienden puedan ofrecer tales descripciones de sus capacidades y su conducta. Para ilustrar lo que he dicho: un hablante competente de una lengua (y un intérprete competente) conoce las condiciones de verdad de un número de oraciones indefinidamente grande. Así, la mayoría de los hablantes del inglés saben que una preferencia de la oración «Montreal está en Canadá» es verdadera si y sólo si Montreal está en Canadá; además, el hablante conoce hechos similares sobre una infinidad de otras oraciones. El hablante no necesita mostrar este conocimiento en palabras. Pero *nosotros* no podemos describir la totalidad del conocimiento que posee el hablante o el intérprete de tal lengua sin que nosotros mismos no tengamos una teoría —una teoría de la verdad, o algo parecido a ella— que sea parte de la descripción del castellano. (Esta descripción de lo que conoce el hablante del castellano no tiene por qué enunciarse en castellano y entonces no suena tan trivial.)

Volviendo ahora a la cuestión con la que empecé: ¿Cuántos hablantes o intérpretes de una lengua debe haber para que haya un hablante? Voy a comenzar con una primera dificul-

tad clara. Para hablar una lengua, las preferencias de uno deben ser consistentes con la definición de alguna lengua. El problema es que las preferencias son finitas, mientras que la definición de una lengua asigna significados a un número infinito de oraciones. Habrá, pues, infinidad de lenguas diferentes que concuerden con todas las preferencias reales de un hablante, pero que difieran con respecto a las oraciones no preferidas. ¿Qué es lo que hace al hablante un hablante de una de esas lenguas en lugar de otra? El problema puede ser aún peor, puesto que si un hablante profiriera todas las oraciones de una lengua (aunque esto es imposible), muchas otras lenguas serían consistentes con todo su comportamiento y sus estados internos, tal y como Quine ha mantenido y yo también acepto.

El hecho de que todas las evidencias públicamente asequibles por lo que respecta a un hablante o grupo de hablantes, incluso si imaginamos que agota todas las evidencias posibles de ese tipo, puedan ser consistentes con muchas lenguas distintas (en el sentido de «lengua» que hemos decretado temporalmente) no debería en sí mismo preocuparnos, puesto que podemos estar de acuerdo en que es suficiente saber que el hablante habla cualquier lengua de entre un conjunto de lenguas empíricamente equivalentes, mientras que las constricciones empíricas definan al conjunto claramente. De modo que de ahora en adelante vamos a llamar simplemente a cualquier lengua de este conjunto «la lengua» de un hablante⁴. Esta estrategia es suficientemente buena para lenguas empíricamente equivalentes si se imagina que las evidencias contienen una preferencia de cada oración que pudiéramos contar como perteneciente a la lengua en cuestión. Pero, por supuesto, tales evidencias no son nunca asequibles. Así pues, habrá

⁴ Considero que la existencia de lenguas empíricamente equivalentes (esto es, lenguas igualmente consistentes con todas las evidencias empíricas posibles) como algo que no es más amenazador para la realidad u objetividad de la interpretación correcta de las preferencias y de sus estados mentales concomitantes de lo que lo es para la realidad u objetividad de la temperatura o la longitud la existencia de diversas escalas que registran temperaturas o longitudes. (Véase el Ensayo 5.)

un sinfín de lenguas consistente con todas las preferencias de un hablante y ninguna de ellas será «la» lengua que el hablante habla.

Puede plantearse el problema de una manera temporal y dirigirlo hacia el intérprete. Si usted (como intérprete) no sabe cómo va a continuar el hablante entonces no sabe qué lengua habla, sin que importe cuánto haya dicho hasta entonces. No servirá de ninguna ayuda mencionar que el hablante ha actuado de acuerdo con lo esperado, o que fue a la misma escuela que usted, o que pertenece a la misma cultura o comunidad, puesto que la cuestión tiene que ver no con el pasado sino con el futuro. Tampoco podemos recurrir a la idea de que el hablante ha dominado un conjunto de convenciones (¿qué convenciones?), o que ha aprendido un conjunto de reglas (¿qué reglas?). Para justificar o explicar el comportamiento lingüístico no puede recurrirse a los conceptos de convención o regla, como no puede recurrirse al concepto de lengua; en el mejor de los casos estos conceptos ayudan a describir (es decir, definir) el comportamiento lingüístico.

Esta dificultad, aunque pudo haber preocupado a Wittgenstein, y ciertamente preocupó a Kripke⁵, me parece que tiene una respuesta relativamente simple. Cuanto más interpretemos con aparente éxito a un hablante como alguien que habla una determinada lengua, una mayor confianza legítima tendremos de que el hablante está hablando esa lengua —esto es, que continuará siendo interpretable como un hablante de esa lengua—. Nuestras expectativas reforzadas están tan bien fundamentadas como lo permiten nuestras evidencias y la inducción común. Estas expectativas son en su mayoría condicionales. Usualmente no sabemos lo que alguien dirá, pero estamos preparados para interpretar cualquiera de entre el gran número de cosas que una persona podría decir. Nuestras disposiciones a interpretar y las disposiciones de un hablante para seguir de cierta manera no son oscuras o misteriosas: son rasgos reales de los cerebros y de los músculos. Es evidente que nuestras creencias sobre lo que es verdadero acerca de

⁵ Saul Kripke, *Wittgenstein on Rules and Private Language*.

otra persona, y por lo tanto, lo que esperamos que esa persona signifique con lo que dice o podría decir, pueden muy bien estar equivocadas. Creo que tales creencias frecuentemente están equivocadas, pero mucho más frecuentemente están en lo cierto, y las cosas de las que estamos en lo cierto usualmente nos sitúan en posición de corregir nuestra comprensión de una preferencia que no pertenece a la lengua que pensábamos que se estaba hablando. En tanto que estamos en lo cierto sobre lo que hay en la cabeza de alguien, y por ello estamos en lo cierto sobre lo que significaría mediante un sinnúmero de cosas que no dice, estamos en lo cierto sobre «la» lengua que habla.

Esta respuesta tan parcial a la pregunta de qué razones puede tener un intérprete para creer que un hablante está hablando una lengua en lugar de otra que sea igualmente compatible con el comportamiento lingüístico que el intérprete ha observado, no depende muy fuertemente de los detalles de cómo explicamos la interpretación exitosa. La aportación principal de la respuesta es que no hay *dos* preguntas, una acerca de las razones para creer que un hablante está hablando una lengua en lugar de otra, y una segunda acerca de cómo formamos expectativas de manera natural; la primera pregunta es simplemente un caso de la segunda.

Sin embargo, hay otro aspecto de la interpretación que es esencial para nuestros propósitos: un intérprete interpreta (correctamente) una preferencia de un hablante solamente si sabe que el hablante pretende que el intérprete asigne ciertas condiciones de verdad a su preferencia (la del hablante)⁶. Una explicación completa de esta tesis requeriría una explicación de la idea de «asignar condiciones de verdad» a una preferencia, y esta idea es sin duda tan difícil de comprender en los aspectos relevantes como el propio concepto de significado.

⁶ Tal y como está, esto es claramente inadecuado. Se puede mejorar añadiendo la condición griceana de que el hablante pretende que el intérprete llegue a las condiciones de verdad correctas a través del reconocimiento del intérprete de la intención del hablante de que se le interprete así. No argumentaré aquí a favor del supuesto de que el conocimiento de las condiciones de verdad es adecuado para la interpretación.

Pero mi propósito aquí no es resolver tal problema, sino sólo el enfatizar, siguiendo a Grice, la importancia capital de la intención en la comunicación. Si como Grice, estuviéramos seguros de que para significar o querer decir algo un hablante debe pretender tener un determinado efecto en un oyente u oyentes específicos, entonces se podría haber mostrado ya que el lenguaje es social en tanto que requiere la existencia de como mínimo dos personas (ya que es defendible que uno no podría pretender tener un efecto en otra persona a menos que tal persona exista). No seguiré esta dirección tan directa y tentadora. Sin embargo, estamos en condiciones de decir que si la comunicación tiene éxito deben existir estas intenciones por parte del hablante, y por lo tanto, *si* la comunicación exitosa es esencial para el significado, estas intenciones son esenciales para el significado. La presencia de intenciones es importante, ya que da contenido a la atribución de errores al permitir la posibilidad de una discrepancia entre intenciones y logros. Una intención, igual que una creencia o una expectativa, no requiere atención o reflexión, y usualmente no se llega a las intenciones mediante el razonamiento consciente. Normalmente, no se atiende a las intenciones mediante sensaciones especiales, ni tampoco usualmente se llega al conocimiento de nuestras propias intenciones mediante inferencia o apoyándonos en la observación. Con todo, la intención tiene un alcance indefinidamente amplio, puesto que las intenciones dependen de la creencia de que uno puede hacer lo que pretende o intenta hacer, y esto requiere que uno no crea que nada impedirá la acción que se intenta realizar. Así pues, parecería que las intenciones tienen precisamente las propiedades necesarias para dar sentido a la idea de que un hablante no ha continuado como antes⁷.

El concepción que he esbozado trata sólo de la interpretación y por ello presupone un entorno social en lugar de ofre-

⁷ En esencia, estas cosas que he señalado sobre la intención ya las ha señalado Crispin Wright al intentar quitarle fuerza, como yo, a la pretensión de Kripke de haber extraído una «paradoja escéptica» esencialmente insoluble del análisis del significado que hizo Wittgenstein. Véase Crispin Wright, «Kripke's Account of the Argument against Private Language».

cer un argumento para ello. Sin embargo, en este momento resultará útil tener en cuenta ciertos aspectos de la concepción que creo que Kripke atribuye a Wittgenstein. (Por razones expositivas atribuiré esta concepción a Kripke.) En realidad, Kripke no dice claramente que la acepte y no estoy seguro de que recoja el punto de vista de Wittgenstein. Así que quizá no refleje el punto de vista de nadie. Kripke se concentra en la idea de seguir una regla. De acuerdo con esta idea hablar una lengua es seguir reglas. Las reglas especifican en qué consiste seguir «de la misma manera»; por ejemplo, cómo utilizar una palabra. Sin embargo, no hay un acto o proceso mental interno de «captar» o de «seguir» una regla, de modo que ningún estudio o conocimiento de lo que hay dentro del hablante revelará si está siguiendo un determinado conjunto de reglas u otro. Los intérpretes simplemente juzgan que un hablante está siguiendo la misma regla que ellos (que sus intérpretes) si el hablante prosigue como ellos lo harían. Dicho en términos de significado: juzgamos que un hablante significa o quiere decir lo que significaríamos o querríamos decir nosotros si profiriéramos las mismas palabras si ese hablante prosiguiera como nosotros lo haríamos⁸.

Deberíamos preguntarnos sobre la adecuación del concepto común de seguir una regla para describir lo que supone hablar una lengua. Cuando hablamos de reglas del lenguaje, usualmente pensamos en las descripciones que los gramáticos o los lingüistas hacen de la práctica real (generalizada e idealizada).

⁸ «... ¿qué quiero decir cuando digo que el instructor juzga que, en ciertos casos, el alumno debe dar la respuesta “correcta”? Quiero decir que el instructor juzga que el alumno ha dado la misma respuesta que él mismo daría... si, en suficientes casos concretos, las inclinaciones de Jones concuerdan con las de Smith, Smith juzgará que verdaderamente Jones está siguiendo la regla» (Kripke, *Wittgenstein on Rules and Private Language*, 90-1). Al siguiente pasaje se le puede dar tal interpretación: «alguien se guía por un indicador de caminos sólo en la medida en que haya un uso estable, una costumbre.... ¿Es lo que llamamos “seguir una regla” algo que podría hacer sólo *un* hombre sólo *una* vez en la vida?... Seguir una regla, hacer un informe, dar una orden, jugar una partida de ajedrez son *costumbres* (usos, instituciones)» (*Philosophical Investigations*, §§ 198, 199). He obviado un aspecto muy importante de la discusión de Kripke, a saber, su afirmación de que la «solución» de Wittgenstein al problema del significado es «escéptica».

zada), 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 reglas⁹. Por supuesto que Wittgenstein trata el significado de manera muy similar a como trata el seguir algún procedimiento como, por ejemplo, la suma en aritmética. Pero hay una distinción clara entre ambos casos que explica por qué normalmente usamos la palabra «regla» en un caso y no en el otro. En el caso de la suma hay un procedimiento explícito para llegar a la respuesta; podemos aprender y describir ese procedimiento y es apropiado llamar regla a este procedimiento o a su descripción. Normalmente no seguimos un procedimiento al hablar; nada en el habla normal corresponde a la suma de una columna de números. Si el concepto de seguir una regla no es lo suficientemente apropiado para describir lo que es significar algo al decir algo es también cuestionable, incluso si estamos de acuerdo en que el uso de una lengua requiere un marco social, si deberíamos aceptar sin dudar la idea de que significar algo exige una convención, costumbre o institución (por contraste con implicarlo algunas veces)¹⁰.

Una cuestión más importante tiene que ver con la idea de que la comunicación lingüística requiere que un hablante siga en la misma dirección que los demás, que para significar algo al hablar, debemos significar la misma cosa y con las mismas palabras que los demás. La explicación que he ofrecido más arriba de la clase de expectativas que se deben satisfacer para que una persona comprenda a otra no sugería que esas personas tuvieran que hablar la misma lengua. Tampoco queda cla-

⁹ Debería ser evidente que la afirmación de que hay reglas internalizadas o implantadas genéticamente es irrelevante en este contexto; las «reglas» de Wittgenstein y Kripke tienen que ver con lo que significan las palabras de las lenguas particulares.

¹⁰ He expresado mi escepticismo acerca del poder explicativo de los conceptos de seguir una regla y de convención en el estudio del lenguaje de manera más amplia en «Communication and Convention», ensayo 18 en *Inquiries into Truth and Interpretation*.

ro por qué esto es necesario. Quizá el lenguaje no habría existido a menos que pudiera depender de la tendencia natural de los animales a imitarse los unos a los otros. Aunque tengo mis dudas puede que fuera así, pero seguramente podría haber sido de otra manera. Si usted y yo fuéramos los únicos hablantes del mundo y usted hablara sherpa y yo inglés, nos podríamos entender, aunque cada uno de nosotros seguiríamos distintas «reglas» (regularidades). Por supuesto que lo que importaría es que cada uno de nosotros deberíamos proporcionarle al otro algo que se pudiera entender como una lengua. Ésta es una intención que deben tener los hablantes; pero llevar a cabo esta intención, aunque pueda requerir un cierto grado de lo que los demás perciben como consistencia, no implica seguir reglas o convenciones compartidas. Podría darse incluso el caso de que, debido a diferencias en nuestras cuerdas vocales, no pudiéramos hacer los mismos sonidos y entonces no podríamos hablar la misma lengua. No conozco ningún argumento que muestre que en tales circunstancias la comunicación no pudiera tener lugar. Así pues, aunque pueda ser verdad que hablar una lengua requiere que haya un intérprete, no se sigue que más de una persona deba hablar la misma lengua. Esto está bien porque si somos precisos sobre lo que constituye una lengua es probable que no haya dos personas que realmente hablen la misma lengua. Concluyo que el criterio de Kripke de lo que es hablar una lengua no puede ser correcto; hablar una lengua no puede depender de hablar como alguien más lo hace (o muchos otros lo hacen)¹¹.

¹¹ Mi madre, como lo hacen las madres entusiastas, registraba la lengua que yo hablaba cuando tenía tres años. Ciertamente no era la lengua que hablaban otras personas de la familia ni, imagino, por nadie más. Sin embargo, ella afirma, correctamente, que me entendía y yo a ella. Y es común que personas que no pueden o no hablan ni escriben francés contesten en inglés cartas escritas en francés.

En «A Nice Derangement of Epitaphs» he defendido que la comunicación no requiere que se comparta una lengua. La misma posición la defiende Noam Chomsky, *Language and Problems of Knowledge*, 36-7. Para un tratamiento más extenso de este tema dirigido directamente a Kripke, véase Chomsky, *Knowledge of Language: Its nature, Origin and Use*, 223-37.

Supongamos entonces que la prueba de que se habla una lengua se modifica para acomodar esta idea; afirmaremos ahora que hablar una lengua no depende de que dos o más hablantes hablen de la misma manera; solamente requiere que cada hablante se haga intencionalmente interpretable al otro (el hablante debe «seguir» más o menos como el otro espera, o al menos como esté equipado para interpretarlo).

Ésta es ciertamente una condición necesaria para la *comunicación* exitosa. Pero ¿por qué es una condición que se deba satisfacer para que se diga que se habla una lengua? ¿Por qué no puede ser que alguien siga del mismo modo —es decir, satisfaga todas las condiciones para ser *interpretable*— sin que realmente sea interpretable?¹²

Es verdad que nuestras evidencias sobre si alguien habla una lengua particular se basan principalmente en el hecho de que sigue como esperamos que un hablante de esa lengua siga. Por supuesto que hay muchas otras clases de evidencias; las ropas del hablante, sus compañeros, su ubicación en la tierra; todas ellas pueden ser pistas de su lengua. Con todo, podemos estar de acuerdo en que el concretar las cosas al final debe depender de los detalles del comportamiento lingüístico. El problema es que la cuestión original no tiene que ver con las condiciones para la comunicación ni con la cuestión de qué evidencias podría tener una persona de que otra estaba hablando una lengua particular; el tema era por qué la única lengua del hablante o la primera no podría ser privada.

Nuestro análisis ha conducido a una modificación o elucidación del concepto de lengua o lenguaje privado: con esto entiendo, no una lengua o lenguaje que solamente habla una persona, sino una lengua o lenguaje que entiende solamente

¹² Kripke parece admitir que se debería juzgar que Robinson Crusoe estaría hablando una lengua en tanto que se le *podría* incluir en una sociedad, incluso si de hecho no está (¿o no ha estado nunca?) en un marco social. Con todo, debería haber una sociedad (o al menos otra persona) para juzgar legítimamente que Robinson Crusoe significa algo mediante los ruidos que emite. Chomsky piensa que al admitir el caso de Robinson Crusoe, Kripke contradice su tesis principal. Quizá sí; pero pienso que Chomsky se equivoca en pensar que es posible el caso puro de Robinson Crusoe. Por caso puro quiero decir un Robinson Crusoe que nunca se ha comunicado con otras personas.

una persona. La cuestión ahora es, ¿por qué no puede haber una lengua o lenguaje que sólo entienda una persona?¹³ La respuesta que Wittgenstein parece ofrecer en el pasaje que sirve de encabezamiento a este ensayo es la siguiente: sin un intérprete no se puede dar contenido a la afirmación de que el hablante se ha equivocado —que ha dejado de seguir en la misma dirección.

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)? Si el comportamiento lingüístico de los demás no proporciona la norma al hablante, ¿qué lo puede hacer? La respuesta es que la intención del hablante de ser interpretado de cierta manera suministra esa «norma»; el hablante no realiza su intención si no habla de tal manera que sea entendido tal como él pretendía. En las circunstancias usuales un hablante sabe que tiene más oportunidades de ser entendido si habla como lo harían sus oyentes, y por tanto pretenderá hablar como piensa que ellos lo harían. Por tanto, no cumplirá una de sus intenciones si no habla como lo hacen los demás. Creo que este simple hecho contribuye a explicar por qué muchos filósofos han ligado el significado de las preferencias de un hablante a lo que los demás significan mediante las mismas palabras (donde el término «los demás» se refiere a una comunidad lingüística, a los expertos, o a una élite de una clase u otra)¹⁴. En mi explica-

¹³ Por supuesto que puede haber una «lengua o lenguaje» que comprenda solamente una persona, por ejemplo un código secreto que se use en un diario. La cuestión es si una primera lengua puede ser privada.

¹⁴ Como ejemplos de esto véase Hilary Putnam, «The Meaning of "Meaning"»; Tyler Burge, «Individualism and the Mental»; Michael Dummett, «The Social Character of Meaning», en *Truth and Other Enigmas*, y «A Nice Derangement of Epitaphs»: Some Comments on Davidson and Hacking», en Ernest Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*.

ción, este vínculo no es ni esencial ni directo; solamente entra en juego cuando el hablante pretende ser interpretado como (ciertos) otros lo harían. Cuando esta intención está ausente, la comprensión correcta de un hablante no se ve afectada por el uso que vaya más allá del alcance que pretende que tenga su voz. (Una intención fallida de hablar «correctamente», a menos que fustre la intención de ser interpretado de cierta manera, no tiene importancia para lo que el hablante significa)¹⁵.

Lo que muestran estas consideraciones, si es que son correctas, es que hay una alternativa más débil y más plausible a la explicación propuesta por Kripke de lo que se requiere para significar algo mediante lo que uno dice. La razón es que aunque la explicación de Kripke propone que la prueba de si un hablante significa algo depende de que haga lo que otros hacen, la misma distinción entre pensar que alguien significa algo y significarlo realmente puede hacerse en términos del éxito de la intención del hablante de ser interpretado de determinada manera. Ambas formas de hacer la distinción dependen del marco social, pero la segunda requiere cosas distintas del hablante¹⁶.

¿Hemos mostrado ahora que no puede haber un lenguaje privado? Seguro que no. Si aceptamos que la propuesta de Kripke es correcta, entonces es verdad que una manera de distinguir pensar que uno significa algo y realmente significarlo requiere que el lenguaje sea público; lo mismo se puede decir de la alternativa que he sugerido, pero no se ha dicho nada definitivo que muestre que no puede haber alguna otra manera de hacer la distinción, una manera que no dependa de un entorno social.

¹⁵ Esta cuestión también se discute en el Ensayo 2.

¹⁶ Hay una idea a la que no he dado cabida. La intención de un hablante falla si no se le interpreta como pretende. Pero sería erróneo decir que tal fallo necesariamente es un fallo en dar el significado de sus palabras que pretendía que el intérprete captara. El segundo fallo depende (de modos que el uso cotidiano no puede fijar definitivamente) de cuestiones como si el hablante estaba justificado en creer que su intérprete podía interpretarle —o que le interpretaría— de la manera que pretendía que lo hiciera.

Si hemos de establecer el carácter esencialmente público del lenguaje, necesitamos un tipo de argumento totalmente distinto. En lo que queda de ensayo voy a proporcionar un argumento así, un argumento que se aplica no solamente al lenguaje sino también a la creencia, la intención y al resto de las actitudes proposicionales. El argumento que sigue no empieza con una duda escéptica a la cual se busca una respuesta, pero termina con lo que podría ser una conclusión de Wittgenstein: un lenguaje es necesariamente un asunto social.

Consideremos en primer lugar una situación primitiva de aprendizaje. Se enseña a una criatura —o de alguna forma ésta aprende— a responder de una manera específica a un estímulo o a una clase de estímulos. El perro oye una campana y se le alimenta; dentro de poco saliva cuando oye la campana. El niño balbucea, y cuando produce un sonido como «mesa» ante la presencia evidente de una mesa, se le recompensa; el proceso se repite y en poco tiempo el niño dice «mesa» en presencia de mesas. El fenómeno de la generalización o de la similitud percibida desempeña un papel esencial en este proceso. Un toque de campana es lo suficientemente parecido a otro como para provocar respuestas similares en el perro, de la misma manera que una presentación de comida es lo suficientemente parecida a otra para provocar la salivación. Si algunos mecanismos discriminativos como éstos no estuvieran en nuestros genes, no podría aprenderse nada. Lo mismo vale para el niño: podemos clasificar los estímulos del niño mediante la similitud de las respuestas que estos estímulos provocan en él¹⁷.

Esto parece sencillo, pero como los psicólogos han señalado hay un problema relativo a los estímulos. En el caso del perro, ¿por qué hemos de decir que el estímulo es el toque de campana? ¿Por qué no podría ser la vibración del aire cerca de los oídos del perro —o incluso la estimulación de sus terminaciones nerviosas? Ciertamente, si se hiciera vibrar el aire

¹⁷ Nada aquí depende de la psicología de aficionados acerca de los balbuceos, las distintas recompensas o la inducción preverbal. Todo lo que importa es el hecho de que la generalización tiene lugar de una manera y no de otra, y de maneras parecidas en distintas personas.

de la misma manera en que la campana lo hace vibrar, no habría diferencia en la respuesta del perro. Y si las terminaciones nerviosas correctas fueran activadas de la manera correcta, tampoco habría diferencia. De hecho, si debemos escoger, parece que es a la causa próxima del comportamiento a la que mejor se puede llamar estímulo, ya que cuanto más causalmente distante está el acaecimiento de su perceptor, más posibilidades hay de que la cadena causal se rompa. ¿Por qué no decir lo mismo sobre el niño: que responde a patrones de estimulación en sus superficies y no a mesas, dado que estos patrones de estimulación siempre producen la respuesta mientras que las mesas la producen solamente en circunstancias favorables?

¿Qué es lo que explica el hecho de que parece tan natural decir que el perro responde a la campana y el niño a las mesas? Nos parece natural porque *es* natural —para nosotros—. Del mismo modo en que el perro y el niño responden de maneras similares a ciertos estímulos, también lo hacemos nosotros. Somos nosotros quienes encontramos natural agrupar las distintas salivaciones del perro; y los toques de campana son los acaecimientos del mundo que están causalmente ligados al comportamiento del perro, que notamos sin esfuerzo y que agrupamos. Nos parecen similares las articulaciones de la palabra «mesa» que hace el niño, y los objetos del mundo que agrupamos juntos de forma natural y que acompañan a estas articulaciones es un conjunto de mesas. Los patrones acústicos y visuales que transcurren a sus respectivas velocidades entre la campana y los oídos del perro, o entre las mesas y los ojos del niño, no los podemos observar fácilmente, y si pudiéramos hacerlo sería difícil para nosotros poder decir qué es lo que los hace similares. (Excepto haciendo trampa, por supuesto: son los patrones característicos de las campanas que sueñan o de las mesas que se ven.) Tampoco observamos la estimulación de las terminaciones nerviosas de otras personas o animales, y si lo hiciéramos podría ser imposible describir de una manera no circular qué es lo que hace que esos patrones sean relevantemente similares de un caso a otro. El problema sería muy parecido al problema (insoluble) de definir las mesas y los toques-de-campana en términos de datos sensoriales sin mencionar las mesas o las campanas.

En nuestra descripción de la situación hay ahora involucrados no dos, sino tres patrones de similaridad. Al niño le parece que las mesas son similares; a nosotros nos parece que las mesas son similares; y nos parece que las respuestas del niño en presencia de las mesas son similares. Ahora tiene sentido que nosotros consideremos las respuestas del niño como respuestas a mesas. Dados estos tres patrones de respuesta podemos asignar un lugar a los estímulos que provocan las respuestas del niño. Los estímulos relevantes son aquellos objetos o acaecimientos que nos parecen similares de forma natural (las mesas) y que se correlacionan con respuestas del niño que nos parecen similares. Es una forma de triangulación: una línea va del niño a la mesa, otra va de nosotros a la mesa y la tercera va de nosotros al niño. «El» estímulo se sitúa donde convergen las líneas que van del niño a la mesa y de nosotros a la mesa. Dada nuestra visión del niño y del mundo, podemos señalar «la» causa de las respuestas del niño. Es la causa común de nuestra respuesta y de la respuesta del niño¹⁸.

Se han establecido las suficientes características como para dar significado a la idea de que el estímulo tiene un lugar objetivo en un espacio común; pero nada en esta descripción muestra que nosotros, los observadores, o nuestros sujetos, el perro y el niño, tengamos esta idea. Sin embargo, hemos hecho un buen recorrido, puesto que si estoy en lo cierto, aunque la clase de triangulación que he descrito no sea suficiente para establecer que una criatura tiene un concepto de un objeto particular o de una clase de objeto, sí que es necesaria si es que ha de haber alguna respuesta a la pregunta siguiente: ¿de qué son conceptos sus conceptos? Si consideramos una sola criatura por sí misma, por complejas que sean sus respuestas no pueden mostrar que está reaccionando a, o pensando en, acaecimientos que se encuentran a determinada distancia en lugar de, por ejemplo, situarse en su propia piel.

¹⁸ En el Ensayo 7 se introduce la metáfora de la triangulación. En el Ensayo 10 se desarrolla la idea de la «causa común». En «Thought and Talk», el Ensayo 11 de *Inquiries into Truth and Interpretation*, y en el Ensayo 7 de este volumen pongo el acento en que el pensamiento es un fenómeno social.

El mundo solipsista puede tener cualquier tamaño; lo que equivale a decir que, desde el punto de vista solipsista, no tiene tamaño alguno, no es un mundo.

Debería remarcar que el problema no es el de verificar a qué objetos o acaecimientos está respondiendo una criatura; lo importante es que sin que haya una segunda criatura que responda a la primera, no puede darse respuesta a la pregunta. Y por supuesto, si no hay respuesta a esta pregunta, no hay respuesta a la pregunta de qué lengua o lenguaje habla la criatura, ya que designar el lenguaje que se está hablando requiere que las preferencias concuerden con los objetos y acaecimientos del mundo (y no, en general, con los acaecimientos que se dan en la superficie de la piel). Así pues, podemos decir, como algo preliminar a la respuesta a la pregunta con la que empezamos, que antes de que alguien pueda hablar una lengua, debe haber otra criatura que interaccione con el hablante. Por supuesto que esto no es suficiente, ya que la mera interacción no muestra en qué modo importa la interacción a las criaturas involucradas. A menos que se pueda decir de las criaturas en cuestión que reaccionan a la interacción, no hay manera de que *ellas* obtengan ventajas cognitivas de la relación a tres que da contenido a *nuestra* idea de que están reaccionando a una cosa y no a otra.

He aquí una parte de lo que pienso que se necesita. La interacción debe ser accesible a las criaturas que interaccionan. Así pues, el niño, al aprender la palabra «mesa» ya ha sentido que las respuestas del instructor son similares (suponen una recompensa) cuando sus propias respuestas (articular la palabra «mesa») lo son. Por su parte, el instructor está entrenando al niño a dar respuestas similares a lo que él (el instructor) percibe como estímulos similares. Para que esto funcione está claro que las respuestas similares innatas del niño y del instructor —lo que agrupan de manera natural— debe ser parecido; de lo contrario el niño respondería a lo que el instructor toma como estímulos similares de una forma que el instructor no encontraría similar. Una condición para ser un hablante es que debe haber otros que sean suficientemente parecidos.

Hasta el momento he dejado de lado una explicación explícita de los conceptos de creencia y de intención que son cla-

ramente esenciales para hablar una lengua. No pienso tratar de introducir estos conceptos en términos de las simples situaciones de condicionamiento que he estado describiendo; el concepto de pensamiento no es reducible a ninguna otra cosa y mucho menos a estos conceptos simples. Todo lo que he intentado mostrar es que la interacción entre criaturas similares es una condición necesaria para hablar una lengua.

A continuación voy a reunir dos ideas. La primera es que si alguien es un hablante de una lengua debe haber otro ser sintiente cuyas respuestas innatas a la similaridad sean suficientemente parecidas a las propias para dar una respuesta a la pregunta: ¿cuál es el estímulo al cual el hablante está respondiendo? Y la segunda es que si las respuestas del hablante son lingüísticas deben ser a sabiendas e intencionalmente respuestas a estímulos específicos. El hablante debe tener el concepto de estímulo —de la campana o de las mesas—. Dado que una campana o una mesa se identifican solamente mediante la intersección de dos (o más) conjuntos de respuestas a la similaridad (casi podríamos decir que líneas de pensamiento), tener el concepto de una mesa o una campana es reconocer la existencia de un triángulo, un ángulo del cual es uno mismo, siendo el segundo otra criatura similar a uno mismo y el tercero un objeto (una mesa o una campana) ubicado en un espacio hecho común de esta manera.

La única forma de saber que el segundo ángulo del triángulo —la segunda criatura o persona— está reaccionando al mismo objeto que uno mismo 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 un intérprete de la otra.

Las precipitadas observaciones de los dos últimos párrafos indican la clase de trabajo que sería necesario para dar una ex-

plicación del habla con significado. Pero no era mi propósito ofrecer una explicación de este tipo; solamente trataba de encontrar un argumento a favor de la idea de que una primera lengua no puede ser privada.

El argumento muestra que la primera lengua no puede ser un lenguaje privado, esto es, un lenguaje que sólo comprenda una criatura, y en esta medida está de acuerdo con el Wittgenstein de Kripke. Pero el argumento toma un camino distinto y es diferente el papel que otorga al aspecto social del lenguaje. Kripke depende de la segunda persona o de una comunidad para incorporar una rutina que el hablante pueda compartir. En contraste, el argumento que he esbozado no requiere (aunque por supuesto que permite) una rutina compartida, sino que depende de la interacción de al menos dos hablantes-intérpretes, puesto que si estoy en lo cierto, sin interacción con una segunda persona no habría manera de decir de qué está hablando el hablante o en qué está pensando y no habría base para afirmar que podría ubicar objetos en un espacio y tiempo objetivos.

Todas estas consideraciones no se aplican solamente al lenguaje, sino que se aplican igualmente al pensamiento en general. 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 éste 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.

La aparición del pensamiento

Según la Biblia, la Palabra estaba allí desde el principio, pero parece más plausible que las palabras y el pensamiento apareciesen más recientemente. Si bien, ¿qué quiere decir que algún aspecto del mundo ha «aparecido»?

La aparición solamente tiene sentido concebida desde un punto de vista, desde un conjunto de conceptos. Dado *nuestro* conjunto de conceptos, podemos comprender la idea de que diferentes conceptos fueran ejemplificados en tiempos diferentes. Los conceptos de la física cuántica se ejemplificaron muy temprano; los conceptos de los diversos elementos se ejemplificaron a lo largo de un periodo posterior. El concepto de vida se ejemplificó bastante recientemente y los conceptos de pensamiento y lenguaje se llenaron sólo hace unos momentos con la llegada de los mamíferos superiores. Esto no quiere decir que las leyes de la física cambiasen, sino que lo que quiere decir es que para describir y explicar el pensamiento necesitamos conceptos que no pueden definirse en el vocabulario de la física (o de ninguna de las ciencias naturales). Así pues, en un claro sentido, la aparición es relativa a un conjunto de conceptos. Los propios conceptos son abstracciones y por ello no se encuentran en un tiempo determinado, pero puede ocurrir que ciertos conceptos se ejemplifiquen solamente con el transcurrir del tiempo. Es innecesario decir que el universo es indiferente a nuestros conceptos. Pero nosotros tomamos muy en cuenta la ejemplificación de los conceptos de los que depende el concepto de pensamiento; son concep-

tos como creencia, deseo, intención, acción intencional, memoria, percepción y todo el resto de actitudes y atributos humanos.

Voy a empezar diciendo por qué según mi punto de vista es tan difícil intentar decir algo interesante y profundo acerca de la aparición del pensamiento. La razón es que hay muchos conceptos que debemos abordar para poder hablar sobre o describir el pensar, actuar por una razón, creer o dudar, todos los cuales dependen unos de los otros. Éste es el *holismo de lo mental*, la interdependencia de varios aspectos de lo mental. Dentro de cada dimensión de lo mental, como, por ejemplo, la creencia, parece claro que es imposible aproximarnos a ella de manera atomista, porque es imposible dar sentido a la idea de tener solamente una o dos creencias. Las creencias no se dan una cada vez; lo que identifica una creencia y hace de ella la creencia que es es su relación (entre otros factores) con otras creencias. Uno no puede creer que está viendo un gato sin creer muchas otras cosas: debe saber lo que es un gato, lo que es ver, y sobre todo, debe reconocer la posibilidad, por remota que sea, de que podría estar equivocado. Algunas personas suponen que un perro podría tener una creencia aislada, pero a mí me parece que por las mismas razones que estoy enumerando, los perros no tienen creencias u otras actitudes proposicionales. No forman juicios.

La razón por la que ni los perros ni cualquier otra criatura pueden tener una creencia aislada, como la de que está viendo un gato, es que lo que identifica una creencia es lo que imprecisamente llamamos *contenido proposicional*. Así pues, para tener una creencia sobre un gato debemos dominar los conceptos que están involucrados en este juicio o creencia. Una criatura no tiene el concepto de gato meramente porque pueda discriminar los gatos de otras cosas de su entorno. Los ratones son muy buenos a la hora de distinguir a los gatos de los árboles, los leones y las serpientes. Pero ser capaz de discriminar los gatos no es lo mismo que tener el concepto de gato. Se tiene el concepto de gato solamente si se puede dar sentido a la idea de *aplicar erróneamente* ese concepto, o de creer o juzgar que algo es un gato cuando no lo es. Para tener el concepto de gato debemos tener el concepto de animal, o al menos de

un objeto físico continuo, el concepto de un objeto que se mueve de determinada manera, de algo que puede moverse libremente en su entorno, de algo que tiene sensaciones. No hay una lista fija de cosas que se tienen que saber sobre los gatos o que se tiene que asociar a la felinidad; pero a menos que tengamos muchas creencias sobre lo que es un gato, no tenemos el concepto de gato.

Se debe tener un gran número de creencias para tener una, por el hecho de que las creencias se individualan e identifican por sus relaciones con otras creencias. Las creencias se apoyan unas a otras y se dan el contenido mutuamente. Las creencias también tienen relaciones lógicas entre ellas. Como resultado de todo ello, a menos que las creencias de uno sean vagamente consistentes entre sí, no hay manera de identificar los contenidos de las creencias. Por ello, una condición para tener creencias es un cierto grado de racionalidad o consistencia.

Esto no significa decir que todo el mundo es perfectamente racional; cada uno de nosotros es capaz de cometer un error lógico, o de abrigar creencias que son inconsistentes con las demás. (Por ejemplo, tanto Frege como Quine escribieron libros basados en lógicas inconsistentes.) Pero hay un límite a cuán inconsistente puede ser una persona y a pesar de ello atribuirle actitudes claramente definidas. Toda creencia debe estar relacionada con otras creencias con las cuales sea consistente si es que se la identifica como algo con un contenido claro. La inconsistencia u otras formas de irracionalidad solamente se dan en el espacio de las razones; las inconsistencias son perturbaciones de la racionalidad, no meramente la ausencia de racionalidad.

Por ello insisto en que para tener algún concepto se debe tener un conjunto plenamente desarrollado de conceptos básicos. Y apenas si he empezado a bosquejar el grado de interdependencia que hay entre las creencias, puesto que además de las relaciones lógicas hay relaciones de apoyo evidencial. Aunque pocos conceptos se puedan definir en términos de su contenido empírico, en algunas cuestiones fácticas una creencia debe contar a favor de la verdad de otras; tales relaciones de apoyo evidencial o inductivo son esenciales para la identidad de muchas creencias.

Además de las relaciones entre creencias hay relaciones entre creencias, por un lado, y actitudes evaluativas, por el otro. Es dudoso que se pudiera decir que una criatura tiene creencias si no tuviera también deseos, puesto que un aspecto esencial de las creencias es que afecten al comportamiento o éste las ponga de manifiesto. Las actitudes evaluativas, incluyendo los deseos, las intenciones, las convicciones morales, lo que se consideran deberes y obligaciones, también son actitudes proposicionales. La gramática de las expresiones que usamos para atribuir tales actitudes no lo hace siempre evidente. Por ejemplo, hablamos de querer una manzana. Si alguien dice: «Quiero una manzana», puede ser de mala educación preguntar: «¿Y para qué la quieres?», pero siempre resulta apropiado. No tiene sentido decir: «Sólo quiero la manzana, no por nada en especial.» Podría replicar: «Muy bien, ahí la tienes. Está allí en el árbol. ¿Qué más quieres?» La respuesta podría ser: «La quiero en la mano; o en la cesta; o fuera del árbol y en el estómago, etc.» El querer cosas y los deseos se dirigen a contenidos proposicionales. Lo que uno quiere es *que* sea verdad que uno tiene la manzana en la mano, o a punto para comerla. De manera similar ocurre con las intenciones. Alguien que pretenda ir a la ópera pretende que sea el caso *que* esté en la ópera.

Las creencias y los deseos conspiran para causar, racionalizar y explicar las acciones intencionales. Actuamos intencionalmente por razones, y nuestras razones siempre incluyen tanto valores como creencias. No actuaríamos a menos que hubiera algún valor o algún fin que esperaríamos conseguir (o algún mal que esperaríamos evitar), y creyéramos, por supuesto, que nuestro curso de acción fuera una forma de realizar nuestro propósito. La teoría de la decisión es una manera de sistematizar las relaciones entre creencias, deseos y acciones. Hace esto imponiendo una pauta compleja, pero claramente definida, de la manera en que interaccionan las creencias y los deseos de las personas. Frecuentemente se ridiculiza a la teoría de la decisión como una descripción falsa de cómo la gente actúa realmente; se ha utilizado mucho ingenio en engañar a sujetos de modo que hagan una serie de elecciones que son inconsistentes con la teoría. La teoría de la decisión parece falsa como descripción del comportamiento real porque no hay

una manera completamente satisfactoria de someterla a prueba; la prueba siempre depende de cómo exactamente se aplica empíricamente la teoría. Sin embargo, la teoría responde a nuestras intuiciones sobre cómo se toman las decisiones reales; en efecto, simplemente hace explícito nuestro aparato de sentido común para explicar la acción intencional. La razón es que todos nosotros, pensemos en ello o no, tomamos decisiones en términos de cómo sopesamos los valores de varias consecuencias posibles de nuestras acciones, y de cuán probable pensamos que es que uno u otro curso de acción lleve a la consecución de esos valores. Comprendemos por qué alguien actúa como lo hace solamente al suponer que valora en mayor o menor medida los posibles resultados de la acción y cuán probable piensa que una determinada acción produce uno u otro resultado. Así pues, aunque no pueda precisar las predicciones de las elecciones reales, la teoría de la decisión corresponde a pesar de todo a nuestras intuiciones sobre cómo se toman las decisiones reales, y por ello es parte de nuestro aparato de sentido común para explicar el comportamiento intencional.

En estas observaciones estoy enfatizando el holismo de lo mental, hasta dónde varios aspectos de lo mental dependen unos de otros. Según he argumentado, no hay creencias sin tener muchas creencias relacionadas, ni creencias sin deseos, ni deseos sin creencias, ni intenciones sin las dos cosas, creencias y deseos. Conceptualmente, las propias acciones pertenecen al reino de lo mental, puesto que un comportamiento cuenta como una acción solamente si hay alguna descripción en la cual es intencional, y por ello se puede explicar como algo que se hace por una razón. Por supuesto que cualquier acción no será intencional en muchas descripciones. Como cualquier objeto o acaecimiento, las acciones se pueden describir de innumerables maneras. Así, si me siento no intencionadamente, desplazo algunas moléculas de aire no intencionadamente; presiono el cojín en el que me siento no intencionadamente; cambio el centro de gravedad de la Tierra, del Sistema Solar y de nuestra galaxia no intencionadamente. Incluso nuestros fallos son intencionales. No es que intentemos fallar, por supuesto; la idea es que cuando fallamos hay, sin

embargo, algo que hacemos a propósito. Si devuelvo la pelota demasiado lejos, una intención ha fallado; con todo, golpee la pelota intencionadamente.

Ya se estará viendo mi estrategia. Estoy ampliando el ámbito de los aspectos mentales de la vida que deben aparecer a la vez. Las acciones intencionales no pueden aparecer antes que las creencias y los deseos, puesto que una acción intencional se explica mediante las creencias y los deseos que la causan; las creencias no pueden aparecer una a una, ya que el contenido de cada creencia depende de su lugar en el nexo con otras creencias, y así sucesivamente. El holismo de lo mental es lo que hace que su aparición sea tan difícil de describir. No puede haber una secuencia en la aparición de las características de lo mental, al menos si estas características se han de describir en el vocabulario mentalista usual. Por supuesto que cada cosa que hay en el universo y su historia pueden describirse en principio en el lenguaje de la física, y por ello cada estadio de la aparición del pensamiento se puede describir en términos físicos. Pero esto no sería una explicación de la aparición de lo mental, ya que no tenemos una manera de hacer corresponder los acaecimientos descritos con el vocabulario de lo físico con acaecimientos descritos con el vocabulario mental y no podemos esperar encontrarla.

La dificultad de describir la aparición de los fenómenos mentales es un problema conceptual: es la dificultad de describir los primeros estadios en la maduración de la razón, los estadios que preceden la situación en que conceptos como los de intención, creencia y deseo tienen una aplicación clara. 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 que está seguido de un estadio en el que sí lo hay. Describir la aparición del pensamiento sería describir el proceso que conduce del primero al segundo de estos estadios. Carecemos de un vocabulario satisfactorio para describir los estadios intermedios. Podemos describir lo que hace un bebé preverbal empleando el lenguaje de la neurología, o en términos conductistas ordinarios podemos describir los movimientos y los sonidos que emite. Usted se puede engañar a sí mismo pensando

que el niño pequeño está hablando si hace sonidos que, si fueran hechos por un usuario genuino del lenguaje, tendrían un significado definido. (Incluso es posible hacer esto con los chimpancés.) Pero las palabras, como los pensamientos, tienen un significado conocido, un contenido proposicional, solamente si se dan en un rico contexto, pues se requiere un contexto así para dar a las palabras o a los pensamientos una ubicación y una función significativa. Si un ratón tuviera cuerdas vocales adecuadas, se le podría entrenar para decir «queso». Pero esa palabra no tendría un significado cuando el ratón la profiriera, ni el ratón comprendería lo que ha «dicho». Los niños pequeños profieren palabras de esta manera; si no lo hicieran así, nunca llegarían a tener lenguaje. Pero si queremos describir lo que pasa en la cabeza del niño pequeño cuando tiene unas cuantas palabras que profiere en situaciones apropiadas, no podremos hacerlo por falta de la clase de palabras adecuadas por nuestra parte. 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 es particularmente evidente cuando hablamos de las «intenciones» y los «deseos» de los animales simples. No tenemos una manera mejor de explicar lo que hacen.

Tampoco es que tengamos una idea clara de qué tipo de lenguaje podríamos utilizar para describir mentes a medio formar; puede que esté implicada una dificultad conceptual profunda o que sea imposible. 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.

A pesar de estas observaciones pesimistas, tengo algunas sugerencias que ofrecer sobre cómo podríamos aproximarnos al problema de decir algo inteligible sobre la aparición del pensamiento. Hay una situación prelingüística, precognitiva, que me parece constituir una condición necesaria para el pensamiento y el lenguaje, una condición que puede existir con independencia del pensamiento, y que por ello puede precederlo. Tanto en el caso de los animales no humanos como en

el caso de los niños pequeños es una condición que puede observarse que se da. La situación básica es una que involucra dos o más criaturas simultáneamente en interacción unas con otras y con el mundo que comparten; es lo que llamo *triangulación*. Es el resultado de una interacción a tres bandas, una interacción que es a dos bandas desde el punto de vista de cada uno de los dos agentes: cada uno de ellos está interaccionando simultáneamente con el mundo y con el otro agente. Dicho de una manera ligeramente distinta: cada criatura aprende a correlacionar las reacciones de las otras con los cambios o los objetos del mundo a los cuales también ella reacciona. Esto se detecta en su forma más simple en un banco de peces, donde cada pez reacciona casi instantáneamente a los movimientos de los demás. Aparentemente ésta es una reacción programada automáticamente. Puede observarse una reacción aprendida en determinados monos que hacen tres sonidos distinguibles dependiendo de si ven que se acerca una serpiente, un águila o un león; los otros monos, quizá sin ver ellos mismos la amenaza, reaccionan a los sonidos de alerta de forma apropiada a los distintos tipos de peligro, ya sea trepando a los árboles, corriendo o escondiéndose. Pero al reflexionar nos damos cuenta de que el comportamiento de estos primates, por muy complejo y con los propósitos que sean, no puede ser debido a creencias, deseos o intenciones proposicionales, ni su modo de comunicación constituye un lenguaje.

Sin embargo, el triángulo que he señalado es esencial para la existencia, y por tanto para la aparición, del pensamiento. La razón es que sin el triángulo hay dos aspectos del pensamiento que no podemos explicar. Estos dos aspectos son la objetividad del pensamiento y el contenido empírico de los pensamientos acerca del mundo.

El pensamiento, el pensamiento proposicional, es objetivo en el sentido de que tiene un contenido que es verdadero o falso independientemente (con raras excepciones) de la existencia del pensamiento o del sujeto que piensa. Además, es éste un hecho del que el sujeto debe dar cuenta; uno no puede creer algo o dudarle sin saber que lo que uno cree o de lo que uno duda puede ser verdadero o falso y que puede equi-

vocarse. ¿De dónde sacamos la idea de que podemos estar equivocados, de que las cosas pueden que no sean como pensamos que son? Wittgenstein sugirió, o al menos así lo interpreto, que no tendríamos el concepto de hacer las cosas bien o mal si no fuera por nuestras interacciones con otras personas. El triángulo que he descrito sirve para la situación interpersonal más simple. En él cada una de dos (o más) criaturas correlaciona sus propias reacciones a los fenómenos externos con las reacciones de la otra. Una vez que las correlaciones están establecidas, cada una de esas criaturas está en posición de esperar el fenómeno externo cuando percibe la reacción asociada de la otra. Lo que introduce la posibilidad de error es el fallo ocasional en lo que se espera; cuando las reacciones no se correlacionan. Wittgenstein expresa esta idea cuando habla de la diferencia entre seguir una regla y meramente pensar que se está siguiendo; dice que seguir la regla (hacer las cosas bien) en el fondo es hacer las cosas como los demás las hacen¹. Por supuesto que los demás a veces pueden estar equivocados. La idea no es que el consenso defina el concepto de verdad, sino que crea el espacio para su aplicación. Si esto es correcto, el pensamiento y el lenguaje son necesariamente sociales.

La interacción social, la triangulación, nos proporciona también la única explicación de cómo la experiencia da un contenido específico a nuestros pensamientos. Sin otras personas con las cuales compartir respuestas a un entorno mutuo, no existe una respuesta a la pregunta de qué es a lo que estamos respondiendo en el mundo. La razón tiene que ver con la ambigüedad del concepto de causa. Es esencial para resolver estas ambigüedades puesto que, en los casos más simples, lo que da contenido a la creencia es lo que la causa. En el presente caso, la causa está doblemente indeterminada: con respecto a la extensión y con respecto a la distancia. La primera ambigüedad tiene que ver con qué parte de la causa total de una creencia es relevante para el contenido. La respuesta breve es que es la parte o el aspecto de la causa total que cau-

¹ Sin embargo, véanse las advertencias del Ensayo 8.

sa típicamente respuestas relevantemente similares. A la vez, lo que hace que las respuestas sean relevantemente similares es el hecho de que otros encuentren que esas respuestas son similares; una vez más son las reacciones sociales compartidas las que hacen asequible la objetividad del contenido. El segundo problema tiene que ver con la ambigüedad de los estímulos relevantes, si son los proximales (digamos que los que se dan en la piel) o los distales. Lo que hace que el estímulo distal sea el determinante relevante del contenido es de nuevo su carácter social; es la causa que se comparte. De este modo se triangula el estímulo; es donde convergen las causas en el mundo.

He defendido que la relación triangular entre los agentes y un entorno al cual reaccionan mutuamente es necesaria para el pensamiento. No es suficiente, como se muestra por el hecho de que puede existir en animales a los cuales no les atribuimos juicios. Por esta razón estamos en posición de decir algo sobre la situación que debe existir para que lo haga el pensamiento, pero es una situación que puede existir independientemente, y por ello puede preceder al pensamiento en el orden de las cosas. Puede existir primero, y seguro que lo hace. De este modo podemos decir que cierta clase de interacción social primitiva es parte de la historia de cómo apareció el pensamiento.

¿Qué más se necesita para el pensamiento? Pienso que la respuesta es el lenguaje. En sí mismo, esto no es de mucha ayuda, ya que es obvio que una criatura que tenga lenguaje puede pensar; el lenguaje es un instrumento para la expresión de los contenidos proposicionales. Con todo, podemos preguntarnos por qué el lenguaje es esencial para el pensamiento. Dicho brevemente, la razón es que a menos que la base del triángulo —la línea entre los dos agentes— se refuerce hasta el punto en que pueda implementar la comunicación de contenidos proposicionales, no hay manera de que los agentes puedan usar la situación triangular para formar juicios acerca del mundo. Solamente cuando ya disponen de lenguaje, las criaturas pueden comprender el concepto de verdad objetiva. Se puede decir mucho más acerca de la relación profunda entre el lenguaje como un instrumento de comunicación inter-

personal y el pensamiento, pero aceptaré sin más esta relación y la usaré para decir algo más sobre la aparición del pensamiento. El lenguaje nos ofrece la oportunidad de hacer comparaciones relativamente tajantes entre distintos sistemas simbólicos lingüísticos y prelingüísticos, y distintas maneras de describirlos.

Para este propósito sugiero que pensemos en las teorías de la semántica formal en analogía con las teorías de la metrización fundamental. En una teoría de la metrización fundamental hay uno o más conceptos primitivos; por ejemplo, en el caso de la metrización del peso, éstos pueden ser la relación ser *al menos tan pesado como* que se establece entre dos cosas y la operación de añadir una cosa a otra (ésta es una función que le hace corresponder a dos cosas su suma). Estos conceptos no están definidos, pero antes de que la teoría puede aplicarse a una colección de objetos, o ser sometida a prueba mediante ellos, se debe indicar un método para dar contenido empírico a los conceptos primitivos. (De este modo, *a* es al menos tan pesado como *b* si el platillo de una balanza de brazos iguales en el que está *a* no se encuentra más elevado que el platillo en el que está *b*.) Un conjunto de *axiomas* especifica entonces las propiedades lógicas de los conceptos primitivos y las entidades a medir. (De este modo, los axiomas para el peso especificarán que la relación de ser al menos tan pesado como es transitiva, y que la suma de *a* y *b* es al menos tan pesada como la suma de *b* y *a*.)

El propósito de la teoría es describir en términos precisos el tipo de estructura que debe tener un conjunto de entidades para asignarles pesos. Si los pesos son las propiedades que esperamos intuitivamente, los axiomas deben resultar suficientes para explicar esas propiedades. Una de las cosas que queremos demostrar (*acerca de* la teoría, no *en* la teoría) es un teorema de representación, que dice que los números pueden asignarse a los objetos que la teoría abarca de una manera que sigue el rastro a las relaciones de peso entre los distintos objetos. La segunda cosa a demostrar es un teorema de unicidad, que dice que si un conjunto de números representa los pesos de los objetos, entonces cualquier otro conjunto de números que se relaciona con el primer conjunto mediante un deter-

minado tipo de transformación representará los pesos igualmente bien. En el caso del peso, la multiplicación mediante cualquier constante positiva transformará una medida en otra; pesar en libras es igual de bueno que pesar en kilos.

Voy ahora a comparar una teoría semántica para un hablante o grupo de hablantes. Las definiciones de verdad de Tarski, cuando se modifican para aplicarlas a los lenguajes naturales, pueden ser vistas como tales teorías. Igual que una teoría para medir el peso, una teoría semántica pondrá restricciones en uno o más conceptos primitivos. El concepto primitivo más importante es el de verdad. (Aquí me desvío de los intereses e intenciones de Tarski. Él *definió* la verdad para lenguajes particulares, pero yo tengo en mente axiomatizaciones de un concepto de verdad general, y por ello indefinible. Este camino se arriesga en aquello que Tarski trató de evitar, esto es, la inconsistencia.) La teoría impone una determinada estructura en cualquier lenguaje que pueda describir; al dar las condiciones de verdad de todas las oraciones de una lengua o lenguaje, define necesariamente las relaciones lógicas entre las oraciones.

Podemos concebir tal teoría como algo que dice qué es lo necesario para entender una lengua o lenguaje; describe lo que un hablante conoce intuitivamente sobre su propia lengua, y puede servir para interpretar lo que tal hablante dice. Puede mostrarse la adecuación de la teoría para estos propósitos demostrando que se pueden representar las propiedades del lenguaje objeto en algún lenguaje que uno entienda. De esta manera, la teoría sirve para validar un manual de traducción. Del mismo modo en que usamos números con una manera conveniente de seguirles el rastro o registrar los pesos de los objetos, seguimos el rastro a los rasgos semánticos de una lengua o lenguaje utilizando oraciones que entendemos. Lo que hace que los números sean tan adecuados para esta tarea es que sabemos exactamente qué estructura tienen. De manera muy parecida, las oraciones de una lengua que entendemos tienen una estructura conocida que puede usarse para entender a otros hablantes.

Diferentes formas de metrización usan diferentes propiedades de los números; hablando laxamente, el modo de medir

es más potente o informativo cuantas más propiedades de los números se utilicen. Por ejemplo, cuando los números se usan para registrar los miembros de un equipo, su única significación es que un número es diferente de otro. La escala de dureza de Mohs se basa en el hecho de que un número es más grande que otro: cualesquiera otros diez números (además de los que ya se usan, 1-10), con tal que se ordenen en magnitud ascendente, servirían igual de bien. Cualquier transformación lineal de la escala de temperatura Fahrenheit serviría, mientras que las escalas de peso y distancia son más potentes al permitir sólo la multiplicación por una constante positiva. En el caso de la temperatura, dos números son arbitrarios, el punto cero y el valor numérico de un intervalo; como resultado, no tiene sentido decir que un cuerpo está el doble de caliente que otro. Por otro lado, un cuerpo puede pesar el doble que otro.

Esto nos proporciona una manera de pensar acerca de la potencia expresiva relativa de las lenguas o lenguajes. Intuitivamente está claro que muchas oraciones son equivalentes respecto a lo que transmiten. Así, tiene sentido preguntar: ¿de cuántas maneras distintas podríamos asignar nuestras propias oraciones a las oraciones de otro lenguaje y captar aún la misma información (esto es, cuántas de estas maneras constituirían manuales de traducción igualmente satisfactorios)? Por supuesto que un manual de traducción no traduce las oraciones de una en una; el manual es una función que correlaciona todas las oraciones de un lenguaje con las oraciones de otro. (Del mismo modo en que usted no puede comparar pesos en libras con pesos en kilos, no puede cambiar los manuales de traducción sin hacerlo notar, puesto que esto destruiría las relaciones entre oraciones.) En consecuencia, lo que debemos considerar son sistemas totales de traducción. La afirmación de que existen tales sistemas empíricamente equivalentes es la tesis de Quine de la indeterminación de la traducción: se podría registrar de diversas maneras todo lo que uno puede llegar a saber sobre lo que otros significan con lo que dicen y, por tanto, todo lo que significan.

Lo que está claro es que cuantas más maneras distintas hay de representar otra lengua en la nuestra, menos poder expre-

sivo tiene esa lengua (desde nuestro punto de vista, por supuesto). Consideremos algunos ejemplos. Todas las lenguas o lenguajes tienen un vocabulario finito básico. Podemos imaginar una lengua o lenguaje que sólo tenga nombres propios y predicados. Tal lengua o lenguaje debe tener un número finito de oraciones; así lo único que necesita nuestra teoría es hacer corresponder una oración de nuestra propia lengua a cada una de las oraciones de la otra lengua o lenguaje. La semántica de tal lengua o lenguaje es también simple: no hay necesidad de romper las oraciones en partes, no hay razón para emplear instrumentos recursivos ni base para atribuir una ontología. Si se considera cuán simple puede ser la semántica de tal lengua o lenguaje uno debe preguntarse si habría de ser considerada como una lengua o lenguaje; después de todo, la expresión «lengua o lenguaje» que no requiera que hallemos ninguna relación entre sus partes (las «palabras») y las entidades del mundo apenas merece esa denominación.

Se da un paso hacia la complejidad al introducir conectivas veritativo-funcionales como los signos de conjunción, negación, alternancia y el condicional material («y», «no», «o», «si... entonces...»). Ahora habrá una infinidad de oraciones, ya que se puede añadir la negación a cualquier oración y obtener una oración, o hacer la conjunción de una oración con cualquier oración y producir otra nueva oración. Esto introduce el elemento de productividad, la posibilidad de producir y entender nuevas oraciones sobre la base de simples reglas sintácticas y semánticas. Sin embargo, la semántica sigue siendo elemental, ya que la semántica para nuestra primera lengua o lenguaje finita es todavía suficiente para proporcionar las condiciones de verdad de las oraciones atómicas, y las tablas de verdad nos dan la semántica de las oraciones moleculares. Las únicas partes del habla que se necesita distinguir son las oraciones más simples y las conectivas oracionales; la ontología todavía está a la espera.

El último estadio requiere dar un salto; introduce la cuantificación, los conceptos expresados mediante las palabras «algún» y «todos». Una vez que avanzamos hasta este estadio hemos llegado a lenguas o lenguajes que se equiparan o empiezan a equiparse a las nuestras en complejidad. La semántica

debe distinguir predicados de diversos órdenes, términos singulares y expresiones que hacen el trabajo de las variables (pronombres y similares), y por supuesto, los cuantificadores. Por primera vez, es necesario encontrar una estructura en las oraciones más simples y correlacionar los predicados y los términos singulares con objetos. En mi opinión es aquí cuando alcanzamos el grado de refinamiento expresivo que asociamos con el pensamiento, puesto que solamente a este nivel tenemos evidencia positiva de que el hablante de la lengua o lenguaje en cuestión puede predicar propiedades de objetos y acaecimientos. Alfred Tarski fue quien aclaró esta cuestión, pues fue él quien primero mostró lo que es necesario para caracterizar la verdad para lenguas o lenguajes que tengan una cuantificación general. Esto lo hizo introduciendo el concepto de *satisfacción*, que es una relación entre predicados de cualquier grado de complejidad y objetos del mundo².

No es necesario ser dogmático acerca de dónde empieza el pensamiento o el lenguaje para apreciar la relativa claridad que la semántica formal y la analogía con la teoría de la medición nos brindan con respecto a la aparición del pensamiento, pues por más problemáticos y vagos que puedan ser nuestros intentos de describir los estadios que conducen al pensamiento y al lenguaje, la semántica y la teoría de la medición nos proporcionan una plataforma desde la cual distinguir los niveles de desarrollo desde un punto de vista objetivo.

² Alfred Tarski, «The Concept of Truth in Formalized Languages».

Lo objetivo

10

Una teoría coherentista de la verdad y el conocimiento

En este ensayo voy a defender lo que bien puede llamarse una teoría coherentista de la verdad y el conocimiento. La teoría que defenderé no está en competencia con una teoría de la correspondencia, sino que depende para su defensa de un argumento que pretende mostrar que la coherencia produce correspondencia.

La importancia del tema es evidente. Si la coherencia es un criterio para la verdad, entonces hay una conexión directa con la epistemología, ya que tenemos razones para creer que muchas de nuestras creencias forman con muchas otras un conjunto coherente, y en ese caso tenemos razones para creer que muchas de nuestras creencias son verdaderas. Cuando esas creencias son verdaderas, entonces parece que se satisfacen las condiciones principales del conocimiento.

Alguien podría tratar de defender una teoría coherentista de la verdad sin defender una teoría coherentista del conocimiento, quizá sobre la base de que una persona que sostiene un conjunto coherente de creencias podría carecer de una razón para creer que sus creencias son coherentes. Esto no es probable, pero puede ser que alguien, aunque tenga creencias verdaderas y buenas razones para sostenerlas, no aprecie la relevancia de las razones en favor de esas creencias. La mejor manera de ver a esa persona sería como alguien que posee un

conocimiento que no sabe que tiene; cree que es un escéptico. En una palabra, es un filósofo.

Dejando a un lado casos anómalos, lo que une a la verdad y al conocimiento es el significado. Si se dan los significados mediante condiciones de verdad objetivas se plantea la cuestión de cómo sabemos que esas condiciones se satisfacen, pues esto parecería requerir una confrontación entre lo que creemos y la realidad, y la idea de tal confrontación es absurda. Pero si la coherencia es un criterio para la verdad, entonces la coherencia es un criterio para juzgar que se satisfacen las condiciones de verdad objetivas y ya no tenemos que explicar el significado sobre la base de una posible confrontación. Mi lema es: correspondencia sin confrontación. Dada una epistemología correcta podemos ser realistas en todas las áreas. Podemos aceptar que las condiciones de verdad objetivas constituyen la clave del significado y aceptar también una concepción realista de la verdad, y podemos insistir en que el saber es conocimiento de un mundo objetivo, independiente de nuestro pensamiento o de nuestro lenguaje.

Puesto que, por lo que yo sé, no existe una teoría que merezca denominarse «la» teoría coherentista, voy a caracterizar el tipo de teoría que pretendo defender. Es evidente que no todo conjunto consistente de oraciones interpretadas contiene sólo oraciones verdaderas, puesto que un conjunto así podría contener sólo la oración consistente s y otro sólo la negación de s , y añadir más oraciones, a la vez que se mantiene la consistencia, no sirve de nada. Podríamos imaginar un sinnúmero de descripciones de estado —descripciones máximamente consistentes— que no describen nuestro mundo.

Mi teoría coherentista concierne a las creencias u oraciones que tiene por verdaderas alguien que las entiende. En este punto no quiero decir que todo posible conjunto de creencias coherente sea verdadero (o que contenga en su mayor parte creencias verdaderas). Me aparto de esto porque no es nada claro qué es lo que es posible. En un extremo, podría sostenerse que el dominio de los conjuntos máximamente consistentes de creencias que son posibles es tan amplio como el dominio de los posibles conjuntos máximamente consistentes de oraciones, y en ese caso no habría ningún mo-

tivo para insistir en que una teoría coherentista defendible trate de creencias y no de proposiciones u oraciones. Pero existen otros modos de concebir lo que es posible creer que justificarían que se dijera no sólo que todos los sistemas de creencias coherentes reales son en gran parte correctos, sino que todos los posibles lo son también. La diferencia entre las dos nociones de lo que es posible creer depende de cuáles sean nuestros supuestos acerca de la naturaleza de las creencias, su interpretación, sus causas, sus sujetos y sus configuraciones. Para mí, las creencias son estados de personas que tienen intenciones, deseos y órganos sensoriales; son estados causados por acaecimientos internos y externos a los cuerpos de quienes las tienen, y que a su vez causan tales acaecimientos. Pero incluso dadas todas estas restricciones, hay muchas cosas que las personas creen y muchas más que podrían creer. A todos esos casos es a los que se aplica la teoría coherentista.

Por supuesto, algunas creencias son falsas. Buena parte de la importancia del concepto de creencia es la brecha potencial que introduce entre lo que se considera verdadero y lo que es verdadero. De modo que la mera coherencia no garantiza que lo que uno cree sea así, sin que importe lo fuertemente que se formule una definición verosímil de coherencia. Todo lo que una teoría coherentista puede mantener es que la mayoría de las creencias de un conjunto total coherente de creencias son verdaderas.

Esta manera de formular mi posición puede como mucho ser tomada como un indicio, puesto que no hay ninguna manera útil de contar creencias y, por tanto, la idea de que la mayor parte de las creencias de una persona son verdaderas no tiene ningún significado claro. Una manera algo mejor de formular la cuestión es decir que existe una presunción a favor de la verdad de una creencia que forma un conjunto coherente con un grupo importante de creencias. Toda creencia de un conjunto total coherente de creencias se justifica por esta presunción, de la misma forma en que se justifica cualquier acción intencionada que realice un agente racional (uno cuyas preferencias, creencias y deseos forman un conjunto coherente en el sentido de la teoría bayesiana de la decisión). De modo que, repitémoslo, si el conocimiento es creencia verda-

dera justificada, entonces parecería que todas las creencias verdaderas de un sujeto que fuera consistente constituirían conocimiento. Como argumentaré, esta conclusión, aunque sea demasiado vaga y demasiado precipitada para ser correcta, contiene un importante núcleo de verdad. Mientras tanto, me limitaré a mencionar los muchos problemas que están reclamando atención. ¿Qué es exactamente lo que exige la coherencia? ¿Qué parte de las prácticas inductivas se debería incluir, qué parte de una teoría verdadera (si es que hay una) del apoyo evidencial debería considerarse? Puesto que no hay nadie que tenga un conjunto completamente consistente de convicciones, ¿cuáles son las creencias con las que el formar un conjunto coherente crea una presunción de verdad? A continuación algunos de estos problemas se presentarán desde una perspectiva mejor.

Debería estar claro que no espero definir la verdad en términos de coherencia y creencia. La verdad es algo hermosamente transparente comparado con la creencia y la coherencia, y la tomo como concepto primitivo. La verdad, tal como se aplica a preferencias de oraciones, muestra el rasgo de desentrecomillado encapsulado en la Convención T de Tarski, el cual basta para fijar su dominio de aplicación. Por supuesto, relativamente a una lengua o lenguaje o a un hablante de modo que en la verdad hay algo más que la Convención T; hay ahí todo lo que puede trasladarse de una lengua o lenguaje a otra u otro, o de un hablante a otro hablante. Lo que revelan la Convención T y las trilladas oraciones que ella declara verdaderas —como «la hierba es verde», dicha por un hablante del castellano, es verdadera si y sólo si la hierba es verde»— es que la verdad de una preferencia depende sólo de dos cosas: de lo que significan las palabras tal como han sido dichas y de cómo está dispuesto el mundo. No hay relatividad ulterior alguna a un esquema conceptual, un modo de ver las cosas o una perspectiva. Dos intérpretes, tan disimilares en cuanto a cultura, lengua y punto de vista como se quiera, pueden estar en desacuerdo acerca de si una preferencia es verdadera, pero sólo si difieren en cómo son las cosas en el mundo que comparten o en qué es lo que significa esa preferencia.

Creo que puedo extraer dos conclusiones a partir de estas sencillas reflexiones. En primer lugar, que la verdad es correspondencia con cómo son las cosas. (No hay ninguna manera directa y no engañosa de formular esto; para hacerlo bien es necesario un desvío a través del concepto de satisfacción en términos del cual se caracteriza la verdad)¹. En consecuencia, si es aceptable una teoría coherentista de la verdad, ésta debe ser consistente con una teoría de la correspondencia. En segundo lugar, una teoría del conocimiento que admita que podemos saber la verdad debe ser una forma de realismo no relativizado, no interno. En consecuencia, si es aceptable una teoría coherentista de la verdad, ésta debe ser consistente con esa forma de realismo. Mi forma de realismo no parece ser ni el realismo interno de Hilary Putnam ni su realismo metafísico². No es el realismo interno porque el realismo interno sostiene que la verdad es relativa a un esquema, y ésta es una idea que no encuentro inteligible³. De hecho, una de las razones más importantes para aceptar una teoría coherentista es la ininteligibilidad de un dualismo que postula un esquema conceptual y un «mundo» que espera que se le haga frente. Pero, ciertamente, mi realismo no es el realismo metafísico de Putnam, pues éste se caracteriza por ser «radicalmente no-epistémico», lo que implica que nuestros pensamientos y teorías mejor investigados y establecidos pueden ser, todos ellos, falsos. Yo creo que la independencia entre la creencia y la verdad requiere únicamente que *cada* una de nuestras creencias pueda ser falsa. Pero, por supuesto, una teoría coherentista no puede admitir que todas ellas puedan estar equivocadas.

Pero ¿por qué no? Quizá sea obvio que la coherencia de una creencia con un grupo importante de creencias aumenta sus posibilidades de ser verdadera, supuesto que haya razones para suponer que ese grupo de creencias sea verdadero, o que lo sea en gran parte. Pero ¿cómo puede la coherencia por sí sola proporcionar las bases para la creencia? Puede ser que lo

¹ Véase mi «True to the Facts», Ensayo 3, en *Inquiries into Truth and Interpretation*, y «Afterthoughts» en este volumen.

² Hilary Putnam, *Meaning and the Moral Sciences*, 125.

³ Véase mi «On the Very Idea of a Conceptual Schema», Ensayo 13 de *Inquiries into Truth and Interpretation*.

mejor que podamos hacer para justificar una creencia sea apelar a otras creencias. Pero entonces el resultado parecería ser que debemos aceptar el escepticismo filosófico, sin que importe lo inalterables que sean en la práctica nuestras creencias.

Esto es escepticismo vestido con uno de sus ropajes tradicionales. Ese escepticismo pregunta: ¿no podrían todas mis creencias ser compatibles y sin embargo ser en conjunto falsas acerca del mundo real? El mero reconocimiento de que es absurdo, o algo aún peor, tratar de *confrontar* nuestras creencias, una a una o en conjunto, con aquello sobre lo que versan no responde a esa pregunta, ni muestra que esa pregunta sea ininteligible. En resumen, incluso una teoría coherentista poco exigente como la mía ha de proporcionarle al escéptico una razón para suponer que las creencias coherentes son verdaderas. El partidario de la teoría coherentista no puede admitir que esa garantía provenga de fuera del sistema de creencias, aunque nada en el interior del mismo proporcione apoyo alguno a menos que pueda mostrarse que descansa, en último término o inmediatamente, en algo que es digno de confianza de forma independiente.

Es natural distinguir las teorías coherentistas de otras teorías haciendo referencia a la cuestión de si la justificación puede o debe llegar a un final. Pero esto no define las posiciones, sino que únicamente sugiere una forma que pueden tomar los argumentos, ya que existen teóricos de la coherencia que sostienen que algunas creencias pueden servir de base para las restantes, mientras que sería posible mantener que la coherencia no es suficiente aunque el dar razones no sea algo que tenga fin. Lo que distingue a una teoría coherentista es simplemente la afirmación de que nada puede contar como razón para mantener una creencia excepto otra creencia. Su partidario rechaza como algo ininteligible la exigencia de un fundamento o fuente de justificación de otro tipo. Como ha dicho Rorty, «nada cuenta como justificación a menos que haga referencia a lo que ya aceptamos, y no hay modo alguno de salirse de nuestras creencias y nuestro lenguaje para hallar una manera de someterlas a prueba distinta de la coherencia»⁴.

⁴ Richard Rorty, *Philosophy and the Mirror of Nature*, 178.

Como pueden ver, en esto estoy de acuerdo con Rorty. Donde diferimos, si es que diferimos, es acerca de si todavía se plantea la pregunta de cómo, dado que no podemos «salir[nos] de nuestras creencias y nuestro lenguaje para hallar una manera de someterlas a prueba distinta de la coherencia», podemos todavía tener conocimiento de un mundo público objetivo que no es de nuestra propia factura y hablar sobre él. Yo creo que esta pregunta todavía se plantea, mientras que sospecho que Rorty no piensa eso. Si ésta es su opinión, entonces debe pensar que estoy cometiendo un error al tratar de responderla. No obstante, ahí va la respuesta.

Mejorará las cosas repasar en este punto muy rápidamente algunas de las razones para abandonar la búsqueda de una base para el conocimiento fuera del dominio de nuestras creencias. Aquí con el término «base» quiero decir específicamente una base epistemológica, una fuente de justificación.

Los intentos que merece la pena tomarse en serio pretenden fundamentar de una manera u otra las creencias en el testimonio de los sentidos: la sensación, la percepción, lo dado, la experiencia, los datos sensoriales, el espectáculo que pasa. Todas esas teorías deben explicar al menos las dos cosas siguientes: ¿cuál es, exactamente, la relación entre la sensación y la creencia que permite que la primera justifique la segunda? y ¿por qué habríamos de creer que nuestras sensaciones son fiables, es decir, por qué habríamos de confiar en nuestros sentidos?

La idea más simple es identificar determinadas creencias con sensaciones. Así, Hume parece no haber distinguido entre percibir una mancha verde y percibir que una mancha es verde. (La ambigüedad de la palabra «idea» fue en esto de gran ayuda.) Otros filósofos señalaron la confusión de Hume, pero trataron de obtener los mismos resultados reduciendo la brecha entre la percepción y el juicio a cero mediante un intento de formular juicios que no vayan más allá de enunciar que existe la percepción, sensación o presentación (sea lo que sea lo que esto pueda significar). Estas teorías no justifican las creencias sobre la base de sensaciones, pero tratan de justificar determinadas creencias afirmando que las mismas tienen exactamente el mismo contenido epistémico que una sensa-

ción. Hay dos dificultades con esta teoría: en primer lugar, si las creencias básicas no exceden en contenido a la sensación correspondiente, no pueden apoyar la inferencia de un mundo objetivo; y, en segundo lugar, no existen tales creencias.

Una pista más plausible es afirmar que no podemos estar equivocados acerca de cómo nos parece que son las cosas. Si creemos que tenemos una sensación, la tenemos; se sostiene que ésta es una verdad analítica, o un hecho acerca de cómo se usa el lenguaje.

Es difícil explicar esta supuesta conexión entre las sensaciones y algunas creencias de un modo que no nos incline al escepticismo acerca de otras mentes, y en ausencia de una explicación adecuada debería haber dudas acerca de qué consecuencias tiene esa conexión para la cuestión de la justificación. Pero, en cualquier caso, no está nada claro cómo, siguiendo por la vía indicada, las sensaciones justifican la creencia en esas sensaciones. La cuestión es más bien que tales creencias no requieren justificación alguna, pues la existencia de la creencia implica la existencia de la sensación y, por lo tanto, la existencia de la creencia implica su propia verdad. A menos que se añada alguna otra cosa, estamos volviendo a otra forma de teoría coherentista.

El énfasis sobre la sensación o la percepción en cuestiones epistemológicas nace del obvio pensamiento siguiente: las sensaciones son lo que conecta el mundo con nuestras creencias, y son candidatas a suministrar justificación porque a menudo nos percatamos de ellas. El problema en el que hemos desembocado es que la justificación parece depender de ese percatarse, que no es sino otra creencia.

Vamos a intentar ir por un rumbo más arriesgado. Supóngase que decimos que las sensaciones mismas, verbalizadas o no, justifican determinadas creencias que van más allá de lo que nos es dado en la sensación. En consecuencia, en ciertas condiciones tener la sensación de ver una luz verde destellando puede justificar la creencia de que una luz verde está destellando. El problema es ver cómo la sensación justifica la creencia. Por supuesto, si alguien tiene la sensación de ver una luz verde destellando, es probable, en determinadas circunstancias, que una luz verde esté destellando. *Nosotros* podemos

decir esto, dado que sabemos de su sensación, pero *él* no puede decirlo, puesto que estamos suponiendo que está justificado sin necesidad de depender de la creencia de que tiene esa sensación. Supóngase que creyera que no tiene esa sensación. ¿Le justificaría todavía la sensación en su creencia en una luz verde destellante objetiva?

La relación entre una sensación y una creencia no puede ser una relación lógica, puesto que las sensaciones no son creencias ni otras actitudes proposicionales. ¿Cuál es entonces la relación? Creo que la respuesta es obvia: esa relación es causal. Las sensaciones causan algunas creencias y en *este* sentido son la base o fundamento de esas creencias. Pero una explicación causal de una creencia no muestra cómo y por qué se justifica esa creencia.

La dificultad de transmutar una causa en una razón acosa de nuevo al antioherentista si trata de responder a nuestra segunda pregunta: ¿qué es lo que justifica la creencia de que nuestros sentidos no nos engañan sistemáticamente? Pues incluso si las sensaciones justifican la creencia en la sensación, todavía no se ve cómo pueden justificar la creencia en acaecimientos y objetos externos.

Según Quine la ciencia nos dice que «nuestra única fuente de información acerca del mundo externo es mediante el impacto de rayos de luz y moléculas en nuestras superficies sensoriales»⁵. Lo que me inquieta es cómo leer las palabras «fuen-

⁵ W. V. Quine, «The Nature of Natural Knowledge», 68. Muchos otros pasajes en la obra de Quine sugieren que conserva la esperanza de asimilar las causas sensoriales a evidencias. En *Word and Object*, 22, escribe que «la excitación de las superficies... agotan nuestros indicios sobre el mundo externo» [pág. 35 de la traducción al castellano]. En *Ontological Relativity*, 75, encontramos que «toda la evidencia que haya podido servir, en última instancia, a cualquiera para alcanzar su imagen del mundo es la estimulación de los receptores sensoriales» [pág. 101 de la traducción]. En la misma página: «Dos principios cardinales del empirismo permanecían siendo... irrefutables... Uno es que la evidencia, cualquiera que ésta sea, que *hay* para la ciencia, es evidencia sensorial. El otro... es que toda inculcación de significados de palabras ha de descansar, en última instancia en la evidencia sensible» [pág. 100 de la traducción]. En *The Roots of Reference*, 37-8, dice Quine que «las observaciones» son básicas «en el sostenimiento de la teoría y en el aprendizaje del lenguaje», y añade: «¿Qué son observaciones? Son visuales, auditivas, táctiles, olfativas. Son senso-

te» e «información». Ciertamente es verdad que acaecimientos y objetos del mundo externo causan que creamos cosas acerca del mundo externo, y buena parte de esa causalidad, si no toda, pasa por los órganos de los sentidos. Sin embargo, la noción de información sólo se aplica no metafóricamente a las creencias generadas en el proceso. De modo que «fuente» ha de leerse simplemente como «causa» e «información» como «creencia verdadera» o «conocimiento». Todavía no puede verse la justificación de las creencias causadas por nuestros sentidos.

El enfoque del problema de la justificación al que hemos estado siguiéndole la pista debe estar equivocado. Hemos tratado de verlo del siguiente modo: una persona tiene todas sus creencias acerca del mundo —es decir, todas sus creencias—. ¿Cómo podemos decir si son verdaderas o si tienden a serlo? Hemos estado suponiendo que sólo es posible decir esto conectando sus creencias al mundo, confrontando algunas de sus creencias, una a una, con lo que los sentidos suministran, o quizá confrontando la totalidad de sus creencias con el tribunal de la experiencia. Una confrontación así no tiene sentido alguno, ya que, por supuesto, no podemos salirnos de nuestra piel para averiguar qué es lo que está causando los sucesos internos de los que nos percatamos. La introducción de entidades o pasos intermedios en la cadena causal, como las sensaciones y observaciones, sólo sirven para hacer más patente el problema epistemológico, pues si los intermediarios meramente son causas, no justifican las creencias que causan, mientras que si proporcionan información, puede que nos mientan. La moraleja es obvia. Puesto que no podemos asegurarnos de que los intermediarios sean verídicos, no deberíamos admitir intermediarios entre nuestras creencias y sus objetos del mundo. Por supuesto, hay intermediarios causales.

riales, evidentemente, y por ello subjetivas... ¿Habremos por ello de decir que, en última instancia, la observación no es la sensación...? No...» [págs. 53-54 de la traducción]. Quine procedió a dejar de hablar de observaciones para pasar a hablar de oraciones observacionales. Pero, por supuesto, las oraciones observacionales, al contrario que las observaciones, no pueden desempeñar el papel de evidencias a menos que tengamos razones para creer que son verdaderas.

De lo que nos debemos guardar es de los intermediarios epistémicos.

Hay concepciones comunes del lenguaje que fomentan la mala epistemología. Esto, desde luego, no es ningún accidente puesto que las teorías del significado se conectan con la epistemología por medio de intentos de responder la pregunta de cómo determinamos que una oración es verdadera. Si conocer el significado de una oración (conocer cómo dar una interpretación correcta de la misma) implica, o consiste en, conocer cómo podría reconocerse que es verdadera, entonces la teoría del significado plantea el mismo problema con el que hemos estado conteniendo, pues dar el significado de una oración requerirá que especifiquemos lo que justificaría aseverarla. Aquí el coherentista sostendrá que no sirve de nada buscar una fuente de justificación fuera de las otras oraciones que se tienen por verdaderas, mientras que el fundamentalista tratará de anclar al menos algunas palabras u oraciones a rocas no verbales. Esta concepción la mantienen, según creo, tanto Quine como Michael Dummett.

Dummett y Quine no están de acuerdo, claro. En particular están en desacuerdo acerca del holismo, es decir, la afirmación de que la verdad de nuestras oraciones debe ser sometida a prueba conjuntamente en lugar de una a una. Y en consecuencia están también en desacuerdo acerca de si existe una distinción útil entre las oraciones analíticas y las sintéticas, y acerca de si una teoría satisfactoria del significado puede admitir la clase de indeterminación que defiende Quine. (En todos estos puntos soy el fiel discípulo de Quine.)

Pero lo que aquí me ocupa es que Quine y Dummett están de acuerdo en un principio básico, que es que, sea lo que sea lo que haya en el significado, hay que remitirlo de algún modo a la experiencia, lo dado o las pautas de estimulación interna, algo intermedio entre la creencia y los objetos usuales acerca de los que versan nuestras creencias. Una vez dado este paso, la puerta al escepticismo está abierta, pues debemos entonces admitir que una gran cantidad de oraciones —quizá la mayoría de ellas— que consideramos verdaderas pueden de hecho ser falsas. Esto es irónico. Tratar de hacer accesible el significado ha hecho que la verdad sea inaccesible. Cuando

el significado se hace epistemológico de esa manera, la verdad y el significado se divorcian necesariamente. Por supuesto, se puede organizar un matrimonio forzado redefiniendo la verdad como lo que estamos justificados en aseverar. Pero con esto no se está casando a la pareja original.

Tomemos la propuesta de Quine de que sea lo que fuere lo que hay en el significado (valor informativo) de una oración observacional, ello se determina mediante las configuraciones de estimulaciones sensoriales que causarían que el hablante asintiera o disintiera de la oración. Ésta es una manera maravillosamente ingeniosa de captar lo que resulta atractivo en las teorías verificacionistas sin tener que hablar de significados, datos sensoriales o sensaciones; por primera vez hizo verosímil la idea de que se podía, y se debía, hacer lo que yo llamo teoría del significado sin necesidad de lo que Quine llama significados. Pero la propuesta de Quine, como otras formas de verificacionismo, se encamina al escepticismo, ya que, claramente, los estímulos sensoriales podrían ser exactamente como son y, sin embargo, el mundo externo ser muy distinto. (Recuérdese el caso del cerebro en la cubeta.)

La manera en que Quine se las arregla sin significados es sutil y complicada. Vincula directamente los significados de algunas oraciones con configuraciones de estímulos (las cuales consituyen también la evidencia, piensa Quine, para asentir a la oración), pero los significados de otras oraciones se determinan mediante la manera en que están condicionadas a las oraciones originales u oraciones observacionales. Los hechos relativos a ese condicionamiento no permiten una división tajante entre oraciones que se consideran verdaderas en virtud del significado y oraciones que se consideran verdaderas en base a la observación. Quine estableció esto mostrando que si una manera de interpretar las preferencias de un hablante era satisfactoria, también lo eran muchas otras. Esta doctrina de la indeterminación de la traducción, como la llamó Quine, no debería verse como algo misterioso o amenazador. No es más misteriosa que el hecho de que la temperatura pueda medirse en grados centígrados o grados Fahrenheit (o en cualquier transformación lineal de esas cifras). Y no es amenazador porque el procedimiento mismo que demuestra el grado

de indeterminación, demuestra al mismo tiempo que todo lo que necesitamos es lo que está determinado.

En mi opinión, borrar la línea entre lo analítico y lo sintético salvó a la filosofía del lenguaje haciendo de ella un tema serio, al mostrar la manera en que podía proseguirse su estudio sin aquello que no puede haber: significados determinados. Sugiero ahora abandonar también la distinción entre las oraciones observacionales y las demás. La razón es que la distinción entre oraciones creencia cuya verdad se justifica por las sensaciones y oraciones creencia cuya verdad se justifica apelando a otras oraciones que se consideran verdaderas es para el coherentista igual de anatema que la distinción entre creencias justificadas por las sensaciones y creencias justificadas únicamente por apelación a otras creencias. De acuerdo con esto, sugiero que abandonemos la idea de que el significado o el conocimiento se fundamenta en algo que cuenta como fuente última de evidencias. Sin duda, el significado y el conocimiento dependen de la experiencia, y la experiencia en último término de la sensación. Pero éste es el «depender» de la causalidad, no de la evidencia o la justificación.

He enunciado ahora mi problema tan bien como me es posible hacerlo. La búsqueda de un fundamento empírico del significado o del conocimiento conduce al escepticismo, mientras que una teoría coherentista de la verdad parece incapaz de proporcionar razón alguna por la que un sujeto crea que sus creencias, son coherentes, son verdaderas. Estamos atrapados entre una respuesta falsa al escéptico y el no darle respuesta alguna.

El dilema no es un verdadero dilema. Lo que se requiere para responder al escéptico es mostrar que alguien con un conjunto de creencias coherente (más o menos) tiene una razón para suponer que sus creencias no son erróneas en su mayor parte. Lo que hemos mostrado es que es absurdo buscar un fundamento justificatorio para la totalidad de las creencias, algo situado fuera de esa totalidad que pueda utilizarse para someter a prueba o para compararse con nuestras creencias. La respuesta a nuestro problema debe consistir entonces en encontrar una *razón* para suponer que la mayoría de

nuestras creencias son verdaderas que no sea una forma de *evidencia*.

Mi argumento tiene dos partes. En primer lugar sostengo que una comprensión correcta del habla, las creencias, los deseos, las intenciones y las otras actitudes proposicionales de una persona conduce a la conclusión de que la mayor parte de las creencias de una persona deben ser verdaderas, y en consecuencia existe una presunción legítima de que una cualquiera de ellas, si forma un conjunto coherente con las demás, es verdadera. A continuación procedo a defender que una persona que tenga pensamientos, y por tanto cualquiera que se pregunte si tiene alguna razón para suponer que en general está en lo cierto acerca de la naturaleza de su entorno, debe saber lo que es una creencia, y cómo, en general, se han de detectar e interpretar las creencias. Al tratarse de hechos perfectamente generales que no podemos dejar de usar cuando nos comunicamos con los demás, o cuando tratamos de comunicarnos con ellos, o incluso cuando meramente pensamos que estamos comunicándonos con ellos, hay un sentido bastante fuerte en que se puede decir que sabemos que existe una presunción en favor de la verdad global de las creencias de cualquier persona, incluyendo las nuestras. En consecuencia, no tiene objeto que alguien pida alguna *otra* consideración tranquilizadora; éso sólo puede añadir algo más al montón de sus creencias. Todo lo que se necesita es que reconozca que las creencias son verídicas por naturaleza.

Puede verse que las creencias son verídicas considerando lo que determina la existencia y contenido de una creencia. Las creencias, como las otras actitudes proposicionales —como se las denomina—, sobrevienen a hechos de diversos tipos, conductuales, neurofisiológicos, biológicos y físicos. La razón para señalar esto no es alentar una reducción definicional o nomológica de los fenómenos psicológicos a algo más básico, y ciertamente no es la de sugerir prioridades epistemológicas. De lo que se trata es más bien de la comprensión. Adquirimos más conocimiento, de un cierto tipo, sobre la naturaleza de las actitudes proposicionales cuando las relacionamos sistemáticamente unas con otras y con fenómenos situados a otros niveles. Dado que las actitudes proposicionales están

profundamente entrelazadas, no podemos llegar a conocer la naturaleza de una de ellas adquiriendo primero una comprensión de alguna otra. Como intérpretes encontramos nuestro propio camino hacia el sistema como un todo, dependiendo mucho de las pautas de interrelaciones.

Tomemos, por ejemplo, la interdependencia de la creencia con el significado. Lo que una oración significa depende en parte de las circunstancias externas que hacen que adquiera cierto grado de convicción y en parte de las relaciones, gramaticales o lógicas, que esa oración tenga con otras oraciones que se consideran verdaderas con diversos grados de convencimiento. Puesto que estas mismas relaciones se traducen directamente a creencias, es fácil ver la manera en que el significado depende de las creencias. Las creencias, no obstante, dependen igualmente del significado, pues el único acceso a la estructura fina y a la individuación de las creencias es a través de las oraciones que los hablantes e intérpretes de los hablantes usan para expresar y describir creencias. Por lo tanto, si queremos iluminar la naturaleza del significado y de la creencia, es preciso comenzar con algo que no asuma ninguna de esas dos cosas. La propuesta de Quine, que seguiré esencialmente, es tomar el *asentimiento provocado* como básico, es decir, la relación causal entre asentir a una oración y la causa de tal asentimiento. Éste es un lugar equitativo para comenzar el proyecto de identificar creencias y significados, puesto que el asentimiento de un hablante a una oración depende tanto de lo que éste pretende que la oración signifique como de lo que cree acerca del mundo. Sin embargo, es posible saber que un hablante asiente a una oración sin saber ni lo que significa esa oración tal como él la dice, ni cuál es la creencia que la misma expresa. Igual de evidente es el hecho de que, en cuanto se haya dado una interpretación a la oración a la que se ha asentido, se habrá atribuido una creencia. Si las teorías correctas de la interpretación no son únicas (no conducen a una única interpretación correcta), lo mismo vale, por supuesto, para la atribución de creencias en tanto en cuanto éstas aparecen vinculadas a la conformidad respecto de oraciones determinadas.

Un hablante que desee que sus palabras sean comprendidas no puede engañar sistemáticamente a sus intérpretes po-

tenciales acerca de cuándo asiente a las oraciones, es decir, de cuándo las considera verdaderas. Por tanto, por cuestión de principio el significado y, por sus conexiones con el significado, también las creencias, están abiertos a la determinación pública. Me aprovecharé de este hecho en lo que sigue y adoptaré la posición de un intérprete radical al inquirir por la naturaleza de las creencias. Todo lo que puede conocerse es lo que un intérprete plenamente informado puede llegar a conocer acerca de lo que quiere decir un hablante, y lo mismo vale para lo que un hablante cree⁶.

El problema del intérprete es que lo que se supone que conoce —las causas de los asentimientos a las oraciones del hablante— es, como hemos visto, el producto de dos cosas que se supone que no conoce, el significado y las creencias. Si conociera los significados conocería las creencias, y si conociera las creencias que expresan las oraciones a las que asiente, conocería los significados. Pero ¿cómo puede él llegar a averiguar ambas cosas al mismo tiempo, puesto que cada una de ellas depende de la otra?

En líneas generales, la solución, como el problema mismo, se debe a Quine. Sin embargo, voy a introducir algunos cambios en la solución de Quine, al igual que lo he hecho en el enunciado del problema. Los cambios son directamente pertinentes para el tema del escepticismo epistemológico.

La interpretación radical (que es en buena medida, pero no totalmente, como la traducción radical de Quine) pretende producir una caracterización de la verdad al estilo de Tarski para la lengua del hablante y una teoría de sus creencias. (La segunda se sigue de la primera al añadirle el conocimiento presupuesto de cuáles oraciones se consideran verdaderas.) Esto añade bien poco al programa de traducción de Quine, puesto que la traducción de la lengua del hablante a la propia lengua, junto con una teoría de la verdad para la propia lengua, da como resultado una teoría de la verdad para el hablan-

⁶ Ahora pienso que, al hacer interpretación radical, es esencial incluir los deseos del hablante desde el comienzo, de manera que los resortes de la acción y la intención, las creencias y deseos por igual, están relacionados con el significado. Pero en el presente ensayo no es necesario introducir este otro factor.

te. Pero trasladar la cuestión de la noción sintáctica de traducción a la noción semántica de verdad sitúa en primer plano las restricciones formales de una teoría de la verdad y pone el énfasis en un aspecto de la estrecha relación entre verdad y significado.

El principio de caridad tiene un papel crucial en el método de Quine y un papel aún más crucial en mi variante del mismo. En cualquiera de los dos casos el principio dirige al intérprete a traducir o interpretar de modo que lea sus propias normas o medidas para la verdad en la pauta de las oraciones que el hablante considera verdaderas. La razón de este principio es hacer inteligible al hablante, puesto que suponer demasiadas desviaciones de la consistencia y la corrección no deja ninguna base común desde la cual juzgar la conformidad o la diferencia. Desde un punto de vista formal, el principio de caridad ayuda a resolver el problema de la interacción del significado y las creencias limitando los grados de libertad que se admiten en las creencias mientras se está determinando cómo interpretar las palabras.

Como ha sostenido Quine, no tenemos más elección que leer nuestra propia lógica en los pensamientos de un hablante. Quine lo afirma con relación al cálculo proposicional, y yo añadiría lo mismo para la teoría cuantificacional de primer orden. Esto conduce directamente a la identificación de las constantes lógicas, junto con la asignación de una forma lógica a cada oración.

Algo así como la caridad opera en la interpretación de las oraciones cuyas causas para asentir van variando según el instante y el lugar; cuando el intérprete encuentra una oración del hablante a la cual el hablante asiente regularmente en condiciones que el intérprete reconoce, el intérprete considera que ésas son las condiciones de verdad de la oración del hablante. Esto sólo es aproximadamente correcto, como vamos a ver en un momento.

Las oraciones y los predicados que están menos directamente integrados con sucesos fáciles de detectar pueden, según el canon establecido por Quine, ser interpretados a voluntad, dadas sólo las restricciones que supone su interconexión con oraciones que están directamente condicionadas al

mundo. Aquí es donde yo extendería el principio de caridad para favorecer interpretaciones que preserven la verdad en la medida de lo posible; pienso que interpretar —cuando se pueda— como verdadero lo que el hablante acepta como verdadero favorece la comprensión mutua, y por tanto una mejor interpretación. Tengo menos opciones que Quine porque no veo cómo trazar desde el comienzo una línea que separe las oraciones observacionales y las oraciones teóricas. Hay varias razones para sostener esto, pero la que es más relevante para el tema presente es que esa distinción se basa en último término en una consideración epistemológica de un tipo al que he renunciado, a saber, que las oraciones observacionales se basan directamente en algo como las sensaciones —configuraciones de estimulaciones sensoriales—. Como he estado sosteniendo, es ésta una idea que conduce al escepticismo. Sin un vínculo directo con tales sensaciones o estimulaciones, la distinción entre las oraciones observacionales y las demás no puede hacerse con fundamentos que sean epistemológicamente significativos. No obstante, sigue habiendo una distinción entre las oraciones cuyas causas de asentimiento varían con las circunstancias observables y aquellas a las que se apega un hablante a pesar de los cambios, y esta distinción ofrece la posibilidad de interpretar las palabras y las oraciones más allá del terreno de la lógica.

Los detalles no vienen aquí al caso. Lo que debería estar claro es que si es correcta la explicación que he dado de cómo se relacionan creencia y significado y cómo los comprende un intérprete, entonces la mayoría de las oraciones que un hablante considera verdaderas —en especial las que él mantiene más tenazmente, las que son más centrales en el sistema de sus creencias— *son* verdaderas, al menos en opinión del intérprete. La razón es que el único método —y por tanto irrecusable— de que puede disponer el intérprete pone automáticamente las creencias del hablante de acuerdo con las normas y medidas de la lógica del intérprete, y por tanto otorga al hablante las verdades sencillas de la lógica. Existen, innecesario es decirlo, diversos grado de consistencia lógica y de otras consistencias, y no ha de esperarse una consistencia perfecta. Lo que es necesario enfatizar es únicamente la

necesidad metodológica de encontrar una consistencia suficiente.

Análogamente, es imposible que un intérprete entienda a un hablante y al mismo tiempo descubra que el hablante está en gran parte equivocado acerca del mundo, ya que el intérprete interpreta las oraciones que se tienen por verdaderas (que es algo que no cabe distinguir del atribuir creencias) de acuerdo con los acaecimientos y objetos del mundo externo que causan que la oración sea considerada verdadera.

Hay una tendencia a pasar por alto lo que creo que es el aspecto importante de este enfoque porque el mismo invierte nuestra manera natural de pensar en la comunicación que se deriva de situaciones en las que la comprensión ya ha sido asegurada. Una vez que se ha asegurado la comprensión podemos a menudo averiguar lo que una persona cree con independencia de lo que le llevó causalmente a creerlo. Esto nos puede conducir a la conclusión crucial —fatal, realmente— de que en general podemos fijar lo que alguien quiere decir o significar independientemente de lo que cree y de lo que causó su creencia. Pero si estoy en lo cierto, no podemos en general identificar primero las creencias y los significados y preguntar luego qué es lo que los causó. La causalidad desempeña un papel indispensable en la determinación del contenido de lo que decimos y creemos. Es éste un hecho que podemos vernos llevados a reconocer al adoptar, como hemos hecho, el punto de vista del intérprete.

La interpretación correcta por parte de un intérprete de lo que una persona dice y de sus actitudes da como resultado que haya un elevado grado de verdad y consistencia en lo que un agente piensa y dice. Pero se trata de la verdad y la consistencia según las normas y medidas del intérprete. ¿Por qué no podría ocurrir que el hablante y el intérprete se comprendieran mutuamente sobre la base de creencias compartidas pero erróneas? Esto puede ocurrir, y sin duda ocurre a menudo. Pero no puede ser la regla, pues imaginemos por un momento a un intérprete que es omnisciente acerca del mundo, y acerca de lo que le lleva y le llevaría causalmente a un hablante a asentir a una oración cualquiera de su repertorio (potencialmente ilimitado). El intérprete omnisciente, utilizando el

mismo método que el intérprete falible, encuentra que el hablante falible es consistente, y está en lo cierto en gran parte. Según sus propias normas, por supuesto, pero puesto que éstas son objetivamente correctas, se ve al hablante falible como alguien que en gran parte está en lo cierto y es consistente según normas objetivas. Si queremos, podemos hacer que el intérprete omnisciente dirija su atención al intérprete falible del hablante falible. Resulta que el intérprete falible puede estar equivocado sobre algunas cosas, pero no en general, por lo que no puede compartir un error universal con el agente que está interpretando. Una vez que estamos de acuerdo con el método general de interpretación que he esbozado, se hace imposible sostener correctamente que nadie pueda estar en su mayor parte equivocado acerca de cómo son las cosas.

Como he señalado anteriormente, existe una diferencia clave entre el método de la interpretación radical que estoy recomendando ahora y el método de la traducción radical de Quine. La diferencia estriba en la naturaleza de la elección de las causas que gobiernan la interpretación. Quine hace depender la interpretación de pautas de estimulaciones sensoriales, mientras que yo la hago depender de los acaecimientos y objetos externos acerca de los cuales se interpreta que versa la oración. Por tanto, la noción de significado de Quine está vinculada a criterios sensoriales, algo que él piensa que puede considerarse como evidencias. Esto conduce a Quine a dar una significación epistémica a la distinción entre las oraciones observacionales y las demás, puesto que se supone que las oraciones observacionales, por estar directamente condicionadas a los sentidos, tienen un tipo de justificación extralingüística. Ésta es la concepción contra la que argumenté en la primera parte de mi ensayo, insistiendo en que las estimulaciones sensoriales son, ciertamente, parte de la cadena causal que conduce a las creencias, pero no pueden, sin confusión, ser considerados como evidencias, es decir, como fuente de justificación de las creencias que aquéllas producen.

Lo que impide el paso hacia el escepticismo global acerca de los sentidos es, en mi opinión, el hecho de que debemos, en los casos más sencillos y metodológicamente más básicos,

considerar a los objetos de una creencia como las causas de esa creencia. Y lo que —como intérpretes— hemos de considerar que son es lo que de hecho son. La comunicación empieza allí donde convergen las causas: una preferencia que usted haga significa lo mismo que una mía si la creencia en su verdad está causada sistemáticamente por los mismos acontecimientos y objetos⁷.

Las dificultades con que se enfrenta esta concepción son evidentes, pero creo que pueden ser superadas. El método se aplica directamente, como mucho, únicamente a las oraciones ocasionales —las oraciones cuyo asentimiento lo causan sistemáticamente cambios comunes y corrientes del mundo. Las demás oraciones se interpretan mediante su condicionamiento a las oraciones ocasionales y a la aparición en las mismas de palabras que también aparecen en las oraciones ocasionales. Entre las oraciones ocasionales, algunas variarán en la credibilidad que tienen no sólo a la vista de las variaciones del entorno, sino también a la vista de la variación en credibilidad que se atribuya a oraciones relacionadas. Sobre esta base se pueden desarrollar criterios para distinguir grados de observacionalidad por razones internas, sin apelar al concepto de una base para las creencias situada fuera del círculo de las mismas.

El problema del error se relaciona con estos problemas y es todavía más fácil de captar, pues incluso en los casos más simples está claro que la misma causa (un conejo pasa corriendo) puede generar diversas creencias en el hablante y en el observador, y alentar así el asentimiento a oraciones que no pueden tener la misma interpretación. Es sin duda este hecho el que hizo que Quine pasara de los conejos a las pautas de estimulaciones como lo que da la clave de la interpretación. No estoy seguro de en qué medida un enfoque es estadísticamen-

⁷ Está claro que la teoría causal del significado tiene poco en común con la teoría causal de la referencia de Kripke y Putnam. Estas teorías tratan de relaciones causales entre nombres y objetos que los hablantes pueden muy bien desconocer. Con ello se incrementa la posibilidad de error sistemático. Mi teoría causal hace lo inverso al relacionar la causa de una creencia con su objeto.

te mejor que el otro. ¿Es la frecuencia relativa con la que pautas idénticas de estimulaciones provocan el asentimiento a «Gavagai» y a «Conejo» mayor que la frecuencia relativa con la que un conejo provoca las mismas dos respuestas en el hablante y el intérprete? No es ésta una cuestión que sea fácil de someter a prueba de un modo convincente. Pero imaginemos que los resultados hablen en favor del método de Quine. Habré de decir entonces lo que habría de decir en cualquier caso, a saber que, incluso al nivel más simple, al problema del error no se le puede hacer frente oración a oración. Lo mejor que podemos hacer es habérnoslas con el error holísticamente, es decir, interpretemos de manera que hagamos a un agente tan inteligible como sea posible dadas sus acciones, sus preferencias y su lugar en el mundo. Acerca de algunas cosas lo consideraremos equivocado, lo cual es el coste necesario de considerarlo acertado en otras ocasiones. En tosca aproximación, considerarlo acertado significa identificar las causas con los objetos de sus creencias, otorgándoles un peso especial a los casos más simples y tolerando el error allí donde pueda ser explicado de la mejor manera.

Supóngase que estoy en lo cierto en que un intérprete debe interpretar de modo que un hablante o agente resulte estar acertado en gran parte acerca del mundo. ¿Cómo le ayuda esto a la persona misma que se pregunta qué razones tiene para pensar que la mayoría de sus creencias son verdaderas? ¿Cómo puede ella aprender cosas acerca de las relaciones causales entre el mundo real y sus creencias que llevan al intérprete a interpretarle como alguien que va en la buena dirección?

La respuesta está contenida en la pregunta. Para dudar o preguntarse acerca de dónde provienen sus creencias, un agente debe saber lo que es una creencia. Ello trae consigo el concepto de verdad objetiva, pues la noción de creencia es la noción de un estado que puede o no mofarse de la realidad. Pero las creencias se identifican también, directa o indirectamente, mediante sus causas. Lo que un intérprete omnisciente sabe lo obtiene un intérprete suficientemente bien si entiende a un hablante, y es ésta precisamente la compleja verdad causal que nos hace ser los poseedores de creencias que

en efecto somos y que fija los contenidos de nuestras creencias. El agente sólo necesita reflexionar acerca de lo que es una creencia para apreciar que la mayoría de sus creencias básicas son verdaderas y, entre sus creencias, aquellas que se mantienen con mayor firmeza y que son coherentes con el grueso principal de sus creencias son las que tienden en mayor medida a ser verdaderas. La pregunta de «¿cómo sé yo que mis creencias son en general verdaderas?» se responde, pues, a sí misma, simplemente porque las creencias son por naturaleza generalmente verdaderas. Parafraseada o ampliada esa pregunta se convierte en: «¿Cómo puedo decir si mis creencias, que son por naturaleza generalmente verdaderas, son generalmente verdaderas?»

Todas las creencias están justificadas en el siguiente sentido: están apoyadas por numerosas otras creencias (de otro modo no serían las creencias que son), y existe una presunción en favor de su verdad. Tal presunción aumenta cuanto mayor y más significativo sea el conjunto de creencias con el cual sea coherente una creencia y, como las creencias aisladas no existen, no existen creencias para las que no haya alguna presunción en su favor. A este respecto, el que interpreta y el interpretado difieren. Desde el punto de vista del intérprete, la metodología hace valer una presunción general de verdad para el conjunto de creencias como un todo, pero el intérprete no necesita presumir que cada creencia particular de alguna otra persona es verdadera. Como he insistido, la presunción general aplicada a los demás no hace que estén globalmente acertados, pero proporciona el trasfondo contra el cual se les puede acusar de error. Pero desde el punto de vista de cada persona debe haber una presunción en grados diversos en favor de cada una de sus propias creencias.

Lamentablemente, no podemos extraer la conclusión pintoresca y agradable de que todas las creencias verdaderas constituyen conocimiento. La razón es que, si bien todas las creencias de un sujeto están en alguna medida justificadas para él, algunas pueden no estar suficientemente justificadas para constituir conocimiento, o no estarlo del modo apropiado. La presunción general en favor de la verdad de las creencias sirve para salvarnos de una forma usual de escepticismo

al mostrar por qué es imposible que todas nuestras creencias sean falsas conjuntamente. Esto deja casi sin empezar la tarea de especificar las condiciones del conocimiento. No he tratado de los cánones de apoyo evidencial (si es que los hay), sino para mostrar que cualquier cosa que cuente como evidencia o justificación para una creencia debe provenir de la misma totalidad de creencias a la que ella pertenece.

REFLEXIONES POSTERIORES

Unos cuantos filósofos ya bastante mayores, categoría en la que podemos incluir a Quine, Putnam y Dummett, y en la que ciertamente estoy incluido, todavía se preocupan seriamente acerca de la naturaleza de la verdad y de sus relaciones o falta de relaciones con el significado y la epistemología. Rorty piensa que esto debería cesar de preocuparnos; él cree que la filosofía ha descubierto el truco de esos rompecabezas o los ha superado, y debería pasar a ocuparse de asuntos de menor peso y mayor interés. Se muestra particularmente impaciente conmigo por no conceder que el antiguo juego ya ha acabado, puesto que encuentra en mi trabajo apoyos útiles para su ilustrada posición; por debajo de mi «retórica pasada de moda» detecta el perfil de una actitud en gran medida correcta.

En «Pragmatism, Davidson, and Truth» sostiene Rorty dos cosas: que mi concepción de la verdad equivale a un rechazo tanto de la teoría coherentista como de la teoría de la correspondencia, y debería propiamente clasificársela como perteneciente a la tradición pragmática, y que no debería pretender que estoy respondiéndole al escéptico cuando realmente le estoy diciendo que desaparezca de la vista. En buena medida estoy de acuerdo con él en ambos puntos.

En nuestra discusión de 1983 en el Encuentro de la División del Pacífico de la Asociación Filosófica Americana, me mostré de acuerdo en cesar de denominar teoría coherentista o teoría de la correspondencia a mi posición si él abandonaba la teoría pragmatista de la verdad. Él ha hecho su parte; rechaza ahora explícitamente tanto a James como a Peirce res-

pecto de la verdad. Tengo la satisfacción de atenerme a mi parte del trato. Si no se hubiera ya publicado, cambiaría ahora el título de «A Coherence Theory of Truth and Knowledge» [«Una teoría coherentista de la verdad y el conocimiento»] y no describiría el proyecto contenido en ese escrito como el de mostrar la manera en que «la coherencia produce correspondencia». Como señala Rorty, basándonos sólo en evidencias internas mi concepción no puede denominarse teoría de la correspondencia. Ya hace mucho tiempo, en 1969 («True to the Facts»)⁸, argumenté que no hay nada que se pueda decir —de manera útil e inteligible— que se corresponda con una oración; esto lo repetí en «Una teoría coherentista de la verdad y el conocimiento». Pensaba entonces que el hecho de que al caracterizar la verdad para una lengua o lenguaje sea necesario poner las palabras en relación con objetos era suficiente para darle a la idea de correspondencia algo a lo que asirse; pero esto me parece ahora un error. El error es, en cierto modo, sólo una denominación inadecuada, pero las inadecuaciones terminológicas encuentran un modo de generar confusiones conceptuales, y así sucede aquí. Las teorías de la correspondencia se han concebido siempre como algo que proporciona una explicación o análisis de la verdad, y esto es algo que una teoría de la verdad al estilo de Tarski no hace. Ahora también rechazaría la observación que generalmente se hace en contra de las teorías de la correspondencia respecto de que no hay manera alguna en que podamos decir nunca que nuestras oraciones o creencias se correspondan con la realidad. Esta crítica es, en el mejor de los casos, engañosa, puesto que nadie ha explicado nunca en qué podría consistir tal correspondencia; y, lo que es peor, se predica sobre la base de la presunción falsa de que la verdad es epistémica de forma transparente.

Lamento también haber denominado a mi teoría «teoría coherentista». Mi énfasis en la coherencia era propiamente sólo un modo de hacer una observación negativa, la de que «cualquier cosa que cuenta como evidencia o justificación

⁸ Reimpreso en *Inquiries into Truth and Interpretation*.

para una creencia debe provenir de la misma totalidad de creencias a la que ella pertenece». Desde luego, esta afirmación negativa ha llevado típicamente a los filósofos que la sostenían a concluir que la realidad y la verdad son constructos del pensamiento; pero esto no me lleva a mí a esta conclusión, y por esta razón, aunque no hubiera otra, no debería haber denominado teoría coherentista a mi teoría. Hay también otra razón de menor peso para no acentuar la coherencia. La coherencia no es nada más que consistencia. Ciertamente que un conjunto de creencias sea consistente dice algo en su favor, pero no hay ninguna posibilidad de que las creencias de una persona no tiendan a ser consistentes, puesto que las creencias se individualizan en parte por sus propiedades lógicas; lo que no es en gran parte consistente con muchas otras creencias no puede ser identificado como una creencia. El contenido principal de «Una teoría coherentista de la verdad y el conocimiento» tiene poco que ver con la consistencia; la tesis importante en favor de la que argumento es que las creencias son intrínsecamente verídicas. Este es el fundamento para mantener que aunque la verdad no es un concepto epistémico, tampoco está totalmente separado de las creencias (como de maneras diferentes lo está tanto para la teoría de la correspondencia como para la teoría coherentista).

Mi énfasis en la coherencia estaba fuera de lugar; denominar «teoría» a mi concepción fue sencillamente un patinazo. En su artículo Rorty acentuaba una actitud minimista hacia la verdad que él pensaba correctamente que él y yo compartíamos. Podría decirse del siguiente modo: la verdad es un concepto tan claro y básico como pueda haberlo. Tarski nos ha dado una idea de cómo aplicar ese concepto general (o de cómo tratar de aplicarlo) a lenguas o lenguajes particulares en el supuesto de que ya lo entendemos; pero desde luego no mostró cómo definirlo en general (más bien demostró que esto no se podía hacer). Cualquier intento ulterior de explicar, definir, analizar o reconstruir ese concepto resultará vacío o estará equivocado; las teorías de la correspondencia, las teorías de la coherencia, las teorías pragmatistas, las teorías que identifican la verdad con la asertabilidad garantizada (quizá en condiciones «ideales» u «óptimas»), las teorías que le piden

a la verdad que explique el éxito de la ciencia, o que sirva como el resultado último de la ciencia o de las conversaciones de una élite, todas estas teorías o bien no le añaden nada a nuestra comprensión de la verdad o tienen contraejemplos evidentes. ¿Por qué habríamos de esperar poder reducir la verdad a algo más claro o más fundamental? Después de todo, el único concepto que Platón logró definir fue el de fango (tierra sucia y agua). La comparación que hizo Putman de diversos intentos de caracterizar la verdad con los intentos de definir «bueno» en términos naturalistas me parece a mí, como a Rorty, acertada. Parece aplicarse también a la identificación por parte de Putman de la verdad con la aseverabilidad justificada idealizada⁹.

Una teoría de la verdad para un hablante, o grupo de hablantes, aunque no sea una definición del concepto general de verdad, nos suministra un sentido sólido de qué es para lo que sirve el concepto; nos permite decir, de una manera compacta y clara, qué es lo que sabe alguien que entiende a ese hablante o a esos hablantes. Tal teoría induce a preguntar cómo podría el intérprete confirmar su verdad —pregunta ésta que sin la teoría no podría articularse—. La respuesta, como trato de mostrar en «Una teoría coherentista de la verdad y el conocimiento», sacará a relucir relaciones esenciales entre los conceptos de significado, verdad y creencia. Si estoy en lo cierto, cada uno de estos conceptos requiere de los demás, pero ninguno está subordinado a los otros, y mucho menos es definible en términos de ellos. La verdad emerge como algo que no está totalmente separado de la creencia (como lo estaría según una teoría de la correspondencia), ni como dependiente de los métodos y capacidades de descubrimiento de los seres humanos (como lo estaría según las teorías epistémicas de la verdad). Lo que salva a la verdad de ser «radicalmente no-epistémica» (en palabras de Putnam) no es que la verdad sea epistémica, sino que la creencia, a través de sus vínculos con el significado, es intrínsecamente verídica.

Finalmente, ¿qué sucede con la recomendación de Rorty de dejar de tratar de responder al escéptico y decirle que desapa-

⁹ Hilary Putman, *Realism and Reason*, pág. xvii.

rezca de la vista? Una respuesta breve sería que al escéptico se le ha dicho esto una y otra vez durante milenios y no parece escuchar; como filósofo que es, quiere un argumento. Siendo algo más explícito: en el ensayo «Pragmatism, Davidson, and Truth» de Rorty encontramos quizá la sugerencia de que un enfoque «naturalista» de los problemas del significado y las actitudes proposicionales dejará al escéptico automáticamente sin espacio de maniobra. Este pensamiento, sea o no de Rorty, está equivocado. La epistemología naturalizada de Quine, debido a que se basa en la premisa empirista de que lo que queremos decir o significar y lo que pensamos se fundamenta conceptualmente (y no meramente de manera causal) en el testimonio de los sentidos, está abierta a los ataques escépticos usuales. Me preocupé mucho en «Una teoría coherrentista de la verdad y el conocimiento» de argumentar en favor de un enfoque alternativo del significado y el conocimiento, y de mostrar que, si esta alternativa fuera correcta, el escepticismo no podría alzar el vuelo. Estoy de acuerdo con Rorty en la medida siguiente; mi propósito no es «refutar» al escéptico, sino delinear lo que pienso que es una explicación correcta de los fundamentos de la comunicación lingüística y de sus consecuencias para la verdad, la creencia y el conocimiento. Si se concede la corrección de esta explicación, se le puede decir al escéptico que desaparezca de la vista.

En lo que Rorty y yo diferimos, si es que lo hacemos, es en la importancia que le damos a los argumentos que conducen a la perdición del escéptico, y en el interés que encontramos en las consecuencias para el conocimiento, la creencia, la verdad y el significado. Rorty quiere demorarse en la cuestión de a dónde han llevado los argumentos: a una posición que nos permite desechar las dudas del escéptico y abandonar en consecuencia el intento de proporcionar una justificación general de las pretensiones de conocimiento, una justificación que ni es posible ni se necesita. Rorty ve la historia de la filosofía occidental como una confusa batalla entre un escepticismo ininteligible e intentos poco convincentes de responder al mismo para la cual no hay vencedor. La epistemología, desde Descartes a Quine, a mí me parece exactamente un capítulo, que en absoluto re-

sulta poco iluminador, de la empresa filosófica. Si este capítulo se acerca al final, será mediante el recurso a modos de análisis y la observancia de normas de claridad que han distinguido siempre a la mejor filosofía y que, con suerte y empeño, continuarán haciéndolo.

11

El contenido empírico

La disputa entre Schlick y Neurath sobre los fundamentos del conocimiento empírico ilustra las dificultades que se presentan al tratar de extraer conclusiones epistemológicas de una teoría verificacionista del significado. Muestra también cómo el asumir la corrección de la ciencia en términos generales no evita automáticamente el escepticismo ni proporciona una respuesta fácil al mismo. Pero aunque ni Schlick ni Neurath llegaron a una explicación satisfactoria del conocimiento empírico, hay en sus escritos indicaciones que prometen una teoría mejor. Siguiendo esas indicaciones y utilizando otras ideas que se deben a Hempel, a Carnap y especialmente a Quine, voy a sugerir cuál es la dirección que creo que debería tomar una epistemología naturalista.

Los positivistas lógicos estaban de acuerdo en que el contenido empírico de una oración interpretada deriva de sus relaciones con un subconjunto de oraciones que informan de, o se basan en, observaciones, experiencias o vivencias. Las dos principales fuentes de dificultades y de discusión se manifestaron inmediatamente. Una de ellas fue la cuestión de cómo caracterizar las relaciones entre las oraciones protocolares y otras oraciones. La historia de los desarrollos y cambios en las ideas acerca de este problema por parte de los positivistas lógicos y sus seguidores ha sido registrada de forma magistral

por Carl Hempel, el cual ha contribuido también mucho a las mismas¹. Éste no es ahora mi tema.

La segunda cuestión era la de cómo deberían formularse las oraciones protocolares, y cuál es su relación con la experiencia y la observación. Éste es el tema que voy a tratar y acerca del cual estaban en desacuerdo Schlick y Neurath, al aceptar Schlick una epistemología fundamentista y Neurath una teoría coherentista. Las diferencias se expresaron en términos bastante fuertes. Neurath describió la posición fundamentista como «relacionada con la creencia en las *vivencias inmediatas* que está presente en la filosofía académica tradicional», y afirmó que «el solipsismo metodológico» (que era la manera en que Carnap denominaba una concepción como la de Schlick) «no sirve de algo por el hecho de que se le haya añadido la palabra “metodológico”»². Schlick, a su vez, calificó la versión de Neurath de la teoría coherentista de «error asombroso»³.

Sea o no un error asombroso, parece que en una determinada época Carnap y Hempel estuvieron de acuerdo con Neurath. En 1935 escribió Hempel: «Pienso que no hay ninguna diferencia esencial entre los enunciados protocolares y otros enunciados», y coincidía con Carnap en sostener que, una vez que se formulaba en el «modo formal» la pregunta de qué oraciones son básicas, la respuesta se convertía en un asunto que debería decidirse por convención. «Esta idea», concluía haciéndose eco de las palabras de Carnap, «elimina de la teoría de la verificación y la verdad del positivismo lógico un resto de absolutismo que se debe a tendencias metafísicas»⁴. Schlick se mofó del convencionalismo diciendo que hacía que la verdad fuera tan relativa «como todas las varas de medir de la física». De manera irónica añadió: «y es esta con-

¹ Véase Carl Hempel, «Empiricist Criteria of Cognitive Significance: Problems and Changes», en *Aspects of Scientific Explanation*.

² Otto Neurath, «Protocol Sentences», citado según Ayer (ed.), *Logical Positivism*, 204, 206 [cfr. págs. 210, 212 de la traducción al castellano]. Todas las referencias subsiguientes a las páginas de *Logical Positivism* aparecen precedidas por las letras LP.

³ Moritz Schlick, «The Foundation of Knowledge», LP 215 [221].

⁴ Carl Hempel, «On the Logical Positivist's Theory of Truth», 58, 59.

cepción junto con sus consecuencias la que se ha ensalzado como algo que elimina el último resto del "absolutismo" en filosofía»⁵.

No obstante, no está del todo claro cuál era la situación. Quince años más tarde, en un artículo cuasihistórico, Hempel afirmaba que «el principio fundamental del empirismo moderno es la tesis de que todo el conocimiento no-analítico se basa en la experiencia». A continuación procedía a explicar lo que quiere decir que el conocimiento se basa en la experiencia: el conocimiento no-analítico puede expresarse mediante oraciones que están confirmadas (de una manera especificada) por medio de oraciones observacionales, las cuales, se «averigua», a su vez, si son verdaderas mediante la observación directa⁶. Esto suena más parecido a Schlick; en realidad, suena a la «fatal confrontación de enunciados y hechos» que Hempel había rechazado con anterioridad⁷.

En una nota añadida a una reimpresión del artículo «Studies in the Logic of Confirmation»⁸, Hempel sugiere una manera de reconciliar esos puntos de vista aparentemente opuestos. La verdad, como él y Carnap habían llegado a reconocer a la luz de la obra de Tarski, es una noción semántica legítima, y no debería tratarse como una cuestión de coherencia. Por otra parte, la confirmación era de dos tipos, *relativa* y *absoluta*. El estudio lógico de la confirmación era el estudio de en qué medida un conjunto de oraciones confirmaba una hipótesis. En este contexto sólo era posible decir que, relativamente a un conjunto de oraciones (provinieran de donde provinieran), una hipótesis se confirmaba o se desconfirmaba. Por otra parte, la confirmación absoluta dependía de una decisión de tipo «pragmático» consistente en tratar determinadas oraciones como verdaderas. Según afirma Hempel, Neurath y Carnap pensaban en la confirmación relativa, que sugiere una teoría coherentista, mientras que

⁵ Moritz Schlick, «The Foundations of Knowledge», LP 213 [219].

⁶ Carl Hempel, «The Empiricist Criterion of Meaning», LP 108-10 [115-17].

⁷ *Íd.*, «On the Logical Positivist's Theory of Truth», 51.

⁸ Nota 49 de «Studies in the Logic of Confirmation», en *Aspectos of Scientific Explanation*.

Schlick, y Hempel en 1950, pensaban en la confirmación absoluta.

La importante distinción de Hempel es reveladora, puesto que a veces parece que los análisis tempranos de las oraciones protocolares vacilaban entre tratar esas oraciones como cualesquiera oraciones de sintaxis especificada, y tratarlas como oraciones aceptadas, quizá sobre la base de la observación o la experiencia. Pero esa distinción no puede reconciliar todas las diferencias. Las diferencias que quedaban eran las siguientes: Schlick sostenía que las oraciones de observación pueden ser establecidas como verdaderas de una vez por todas, mientras que Neurath lo negaba; Schlick afirmaba que podía decirse inteligiblemente que una oración puede compararse con la realidad, mientras que Neurath lo negaba. Había también diferencias en la cuestión de acerca de qué tratan los protocolos y en la cuestión de si informan sobre algo privado o sobre algo público. Es evidente que estos diversos puntos se relacionan estrechamente unos con otros.

Una manera de enfocar nuestro problema principal es preguntar cuál es la naturaleza de las *evidencias*: ¿consisten éstas en objetos, acaecimientos, hechos, vivencias, sensaciones, creencias, proposiciones u oraciones? Casi todas estas respuestas posibles pueden hallarse en los escritos del Círculo de Viena.

Observo dos trozos de papel verde [escribe Schlick] y determino que tienen el mismo color. La proposición que afirma la identidad de color se verifica, entre otras maneras, mediante el hecho de que al mismo tiempo yo tengo dos vivencias idénticas en cuanto al color. La proposición: «hay dos manchas del mismo color ante mí ahora» no puede reducirse a otras; la verifica el hecho de que describe lo dado⁹.

¿A qué se debe exactamente la verificación? Schlick dice que al hecho de que él tiene determinadas experiencias que son verídicas. Pero si esas mismas experiencias no fueran verídicas, ¿verificarían todavía la misma proposición? El hecho

⁹ Moritz Schlick, «Positivism and Realism», *LP* 92-3 [99].

de que yo vea un trozo de papel implica que hay un trozo de papel, pero aquí no nos movemos de las evidencias a la hipótesis de un modo interesante; meramente deducimos una proposición que está implicada. En otro lugar, Schlick insiste en que debemos comenzar con enunciados que «tienen su origen» en oraciones observacionales; y esto lo elabora así: «derivan, como podría confiadamente decirse al modo tradicional de hablar, “de la experiencia”»¹⁰. En el mismo ensayo declara que «está claro —y hasta donde sé no lo discute nadie— que el conocimiento en la vida y en la ciencia *comienza* en *algún* sentido con la confirmación de hechos y que los “enunciados protocolares” en los que esto sucede están en el mismo sentido al *comienzo* de la ciencia»¹¹. En otro pasaje dice que todo se remonta a lo que se «observa inmediatamente»¹². Como dijo Ayer, algunas proposiciones pueden ser «confrontadas directamente con los hechos».

Voy a tratar de formular de otra manera la cuestión aparentemente problemática del estatuto epistemológico de las evidencias. Decimos que las leyes resultan confirmadas mediante sus ejemplificaciones positivas; de modo que las ejemplificaciones positivas constituyen, presumiblemente, evidencias en favor de las leyes que ellas confirman. Supóngase, en aras de la claridad en una de las direcciones, que identificamos las leyes con oraciones condicionales universalmente cuantificadas —oraciones, innecesario es decirlo, interpretadas—. ¿Qué constituye una ejemplificación positiva? Supongamos que la ley tiene la forma « $(x)(Fx \rightarrow Gx)$ ». Entonces, según afirma Hempel, es razonable suponer que un *objeto* que es F y G confirma la ley¹³. (Goodman habla a menudo de este modo en *Fact, Fiction, and Forecast*.) Hempel explica entonces que en lugar de ver la confirmación como una «relación entre un objeto o un conjunto ordenado de objetos, como representantes de las evidencias, y una oración como representante de la hi-

¹⁰ Moritz Schlick, «The Foundations of Knowledge», LP 215 [221].

¹¹ *Ibid.*, 210 [215].

¹² *Ibid.*, 220 [225].

¹³ Carl Hempel, «Studies in the Logic of Confirmation», en *Aspects of Scientific Explanation*, 14.

pótesis», la tomará como una relación entre una oración que describe la evidencia, y la hipótesis. De este modo, «Las evidencias aducidas para apoyar una teoría científica se expresan siempre en oraciones, las cuales poseen frecuentemente el carácter de informes de observaciones... Las evidencias... consisten, en último análisis, en datos accesibles a lo que de modo impreciso se denomina observación directa y tales datos pueden expresarse en forma de “informes de observaciones”»¹⁴. Lo que es accesible a la observación son objetos y acaecimientos. Éstos no son lo mismo que los hechos, ni, por supuesto, que las oraciones. En un sentido impreciso, las oraciones pueden expresar hechos (es decir, proposiciones verdaderas) y describir objetos. Nada de lo que digo es una crítica a la ejemplar obra de Hempel sobre la confirmación como relación entre oraciones. Estoy utilizando la distinción entre varias maneras de describir las ejemplificaciones positivas de las leyes (o de oraciones nómicas) a fin de enfatizar las muy diferentes maneras en que es natural hablar de evidencias.

Tal vez no sea extraño denominar a un cuervo negro —un pájaro real— una *ejemplificación* de una ley, pero parece extraño decir que el pájaro en cuestión constituye *evidencia* a favor de la ley. Como mucho esto parece un modo abreviado de decir que es el *hecho* de que este pájaro sea un cuervo negro lo que constituye la evidencia; o podríamos hablar de la verdad de la proposición, o de alguna oración apropiada. Sin embargo, hasta el momento no hemos tocado la cuestión epistemológica, la cuestión de qué significa que *alguien* tenga una razón para aceptar la ley, que *posea* evidencia. Ni la existencia de un cuervo negro ni la verdad de la proposición o de la oración que dice que hay un cuervo negro por sí misma le proporcionan a nadie una razón para creer que hay un cuervo negro, y mucho menos una razón para creer que todos los cuervos son negros. Para que alguien tenga una razón para creer que todos los cuervos son negros es necesario que *crea*, por ejemplo, que ahí hay un cuervo negro.

¹⁴ *Ibíd.*, 21-2.

Nos encontramos así en un camino muy transitado. Por supuesto no es suficiente con creer que ahí hay un cuervo negro, ni tampoco es suficiente con que la creencia haya también de ser verdadera, pues esas dos condiciones conjuntamente no constituyen evidencia a menos que la persona tenga una razón adecuada para mantener esa creencia. Si la razón debe ser otra creencia, nos enfrentamos a un regreso infinito o a un círculo. Un regreso haría imposible el conocimiento, mientras que un círculo conduciría a las dificultades de una teoría coherentista pura del conocimiento. En breve volveré sobre esto último.

En este punto llegamos a los diversos intentos de hallar estados mentales que sirvan para salvar o eliminar la brecha entre la sensación, en la que no se suscita la cuestión de la verdad, y el juicio, que verosímilmente es una fuente de evidencia. Se ha propuesto que cuasioraciones como «Negro ahí ahora» expresan tales estados mentales. Y quizá nos convenceremos de que existen tales estados mentales si pasamos por alto el hecho de que se ha omitido el verbo (puesto que ponerlo llevaría las cosas demasiado lejos en la dirección del juicio) y que las palabras «ahí» y «ahora» no puedan comprenderse excepto en tanto que involucran una referencia a un agente. En cualquier caso, el intento de basar la ciencia en tales estados mentales está condenado de antemano, puesto que nadie ha conseguido nunca mostrar la manera de basar el conocimiento de un mundo común, objetivo, en tales «evidencias». Incluso Schlick, que de algún modo tenía la esperanza de respaldar las oraciones protocolares de la forma «A experimenta negro en el instante t » mediante lo que expresa «Negro ahí ahora» cuando lo dice o piensa A en el instante t , sea esto lo que sea, no creía en una construcción de la ciencia o de un mundo público que estuviera basada en las «vivencias inmediatas».

Debería ser evidente que el recurrir a la percepción no puede en modo alguno aclarar la cuestión de qué es lo que constituye la fuente última de las evidencias de una persona, pues si decimos que la percepción consiste en una sensación causada por un acaecimiento que se da en mundo (o en el cuerpo del perceptor), el hecho de la causalidad no puede darse

aparte de la sensación, y la sensación no puede servir como evidencia a menos que cause una creencia. Pero ¿cómo sabe uno que la creencia fue causada por una sensación? Únicamente otras creencias pueden servir de ayuda. Si la percepción se expresa mediante locuciones como «A percibe que hay un cuervo negro», entonces esto puede realmente servir como evidencia. Esto no resuelve el problema, sino que sólo lo traslada al concepto de percepción, puesto que percibir que hay un cuervo negro no consiste sino en que un cuervo negro cause, *del modo adecuado*, la creencia de que hay un cuervo negro.

Al leer los escritos tempranos de los miembros del Círculo de Viena impresiona el modo embarazoso con que se refieren a la experiencia, a lo que es dado inmediatamente, a lo que se observa directamente. Así, Schlick dice que todo significado se remite a situaciones de ostensión, «y esto significa, en un sentido obvio, hacer referencia a la “experiencia” o “posibilidad de verificación”¹⁵. En «The Turning Point in Philosophy» dice: «El acto de verificación en el que acaba la senda que conduce a la solución es siempre del mismo tipo; consiste en que se da un hecho definido, confirmado por la observación, por medio de la vivencia inmediata»¹⁶; pero en «Positivism and Realism» expresa dudas graves acerca de términos como «lo dado» (*das Gegebene*) y le preocupa que si utilizáramos la palabra «vivencia» «estaríamos presuponiendo una distinción entre lo que puede experimentar tal vivencia y lo que es experimentado»¹⁷.

Hay, pues, buenas razones para concluir que la idea de comparar nuestras creencias con la realidad o de confrontar nuestras hipótesis con la observación no tiene un significado claro. Por descontado, esto no quiere decir que se niegue que haya un sentido corriente en el que llevamos a cabo experimentos y anotamos los resultados, o descubrimos en nuestras ocupaciones cotidianas que algunas de nuestras creencias son verdaderas y otras falsas. Lo que debería negarse es que estos

¹⁵ Moritz Schlick, «Meaning and Verification», 148.

¹⁶ *Íd.*, «The Turning Point in Philosophy», LP 56 [62].

¹⁷ *Íd.*, «Positivism and Realism», LP 84 [90].

acaecimientos mundanos hayan de analizarse como algo que involucra evidencias que no tienen carácter proposicional —evidencias que no son creencias de alguna clase—. ¡No es de extrañar que Neurath y Carnap se sintieran atraídos por la idea de una teoría coherentista!

Por supuesto, las teorías coherentistas del conocimiento no están libres de dificultades, y, como hemos visto, Schlick estuvo presto a señalarlas. Voy a detenerme aquí por un momento para hacer la distinción obvia entre una teoría coherentista del conocimiento y una teoría coherentista de la verdad. En su artículo de 1935 «On the Logical Positivist's Theory of Truth», Hempel apenas había distinguido entre las dos; comprensiblemente, puesto que no conocía el método de Tarski para definir semánticamente la verdad. Se inclinaba por tanto a pensar que el único sentido que podemos darle a la expresión «La oración O es verdadera» es « O está confirmada en muy alto grado por los informes de observación aceptados»¹⁸. Pero el concepto de estar confirmado en un grado muy alto por los informes de observación aceptados pertenece más bien al dominio de la epistemología; y cuando se le empareja con la idea de que los enunciados protocolares «únicamente pueden ser caracterizados mediante el hecho histórico» de que son aceptados (es decir, se cree que son verdaderos), conduce directamente a una teoría coherentista del conocimiento. Ésta es la teoría a la que se oponía Schlick.

Schlick estaba de acuerdo con Neurath en que los protocolos pertenecen, de un modo general, a las hipótesis de la ciencia. Son objetivos y por lo tanto se entienden de un modo intersubjetivo y son contrastables. Versan *acerca* de observaciones o de experiencias, pero no tratan de *expresarlas*. Toman la forma, al menos aproximadamente, de oraciones como « A vio un cuervo negro en el instante t ». Está claro que no se puede estar seguro de la verdad de tales oraciones —ni siquiera A en el instante t puede tener la certeza de que está viendo un cuervo negro—; o, dejando aparte la cuestión de la certeza

¹⁸ Carl Hempel, «Studies in the Logic of Confirmation», en *Aspects of Scientific Explanation*, 42.

za, está claro que cualquier persona que juzgue que esa oración es verdadera puede estar equivocado. Ulteriores evidencias son siempre pertinentes, y pueden llegar a pesar más que la evidencia de un determinado momento. En lo que Schlick estaba en desacuerdo con Neurath es en la cuestión de si hay fundamentos incuestionables en base a los cuales juzgamos que los protocolos son verdaderos¹⁹.

La objeción a la teoría coherentista de Neurath era la objeción estándar a todas esas teorías; la consistencia no es suficiente, puesto que nos deja sin base alguna para elegir entre diversas teorías en conflicto. Quizá una teoría de este tipo elimina los últimos restos del absolutismo en filosofía, como decía Schlick de las afirmaciones de Neurath, pero nos deja sin base alguna para juzgar lo que es verdadero. Schlick insistió en que hemos de tener, y tenemos, bases indudables para elegir algunas oraciones como verdaderas en lugar de otras. Hay enunciados que *no* son oraciones protocolares, los cuales «expresan hechos de la propia “percepción”» («o como quera- mos decirlo», añade). Schlick admite entonces que «a pesar del hecho de que los enunciados de este tipo parecen tan simples y claros, los filósofos se han encontrado en un laberinto imposible desde el instante mismo en que han intentado utilizarlos como fundamento de todo el conocimiento». Pero piensa que podemos librarnos de todas esas conocidas dificultades si tenemos presente que «son nuestros *proprios* enunciados los únicos que en último término desempeñan un papel decisivo».

Hay que admitir que el intento de Schlick de explicar su teoría finaliza en la oscuridad. Las «oraciones» de observación que constituyen el «criterio último» de todo conocimiento no son en realidad oraciones, al ser siempre de la forma «Ahí ahora tal y cual». No se pueden anotar esas oraciones (puesto que en un instante pierden su certeza); expresan un «sentimiento de *consumación*, una satisfacción muy característica: nos sentimos *satisfechos*». «No se puede construir una estructu-

¹⁹ Moritz Schlick, «The Foundations of Knowledge», LP 213 [215]. Las siete citas que vienen a continuación provienen de LP 218-25 [219-25].

ra sostenible lógicamente sobre tales confirmaciones, pues en el instante en que se comienza a construir ya se han ido.» Finalmente, «la ocasión en que se comprenden [los enunciados de observación] es al mismo tiempo aquella en que se verifican: capto su significado al mismo tiempo que capto su verdad». Se puede simpatizar con Neurath por rechazar este último paso. Pero entonces lo que queda es una teoría coherentista.

Hempel denomina a la posición de Neurath-Carnap teoría coherentista «comedida». La razón es que Neurath y Carnap nos proporcionan un criterio para elegir una teoría científica entre aquellas que son consistentes. El criterio es que se trata de la teoría consistente que maximice el acuerdo con los enunciados que históricamente son considerados verdaderos «por la humanidad, especialmente los científicos de nuestro círculo cultural»²⁰. Al final las oraciones protocolares no tienen un lugar privilegiado; como cualesquiera otras pueden ser abandonadas si entran en conflicto con demasiadas cosas que consideramos verdaderas.

De esta suerte resulta, más bien de manera sorprendente, que tanto Schlick como Neurath sostenían concepciones que podrían denominarse teorías coherentistas «comedidas». Estaban de acuerdo en que cualesquiera cosas que pertenezcan al corpus científico, incluyendo las oraciones protocolares, han de considerarse sólo como establecidas de manera provisional en cualquier estadio dado de progreso de la ciencia, y todas las oraciones son susceptibles de revisión a la luz de nuevas evidencias. Del mismo modo estaban de acuerdo en que cuando se requiere proceder a la revisión, no hay reglas estrictas para decidir en qué lugar ha de efectuarse esa revisión; es una cuestión de «decisión». La única diferencia importante concernía a la cuestión de cómo la configuración global de las oraciones que la ciencia acepta en un momento dado se relaciona con la experiencia, la observación o el mundo real. Y a este respecto sus respuestas fueron insuficientemente claras o satisfactorias.

²⁰ Carl Hempel, «On the Logical Positivist's Theory of Truth», 57.

La respuesta de Schlick era insatisfactoria porque finalizaba con algo tan privado que incluso su significado podría especificarse sólo de manera relativa a un individuo y a un instante dados. No se explicaba cómo una base así puede proporcionar garantías a las creencias acerca de un mundo objetivo público. Neurath rechazaba la idea de una confrontación entre una creencia acerca del mundo y el mundo mismo, al igual que la idea de una base subjetiva incorregible para el conocimiento científico. Pero su sugerencia acerca de cómo hacer que una teoría coherentista sea «comedida» no es nada atractiva. Sugiere que comencemos con protocolos de la forma «A ve un cuervo negro en el instante t », y rechaza la idea de que tales protocolos constituyan la base del conocimiento de A en mayor medida que el conocimiento de B. Esto garantiza el aspecto intersubjetivo del lenguaje de la ciencia hasta llegar a los propios protocolos (puesto que, como dijo Neurath, «*todo lenguaje como tal es intersubjetivo*»²¹ —claramente una chanza dirigida a Schlick y sus enunciados de observación, cuyo significado se le revela sólo a una persona y sólo por un instante—). Neurath imagina que todos los protocolos se echan a una gran máquina; suena un timbre si surge una contradicción y alguna cosa debe ser expulsada, ya sea uno o más de los protocolos, ya una ley u otro enunciado teórico. Pero «*quién sea el que rehace la máquina o de quién sean las oraciones protocolares que se echan a la máquina no tiene ninguna importancia en absoluto*»²².

Hay aquí una dificultad obvia. Si a las oraciones protocolares sólo se las conoce por su forma, introducirlas *todas* en la máquina querrá decir que se introduce en ésta cada una de las oraciones *juntamente con* su negación y también tanto «A ve un cuervo negro en el lugar s en el instante t » como «B ve un cuervo negro en el lugar s en el instante t ». De esto no puede salir ninguna base para la ciencia, sino sólo sistemas consistentes sin fin. Por otra parte, si los protocolos se limitan a oraciones que expresan *creencias*, entonces *sí* que importará quién

²¹ Otto Neurath, «Protocol Sentences», LP 205 [211].

²² *Ibid.*, 207 [213].

es el que arregla la máquina y a quién pertenecen los protocolos que se introducen y se extraen, ya que cada persona soportará las oraciones que acepta (sean protocolos o no) de acuerdo con la fuerza de sus creencias —es esto lo que quiere decir que son *sus* creencias—. Una les asignará un peso a los protocolos de otras personas en la medida en que piense que son verdaderos, que es justo lo que mantenía Schlick.

No me detendría tanto en los conocidos problemas epistemológicos que acosan a Schlick y Neurath y a sus seguidores si pensase que la filosofía los ha rechazado ahora o los ha resuelto. Por el contrario, pienso que los miembros del Círculo de Viena y sus amigos enfatizaron de un modo particularmente útil un problema central de la epistemología que no ha sido resuelto, aunque lo hicieran en parte a pesar de ellos mismos. Y pienso que podemos hallar un buen número de ideas e intuiciones en los escritos que he estado examinando que apuntan en la dirección de una nueva manera de ver el viejo problema.

El problema central en cuestión puede ser enunciado como un dilema en la teoría del conocimiento. Cada persona tiene una red compleja de creencias. El conocimiento requiere como mínimo estas dos cosas: que algunas de estas creencias sean verdaderas en relación con el mundo público y que cada persona posea razones adecuadas para sostener esas creencias. Estoy muy bien dispuesto a suponer que se satisface la primera condición; el problema difícil concierne a la segunda cuestión, el modo en el que el sistema de creencias se relaciona con el mundo, no meramente de manera semántica, sino epistemológicamente. El dilema se origina porque si tomamos como vínculo conector algo autoevidente (como los enunciados o acaecimientos de observación de Schlick), se trata de algo tan privado que carece de conexión alguna con las oraciones del lenguaje público, sólo las cuales pueden expresar afirmaciones científicas o incluso meramente objetivas. Pero si comenzamos con oraciones o creencias que ya pertenezcan a un lenguaje público (o que puedan expresarse en él), no encontraremos ninguna manera inteligible de basar todo ello en algo que sea autoevidente (éste es el problema de Neurath). En resumen, los fundamentos del conocimiento deben

ser subjetivos y objetivos al mismo tiempo, debe haber seguridad acerca de ellos y sin embargo debe podérselos cuestionar.

Los problemas que estoy repasando pertenecen, como todos sabemos, a los fundamentos de la epistemología, y de una forma u otra esos problemas son antiguos. Los positivistas lógicos —como se siente enseguida en sus escritos— reaccionaron con impaciencia ante estos problemas, que sentían que estaban en la frontera de lo carente de sentido, o que debían ser resueltos mediante «convenciones» o «decisiones». Esta actitud nos parece ahora, al releer a estos osados clásicos, despectiva. Pero la precipitada convicción de que una línea clara y correcta se abriría de una manera u otra a la vista de todo ese entusiasmo y potencia intelectual produjo algunas pistas profundamente nuevas y valiosas. Lo que en la superficie pueden parecer ahora intentos ingenuos y fallidos contenía intuiciones profundamente sugerentes de ideas radicales nuevas.

Dada la tendencia de los positivistas a permanecer dentro del vocabulario de las ideas y las hipótesis intersubjetivas, no es sorprendente que Neurath, Schlick, Hempel y Carnap flirtearan con alguna forma de teoría epistemológica coherentista. Pero las teorías coherentistas se han visto siempre complicadas por no distinguir entre teorías coherentistas de la verdad y teorías coherentistas del conocimiento. Esto es bastante natural, puesto que si el conocimiento que versa sobre lo verdadero no exige más que coherencia en las creencias, ¿cómo puede la verdad necesitar algo más que un conjunto coherente de proposiciones? De este modo se le invita a uno a descartar la diferencia entre la coherencia de creencias y la coherencia de oraciones (o enunciados, o proposiciones) para someter a prueba el conocimiento y la verdad. Pero aunque encontramos esta confusión en los escritos de Schlick y Neurath, también hallamos momentos en que la distinción se traza con claridad. Neurath, recordémoslo, tenía la esperanza de destruir cualquier aura de subjetividad en los enunciados protocolares al insistir que aunque pueda ser un «accidente histórico» que una persona esté más inclinada a aceptar sus propios protocolos que los de algún otro, de hecho ambos han

de ser aceptados al mismo nivel. Esto le conduce a la imagen de la máquina clasificadora impersonal en la cual se introducen las oraciones protocolares. Al proponer esta idea, Neurath nunca indica que lo único que haya de contar sean las oraciones que alguien considera verdaderas. Es este hecho el que hizo que Schlick explotara: «El asombroso error de la "teoría coherentista" sólo puede explicarlo el hecho de que sus defensores y expositores estaban pensando únicamente en los enunciados tal como éstos se dan realmente en la ciencia»²³. A lo que se refiere Schlick aquí es a tal como se presentan como *aseveraciones*. Sin embargo, a pesar de mucho de lo que encontramos en el artículo de Neurath acerca de las oraciones protocolares, ¿puede acusársele realmente de haber olvidado las diferencias entre la coherencia de un conjunto cualquiera de oraciones y la coherencia de un conjunto de oraciones que se tiene por verdadero? La famosa metáfora del barco que debe ser reconstruido en alta mar pieza por pieza demuestra que era consciente de que lo que está en juego son creencias, no meras oraciones, ya que si estuviéramos tratando de un conjunto cualquiera de oraciones nada nos impediría ponerlas a todas ellas simultáneamente en dique seco. Y Hempel, como mencioné anteriormente, deja clara la distinción.

Dado que las objeciones usuales a las teorías coherentistas del conocimiento proceden en paralelo con las objeciones usuales a las teorías coherentistas de la verdad, no es evidente de forma inmediata por qué es tan importante distinguir entre ellas. Pero, desde luego, las creencias no son arbitrarias, ni histórica ni causalmente; incluso si nuestras *razones* para nuestras creencias son siempre otras creencias, algunas veces las *causas* se sitúan en otra parte. Algo de la importancia de esta observación debe, como veremos, atribuírsele al positivismo lógico.

Los positivistas lógicos preferían hablar de oraciones o «enunciados» en lugar de creencias, y podemos fácilmente hacer el cambio mientras recordemos que las oraciones que se corresponden con creencias son: (1) oraciones que alguien

²³ Moritz Schlick, «The Foundations of Knowledge», LP 215 [221].

considera verdaderas, y (2) oraciones que tienen una interpretación. Algún otro puede saber lo que creo si sabe qué oraciones considero verdaderas y lo que esas oraciones significan. Voy a examinar algunas de las ideas de los positivistas lógicos a la luz de estas simples consideraciones.

Tenemos la idea de Schlick de que las oraciones observacionales se entienden en el acto de captar su verdad. Esto puede muy bien parecer extremo u oscuro, pero se relaciona con la doctrina correcta de que un intérprete se ve constreñido a tomar las atribuciones de actitudes de primera persona del presente como presuntamente verdaderas. Si un hablante considera verdaderas oraciones (del castellano) como «Creo que veo ahora un cuervo negro», se requiere de un intérprete que asigne a priori una elevada probabilidad a su verdad. Esto significa lo siguiente: siempre que sea posible, intérpreten esas oraciones de modo que resulten verdaderas.

Hempel comenta el hecho de que los enunciados protocolares «producidos» por personas diferentes podrían no admitir la construcción de un sistema único de enunciados científicos. Añade: «pero afortunadamente esta posibilidad no se da; de hecho, la mayor parte con mucho de los científicos llegarán a un acuerdo más pronto o más tarde, de modo que, como hecho empírico, un sistema de enunciados coherentes y de teorías en perpetuo crecimiento y expansión resulta a partir de sus enunciados protocolares»²⁴. De nuevo debemos suponer que los enunciados protocolares no son cualesquiera oraciones que se escriban o se profieran; deben ser oraciones cuyos hablantes creen que son verdaderas, o al menos que un oyente cree que el hablante las considera verdaderas. Pero realmente es extraño considerar como meramente «afortunado» que haya un elevado grado de consenso; y ¿por qué habríamos de esperar que el acuerdo se incremente con el tiempo?

Schlick tiene un análisis todavía más sorprendente de la posibilidad de que alguien pudiera descubrir que todas sus observaciones no apoyan en modo alguno las aseveraciones

²⁴ Carl Hempel, «On the Logical Positivist's Theory of Truth», 57.

sobre el mundo que otras personas han hecho. Dice que en estas circunstancias uno no sacrificaría simplemente los propios enunciados protocolares, al contrario de lo que haría la máquina de protocolos de Neurath. En lugar de eso, uno se atendería a un

sistema de conocimiento en el cual las propias observaciones se ajustaran sin mutilación. Y yo puedo siempre construir tal sistema. Sólo es necesario ver a los demás como locos soñadores en cuya locura hay un método notable, o —para expresarlo más objetivamente— diría que los otros viven en un mundo distinto del mío... En cualquier caso, no importa cuál sea la imagen del mundo que yo construya, siempre sometería a prueba su verdad en los términos de mi propia experiencia²⁵.

Admitir esto es algo extraordinario para quien ha objetado que una teoría coherentista nos deja con un relativismo inaceptable.

Sin embargo, en otro lugar encontramos en Schlick una manera bastante distinta de ver la posibilidad de un desacuerdo a gran escala. Señala un contraste básico entre un desacuerdo acerca de si dos trozos de papel son del mismo color y un desacuerdo acerca de qué color son los dos trozos. Con respecto al primero dice que «en virtud del uso lingüístico la proposición expresa justo esa vivencia» —esto es, la vivencia de la identidad—. Pero en el caso del color no hay manera objetiva —es decir, no hay manera— de decir si usted y yo experimentamos el mismo color. Schlick dice que incluso si todos los juicios que usted exprese sobre el color coinciden completamente con los míos, yo no puedo inferir a partir de aquí que usted experimenta la misma cualidad. Mientras el orden interno de sus vivencias coincida con el mío, nos entenderemos perfectamente el uno al otro²⁶.

Al final Schlick parece rechazar la idea de que las vivencias pueden ser distintas de un modo que es imposible detectar. «... el enunciado de que distintos individuos tienen la misma

²⁵ Moritz Schlick, «The Foundation of Knowledge», *LP* 219 [225].

²⁶ Parfraseo aquí a Schlick, «Positivism and Reslism», *LP* 93 [99].

experiencia tiene su único sentido verificable en el hecho de que todas las aseveraciones del mismo... muestran ciertos acuerdos... el enunciado no *significa* nada más que esto»²⁷. No es fácil determinar por este pasaje si Schlick piensa que las vivencias podrían ser cualitativamente distintas aunque no pudiéramos, por principio, averiguarlo o afirmarlo con significado, o si piensa que no podría darse una situación así. La afirmación radical —que no es imposible extraer de la lectura del ataque de Schlick a la teoría coherentista— es que el acuerdo interpersonal, y por tanto la objetividad, están imbricadas en el modo en que determinamos los significados de las preferencias de otras personas y por tanto los contenidos de sus creencias.

He mencionado anteriormente la observación de Hempel de que «afortunadamente» los enunciados protocolares de diferentes personas permiten la construcción de un sistema único de ciencia. A esto añade que Carnap «quizá nos ha proporcionado una posibilidad de explicar este afortunado hecho»²⁸. Esta posible explicación estriba en el hecho de que «a los científicos jóvenes se les condiciona» a producir enunciados protocolares verdaderos, y añade: «Quizá el hecho del condicionamiento general y congruente de los científicos pueda explicar en un cierto grado el hecho de un único sistema de ciencia»²⁹.

Sería algo misterioso que a las personas se les enseñara lo que significan diversas oraciones y *luego* se las condicionara a «producir» las verdaderas; esto equivaldría a enseñarles, por un lado, cómo ser mejores observadores y, por el otro, cómo ser honestos. Pero la situación puede verse más bien como que de lo que se trata es de condicionar personas, como ciertamente hacemos, a considerar verdaderas ciertas oraciones cuando se dan determinadas condiciones públicamente observables, y a fijar la interpretación de las preferencias de las oraciones de acuerdo con el éxito de ese condicionamiento.

²⁷ *Ibid.*, 93 [100].

²⁸ Carl Hempel, «On the Logical Positivist's Theory of Truth», 37.

²⁹ *Ibid.*, 58.

Esto explicaría de un modo natural el acuerdo interpersonal sobre los rasgos principales del entorno.

Carnap parece claramente tomar este derrotero en un determinado momento. En «Psychology in Physical Language» (escrito en 1932) Carnap rechaza de plano la idea de Neurath de que uno deba o pueda tratar las oraciones protocolares de otro igual que las mías.

Hablando en general [escribe], las oraciones protocolares habladas, escritas o impresas de un psicólogo, cuando se basan en la llamada introspección, han de ser interpretadas por el lector, y figurar así en una ciencia inter-subjetiva, *no principalmente como oraciones científicas, sino como hechos científicos*. La confusión epistemológica de la psicología contemporánea procede, en gran parte, de la confusión de hechos en forma de oraciones con las oraciones mismas consideradas como partes de la ciencia³⁰.

Las conclusiones que se nos permite extraer del hecho de que alguna otra persona profiere una oración no son consecuencias deductivas que manen de la oración ya interpretada, sino más bien el tipo de conclusión que podemos extraer de la observación de los movimientos de un voltímetro o de los movimientos de una gota de lluvia. La cuestión no es que las oraciones que profieren los demás no signifiquen nada, sino que no podemos dar por sentado que conocemos de antemano *lo que* significan; se requiere explícitamente una interpretación o bien es implícitamente asumida.

Creo que siguiendo esta pista, junto con otras varias sugerencias que se extraen de pasajes que he estado citando, podemos descubrir los perfiles de una concepción correcta de los fundamentos del conocimiento empírico, una concepción que reconcilia la teoría coherentista de Neurath con la insistencia de Schlick en un vínculo básico con la experiencia y la observación.

En adelante, aunque estaré apoyándome de muchas y obvias maneras en ideas de Schlick, Neurath, Hempel y Carnap,

³⁰ Rudolph Carnap, «Psychology in Physical Language», LP 195 [200-1].

estaré enunciando mi propia posición. Se trata de una posición profundamente influida por Quine, aunque no sea la posición de Quine.

Neurath estaba en lo cierto al rechazar la inteligibilidad de comparar oraciones o creencias con la realidad. Experimentamos y observamos, pero esto no es «comparar» en ningún sentido que no sea metafórico, pues nuestra experimentación no produce ningún fruto epistemológico excepto en que *causa* que le añadamos algo a nuestras creencias, nos apeguemos a ellas o las abandonemos. Esta relación causal no puede ser una relación de *confirmación* o *desconfirmación*, dado que la causa no es una proposición o una creencia, sino sólo un acaecimiento que se da en el mundo o en nuestro aparato sensorial. Tampoco pueden esos acaecimientos por sí mismos considerarse como evidencias, a menos, por supuesto, que causen en nosotros alguna creencia, y entonces es la creencia la que propiamente ha de considerarse como la evidencia, no el acaecimiento.

Neurath estaba también en lo cierto al decir que, dada esta situación, podemos también perfectamente admitir que los protocolos, como cualesquiera otras proposiciones de la ciencia o del sentido común, pueden estar equivocados; estamos preparados a hacer apaños cuando hacer apaños es lo mejor que se puede hacer. Como observó Hempel, no se asigna ninguna prioridad epistemológica a los protocolos —son como todo lo demás—. Todo esto constituye, por supuesto, la pista que luego habría de explotar Quine al argumentar contra la distinción analítico-sintética.

Acabamos, por tanto —como enfatizó Neurath—, en una situación en la que nuestras únicas evidencias en favor de una creencia la constituyen otras creencias; no es ésta únicamente la situación *lógica*, sino también la situación pragmática. Y puesto que ninguna creencia es autoevidente, ningún conjunto de creencias puede suministrar una base cierta para el resto. ¿Cómo podemos entonces escapar a la objeción de Schlick de que esto convierte a «cualquier cuento de hadas en algo tan verdadero como un informe histórico»? Schlick concluye: «De este modo se muestra que la teoría coherentista es lógicamente imposible... ya que por medio de ella yo puedo

llegar a cualquier número de sistemas consistentes de enunciados que son incompatibles entre sí»³¹. No está nada claro qué es lo que significa decir que yo podría «llegar» a diversos sistemas, puesto que no me invento mis creencias; la mayoría de ellas no son voluntarias. Con todo, parecería aún que el punto crucial de la crítica sigue en pie en la forma de reto, un reto a que se diga cuál es la razón que yo tengo para considerar verdaderas la mayor parte de mis creencias.

La clave de la respuesta estriba, según creo, en generalizar las dos sugerencias de Carnap, de que estamos condicionados para producir (considerar verdaderas) oraciones específicas cuando se dan determinadas condiciones, y que no podemos utilizar los enunciados de otras personas como evidencias hasta haberlos interpretado. Carnap dijo esto sólo para las oraciones protocolares, pero lo mismo se debería decir para todo el lenguaje.

Como enfatizó Neurath, el lenguaje es por su naturaleza intersubjetivo; lo que las palabras de otra persona significan en una ocasión determinada es siempre algo que podemos en principio averiguar a partir de indicios públicos. Consideremos cómo descubrimos lo que significa alguna oración sencilla, por ejemplo, «Hay una mesa» o «Aquí hay un trozo de papel verde». Nuestra evidencia básica es que el hablante se ve causalmente llevado a asentir (no sólo en esa ocasión, sino en general) a estas oraciones por la presencia de mesas o trozos de papel verde, mientras que la ausencia de estos objetos le lleva causalmente a disentir (en general) de las mismas oraciones. No estoy concibiendo el asentimiento y disentimiento como actos de habla públicos, sino como actitudes hacia las oraciones que se revelan a veces en el habla y a veces de otros modos. Mi observación principal es que nuestra metodología básica para interpretar las palabras de los demás lleva necesariamente a que la mayor parte del tiempo las oraciones más simples que los hablantes tienen por verdaderas *son* verdaderas. No es el *hablante* el que debe llevar a cabo la hazaña imposible de comparar su creencia con la realidad; es el *intérprete*

³¹ Moritz Schlick, «The Foundation of Knowledge», LP 216 [222].

te el que debe tener en cuenta la interacción causal entre el mundo y el hablante para averiguar lo que quiere decir el hablante, y por tanto lo que el hablante cree. Un hablante no puede hacer nada mejor que procurar que su sistema de creencias sea coherente, ajustando ese sistema tan racionalmente como pueda a medida que nuevas creencias se le imponen. Pero no hay ninguna necesidad de temer que esas creencias puedan ser sólo un cuento de hadas, ya que se entiende correctamente que las oraciones que expresan esas creencias, y las creencias mismas, versan acerca de las cosas públicas y eventos que causan tales creencias, y por lo tanto deben ser en su mayor parte verídicas. Cada persona sabe esto, puesto que conoce la naturaleza del habla y de las creencias. Por descontado, esto no le dice *cuáles* de sus creencias y oraciones son verdaderas, pero le asegura que su imagen de conjunto del mundo entorno es como la que tienen otras personas y que es correcta a grandes rasgos.

Neurath, Carnap y Hempel están en lo cierto, según creo, al abandonar la búsqueda de un tipo fundamental de evidencia sobre el que basar nuestro conocimiento del mundo. No podemos disponer de nada de eso, ni tampoco hace ninguna falta. Lo que quizá no llegaron a apreciar es *por qué* no hace ninguna falta. No es necesario porque las relaciones causales entre nuestras creencias y el habla, de una parte, y el mundo, de la otra, proporcionan la interpretación de nuestra lengua y de nuestras creencias. En este sentido bastante especial, la «experiencia» es la fuente de todo conocimiento. Pero éste es un sentido que no nos anima a encontrar un puente mental o inferencial entre los acaecimientos externos y las creencias ordinarias. De acuerdo en que hay un puente: un puente causal que implica a los órganos sensoriales. El error está, como vio Neurath, en tratar de transformar este puente causal en un puente epistemológico que tuviera como imposibles arcadas a los datos sensoriales, lo dado sin interpretar, u oraciones que no pueden ser escritas.

Hay por supuesto algunas creencias que poseen un alto grado de certeza, y en algunos casos su contenido crea una presunción en favor de su verdad. Éstas son creencias acerca de nuestras actitudes proposicionales presentes. Pero la relativa

certeza de estas creencias no las hace apropiadas para constituir el fundamento del conocimiento empírico. Éste procede más bien de la naturaleza de la interpretación. Como intérpretes tenemos que tratar las autoadscripciones de creencias, dudas, deseos y cosas parecidas como privilegiadas; éste es un paso esencial en la interpretación de lo que queda en lo que la persona dice y piensa. Los fundamentos de la interpretación no son los fundamentos del conocimiento, aunque reconocer la naturaleza de la interpretación puede conducir a apreciar la naturaleza esencialmente verídica de la creencia.

12

Epistemología y verdad

Todo el mundo está de acuerdo en que lo que se sabe ha de ser verdadero, pero más allá de esto hay poco acuerdo acerca del estatuto epistémico de la verdad. Muchos filósofos han sostenido en décadas recientes que la verdad es un concepto epistémico; incluso cuando no han mantenido explícitamente esta tesis, han sostenido otras cosas que la implicaban. Las teorías coherentistas de la verdad se ven generalmente impulsadas por un motor epistémico, como lo están las caracterizaciones pragmáticas de la verdad, el antirrealismo de Dummett y Crispin Wright, la idea de Peirce de que la verdad es lo que acabe finalmente creyendo la ciencia, la afirmación de Richard Boyd de que la verdad es lo que explica el éxito de la ciencia y el realismo interno de Hilary Putnam. También Quine, al menos algunas veces, ha mantenido que la verdad es interna a la teoría del mundo y por ello dependiente en esa medida de nuestra postura epistemológica. El relativismo acerca de la verdad quizá sea siempre un síntoma de infección por el virus epistemológico; en todo caso esto es lo que sucede con Quine, Goodman y Putnam.

La idea intuitiva de que la verdad, aparte de algunos casos especiales, es enteramente independiente de nuestras creencias se opone aparentemente a todas esas opiniones; esto se formula a veces diciendo que nuestras creencias podrían ser exactamente como son y sin embargo la realidad —y por tanto la verdad acerca de la realidad— ser muy diferente. Según esta intuición, la verdad es «radicalmente no-epistémica» (así

es como Putnam caracteriza el «realismo transcendental»), o «transcendente a las evidencias» (Michael Dummett).

Si fuéramos a buscar una descripción breve de estas dos concepciones de la verdad, muy bien podrían ocurrírse nos las palabras «subjetivo» y «objetivo»; la afirmación de un vínculo esencial con la epistemología introduce una dependencia de la verdad con respecto a lo que puede ser verificado por criaturas racionales finitas, mientras que la negación de tal dependencia hace a la verdad objetiva. Posiblemente las palabras «inmanente» y «transcendente» sugieren también la misma división básica. Y puesto que la cuestión de la verdad no puede separarse de las cuestiones acerca de la naturaleza de la realidad, los términos «antirrealismo» y «realismo» se han utilizado a menudo para hacer una distinción estrechamente relacionada. Por razones que se irán aclarando a medida que vaya avanzando, yo prefiero «subjetivo» y «objetivo»; pero me doy cuenta de que algunas de las partes de la disputa encontrarán tendenciosa esta terminología.

En la presente charla esbozo un modo de enfocar el concepto de verdad que creo que socava la disputa entre las dos concepciones de la verdad. No pretendo reconciliar estas dos concepciones, sino cuestionar su claridad y al hacerlo así dejarnos en libertad para abandonar la búsqueda de la «correcta» teoría de la verdad. El hecho de que ambas concepciones induzcan al escepticismo sugiere como mínimo que las dos están equivocadas de modo fundamental, aunque respondan, sin duda, a poderosas intuiciones. Las teorías subjetivas son escépticas al modo en que lo son el idealismo y muchas versiones del empirismo; son escépticas no porque hagan de la realidad algo incognoscible, sino porque reducen la realidad a mucho menos de lo que creemos que existe. Por otra parte, las teorías objetivas parecen poner en duda no sólo nuestro conocimiento de aquello que es «transcendente a las evidencias», sino todo el resto de lo que pensamos que sabemos, pues tales teorías niegan que haya vínculo alguno entre la creencia y la verdad.

Vamos a considerar en primer lugar las definiciones de verdad en el estilo de Tarski, con la vista puesta en tomar una decisión acerca de si son objetivas o subjetivas, o ninguna de las

dos cosas¹. La primera cosa que debe quedar clara es que Tarski no definió ni mostró cómo definir el concepto de verdad; lo que hizo fue definir la verdad para una lengua o un lenguaje² específico e indicar cómo definir la verdad para otras lenguajes o lenguajes de un tipo específico. La idea básica de Tarski fue utilizar el hecho aparentemente trivial de que las oraciones de la forma «“La nieve es blanca” es verdadera en L si y sólo si la nieve es blanca» han de ser verdaderas si la oración citada es una oración del lenguaje utilizado para enunciar esa perogrullada. Llamemos oraciones-T a estas perogrulladas. Tarski señaló que una definición de «es verdadera en L» que implicase una oración-T para cada oración de L seleccionaría exactamente las oraciones verdaderas de L; y señaló además que la definición sería igual de correcta si el predicado de verdad estuviera en una lenguaje que no contuviera a L y las oraciones-T implicadas tuvieran a la derecha del bicondicional una traducción de la oración citada en lugar de la oración misma. El trabajo difícil consistía en proporcionar una definición correcta, dadas las restricciones impuestas por Tarski. (No hay nada en este resumen que sea totalmente exacto, pero las inexactitudes son irrelevantes para la discusión.)

Está claro que ninguna definición en el estilo de Tarski es una definición de la verdad; cada una de esas definiciones es una definición de la verdad *en una lengua o lenguaje determinado*. Si pensamos que es el mismo concepto de verdad el que está involucrado en cada una de las definiciones de Tarski, esto sólo puede ser debido a que tenemos una comprensión de ese concepto que es independiente del trabajo de Tarski³. Efectivamente, esto es algo completamente claro; la condición de Tarski acerca de las definiciones satisfactorias de la verdad para lenguajes particulares descansa en el reconocimiento de que las oraciones-T son obviamente verdaderas, y este reconocimiento depende de nuestra comprensión previa

¹ Alfred Tarski, «The Concept of Truth in Formalized Languages».

² Véase la nota 3 del Ensayo 3.

³ Por lo que sé, esta observación la hizo por primera vez Max Black, «The Semantic Definition of Truth». Véase también Michael Dummett, «Truth», en *Truth and Other Enigmas*.

del concepto de verdad. La trivialidad de las oraciones-T es una garantía de su verdad y esta garantía la traspasa el trabajo de Tarski a las «definiciones parciales» de la verdad que éste proporciona. Las oraciones-T son trivalmente verdaderas sólo si sabemos que la oración descrita es idéntica a, o es una traducción de, la oración que da sus condiciones de verdad; pero esto es suficiente para mostrar el sentido en que tenemos una comprensión firme del concepto general de verdad.

A las definiciones de Tarski se las denomina a veces como «meramente traductoras» o «desentrecomilladoras». Son realmente desentrecomilladoras en el caso especial en el que se define la verdad para una lengua o lenguaje en una lengua o lenguaje que lo contenga; en tales casos, las oraciones-T muestran la manera de eliminar las comillas obteniendo así una oración que puede demostrarse que tiene el mismo valor de verdad. Pero sólo se sabe que este rasgo está presente cuando se sabe que la lengua o lenguaje objeto está contenido en el metalenguaje. La idea de que tales definiciones son traductoras descansa en un error. Si se sabe cómo traducir un lenguaje o lengua L a la propia lengua, no se sigue de ello que se pueda producir automáticamente una definición de la verdad. La traducción relaciona las lenguas o los lenguajes unos con otros; la verdad, incluso en las ediciones limitadas de Tarski, relaciona una lengua o lenguaje con el mundo.

Se llega a las definiciones de Tarski en varios pasos. Se define la noción de expresión bien-formada (en particular una oración o sentencia cerrada); se da luego una definición recursiva de una relación de satisfacción (la satisfacción es una versión muy generalizada de la referencia); se define la verdad sobre la base de los conceptos de oración y satisfacción, y las definiciones recursivas se convierten en definiciones explícitas. Para los propósitos presentes es mejor olvidarse del último paso. Lo que entonces nos queda no es propiamente una definición en absoluto, sino una caracterización recursiva de un concepto no definido, el de satisfacción; y dado que la verdad se define en términos de satisfacción, que es un concepto semántico, no se puede decir ya que la verdad haya sido definida en términos no-semánticos, ni siquiera para lenguas o lenguajes particulares. Miremos entonces el resultado como

una teoría axiomatizada que tiene sin definir las nociones semánticas básicas.

Cuál sea el concepto semántico que tomemos como básico es algo que podemos elegir. Como mostró Tarski, la verdad puede definirse sobre la base de la satisfacción; pero la satisfacción también puede considerarse como aquel concepto, sea el que fuere, que produzca una explicación correcta de la verdad. La obra de Tarski, a pesar de las apariencias, es en último término neutra respecto a la cuestión de cuál es el concepto semántico básico. Como hemos visto, él supone una comprensión del concepto de verdad, y muestra cómo puede implementarse en detalle esta intuición para lenguajes o lenguas particulares. Esa implementación requiere la introducción de la relación de satisfacción, que relaciona palabras y cosas (un poco indirectamente). El relato de la verdad dibuja la pauta de las formas lógicas de una lengua o lenguaje —la gramática adecuadamente concebida— y por tanto la pauta de las dependencias lógicas. No hay modo de relatar esto (en una lengua o lenguaje de una riqueza apreciable) sin asignar papeles semánticos a las partes de las oraciones.

Existe una fuerte tradición que ve las cosas de modo diferente. De acuerdo con esta tradición no podríamos nunca llegar a entender las oraciones en su serie infinita o indefinidamente grande a menos que entendamos las palabras que las forman. Por lo tanto, las propiedades semánticas de las palabras deben aprenderse antes de que entendamos las oraciones, y las propiedades semánticas de las palabras tienen una prioridad conceptual porque son ellas las que explican las propiedades semánticas—sobre todo las condiciones de verdad— de las oraciones. Creo que esta línea argumental, que comienza con una perogrullada, acaba con una conclusión falsa; de modo que hay algo que debe estar equivocado en ese razonamiento. El error consiste en confundir el orden apropiado de explicación cuando la teoría está ya formada con la explicación de por qué es correcta la teoría. La teoría es correcta porque produce oraciones-T correctas; se pone a prueba su corrección enfrentándola a nuestra comprensión del concepto de verdad aplicado a oraciones. Puesto que las oraciones T no dicen nada acerca de referencia, ni de satisfac-

ción, ni de expresiones que no son oraciones, el examen de la corrección de la teoría es independiente de nuestras intuiciones relativas a estos conceptos. Sin embargo, una vez que tenemos la teoría, podemos explicar la verdad de las oraciones sobre la base de su estructura y de las propiedades semánticas de sus partes. La analogía con las teorías de la ciencia es impresionante: para organizar y explicar lo que observamos directamente postulamos objetos y fuerzas no observados o indirectamente observados, pero la teoría se pone a prueba por lo que se observa directamente.

La perspectiva acerca del lenguaje y la verdad que hemos adquirido es la siguiente; lo que está más abierto a la observación son las oraciones y sus usos, y la verdad es el concepto semántico que entendemos mejor. La referencia y las nociones semánticas relacionadas, como la satisfacción, son teóricas en comparación con aquél (como lo son las nociones de palabra, predicado, conectiva oracional y demás); no se suscita cuestión alguna sobre su «corrección» que vaya más allá de si producen o no una explicación satisfactoria de las condiciones de verdad de las oraciones y de las relaciones entre oraciones.

Estas observaciones nos invitan a concentrar nuestra atención en la importancia capital que tiene el concepto de verdad para la comprensión del lenguaje; es nuestra comprensión implícita de los contenidos de este concepto lo que nos permite dar un sentido a la cuestión de si una teoría de la verdad para un lenguaje o lengua es correcta. No hay ninguna razón para buscar una explicación de alguna relación referencial que sea anterior o independiente. La otra consecuencia principal de estas reflexiones es que proporcionan una oportunidad para decir con notable precisión lo que como explicación de la verdad falta en una teoría de la verdad al estilo de Tarski.

Lo que falta es la conexión con los usuarios de una lengua o lenguaje. Nada en el mundo podría considerarse como una oración, y el concepto de verdad no se aplicaría por tanto a nada si no hubiera criaturas que usaran oraciones profiriendo o inscribiendo ejemplares concretos de ellas. Una explicación completa de la verdad debe, por tanto, ponerla en relación con

la interacción lingüística real. Dicho con mayor precisión: una teoría de la verdad al estilo de Tarski es verdadera para una lengua o lenguaje dado sólo si las oraciones de esa lengua o lenguaje tienen un significado anterior a la construcción de la teoría (de lo contrario la teoría no es una teoría en el sentido corriente, sino una descripción de un lenguaje posible). O, volviendo a la forma de definición propugnada por Tarski: si se puede plantear la cuestión de si la definición es correcta para una determinada lengua o lenguaje, esta lengua o lenguaje deberá tener una vida independiente de la definición (de lo contrario la definición es meramente estipulativa: especifica una lengua o lenguaje, pero no es verdadera acerca de éstos). De modo que para plantear el tema de la corrección la teoría debe verse como verdadera o falsa empíricamente. La única excepción es el caso que consideramos antes, en el que se supone que el lenguaje objeto se contiene en el metalenguaje; en este caso el examen es formal, no está en juego ninguna cuestión empírica. Se suscita, sin embargo, el problema de cuándo está justificado ese supuesto. Si tengo una teoría de la verdad para su lengua, el supuesto sería obviamente correcto sólo si fuera evidente que usted y yo queremos decir lo mismo mediante todas las oraciones que decimos o podríamos decir. Dudo de que esto sea nunca evidente, pero en cualquier caso el supuesto es, naturalmente, empírico. Se plantea incluso la duda de si hay garantías de que un examen formal de la corrección sea todo lo que se necesita para decir si mi teoría de la verdad, formulada con mis propias palabras, es correcta para mi propia lengua; ésta es, sin embargo, una cuestión que no discutiré aquí.

Si supiéramos lo que hace verdadera a una teoría de la verdad para un hablante o grupo de hablantes, podríamos decir plausiblemente que entendemos el concepto de verdad; y si pudiéramos decir lo que hace verdadera a tal teoría podríamos dar una explicación explícita, si bien no una definición, de verdad. Debemos suponer que las *evidencias* últimas —por contraposición a un criterio— de la corrección de una teoría de la verdad residen en hechos sobre la manera en que los hablantes usan la lengua que está a nuestra disposición. Cuando digo a nuestra disposición, quiero decir que son pública-

mente asequibles —asequibles no sólo en principio, sino asequibles de hecho a cualquiera que pueda entender al hablante o hablantes de esa lengua—. Puesto que todos nosotros entendemos a algunos hablantes de algunas lenguas, debemos de poseer todos nosotros evidencias adecuadas para atribuir condiciones de verdad a las proferencias de algunos hablantes; todos nosotros tenemos, por tanto, una comprensión competente del concepto de verdad en su aplicación al comportamiento verbal de los demás.

¿Hemos resuelto ya el problema de si la verdad es radicalmente no epistémica, como afirman los objetivistas, o fundamentalmente epistémica, como declaran los subjetivistas? Podría parecerlo así, puesto que hemos seguido una línea argumentativa que conduce a la conclusión de que es el modo en que se usa una lengua o lenguaje lo que decide si una teoría de la verdad para esa lengua o lenguaje es correcta.

Pero en realidad el asunto no está resuelto, pues los objetivistas pueden sostener que la cuestión de si la teoría es verdadera para una lengua dada o para un grupo de hablantes es, ciertamente, empírica, pero únicamente porque la cuestión de qué es lo que significan las palabras es empírica; la cuestión de la verdad —puede afirmarse— sigue sin respuesta, ni por parte de la teoría ni de ningún otro modo.

¿Contiene la teoría misma la respuesta? Sí, si una teoría de la verdad al estilo de Tarski es una teoría de la correspondencia, pues entonces la forma de la teoría define, efectivamente, que la verdad es correspondencia con la realidad —la forma clásica del objetivismo con respecto a la verdad—. Lamento decir que yo mismo sostuve en el pasado que las teorías de la clase que Tarski nos enseñó a producir eran teorías de la correspondencia de un determinado tipo. Dije esto por la razón de que no hay ningún modo de dar una teoría así sin introducir un concepto como el de referencia o satisfacción que relaciona expresiones con objetos del mundo⁴.

⁴ Tarski parece haber visto sus definiciones de verdad como una realización o concreción de la idea de correspondencia. Acepté neciamente este término en «True to the Facts», el Ensayo 3 de *Inquiries into Truth and Interpretation*.

Ahora creo que cometí un error al denominar teorías de la correspondencia a esas teorías. Voy a explicar por qué creo que fue un error. La objeción usual acerca de las teorías de la correspondencia es que no tiene sentido alguno afirmar que es de algún modo posible comparar las propias palabras o creencias con el mundo, puesto que tales intentos habrán de acabar siempre con la adquisición de más creencias. Esta objeción la expresó, por ejemplo, Otto Neurath, quien por esta razón adoptó una teoría coherentista de la verdad; Carl Hempel expresó la misma objeción, hablando de «fatal confrontación de los enunciados con los hechos»⁵. Richard Rorty ha insistido repetidamente —con Dewey— en que una concepción de la verdad como correspondencia convierte en inútil al concepto de verdad⁶. Yo mismo he dicho algo muy parecido⁷.

Pienso ahora que esta objeción contra las teorías de la verdad como correspondencia es un error. Una razón de por qué es un error es que depende de haberse puesto ya de parte de los subjetivistas; sería una objeción legítima tan sólo si la verdad fuera un concepto epistémico al modo en que los subjetivistas afirman que lo es. Si ésta fuera la única razón para rechazar las teorías de la correspondencia, el objetivista podría replicar simplemente que su posición no se ve afectada; él siempre mantuvo que la verdad era independiente de nuestras creencias o de nuestras capacidades para llegar a ella.

La objeción real a las teorías de la correspondencia es más sencilla; se trata de que no hay nada con lo que puedan corresponderse las oraciones. Esta observación la hizo hace mucho tiempo C. I. Lewis, al retar al partidario de la teoría de la correspondencia a *localizar* el hecho o parte de la realidad, o del mundo, con el que se corresponde una oración verdadera. Se pueden localizar objetos individuales si sucede que la oración correspondiente los nombra o describe, pero incluso

⁵ Véase Otto Neurath, «Protocol Sentences», y Carl Hempel, «On the Logical Positivist's Theory of Truth», 51.

⁶ Richard Rorty, *Consequences of Pragmatism*, introducción; también su «El pragmatismo, Davidson y la verdad».

⁷ Por ejemplo, en el Ensayo 10 de este volumen.

esa localización tiene sentido de manera relativa a un marco de referencia, por lo que el marco de referencia, según cabe presumir, tiene que ser incluido en aquello con lo que se corresponda una oración verdadera. Seguir esta trayectoria de pensamiento le llevó a Lewis a concluir que si las oraciones verdaderas se corresponden con alguna cosa, ésta ha de ser el universo como un todo, y, por tanto, todas las oraciones verdaderas se corresponden con la misma cosa. Como sabemos, Frege llegó independientemente a la misma conclusión mediante una línea de razonamiento muy parecida. El argumento de Frege, si Alonzo Church está en lo cierto, puede formalizarse: comenzando por el supuesto de que una oración verdadera no puede ponerse en correspondencia con algo distinto mediante la sustitución de términos singulares correferenciales, o mediante la sustitución de oraciones lógicamente equivalentes, se puede mostrar que si las oraciones verdaderas se correspondieran con alguna cosa, todas ellas se corresponderían con la misma. Pero esto significa trivializar por completo el concepto de correspondencia; la *relación* no tiene ningún interés si sólo hay una cosa con la que corresponderse, puesto que, como en cualquier caso así, esa relación puede perfectamente reducirse a un predicado monádico: «*x* se corresponde con el universo». Análogamente, «*x* se corresponde con (o nombra) lo Verdadero» o «*x* se corresponde con los hechos» puede leerse de manera menos engañosa como «*x* es verdadera». Peter Strawson hizo esta observación en su famosa discusión sobre la verdad con J. L. Austin. Señaló que parte de la oración puede corresponderse con parte del mundo (es decir, referirse a ellas), pero añade que no hay nada más en el mundo con lo que una oración pueda ponerse en relación. Correctamente pasa de aquí a decir que «aunque ciertamente decimos que un enunciado se corresponde (se ajusta a, es corroborado por, es conforme a) los hechos», esto no es más que «una forma de decir que es verdadero»⁸.

La objeción correcta contra las teorías de la correspondencia no es, por tanto, que la verdad no sea algo a lo que noso-

⁸ Peter Strawson, «Truth», 194-5.

tros los humanos podamos nunca aspirar legítimamente; la objeción real es más bien que tales teorías no proporcionan entidades con las que pueda decirse que se corresponden los vehículos de la verdad (tanto si consideramos que éstos son enunciados como oraciones o proferencias). Como dije una vez: «Nada, ninguna *cosa*, hace que los enunciados sean verdaderos.» Si esto es cierto, y estoy convencido de que lo es, deberíamos también poner en cuestión el popular supuesto de que a las oraciones, a sus ejemplares hablados, a entidades oracionales o a configuraciones de nuestros cerebros se les pueda con propiedad denominar «representaciones», puesto que no hay nada que ellas puedan representar. Si prescindimos de los hechos como entidades que hacen verdaderas a las oraciones, deberíamos al mismo tiempo prescindir de las representaciones, pues la legitimidad de cada una de esas cosas depende de la legitimidad de la otra.

He dicho que lamentaba haber dicho alguna vez que una teoría de la verdad al estilo de Tarski era una forma de teoría de la correspondencia, y lo lamento. Mi razón para decirlo no era que hubiera cometido el error de suponer que las oraciones o las proferencias de oraciones se correspondían con alguna cosa. Pero estaba todavía influenciado por la idea de que la verdad es objetiva; por la idea de que la verdad, y por tanto la realidad, son (excepto para casos especiales) independientes de lo que una persona cualquiera crea o pueda saber. Es así como anuncié mi posición como una variedad de realismo: realismo con respecto al «mundo externo», con respecto al significado y con respecto a la verdad.

Si escogí mal mis palabras, ello se debió a que sugieren una aprobación decidida de la posición o del supuesto de que hay una tesis positiva clara que puede ser aceptada. Aunque todo lo que estaba legitimado a sostener, todo lo que mi posición implicaba en realidad con respecto al realismo y la verdad era la declaración negativa de que el subjetivismo es falso. Si tiene algún contenido, la concepción objetiva de la verdad debe basarse en la correspondencia, la correspondencia aplicada a oraciones, creencias o proferencias, es decir, entidades con un carácter proposicional, y no puede hacerse que esa correspondencia resulte inteligible. En la medida en que el realismo no

es sino la versión ontológica de una teoría de la correspondencia, debo también rechazarlo. Todo lo que pretendía al decir que mi posición era una forma de realismo era rechazar el antirrealismo; lo que me preocupaba era rechazar la doctrina de que la realidad o la verdad dependan de nuestras capacidades epistémicas. De nuevo, sin embargo, debo precisar que no estoy ni aceptando ni rechazando el lema realista-objetivista de que lo real y lo verdadero son «independientes de nuestras creencias». El único sentido positivo evidente que podemos darle a esta frase que concuerda con las intenciones de quienes la valoran deriva de la idea de correspondencia, y esta idea debe ser rechazada. Pero rechazarla no quiere decir que no exista ninguna relación entre creencia y verdad; como hemos visto, alguna conexión debe haber si es que hemos de relacionar la verdad de las preferencias con el contexto humano de las mismas. La cuestión es cuál pueda ser esa conexión.

Las diversas formas de subjetivismo —es decir, de las concepciones que consideran que la verdad es un concepto epistémico— conectan de diversas maneras los pensamientos, deseos e intenciones humanos con la verdad, y no puedo pretender hacer justicia a todas esas concepciones. Lo mejor que puedo hacer es explicar por qué, a pesar de las diferencias, tiene sentido agrupar muchas de ellas en la misma clase, e indicar las razones para no estar satisfecho con las mismas.

He clasificado las teorías coherentistas como subjetivas y esto requiere una explicación. Una teoría coherentista pura de la verdad sostiene, supongo, que todas las oraciones de un conjunto consistente de oraciones son verdaderas. Dudo de que nadie nunca haya sostenido tal teoría, pues es obviamente falsa. Quienes propusieron teorías coherentistas, por ejemplo Neurath y Carnap (en un cierto momento), precisaron por lo general que eran conjuntos de *creencias* o de oraciones que se consideran verdaderas aquellas cuya consistencia era suficiente para hacer a éstas verdaderas; ésta es la razón por la que clasifico las teorías coherentistas con otras concepciones subjetivas; vinculan directamente la verdad con los contenidos de las creencias. Pero si no se añade algo más, esta concepción parece tan equivocada como sostuvo Schlick que lo

era (dijo de ella que era un «error asombroso»)⁹. La objeción más obvia es que son posibles muchos conjuntos consistentes de creencias que no son consistentes entre sí.

Existen teorías a las que no se concibe generalmente o con propiedad como teorías coherentistas que tienen un inconveniente muy parecido. Quine sostiene que la verdad de algunas oraciones, denominadas por él oraciones observacionales, está directamente vinculada a los inputs sensoriales; otras oraciones derivan su contenido empírico de sus conexiones con las oraciones observacionales y de sus relaciones lógicas recíprocas. La verdad de estas otras oraciones (llamémoslas oraciones teóricas) no depende más que de lo bien que sirvan para explicar o predecir las oraciones observacionales sobre la base de otras oraciones observacionales. Quine sostiene (de manera plausible) que podría haber dos estructuras teóricas con la misma capacidad para dar cuenta de todas las oraciones observacionales y tales que, sin embargo, ninguna de ellas pueda ser reducida a la otra (cada una de esas teorías contiene al menos un predicado que no puede definirse utilizando los recursos de la otra teoría). Quine ha adoptado en diferentes momentos dos maneras distintas de concebir esta situación. De acuerdo con una de las tesis, ambas teorías son verdaderas. Contra esto no tengo ninguna objeción. De acuerdo con la otra concepción, alguien que habla o piensa en un momento determinado funciona con una teoría, y para él la teoría que usa en ese momento es verdadera y la otra teoría es falsa. Si cambia a la otra teoría, entonces *ésta* es verdadera y la otra falsa. Ésta es al parecer la posición a la que se ha referido Quine cuando ha dicho que la verdad es «inmanente»¹⁰. No debería confundirse esta concepción de la inmanencia o relatividad de la verdad con el sentido pedestre en el que la verdad de las oraciones es relativa a la lengua o lenguaje en el que se

⁹ Moritz Schlick, «The Foundations of Knowledge».

¹⁰ W. V. Quine, *Ontological Relativity and Other Essays*. Para el problema de Quine acerca de las teorías empíricamente equivalentes, irreducibles la una a la otra, véase su «On Empirically Equivalent Systems of the World», 313-28, *Theories and Things*, 29-30, y Quine en Hahn y Schilpp (eds.), *The Philosophy of W. V. Quine*, 156-7.

dan. Las dos teorías de Quine pueden pertenecer a, y estar enunciadas en, el mismo lenguaje o lengua. No es fácil ver cómo la misma oración (sin elementos indéxicos), con su significado inalterado, pueda ser verdadera en relación con una persona y no con otra, o en relación con una persona dada en un momento dado y no en otro. La dificultad parece ser debida al intento de introducir consideraciones epistemológicas en el concepto de verdad.

El «realismo interno» de Putnam también hace inmanente a la verdad, aunque no, como lo hace Quine, inmanente a una teoría, sino a toda la lengua y todo el esquema conceptual que una persona acepta. De nuevo, por supuesto, si todo lo que esto significa es que la verdad de las oraciones o preferencias es relativa a un lenguaje o lengua, esto es algo conocido y trivialmente correcto. Pero Putnam parece tener algo más en mente, por ejemplo que una oración de usted y una mía pueden contradecirse la una a la otra y sin embargo ambas ser verdaderas «para el hablante». Es difícil ver en qué lengua o lenguaje pueda formularse de forma persuasiva esta afirmación. El origen de la dificultad es otra vez la necesidad que se siente de hacer accesible la verdad. Putnam es muy claro acerca de esto. Él identifica la verdad explícitamente con la aseverabilidad justificada idealizada. Considera que esto es una forma de realismo porque hay «una cuestión objetiva acerca de cuál sería el veredicto si las condiciones fueran lo suficientemente favorables, un veredicto hacia el cual “convergerían” las opiniones si fuéramos razonables». Añade que su teoría constituye «una forma *humana* de realismo, una creencia de que existe una cuestión objetiva acerca de lo que podemos aseverar correctamente, por contraposición con lo que podría aseverarse correctamente desde la perspectiva del ojo de Dios, tan cara al realista metafísico clásico»¹¹. Uno sospecha que si se hicieran explícitas las condiciones en las que alguien está justificado de forma ideal en aseverar algo, se vería claro que o bien dejan espacio para la posibilidad de error o bien son tan ideales como para no usar de ninguna manera la cone-

¹¹ Hilary Putnam, *Realism and Reason*, pág. xviii.

xión con las capacidades humanas. Es también notable que Putnam parezca no tener ningún argumento para su posición excepto que la alternativa (el «realismo metafísico», es decir, la teoría de la correspondencia) es inaceptable. No argumenta que no pueda haber ninguna otra posición.

Putnam describe su posición acerca del punto principal —el estatuto epistemológico de la verdad— como próxima a la de Michael Dummett. Una diferencia es que Putnam está menos seguro que Dummett de que la verdad se limite a lo que pueda ser efectivamente averiguado, y por tanto está menos seguro de que el principio de bivalencia deba abandonarse; esto explica tal vez por qué Putnam considera a su posición como una forma de realismo mientras que Dummett denomina antirrealismo a su posición. Putnam piensa también que él difiere de Dummett respecto a vincular la verdad con la aseverabilidad justificada *idealizada* en lugar de la aseverabilidad justificada; pero creo que aquí una lectura detallada de Dummett mostraría que tiene una idea muy parecida; ciertamente Crispin Wright la tiene. Si Dummett insiste en algo parecido a las condiciones ideales de Putnam, entonces pienso que se aplica a su posición una crítica dirigida a Dummett que en cierto momento fue formulada por Putnam; si la verdad depende simplemente de la aseverabilidad justificada, la verdad puede «perdersé», y esto debe estar equivocado¹². Dummett dice que está de acuerdo en que la verdad no puede perderse, pero no logra dar una idea clara de cómo la aseverabilidad justificada puede ser a la vez una propiedad fija y una que depende de la capacidad real de los hablantes humanos de reconocer que se satisfacen determinadas condiciones. Las capacidades reales crecen y se desvanecen, y varían de una persona a otra; la verdad no.

¿Por qué acepta Dummett esta concepción de la verdad? Hay diversas razones, pero una de ellas parece ser la siguiente: Hemos visto que una teoría de la verdad en el estilo de Tarski no define ni tampoco caracteriza plenamente la ver-

¹² Hilary Putnam, «Reference and Understanding» y «Respuesta al comentario de Dummett», en Margarit (ed.), *Meaning and Use*.

dad; no hay modo alguno de saber si la teoría se aplica a un hablante o grupo de hablantes a menos que algo se le añada que relacione la teoría con los usos humanos del lenguaje. Dummett piensa que el modo (quizá el único modo) de hacer esto es hacer que la verdad sea algo que los seres humanos puedan reconocer. El uso humano del lenguaje debe ser una función del modo en que los hablantes entienden el lenguaje, de manera que si la verdad ha de desempeñar un papel en la explicación de en qué consiste entender una lengua o lenguaje, debe haber una situación posible en la que pueda decirse que una persona posee «evidencias concluyentes» de que un determinado enunciado es verdadero. Uno puede apreciar la fuerza de esta idea y al mismo tiempo hacérselo difícil de aceptar. Ya he dado mi razón principal para rechazarla: que es vacía o que hace de la verdad una propiedad que puede perderse. Pero hay otras intuiciones fuertes que habrían de sacrificarse también si Dummett estuviera en lo cierto. Una es la conexión de la verdad con el significado; según la concepción de Dummett podemos comprender una oración como «Nunca se construirá una ciudad en este lugar» sin tener idea alguna de en qué consistiría que esta oración fuera verdadera (puesto que esa oración, o una preferencia de la misma, no tiene valor veritativo para Dummett). Otra es la conexión de la verdad con las creencias: según la concepción de Dummett, puedo entender y creer que nunca se construirá una ciudad en este lugar, pero mi creencia carecerá de valor veritativo.

Podría verme tentado de ir en la misma dirección que Dummett si pensara que debo elegir entre lo que Putnam denomina realismo trascendental (es decir, la concepción según la cual la verdad es «radicalmente no-epistémica» —nuestras creencias y teorías mejor investigadas y mejor establecidas podrían todas ellas ser falsas—) y la identificación que Dummett hace de la verdad con la aseverabilidad justificada, puesto que considero incomprensible la primera de estas concepciones —esencialmente la teoría de la correspondencia—, mientras que la de Dummett la considero meramente falsa.

Pero no veo ninguna razón para suponer que el realismo y el antirrealismo, explicados en términos del carácter radical-

mente no-epistémico o el carácter radicalmente epistémico de la verdad, sean los únicos modos de llenar de contenido una teoría de la verdad y el significado. Voy a sugerir una manera que me parece distinta de las dos de las que hemos estado hablando. En primer lugar, yo doy por sentado que existen vínculos ineludibles y evidentes entre los conceptos de verdad, creencia y significado. Si una oración mía *s* significa que P, y yo creo que P, entonces yo creo que *s* es verdadera. Lo que le dota de contenido a mi creencia, y a la oración de su significado, es mi conocimiento de lo que se requiere para que la creencia o la oración sean verdaderas. Puesto que las creencias y la verdad están relacionadas de este modo, las creencias pueden considerarse como las actitudes humanas que conectan una teoría de la verdad con los asuntos humanos.

Observábamos que las evidencias para la corrección de una teoría de la verdad deben producirse en el nivel de las oraciones-T; éstas constituyen también las predicciones contrastables de la teoría. Las evidencias respecto a que las proferencias de una determinada oración por parte de una hablante son verdaderas en las condiciones que recoge una oración-T consisten en hechos acerca de las condiciones que hacen que el hablante tenga a esa oración por verdadera y que considere que esa oración expresa una de sus creencias. Por descontado, los hablantes algunas veces consideran que son verdaderas oraciones que no lo son; estas oraciones expresan creencias erróneas. De modo que la relación entre creencias y verdad no es una cosa simple. Pero puede verse que existe y que es de importancia capital para el significado y la verdad en los casos más sencillos y en el papel que éstos tienen en la comprensión de los demás casos. Los casos más sencillos son, por ejemplo, aquellos en los que lo que la conspicua presencia de libros o de cosas amarillas es lo que causa que se consideren verdaderas oraciones como «Eso es un libro» o «Esto es amarillo». Esto constituye evidencia de que esas oraciones *son* verdaderas exactamente cuando se está en presencia de libros o de cosas amarillas; la razón es que lo que determina el significado de esas oraciones es lo que rutinariamente las hace verdaderas. Está claro que esta explicación debe ser elaborada en diversos respectos para dar cuenta de la posibilidad de error;

hay que añadir muchas más cosas acerca de oraciones que estén más alejadas de las observaciones directas; se deben tener en cuenta las condiciones de verdad de oraciones que se consideran verdaderas sin importar lo que se observe, etc. He explorado estas cuestiones en otro lugar. La cuestión importante en el presente contexto es que es el papel de la creencia lo que conecta verdad y significado; y si ello es así, la verdad no es ni radicalmente no epistémica, ni radicalmente epistémica.

El enfoque a la verdad que yo recomiendo no es epistémico en el sentido de Putnam o de Dummett, ni lo es en el sentido de Peirce, de James o de Dewey, pues no se hace ningún intento de vincular la verdad en general con lo que puede ser averiguado por criaturas racionales. Pero tampoco es este enfoque radicalmente no epistémico, pues no se sigue de mi concepción que puedan ser falsas todas las creencias bien comprobadas y generalmente aceptadas. Si prestamos mucha atención a lo que les dota de contenido a nuestras creencias y de significado a nuestras oraciones, veremos que es imposible que la mayoría de nuestras creencias de percepción sean falsas. Lo que consideramos verdadero —lo que creemos— determina lo que queremos decir o significamos, y, por tanto, indirectamente cuándo son verdaderas nuestras oraciones. Creer algo no hace que sea verdadero, pero crea una presunción de que lo es.

He de hacer dos observaciones adicionales y extraer una conclusión. La primera observación es que considerar verdadera una oración o considerarla como expresión de una creencia no es un *uso* del lenguaje, de modo que puede parecer que no proporciona evidencias relativas a lo que las personas quieren decir o significan mediante lo que dicen. Esta conclusión no se sigue. Hay relaciones obvias entre considerar verdadera una oración y el comportamiento lingüístico (y de otros tipos). Si no pudiéramos a menudo desentrañar cuándo un hablante considera verdadera una oración a partir de su comportamiento lingüístico, no podríamos interpretar su habla. Si he elegido la actitud de la creencia en lugar de actos de aseveración como lo que proporciona la llave de acceso a la verdad no es porque piense que el uso real del lenguaje carezca de importancia, sino porque no conozco ninguna

relación simple entre las intenciones con las que proferimos o escribimos las oraciones y lo que nuestras palabras significan. Nuestros actos de habla revelan nuestras actitudes subyacentes hacia las oraciones; pero a menudo indirectamente.

La segunda observación es que el modo en que enfoco el concepto de verdad, al igual que la manera de enfocar las cuestiones acerca del significado, asume que lo que queremos es una explicación de en qué consiste entender una lengua o lenguaje. Estoy de acuerdo en esto con Putnam, Quine, Dummett y otros. Pero en un aspecto importante Quine y yo diferimos de Dummett, y quizá de Putnam (aunque estoy menos seguro acerca de Putnam). En cualquier caso Dummett piensa que una explicación de en qué consiste entender una lengua o lenguaje exige una descripción de en qué consiste captar o «asir» los significados de las expresiones. Nunca he estado seguro de entender esta metáfora y por lo tanto en lugar de ello me he preguntado qué es lo que sería suficiente conocer para entender una preferencia cualquiera de un hablante (mi respuesta: una teoría de la verdad para la lengua de ese hablante), añadiéndole una explicación del modo en que alguien podría llegar a saber que una teoría así es verdadera. El enfoque de Dummett mira el lenguaje desde el punto subjetivo, desde la perspectiva del hablante. Mi enfoque mira el lenguaje desde el comienzo como un intercambio social y se concentra, por tanto, en lo que pueda aprender una persona acerca de otra en un mundo compartido.

La conclusión que he prometido es la siguiente. Todos los intentos de caracterizar la verdad que van más allá de dar un contenido empírico a una estructura del tipo de las que Tarski nos enseñó a describir son vacíos, falsos o llenos de confusiones. No deberíamos decir que la verdad sea correspondencia, coherencia, aseverabilidad justificada, aseverabilidad justificada idealmente, aquello que las personas adecuadas aceptan en la conversación, lo que la ciencia acabará sosteniendo, o lo que explica el éxito de la ciencia o de nuestras creencias corrientes. En la medida en que el realismo y el antirrealismo dependen de alguna de estas concepciones de la verdad deberíamos negarnos a aceptar ninguno de los dos. El realismo, con su insistencia acerca de la correspondencia radicalmente no episté-

mica, le exige más a la verdad de lo que nosotros podemos entender; el antirrealismo, con su limitación de la verdad a aquello que puede ser averiguado, quita innecesariamente a la verdad el papel que tiene como norma o medida intersubjetiva. Si queremos decir la verdad acerca de la verdad, no deberíamos decir más de lo necesario.

13

La externización de la epistemología

Nadie más puede saber de la misma manera en que nosotros sabemos lo que creemos, tememos, queremos, valoramos y pretendemos. Sabemos cómo nos parece que son las cosas, qué apariencia nos muestran, qué sensaciones nos provocan, cómo nos huelen o nos suenan. Estas cosas las conocemos de una manera en que nunca podremos conocer las cosas del mundo entorno. Tanto si a veces estamos equivocados acerca de los contenidos de nuestras propias mentes como si no lo estamos, tanto si podemos como si no dudar acerca de nuestras propias sensaciones y pensamientos, una cosa es verdadera acerca de esas creencias: no pueden estar equivocadas de un modo general. Si pensamos que tenemos un determinado pensamiento o sensación existe la fuerte presunción de que estamos en lo cierto.

La autoridad especial que es propia del conocimiento de primera persona ha parecido a muchos filósofos que lo capacita de un modo único para la tarea de proporcionar un fundamento al resto del conocimiento y, en particular, por supuesto, a nuestro conocimiento del «mundo externo» y de las mentes de los demás. Se piensa que tal conocimiento necesita un fundamento precisamente porque no existe presunción de que nuestras creencias acerca del mundo o de las mentes de los demás sean verdaderas.

No tengo ninguna intención de repasar las dificultades en que nos ha metido esta imagen de la mente y su lugar en la

naturaleza, ni de recordarles las maneras ingeniosas pero poco convincentes que se han propuesto de sortear sus evidentes fallos. En este ensayo describiré una alternativa, una alternativa que creo correcta.

Quine ha propuesto que se naturalice la epistemología¹. Lo que quería decir con esto es que la filosofía debería abandonar el intento de proporcionar un fundamento al conocimiento o de justificarlo de otro modo, y en lugar de esto debería dar una explicación de cómo se adquiere. Sus críticos han aducido que, al renunciar a la tarea normativa tradicional de la epistemología, Quine simplemente ha cambiado de tema. Sospecho que él está de acuerdo y que ello es parte de lo que pretendía. Desde luego, no está en absoluto clara la distinción entre describir y justificar, entre una explicación empírica de la génesis del conocimiento y un enunciado de las normas que deben satisfacer las creencias para constituir conocimiento, como quizá sea evidente por el hecho de que la epistemología naturalizada de Quine, aunque no intente seriamente responder al escéptico, puede ser reconocida como una forma bastante convencional de empirismo². Y en cualquier caso, ¿cómo se pueden describir el conocimiento o sus orígenes sin decidir lo que es el conocimiento? La respuesta de Quine a esta pregunta es que debemos aceptar lo que la ciencia y un sentido común ilustrado nos dictan sin tratar de justificarlo; sin embargo, su explicación de cómo llegamos a este conocimiento es precisamente el tipo de explicación que se ha interpretado tradicionalmente como un intento de justificación.

Yo no acepto la explicación quineana de la naturaleza del conocimiento, que es esencialmente una explicación cartesiana y de primera persona. Pero considero afín la manera en que Quine enfoca la epistemología, que decididamente es de tercera persona, y la cuestión de en qué medida la naturalización de la epistemología incita o abarca tal enfoque.

¹ W. V. Quine, «Epistemology Naturalized», en *Ontological Relativity and Other Essays*.

² Véase el Ensayo 10 de este volumen y mi «Meaning, Truth and Evidence» para esta caracterización de la epistemología de Quine.

Acepto perfectamente considerarme un epistemólogo naturalizado.

Existe al menos la presunción de que estamos en lo cierto acerca de los contenidos de nuestras propias mentes, de modo que, en los casos en que lo estamos, tenemos conocimiento. Pero cualquier caso particular de ese conocimiento es lógicamente independiente de nuestras creencias acerca del mundo externo, de forma que no puede proporcionarnos un fundamento para la ciencia y las creencias de sentido común. Así es como razonan escépticos como Hume, y pienso que tienen razón: el conocimiento de los contenidos de nuestras propias mentes no puede constituir la base para el resto de nuestro conocimiento. Si esto es correcto, entonces nuestras creencias acerca del mundo deben aguantarse por sí mismas, si es que han de ser consideradas como conocimiento. Sin embargo, a muchos filósofos les ha parecido evidente que si cada una de nuestras creencias acerca del mundo, tomada por sí sola, puede ser falsa, no hay razón alguna por la que todas las creencias no pudieran ser falsas.

Este razonamiento es falaz. Del hecho de que cualquiera de los billetes que tengo en el bolsillo pueda tener el número de serie más alto no se sigue que todos los billetes que tengo en el bolsillo puedan tener el número de serie más alto, o del hecho de que cualquiera pueda ser elegido presidente, no sigue que puedan ser elegidos todos. Tampoco podría ocurrir que todas nuestras creencias acerca del mundo fueran falsas. Supongamos que creo ver a un ratón ocultándose tras una silla. Está claro que esta creencia podría ser errónea. ¿Pero sería errónea esta creencia si yo no creyera con verdad que un ratón es un mamífero pequeño de cuatro patas o una silla un objeto que está hecho para sentarse? Tal vez. No podemos decir exactamente qué otras creencias verdaderas debo yo tener para poseer una creencia determinada que sea falsa. Pero parece claro que la identificación de una creencia del tipo que sea, verdadera o falsa, descansa en un trasfondo de creencias verdaderas, ya que un concepto —como el de ratón o el de silla— no puede seguir siendo el mismo concepto independientemente de qué creencias figuren en él. Es posible tratar de evitar esta conclusión argumentando que una creencia

como la de que un ratón es un animal pequeño de cuatro patas es verdadera sólo en virtud del concepto —una verdad analítica— de modo que realmente no trata acerca del mundo. Todavía sería posible decir que todas nuestras creencias acerca del mundo podrían ser falsas. No dispone de esta opción alguien que, como yo, no piensa que pueda trazarse una clara línea divisoria entre las verdades analíticas y las sintéticas. Pero incluso si existen verdades indubitavelmente analíticas, no es verosímil que éstas sirvan para eliminar la multitud de casos en que los conceptos se individualizan mediante criterios empíricos múltiples.

Por tanto, es imposible que todas nuestras creencias acerca del mundo sean falsas, debido al carácter holista de las creencias empíricas. Sin embargo, la reflexión acerca del razonamiento que conduce a esta tesis sugiere que posee unas limitaciones inquietantes. Lo que muestra es que no podemos abrigar creencias particulares falsas (ni verdaderas) acerca de objetos individuales a menos que poseamos muchas creencias verdaderas acerca de la naturaleza de tales objetos. Esto deja abierta la posibilidad de que podamos estar equivocados acerca de todas nuestras creencias particulares relativas a lo que existe en el mundo y éste sería un escepticismo muy extremo, aunque no completamente total.

Pero ¿podría ocurrir esto? Supongamos que sé lo que es un ratón y lo que es una silla. ¿Cómo llegué yo a mi creencia falsa de que vi ocultarse a un ratón tras la silla? Presumiblemente interpreté lo que vi o experimente tal como lo hice porque satisfacía mis criterios para la presencia de un ratón y de una silla (y los de ocultarse). Pero acabamos de decidir que no puedo estar en general equivocado acerca de esos criterios. De modo que mi error pudo originarse únicamente del siguiente modo: la escena que presencié justificaba mi creencia, aunque esa creencia fuera falsa. Esto realmente puede ocurrir. Sin embargo, es difícil ver cómo podría ocurrir siempre o incluso a menudo, ya que, ¿cómo podría ser que las escenas que presencio justifiquen creencias ninguna de las cuales sea verdadera?

Ningún escéptico resultará convencido por este argumento. Trataré de hacerlo más persuasivo. Permítanme por el mo-

mento dar por sentado algo que supone caer en una petición de principio con respecto a la cuestión de un mundo externo poblado por otras personas; supongamos que tenemos un hablante y alguna otra persona que trata de entender sus palabras. Cada vez que aparece un ratón por allí cerca, con buena luz y con el hablante orientado en la dirección del ratón, etc., el hablante profiere lo que al oyente le suena como la misma expresión: «Mouse». Cuando hay poca luz o el hablante no presta atención la respuesta está menos firmemente correlacionada con las apariciones de algún ratón. Creo que, a menos que haya muchas evidencias contra una interpretación tal, un intérprete competente pensará que lo que el hablante pretende decir mediante sus palabras y lo que cree es que está en presencia de un ratón. Lo que recomienda a esta interpretación es el hecho de que la presencia de un ratón ha provocado que, en cada caso, el hablante profiriera intencionadamente la misma expresión, «Mouse», y lo hiciera con un talante afirmativo. Por descontado que un intérprete puede cometer un error; puede estar equivocado acerca de la intención del hablante, o acerca de cuál es la causa exclusiva y regular de la respuesta del hablante. Por ejemplo, el intérprete se equivocaría si una ardilla provocara la misma respuesta en el hablante, pues entonces los ratones no serían la causa exclusiva y regular de las preferencias de «Mouse». Pero supuesto que el intérprete no comete ninguno de esos errores (ni algunos otros que se mencionarán), su interpretación es correcta.

Supongamos que el hablante ve un ratón mecánico muy bien hecho y emite su «Mouse». El intérprete debe decidir si la palabra y el concepto del hablante incluyen los ratones mecánicos o si hay ahí un error por parte del hablante. Decidir esto no es algo que esté fuera del alcance de un intérprete. Puede, por ejemplo, hacer patente el mecanismo al hablante y registrar su respuesta. El intérprete cuidadoso habrá de aprender el modo en que el hablante responde a muchas otras clases de objetos antes de llegar a una conclusión interpretativa definitiva.

Es patente que el asunto es sutil y complicado. Pero en tanto nos adheramos a la intuición básica de que en los casos más simples las palabras y los pensamientos se refieren a lo que los

causa, está claro que no puede ocurrir que la mayoría de nuestras creencias más sencillas acerca de lo que existe en el mundo sean falsas. La razón es que no formamos primero conceptos para luego descubrir qué es a lo que se aplican; más bien en los casos básicos la aplicación determina el contenido del concepto. Un intérprete que empiece desde cero —uno que no entienda ya la lengua de un hablante— no puede descubrir independientemente acerca de qué versan las creencias de un agente y luego preguntarse si son verdaderas. Esto se debe a que las situaciones que causan normalmente una creencia determinan las condiciones en las que ésta es verdadera.

Estas observaciones se aplican directamente sólo a algunas creencias. Alguien que haya aprendido por los libros qué pinta tiene un guanaco puede que nunca se haya visto llevado a asentir a «Eso es un guanaco» por la visión de un guanaco y puede no obstante estar dispuesto a asentir cuando ve uno (al haber visto quizá fotografías o dibujos de guanacos). O, considerando un caso más difícil, alguien puede saber, en algún sentido razonable, lo que es un guanaco, y que no se trata de una llama, y sin embargo verse llevado a asentir a «Eso es un guanaco» en presencia de una llama. En ambos casos, los contenidos de la creencia de que se está en presencia de un guanaco no se determinan por verse expuesto a guanacos, sino por haber adquirido otras palabras y conceptos, como los de llama, animal, camello, domesticado, etc. Sin embargo, en algún momento hemos de llegar a las exposiciones directas que anclan el pensamiento y el lenguaje al mundo.

El enfoque general a la cuestión de la identificación de los estados mentales que estoy propugnando tiene algo en común con lo que hoy en día se denomina algunas veces externismo o (por parte de Tyler Burge) antiindividualismo; pero se diferencia en algunos rasgos importantes de las versiones del externismo mejor conocidas. De manera que llegado a este punto quisiera distinguir mi versión del externismo de otras formas de externismo, en particular las de Hilary Putnam y Tyler Burge.

Vuelvo a la autoridad de primera persona, es decir, al hecho de que cada persona sabe lo que piensa sin apelar o recurrir a

evidencias, y conoce así lo que está en su propia mente de un modo en que nadie puede conocer lo que está en la mente de otra persona. Muchos han supuesto que el externismo, el cual sostiene que los contenidos de las actitudes proposicionales de una persona se determinan en parte mediante factores que la persona puede desconocer, no puede ser reconciliado con la autoridad de primera persona. Así, Putnam decidió que, dado que el externismo (en una cierta forma) es verdadero, los significados «no pueden estar en la cabeza», y sostuvo que lo mismo vale para los pensamientos en general³. Otros han llegado a la misma conclusión⁴. Pero esa conclusión no se sigue lógicamente, al menos para el tipo de externismo que he descrito. He argumentado esto en detalle en otro lugar⁵, pero el razonamiento puede resumirse del siguiente modo: un intérprete puede descubrir, o suponer correctamente sobre la base de evidencias indirectas, cuáles son los factores externos que determinan el contenido del pensamiento de otra persona; pero puesto que estos factores determinan tanto los contenidos de los propios pensamientos como los contenidos de los pensamientos que uno cree que tiene (al ser éstos el mismo pensamiento), no hay lugar para un error acerca de los contenidos de nuestros propios pensamientos del tipo que puede darse con respecto a los pensamientos de otros.

Aunque no creo que la forma principal de externismo que Putnam acepta⁶ amenace la autoridad de primera persona, por otras razones no estoy enteramente satisfecho con su tesis. El externismo de Putnam se aplica principalmente a palabras de género natural, como «agua» y «leopardo». La idea es que si yo aprendo la palabra «agua» mediante experiencias de H₂O, esa palabra ha de referirse en adelante únicamente a sustancias con esa misma microestructura. Aunque, como dije

³ Hilary Putnam, «The Meaning of "Meaning"», 227 [pág. 144 de la traducción al castellano].

⁴ Ejemplos de ello se encontrarán en John Searle, *Intentionality*; y Andrew Woodfield, «Introduction», en Woodfield (ed.), *Thought and Object*, pág. viii.

⁵ En el Ensayo 2 de este volumen.

⁶ Tengo en mente el tipo de externismo inspirado por los ejemplos de la Tierra Gemela en lugar de los que están motivados por la «división del trabajo lingüístico».

antes, estoy de acuerdo en que la causa usual de mi utilización de la palabra determina lo que significa, no veo por qué la identidad de microestructura ha de ser necesariamente la similitud relevante que determine la referencia de mi palabra «agua». (Enseguida diré lo que pienso que controla la similitud relevante.) También pienso que no hay ninguna razón para limitar el externismo a una categoría de palabras, o a unas cuantas. Creo que es característico del lenguaje y del pensamiento con completa generalidad que sus vínculos con el mundo proceden de las conexiones causales de que he estado hablando.

Tyler Burge ha argumentado en favor de dos formas de externismo. En sus artículos relativamente más tempranos⁷ se concentró en la idea de que los significados de las palabras de una persona y los contenidos de los pensamientos de esa persona dependen en parte de las prácticas lingüísticas de su comunidad, incluso en los casos en los que el individuo está equivocado acerca de las prácticas pertinentes. Los artículos posteriores⁸ pusieron el énfasis en los modos en que los contenidos de las proferencias y pensamientos dependen de la historia causal del individuo, en particular en relación con la percepción. Podemos llamar a estos dos tipos de externismo externismo social y externismo perceptual. Si atendemos al modo en que Burge ha argumentado en su favor, creo que son en gran medida independientes, en el sentido de que ninguno de ellos implica al otro.

Hay tres razones por las que no me impresionan los argumentos en favor del externismo social.

En primer lugar, me parece falso que nuestras intuiciones hablen fuertemente a favor de comprender e interpretar el habla y los pensamientos de un agente en términos de lo que otros quieren decir mediante las mismas palabras. Para empezar, existe el problema de decidir qué grupo es el que ha de deter-

⁷ Tyler Burge, «Individualism and the Mental», «Other Bodies», «Two Thought Experiments Reviewed» y «Individualism and Psychology».

⁸ Por ejemplo, «Cartesian Error and the Objectivity of Perception», «Intellectual Norms and Foundations of Mind» y «Individualism and Self-Knowledge».

minar las normas. Y, lo que es más importante, entendemos mejor a un hablante cuando lo interpretamos como él pretende que se le interprete; esto explicará sus acciones mucho mejor que si suponemos que él quiere decir y piensa lo que alguna otra persona que usara sus palabras «correctamente» podría querer decir y pensar.

En segundo lugar, creo que existe un conflicto entre el externismo social de Burge, el cual vincula el significado del hablante a los usos de una élite de los que puede no ser consciente, y la autoridad de primera persona.

En tercer lugar, desconfío en general de los experimentos mentales que pretenden revelar qué es lo que diríamos en condiciones que nunca se dan. Mi versión del externismo depende de lo que creo que son nuestras prácticas reales⁹.

Paso ahora al externismo perceptual de Burge. Burge hace dos afirmaciones importantes. La primera es que podemos tener conocimiento perceptual —por ejemplo, de que estamos viendo una vaca— sin tener que averiguar independientemente que no nos engañamos de una manera u otra (un caballo pintado para que parezca una vaca, un holograma, etc.). Cuando se dan esos casos de conocimiento perceptual, el contenido de nuestros pensamientos lo determina parcialmente la causa de los mismos. Burge dice lo siguiente:

para concebir algo como agua... se debe estar en alguna relación causal con el agua, o al menos en alguna relación causal con otras sustancias particulares que nos permitan teorizar adecuadamente sobre el agua. En los casos normales el agua se ve y se toca. Estas relaciones ilustran el tipo de condiciones que hacen posible concebir algo como agua... Para *pensar* que el agua es un líquido no se necesita *conocer* las complejas condiciones que deben darse para pensar ese pensamiento¹⁰.

Es difícil decidir, a juzgar por este pasaje, si los contenidos del pensamiento de que esto es agua, los determina —en el caso en que el pensamiento es verdadero— el hecho de que

⁹ Véase el Ensayo 2 de este volumen.

¹⁰ Tyler Burge, «Individualism and Self-Knowledge», 653-4.

es el agua lo que causa ese pensamiento, o si en este caso, como en los casos en los que el pensamiento es falso, los contenidos los determina el caso normal, el desarrollo general del contacto causal del sujeto pensante con su entorno. No veo ninguna dificultad obvia en aceptar y combinar estas dos tesis, aunque desde luego debe haber una distinción entre el papel que tiene la presencia real de agua cuando hace que piense que esto es agua, y el papel que la historia de mis relaciones con el agua desempeña en hacer que un pensamiento falso de que estoy viendo agua sea, sin embargo, un pensamiento acerca del agua. Es claro que esta última tesis es la que es esencial para el externismo. Burge lo dice del modo siguiente: «Las naturalezas de tales estados las determinan en parte las relaciones normales entre la persona... y el entorno. El error se determina frente a un trasfondo de interacciones normales»¹¹.

Simpatizo tanto con la tesis epistemológica de que el conocimiento perceptual no requiere que sepamos independientemente que efectivamente se dan las condiciones que lo posibilitan, como, por supuesto, con la tesis (más importante en el contexto presente) de que los contenidos de nuestros pensamientos y de lo que decimos lo determina parcialmente la historia de las interacciones causales con el entorno.

El estar de acuerdo sobre este punto se le hace natural a quien como yo ha estado insistiendo durante treinta años en que los contenidos de las oraciones que aprendemos más tempranamente y que son más básicas («Mamá», «Perrito», «Rojo», «Fuego», «Gavagai») lo determina por fuerza aquello que hay en el mundo que provoca causalmente que las consideremos verdaderas. Es aquí, según he afirmado por mucho tiempo, donde se establecen los vínculos entre lenguaje y mundo, y donde se fijan las restricciones más importantes acerca del significado; y dadas las estrechas relaciones entre pensamiento y lenguaje, observaciones análogas valen para los contenidos de las actitudes.

Burge y yo estamos esencialmente de acuerdo, por tanto, acerca del externismo perceptual. Pero existe un asunto estre-

¹¹ «Cartesian Error and the Objectivity of Perception», 125.

chamente relacionado en el que diferimos, el cual concierne a la cuestión con la que comencé el presente ensayo, a saber, la cuestión de si el externismo proporciona una respuesta a ciertos tipos de escepticismo. Dice Burge: «no hay ningún argumento fácil contra el escepticismo a partir del antiindividualismo y de la autoridad del autoconocimiento»¹². Estoy de acuerdo en un sentido, pero sólo porque no veo qué es lo que la autoridad de primera persona tiene que ver con ello. Pero a mí me parece que, *si se acepta el externismo en la percepción, sí que hay un argumento fácil contra el escepticismo global de los sentidos del tipo que Descartes, Hume, Russell y muchísimos otros han pensado que necesita una respuesta. Dice Burge:*

La mayoría de las representaciones perceptuales se forman y obtienen su contenido a través de la interacción regular con el entorno. Ellas representan —en algún sentido complejo del término «normalmente»— aquello de donde normalmente provienen y aquello a lo que normalmente se aplican. No tiene sentido atribuir un error perceptual sistemático a un ser cuyas representaciones perceptuales pueden explicarse como resultado de interacciones regulares con un entorno físico¹³...

Puede parecer que esto está suponiendo *más* que el mero externismo. Pero no veo cómo. Burge considera un caso, que es muy parecido a los casos usuales de cerebro-en-cubeta, y muestra por qué tal cerebro no puede ser engañado acerca del entorno (por mucho tiempo). Apruebo este argumento, al haberlo utilizado a menudo yo mismo. Si alguna cosa causa sistemáticamente determinadas experiencias (o respuestas verbales), los pensamientos y preferencias son acerca de eso precisamente. Esto excluye el error sistemático. Si no hay nada que sistemáticamente cause las experiencias, no hay contenido alguno acerca del cual estar equivocado. Citándome a mí mismo: «Lo que obstruye el camino al escepticismo global de los sentidos es, en mi opinión, el hecho de que en los casos más sen-

¹² «Individualism and Self-Knowledge», 655 n.

¹³ «Cartesian Error and the Objectivity of Perception», 131.

cillos y metodológicamente más básicos debemos considerar que los objetos de una creencia son las causas de esa creencia»¹⁴. Todo el que acepte el externismo en la percepción sabe que no podemos estar sistemáticamente equivocados acerca de si existen cosas tales como vacas, personas, agua, estrellas y chicle. Al saber por qué es esto así, debe reconocer situaciones en las que esté justificado en creer que está viendo agua o está viendo una vaca. En los casos en que está en lo cierto, esa persona sabe que está viendo agua o está viendo una vaca.

He rechazado el externismo social tal como lo defiende Burge. ¿Pienso por ello que los factores sociales no tienen ningún papel en el externismo de lo mental? No, en absoluto. Pero yo introduciría el factor social de una manera que lo conecte directamente al externismo perceptual, ubicando así el papel de la sociedad dentro del nexo causal que incluye la interacción entre las personas y el resto de la naturaleza.

Burge y yo estamos de acuerdo en que la causa de determinados estados mentales es relevante para el contenido de esos estados. Y estamos de acuerdo en que hay un tipo de caso que es especialmente importante: un ejemplo del mismo es el modo en que el hecho de que un determinado estado mental ha sido causado típicamente por la percepción de vacas nos permite pensar «Hay una vaca» incluso cuando no estamos en presencia de vaca alguna. Pero aquí se plantea un problema. Lo que determina el contenido de tales pensamientos básicos (y lo que queremos decir mediante las palabras que utilizamos para expresarlos) es lo que típicamente ha causado pensamientos similares. Pero ¿qué es lo que los ha causado típicamente? Hay muchas alternativas, como, por ejemplo, acaecimientos que sucedieron antes de todas las vacas, o acaecimientos espacialmente más cercanos al sujeto que cualquier vaca. En un determinado momento, Burge dice que las causas que anteceden a la causa «natural» o que son intermedias entre la causa natural y el efecto mental («como, por ejemplo, los rayos de luz que llegan a la retina») complicarían la descripción del contenido «de una manera que no ha sido nunca ar-

¹⁴ Ensayo 10 en este volumen.

ticulada plenamente»¹⁵. Pero ¿para quién sería complicado? ¿Articulado por quién? Por supuesto, todos esos rayos de luz que llegan a la retina tienen *algo* en común; y lo mismo vale para los otros tipos de causas. Somos nosotros los humanos para quienes estas clasificaciones son complicadas e imposibles de articular (excepto, desde luego, recurriendo a trampas bien conocidas). Somos nosotros los que agrupamos las apariciones de vacas en la misma clase, de un modo más o menos natural, o con un mínimo de aprendizaje. Incluso así se necesita otra clasificación para completar esto, pues la clase de causas pertinentes se define a su vez por la similitud en las respuestas: agrupamos las causas de las respuestas de alguien, sean verbales o de otro tipo, porque encontramos parecidas esas respuestas. ¿Qué es o qué hace que estos parecidos sean los aquí pertinentes? De nuevo, la respuesta es obvia: somos nosotros, debido a la manera en que estamos hechos (la evolución tuvo que ver con esto), los que encontramos natural y fácil poner en la misma clase esas respuestas. Si no fuera así no tendríamos razón alguna para sostener que los demás responden a los mismos objetos y acontecimientos (esto es, causas) que nosotros. Puede ser que ni siquiera las plantas podrían sobrevivir en nuestro mundo si no reaccionaran hasta cierto punto de maneras que encontramos parecidas a acontecimientos y objetos que consideramos parecidos. Claramente, esto vale para los animales, y, por supuesto, se hace más patente cuanto más parecido a nosotros es el animal.

Por tanto, la identificación de los objetos del pensamiento descansa sobre una base social. Sin una criatura que observe a otra, no tendría lugar la triangulación que ubica en un espacio público a los objetos pertinentes. Con esto no quiero decir que una criatura que observa a otra proporcione a cualquiera de las dos el concepto de objetividad; la presencia de dos o más criaturas que interaccionan una con otra y con un entorno común es como mucho una condición necesaria para un concepto así. Únicamente la comunicación puede proporcionar ese concepto, pues tener el concepto de objeti-

¹⁵ «Cartesian Error and the Objectivity of Perception», 126-7.

vidad, el concepto de objetos y acaecimientos que ocupan un mismo mundo compartido, de objetos y acaecimientos cuyas propiedades y existencia es independiente de nuestro pensamiento, requiere que nos percatemos del hecho de que compartimos pensamientos y un mundo con los demás.

Por esta razón no podemos resolver la cuestión de los contenidos de los estados mentales desde el punto de vista de una sola criatura. Tal vez como mejor se vea esto sea pensando en cómo aprende una persona de otra a hablar y a pensar en las cosas corrientes. Dicho de manera muy simplificada, un aspecto básico de ese aprendizaje puede describirse del siguiente modo: se recompensa al aprendiz, a propósito o no, cuando emite sonidos o respuestas de otro tipo de una manera que el instructor encuentra apropiada en casos que él clasifica del mismo modo. Con posterioridad, por medio de situaciones que el aprendiz instintivamente clasifica del mismo modo, se hace que el aprendiz emita sonidos semejantes. Por descontado que es posible hacer correcciones. El éxito en el primer nivel se logra en la medida en que el aprendiz responde con sonidos que el instructor encuentra parecidos a situaciones que él encuentra parecidas. El instructor está respondiendo a dos cosas: la situación externa y las respuestas del aprendiz. El aprendiz está respondiendo a dos cosas: la situación externa y las respuestas del instructor. Todas estas relaciones son causales. Es así como se forma el triángulo esencial que hace que sea posible la comunicación acerca de objetos y acaecimientos compartidos. Pero es también este triángulo el que determina el contenido de las palabras y los pensamientos del aprendiz cuando éstos se hacen lo suficientemente complejos como para merecer esa denominación. El papel del instructor en la determinación del contenido de las actitudes del aprendiz no es sólo el «determinar» de la causalidad, pues además de ser una causa de esos pensamientos, lo que hace que un aspecto particular de la causa de las respuestas del aprendiz sea el aspecto que les dote del contenido que tienen es el hecho de que este aspecto de la causa es algo compartido por instructor y aprendiz. Sin este compartir, no habría fundamento alguno para seleccionar una causa y no otra como la causa que fija el contenido. Una criatura que no pue-

da comunicarse puede ser vista por nosotros como algo que responde a un mundo objetivo; pero no estamos justificados en atribuirle pensamientos acerca de nuestro mundo (o ningún otro).

Esto me hace volver a la autoridad de primera persona. Sostenía al comienzo que la autoridad de primera persona tiene dos rasgos estrechamente relacionados: uno es el hecho de que produce conocimiento que no se basa en inferencias o evidencias; el otro es la asimetría entre cómo conocemos nuestras propias mentes y cómo conocemos las mentes de los demás. Es notable que, por ejemplo, la explicación que da Burge de la autoridad de primera persona no explica el hecho de que la misma cosa, a saber, lo que está en la mente de una persona, pueda ser conocida también por otra aunque de una manera muy distinta. Según creo, la parte de la explicación que falta se da cuando se reconoce la manera en que la interacción con las otras personas determina parcialmente los contenidos de los estados mentales. El conocimiento de la propia mente es personal. Pero lo que individualiza a un estado es al mismo tiempo lo que lo hace accesible a los demás, pues el estado se individualiza mediante la interacción causal entre tres elementos: el sujeto pensante, aquellos otros con quien se comunica y el mundo objetivo que ellos saben que comparten.

Tres variedades de conocimiento

En su mayor parte sé lo que pienso, quiero y me propongo, y cuáles son mis sensaciones. Sé bastante acerca del mundo que me rodea, la ubicación, tamaño y propiedades causales de los objetos del mismo. A veces sé también lo que hay en la mente de otras personas. Cada uno de esos tres tipos de conocimiento empírico tiene sus características distintivas. Lo que sé acerca de los contenidos de mi propia mente lo sé en general sin apelar a evidencias o investigación. Hay excepciones, pero la primacía del autoconocimiento no mediado la atestigua el hecho de que desconfiamos de las excepciones hasta que podemos reconciliarlas con lo no mediado. Por otra parte, mi conocimiento del mundo externo a mí depende del funcionamiento de mis órganos sensoriales, y esta dependencia causal respecto de los sentidos hace que mis creencias acerca del mundo de la naturaleza estén abiertas a un tipo de incertidumbre que raramente se da en el caso de las creencias acerca de nuestros propios estados mentales. Muchas de mis percepciones simples de lo que sucede en el mundo no se basan en ulteriores evidencias; simplemente mis creencias perceptivas las causan directamente acaecimientos y objetos de mi entorno. Pero mi conocimiento de los contenidos proposicionales de otras mentes no es nunca inmediato en este sentido; no tendría acceso a lo que los demás piensan y valoran si no pudiera percatarme de su comportamiento.

Por supuesto, las tres variedades de conocimiento conciernen a aspectos de la misma realidad; en lo que se diferencian es en el modo de acceso a esa realidad.

Las relaciones entre los tres tipos de conocimiento empírico, en particular las cuestiones de prioridad conceptual, han encabezado por largo tiempo la lista de las preocupaciones epistemológicas de los filósofos, y ellas son aquí mi tema. Muchos enfoques conocidos de la cuestión de cómo se relacionan los tres tipos de conocimiento consideran primario al autoconocimiento, quizá debido a su carácter directo y a su relativa certeza, intentando posteriormente derivar a partir de él el conocimiento del mundo externo; como último paso tratan de basar el conocimiento de otras mentes en observaciones de la conducta. Ésta no es, innecesario es decirlo, la única dirección que puede tomar esa derivación; en su lugar se puede aceptar el conocimiento del mundo externo como básico, al menos en alguna de sus manifestaciones, y tratar de relacionar con él o reducir a él las otras formas de conocimiento. La elaboración de estas propuestas reductivas y la demostración de su fracaso forma buena parte de la historia de la filosofía que va de Descartes al presente. Si muchos filósofos se han alejado de estos problemas en tiempos recientes no es porque se piense que los problemas han sido resueltos, sino porque parecen insolubles. Por supuesto, siempre queda la triste esperanza de que los propios problemas sean ilusorios.

Esto no puede ser así. Existen razones que nos fuerzan a aceptar la tesis de que ninguna de las tres formas de conocimiento es reducible a una de ellas o a las dos restantes. Aquí daré mis propias razones para creer esto, pero interpreto que la imposibilidad de encontrar maneras efectivas de reducción está clara a partir del rechazo casi universal de los programas reduccionistas usuales. El escepticismo en varias de sus formas conocidas constituye el tributo que a pesar nuestro rendimos a la clara imposibilidad de unificar las tres variedades de conocimiento: una forma de escepticismo se origina en la dificultad de explicar nuestro conocimiento del mundo externo sobre la base de nuestro conocimiento de nuestras propias mentes; otra reconoce que nuestro conocimiento de las otras

mentes no puede sólo consistir en lo que podamos observar desde fuera. La imposibilidad de resolver el problema mente-cuerpo constituye otro de esos tributos.

Es notable en qué medida los filósofos, incluidos los que se han mostrado escépticos acerca de la posibilidad de justificar creencias acerca del mundo externo, han dejado de lado estas dudas cuando se han puesto a pensar en el problema de las otras mentes; es notable puesto que este último problema puede plantearse sólo si es posible el conocimiento del comportamiento, y por tanto del mundo externo. Mantener separados estos problemas tiene el desafortunado efecto de oscurecer el hecho de que ambos descansan en un supuesto común. Este supuesto es que la verdad acerca de qué es lo que una persona cree sobre el mundo es lógicamente independiente de la verdad de esas creencias. Ciertamente esto parece ser así, porque claramente parece que la totalidad de las creencias y experiencias subjetivas de una persona es lógicamente compatible con la falsedad de cualquiera de esas creencias, de modo que por más conocimientos que se posean acerca de los contenidos de la propia mente, tales conocimientos no aseguran la verdad de una creencia acerca del mundo externo. La independencia lógica de lo mental opera igualmente en la otra dirección: ninguna cantidad de conocimientos acerca del mundo externo implica la verdad acerca de lo que hay en la mente. Si existe una barrera lógica o epistémica entre la mente y la naturaleza, no sólo nos impide ver el exterior, sino que también obstruye la visión del interior desde fuera.

Algunas veces se piensa que si separamos el problema de saber qué es lo que hay en una mente del problema de saber cualquier cosa externa a nosotros, entonces el problema del conocimiento de las otras mentes se soluciona cuando nos damos cuenta de que es parte del concepto de estado o acaecimiento mental el que determinadas formas de comportamiento u otros signos externos cuenten como evidencias en favor de la existencia de tal estado o acaecimiento mental. Sin duda es verdad que es parte del concepto de estado o acaecimiento mental que el comportamiento proporciona evidencias de tales estados o acaecimientos. Lo que no está claro es cómo esto constituye una respuesta al escéptico, pues el he-

cho de que la conducta proporciona evidencias acerca de qué es lo que hay en la mente no nos da ninguna explicación de la asimetría entre el conocimiento indirecto que tenemos de las otras mentes y el conocimiento directo que tenemos de nuestra propia mente. La mencionada solución insiste en que las evidencias conductuales pueden ser suficientes para atribuir de forma justificada estados mentales a los demás mientras que reconoce que tales evidencias son en general irrelevantes para la autoadscripción de esos mismos estados. Pero si no se nos da explicación alguna de esta notable asimetría, deberíamos concluir que en realidad hay dos tipos de conceptos: conceptos mentales que se aplican a los demás, y conceptos mentales que se nos aplican a nosotros mismos. Si los estados mentales de los demás se conocen únicamente a través de sus manifestaciones conductuales y otras manifestaciones externas, mientras que esto no es verdad para nuestros propios estados mentales, ¿por qué habríamos de pensar que nuestros propios estados mentales son de alguna manera parecidos a los de los demás? También podríamos preguntarnos por qué, de ser satisfactoria esa respuesta al problema del conocimiento de las otras mentes, no habríamos de aceptar una solución análoga para el problema de nuestro conocimiento del mundo externo. Sin embargo, hay un reconocimiento ampliamente compartido de que esta respuesta al escepticismo general es inaceptable. ¿Distinguimos entre los problemas en cuestión porque suponemos que mientras que no tenemos acceso al mundo externo si no es a través de la experiencia, podemos no obstante extrapolar inteligiblemente, extrayendo conclusiones acerca de las experiencias de los demás, dado que tenemos acceso a la experiencia en nuestro propio caso? Pero suponer esto es caer en una petición de principio, puesto que se está dando por sentado sin argumentarlo que lo que llamamos estados mentales de los demás son parecidos a lo que identificamos como estados mentales en nosotros mismos.

He estado repasando estos problemas y perplejidades porque quisiera primero de todo enfatizar lo extraño que resulta el que tengamos tres variedades irreducibles de conocimiento empírico. Necesitamos tener una panorámica de conjunto que

no sólo acomode los tres modos de conocimiento, sino que dé sentido a sus relaciones mutuas. Sin esa panorámica general debería preocuparnos profundamente el hecho de que el mismo mundo nos sea conocido de tres modos tan diferentes. Y, en segundo lugar, es esencial darse cuenta de en qué medida están interrelacionados problemas que usualmente se han considerado de uno en uno. Hay tres problemas básicos: cómo puede una mente conocer el mundo de la naturaleza, cómo le es posible a una mente conocer a otra, y cómo es posible conocer los contenidos de nuestras propias mentes sin recurrir a observaciones o evidencias. Sostendré que es un error suponer que estas cuestiones pueden reducirse a dos o que se puedan tratar aisladamente.

Para tratar de formarnos una panorámica de las relaciones entre las tres clases de conocimiento hay mucho más que hacer que mostrar *que* son mutuamente irreducibles; debemos ver *por qué* son irreducibles. Esto implicará a su vez sacar a la luz los papeles conceptuales respectivos que tienen cada una de esas formas de conocimiento, y por qué cada una de las tres variedades de conocimiento es indispensable —por qué no podríamos prescindir de ninguna de ellas—. Por supuesto, si estoy en lo cierto respecto a que cada una de las tres variedades de conocimiento es indispensable, ha de desecharse el escepticismo sobre los sentidos y el escepticismo acerca de las otras mentes, pues el escéptico cartesiano o humeano acerca del mundo externo sostiene que es completamente evidente que podemos prescindir del conocimiento del mundo de la naturaleza: lo que conocemos acerca de nuestra propia mente es autosuficiente, y puede que sea todo el conocimiento que tenemos. El escéptico acerca de las otras mentes está igualmente convencido de que podemos prescindir del conocimiento de otras mentes —ello debe ser posible si siempre estamos inseguros de si lo tenemos.

Al principio puede parecer que nos las podríamos apañar fácilmente sin palabras que expresaran nuestras creencias acerca de los estados mentales de los demás o de nosotros mismos. Creo que esto es imaginable, pero el asunto que me ocupa es primariamente epistémico, no lingüístico. La cuestión epistémica es si nos las podríamos arreglar sin conoci-

miento de mentes, tanto la nuestra como la de otros. Argumentaré que no. Lo que no podríamos hacer es apañárnoslas sin una manera de expresar, y por tanto de comunicar, nuestros pensamientos acerca del mundo natural. Pero si podemos hacer esto, la transición a ser también capaces de atribuir pensamientos de un modo verbal es relativamente sencilla y sería asombroso que no se diese ese paso. Con respecto a nuestros propios pensamientos, no es más que la diferencia entre decir aseverativamente «La nieve es blanca» y «Creo que la nieve es blanca». Las condiciones de verdad de estas aseveraciones no son las mismas, pero cualquiera que entienda la primera de ellas conoce las condiciones de verdad de la segunda, incluso si carece de un dominio de una oración con esas condiciones de verdad. Esto se debe a que todo el que entiende el lenguaje puede reconocer aseveraciones y sabe que quien hace una aseveración se representa a sí mismo como alguien que cree lo que dice. De manera similar, quien le dice a García que la nieve es blanca conoce las condiciones de verdad de «García cree que la nieve es blanca» (incluso si no sabe castellano ni tiene un modo de expresar las creencias).

La creencia es una condición del conocimiento. Pero para tener una creencia no es suficiente con discriminar entre aspectos del mundo, comportarse de modos distintos en circunstancias distintas; esto lo hace un caracol o una litorina. Tener una creencia exige además apreciar el contraste entre creencia verdadera y creencia falsa, entre la apariencia y la realidad, el mero parecer y el ser. Por supuesto, podemos decir que un girasol ha cometido un error si se gira hacia una luz artificial como si fuera el sol, pero no suponemos que el girasol pueda pensar que ha cometido un error, de modo que no atribuimos una creencia al girasol. Quien tenga una creencia acerca del mundo —o acerca de cualquier otra cosa— debe captar el concepto de verdad objetiva, de lo que es el caso independientemente de lo que él o ella piensa. Debemos preguntar, por tanto, por la fuente del concepto de verdad.

Wittgenstein nos puso en la pista de la única respuesta posible a esta pregunta, tanto si su problema era tan amplio como el nuestro como si no, y tanto si creía que hay respues-

tas a los problemas filosóficos como si no. La fuente del concepto de verdad objetiva es la comunicación interpersonal. El pensamiento depende de la comunicación. Esto se concluye inmediatamente si suponemos que el lenguaje es esencial para el pensamiento y estamos de acuerdo con Wittgenstein en que no puede haber un lenguaje privado¹. El argumento fundamental contra los lenguajes privados es que, a menos que un lenguaje sea compartido, no hay modo de distinguir entre usar el lenguaje correctamente y usarlo incorrectamente; únicamente la comunicación con otros puede proporcionar un control objetivo. Si únicamente la comunicación puede proporcionar un control sobre el uso correcto de las palabras, únicamente la comunicación puede suministrar una norma de objetividad en otros dominios, o al menos eso es lo que voy a argumentar. No tenemos razones para atribuir a una criatura la distinción entre lo que se piensa que es el caso y lo que es el caso a menos que esa criatura posea la norma que un lenguaje compartido proporciona; y sin tal distinción no hay nada que pueda claramente denominarse pensamiento.

En la comunicación, lo que deben compartir un hablante y el intérprete de un hablante es una comprensión de lo que el hablante quiere decir mediante lo que dice. ¿Cómo es esto posible? Podría ayudarnos conocer cómo vino a la existencia el lenguaje por primera vez, o si al menos pudiéramos explicar cómo aprende un individuo su primera lengua, cuando otros de su entorno ya son lingüísticamente competentes. A falta de ese conocimiento o esa explicación lo que podemos hacer es preguntar cómo un intérprete competente (uno que

¹ Por supuesto que puede existir un código privado basado en un lenguaje adquirido públicamente. No tengo ni idea de cuán ampliamente Wittgenstein pretendía que se interpretara su tesis acerca de los lenguajes privados; tal vez quería que su argumento se aplicara sólo a conceptos que son necesariamente privados. Pero, como Saul Kripke en *Wittgenstein on Rules and Private Language*, pienso que el argumento se aplica al lenguaje con toda generalidad y por tanto (yo diría) al pensamiento proposicional. Pero mientras que acepto la idea de que la comunicación es la fuente de la objetividad, no pienso que la comunicación dependa de que los hablantes utilicen las mismas palabras para expresar los mismos pensamientos.

posea recursos conceptuales adecuados y una lengua propia) podría llegar a entender al hablante de una lengua que le es completamente extraña. Una respuesta a esta pregunta debería revelar algunos rasgos importantes de la comunicación e iluminar aquello que hace posible una primera entrada en el lenguaje.

Un intérprete intrépido que trabaje sin ayuda de un bilingüe trata de asignar contenido proposicional a las preferencias de un hablante. Efectivamente, asigna una oración propia a cada una de las oraciones del hablante. En la medida en que hace bien las cosas, la oración del intérprete proporciona las condiciones de verdad de las oraciones del hablante, y por tanto suministra la base para la interpretación de las preferencias del hablante. Puede concebirse el resultado como una caracterización recursiva, por parte del intérprete, de la verdad de las oraciones, y por tanto de las preferencias efectivas y potenciales del hablante.

Un intérprete no puede observar directamente las actitudes proposicionales de otra persona; las creencias, deseos e intenciones, incluyendo las intenciones que determinan parcialmente los significados de las preferencias, son invisibles al ojo desnudo. El intérprete puede, sin embargo, atender a las manifestaciones externas de esas actitudes, incluyendo las preferencias. Puesto que todos somos capaces de descubrir a partir de tales manifestaciones lo que un agente piensa y lo que quiere decir, debe existir una relación inteligible entre las evidencias y las actitudes. ¿Cómo salvamos el espacio entre una cosa y otra? Sólo conozco un modo: un intérprete puede percibir, con la suficiente frecuencia, que un agente tiene un cierto tipo de actitud hacia un objeto o acaecimiento que ese intérprete percibe. Si el intérprete pudiera *individuar* de este modo directamente las actitudes de otro, el problema estaría resuelto, pero únicamente bajo el supuesto de que el oyente pueda leer las mentes. Sin embargo, no se caería en petición de principio si se supusiera que el intérprete puede detectar una o más actitudes no individuativas. Son ejemplos del tipo de actitudes especiales que tengo en mente las siguientes: el considerar verdadera una oración en un determinado momento de tiempo, el querer que una oración sea verdadera o

preferir que una oración sea verdadera en lugar de otra. El supuesto de que podemos detectar tales actitudes no constituye una petición de principio respecto a la cuestión de cómo dotamos de contenido a las actitudes, puesto que una relación como la de considerar verdadera, entre un hablante y una preferencia, es una relación extensional que puede saberse que se da sin saber lo que significa la oración. Estas actitudes son no-individuativas porque aunque son psicológicas por su naturaleza, no distinguen los diferentes contenidos proposicionales que expresan preferencias distintas.

En *Palabra y objeto*, Quine recurrió a las actitudes no-individuativas que provocan asentimiento. Puesto que una persona asiente a una preferencia, o considera verdadera una oración, en parte debido a lo que cree y en parte a lo que la preferencia o la oración significa en su lengua, el problema de Quine era separar estos dos elementos sobre la base de evidencias que combinan los dos factores. Si la separación tiene éxito, el resultado es una teoría acerca tanto de las creencias como del significado para tal hablante, puesto que debe proporcionar una interpretación de las preferencias de un hablante, y si se sabe que el hablante asiente a la preferencia y lo que quiere decir al emitir palabras, se sabe también lo que cree.

El proceso de separar significado y opinión recurre a dos principios clave que deben aplicarse si un hablante puede ser interpretado: el Principio de Coherencia y el Principio de Correspondencia. El Principio de Coherencia lleva al intérprete a descubrir un cierto grado de consistencia lógica en el pensamiento del hablante; el Principio de Correspondencia lleva al intérprete a considerar que el hablante está respondiendo a los mismos rasgos del mundo a los que él (el intérprete) respondería en circunstancias similares. Ambos principios pueden denominarse principios de caridad (y se los ha denominado así); un principio dota al hablante de una módica cantidad de lógica y el otro le dota en un cierto grado con lo que el intérprete considera que son creencias verdaderas acerca del mundo. La interpretación con éxito inviste necesariamente a la persona interpretada de una racionalidad básica. Se sigue de la naturaleza de la interpretación correcta que hay una medida in-

terpersonal de consistencia y correspondencia con los hechos que se aplica tanto al hablante como al intérprete de ese hablante, a sus preferencias y a sus creencias.

Hay dos cuestiones ahora que nos impiden el paso. La primera es: ¿por qué habría una medida interpersonal de constituir una medida objetiva? Es decir, ¿por qué habría de ser verdad aquello en lo que las personas están de acuerdo que es verdad? La segunda es: incluso si sucede que la comunicación presupone una medida o norma objetiva de la verdad, ¿por qué habría de ser ésta la única manera de establecer esa norma?

Una manera de responder estas preguntas es como sigue. Todas las criaturas clasifican objetos y aspectos del mundo en el sentido de que tratan unos estímulos como más parecidos que otros. En esa actividad clasificadora el criterio es la similitud de respuesta. No hay duda de que la evolución y el aprendizaje subsiguiente explican estas pautas de comportamiento. Pero ¿desde qué punto de vista se les puede llamar pautas? El criterio en base al cual puede decirse que una criatura trata determinados estímulos como similares, como pertenecientes a la misma clase, es la similitud de las respuestas de esa criatura a tales estímulos; pero ¿cuál es el criterio de similitud de las respuestas? *Este* criterio no puede derivarse de las respuestas de la criatura en cuestión; sólo puede provenir de las respuestas de un observador a las respuestas de la criatura. Y sólo cuando un observador correlaciona conscientemente las respuestas de otra criatura con los objetos y acaecimientos del mundo del observador existe una base para decir que esa criatura está respondiendo a esos objetos y acaecimientos en lugar de a otros objetos y acaecimientos. Como si fuéramos intérpretes del comportamiento verbal del hablante de una lengua que nos es totalmente ajena, agrupamos diversos actos verbales del hablante: «Madre», «Nieve», «Mesa», cuando éste los repite como oraciones de una sola palabra, suenan parecidos si estamos en la sintonía adecuada. Cuando descubrimos tipos de objetos o acaecimientos del mundo que podemos correlacionar con las preferencias de un hablante, nos encontramos en el camino de la interpretación de la conducta lingüística más simple.

Si estamos enseñando una lengua a alguien la situación se vuelve más compleja, pero más claramente interpersonal. Lo que parece que es básico es lo siguiente: un observador (o un instructor) encuentra (o infunde) una regularidad en el comportamiento verbal del informante (o aprendiz) que aquél puede correlacionar con acontecimientos y objetos del entorno. Por supuesto, todo esto puede tener lugar sin que haya un pensamiento desarrollado por parte del que aprende, pero constituye la base necesaria para atribuir pensamientos y significados a la persona observada, pues hasta que se complete el triángulo que conecta a las dos criaturas y a cada una de ellas con rasgos comunes del mundo no puede haber respuesta alguna a la pregunta de si una criatura, al discriminar entre unos estímulos y otros, está discriminando entre estímulos que si sitúan en las superficies sensoriales o en algún lugar más externo, o más interno. Sin este compartir reacciones a estímulos comunes, el pensamiento y el habla no tendrían ningún contenido en particular, es decir, no tendrían ningún contenido en absoluto. Son necesarios dos puntos de vista para asignar una ubicación a la causa de un pensamiento, y para, de ese modo, definir su contenido. Podemos concebir esto como una cierta forma de triangulación: cada una de las dos personas reacciona de modo diferente a los estímulos sensoriales que provienen de una determinada dirección. Proyectando hacia fuera las líneas entrantes, la causa común se encuentra en su intersección. Si las dos personas se dan cuenta entonces de las reacciones del otro (reacciones verbales en el caso del lenguaje), cada una de ellas puede correlacionar estas reacciones observadas con sus estímulos procedentes del mundo. Se ha determinado una causa común. El triángulo que da contenido al pensamiento y al habla está completo. Pero se requieren dos personas para hacer un triángulo.

Hasta que se establezca una línea de base mediante la comunicación con algún otro no hay motivo para decir que los propios pensamientos o las propias palabras poseen un contenido proposicional. Si esto es así, está claro entonces que el conocimiento de otra mente es esencial a todo pensamiento y todo conocimiento. Sin embargo, el conocimiento de otra mente es posible únicamente si se posee conocimiento del

mundo, pues la clase de triangulación que es esencial para el pensamiento requiere que los que se comunican reconozcan que ocupan posiciones en un mundo compartido. Por tanto, el conocimiento de otras mentes y el conocimiento del mundo dependen el uno del otro; ninguno es posible sin el otro. Con toda seguridad, Ayer estaba en lo cierto cuando dijo: «Únicamente con el uso del lenguaje entran plenamente en escena la verdad y el error, la certeza y la incerteza»².

No es posible el conocimiento de los contenidos proposicionales de nuestras propias mentes sin las otras formas de conocimiento, puesto que no existe pensamiento proposicional sin comunicación. Sucede también que no estamos en situación de atribuir pensamientos a los demás a menos que sepamos lo que pensamos, puesto que atribuir pensamientos a los demás consiste en establecer una correspondencia entre el comportamiento verbal y otros tipos de comportamiento de los demás con nuestras propias proposiciones u oraciones con significado. De este modo, el conocimiento de nuestras propias mentes y el conocimiento de las mentes de los demás dependen el uno del otro.

Debería estar claro ahora qué es lo que asegura que el modo en que concebimos el mundo es en gran parte correcto en sus rasgos más simples. La razón es que los estímulos que causan nuestras respuestas verbales más básicas determinan también lo que esas respuestas verbales significan y el contenido de las creencias que las acompañan. La naturaleza de la interpretación garantiza tanto que un gran número de nuestras creencias perceptuales más simples sean verdaderas como que la naturaleza de esas creencias le sea conocida a los demás. Es evidente que a muchas creencias se les da contenido por sus relaciones con otras creencias, o que están causadas por sensaciones engañosas; cualquier 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

² A. J. Ayer, *The Problem of Knowledge*, 54.

nuestras creencias y las hace inteligibles, tanto si son verdaderas como si son falsas.

El supuesto de que la verdad de lo que creemos es lógicamente independiente de qué es lo que creemos se revela como ambiguo. Es cierto que cualquier creencia particular puede ser falsa; pero en el marco y en el tejido de nuestras creencias debe haber lo suficiente que sea verdadero como para proporcionar contenido a lo restante. Las conexiones conceptuales entre nuestro conocimiento de nuestras propias mentes y nuestro conocimiento del mundo de la naturaleza no son de naturaleza definicional, sino que son holísticas. Lo mismo vale para las conexiones conceptuales entre nuestro conocimiento de la conducta y nuestro conocimiento de las otras mentes.

Por tanto, no existen «barreras», sean éstas lógicas o epistémicas, entre las tres variedades de conocimiento. Por otro lado, el modo mismo en que cada una depende de las demás muestra por qué ninguna puede ser eliminada o reducida a las otras.

Como se ha hecho notar anteriormente, concebimos a un intérprete que pretende comprender a un hablante como alguien que establece correspondencias entre oraciones propias y preferencias y estados mentales del hablante. La totalidad de las evidencias de que pueda disponer un intérprete no determina una teoría de la verdad única para un hablante dado, y no sólo porque las evidencias de que realmente se dispone son finitas en número mientras que la teoría tiene una infinidad de consecuencias contrastables, sino porque todas las evidencias posibles no pueden reducir a uno el número de teorías aceptables. Dada la riqueza de la estructura representada por el conjunto de las propias oraciones y la naturaleza de las conexiones entre los miembros de este conjunto y el mundo, no debería sorprendernos que haya muchas maneras de asignar nuestras propias oraciones a las oraciones y pensamientos de otro que captan todo lo que es significativo.

La situación es parecida a la de la medición del peso o la temperatura mediante la asignación de números a objetos. Incluso suponiendo que no hubiera errores de medida y que se hicieran todas las observaciones posibles, una asignación de

números a objetos que registre correctamente sus pesos no es única; dada una asignación así, se puede producir otra multiplicando todos los números por una constante positiva. En el caso de la temperatura ordinaria (no la temperatura absoluta), cualquier asignación correcta de números puede convertirse en otra mediante una transformación lineal. Puesto que existen muchas maneras distintas pero igualmente aceptables de interpretar a un agente, si queremos podemos decir que la interpretación o la traducción es indeterminada, o que no hay asunto objetivo respecto a qué es lo que alguien quiere decir con las palabras que profiere. De la misma forma podríamos hablar de la indeterminación del peso o la temperatura. Pero normalmente acentuamos lo positivo al dejar claro qué es lo que permanece invariante en la transformación de una asignación de números a otras, pues lo que tiene significado empírico es lo que permanece invariante. Lo invariante *es* el asunto objetivo. Podemos permitirnos mirar la traducción y el contenido de los estados mentales a la misma luz³.

En el pasado creía que la indeterminación de la traducción proporcionaba una razón para suponer que no hay leyes estrictas que conecten los conceptos mentales y físicos, y que por ello daba apoyo a la afirmación de que los conceptos mentales no son tampoco reducibles nomológicamente a los conceptos físicos. Estaba equivocado; la indeterminación se presenta en ambos dominios. Pero una fuente de indeterminación en el caso de lo mental es que la línea que separa las verdades empíricas y las verdades que se deben al significado en general no puede definirse claramente sobre fundamentos conceptuales, y los fundamentos conceptuales son todo lo que tenemos para determinar lo que los hablantes quieren decir. Es aquí donde comienza a salir a la superficie la diferencia irreducible entre los conceptos mentales y los conceptos físicos; los primeros requieren —al menos en la medida en que son intencionales por naturaleza— que el intérprete considere cuál es la mejor manera de hacer inteligible a la criatura que

³ Acepto aquí la tesis quineana de la indeterminación de la traducción y la extendiendo a la interpretación del pensamiento en general. La analogía con la medición es mía.

se está interpretando, es decir, como una criatura dotada de razón. Como consecuencia, un intérprete debe separar el significado de la opinión en parte sobre bases normativas, decidiendo qué es lo que, desde su punto de vista, maximiza la inteligibilidad. En esta empresa el intérprete no tiene, desde luego, ninguna otra medida o norma de racionalidad a la que recurrir que la suyas propias. Cuando tratamos de entender el mundo como lo hacen los físicos utilizamos necesariamente nuestras propias normas, pero no pretendemos descubrir racionalidad en los fenómenos.

¿De qué modo impide el elemento normativo de los conceptos mentales su reducción a conceptos físicos? Posiblemente es obvio que una reducción mediante definiciones es imposible; pero ¿por qué no puede haber leyes —leyes estrictas— que conecten cada acaecimiento o estado mental con acaecimientos o estados descritos en el vocabulario de una física avanzada? Al escribir acerca de esto hace veinte años dije, efectivamente, que únicamente pueden esperarse leyes de conexión estrictas cuando los conceptos que las leyes conectan se basan en criterios del mismo tipo, de modo que una ley estricta no podría combinar conceptos normativos y no-normativos⁴. Esta respuesta todavía me parece correcta hasta donde alcanza, pero es comprensible que los que la han examinado críticamente la hayan encontrado poco concluyente. Quisiera ahora añadir algunas otras consideraciones.

Una consideración a añadir es la siguiente: las leyes estrictas no utilizan conceptos causales, mientras que la mayoría, si no es que todos, los conceptos mentales son irreduciblemente causales. Por ejemplo, una acción debe ser intencional bajo alguna descripción, pero es intencional sólo si está causada por factores mentales, como las creencias y deseos. Las creencias y deseos se identifican en parte por el tipo de acciones que son propensas a causar, supuestas las condiciones adecuadas. Muchos de los conceptos que figuran en las explicaciones de sentido común son causales del modo descrito. Un accidente fue causado por el hecho de que la carretera estaba

⁴ En «Mental Events», Ensayo 11 de *Essays on Actions and Events*.

resbaladiza; algo es resbaladizo si causa que los objetos apropiados resbalen en circunstancias apropiadas. Explicamos por qué el ala de un avión no se rompe al combarse haciendo notar que está hecha de materiales elásticos; un material es elástico si existe algo en él que hace que, en circunstancias apropiadas, vuelva a su forma original después de su deformación. Tales explicaciones no se prestan a ser precisadas por dos razones: no podemos formular en detalle cuándo son apropiadas las circunstancias y la apelación a la causalidad pasa por alto lo que una explicación a escala real pondría de manifiesto. Las descripciones de los objetos, estados y acaecimientos que se requieren para ejemplificar leyes estrictas, sin excepciones, no contienen conceptos causales (lo que no quiere decir que leyes que únicamente contengan conceptos no-causales no sean leyes causales).

En el caso de propiedades causales como la elasticidad, la deslizabilidad, la maleabilidad o la solubilidad tendemos a pensar correcta o incorrectamente que lo que dejan sin explicar puede ser explicado (o ya lo ha sido) con el avance de la ciencia. No estaríamos cambiando de tema si abandonáramos el concepto de elasticidad en favor de una especificación de la microestructura de los materiales del ala del avión que hace que vuelva a su forma original cuando se ve expuesta a determinadas fuerzas. Los conceptos y explicaciones mentales no son así. Apelan a la causalidad porque, como el propio concepto de causalidad, están diseñados para singularizar, de entre todas, las circunstancias que conspiran para causar un acaecimiento dado, precisamente los factores que satisfacen algún interés explicativo particular. Por ejemplo, cuando queremos explicar una acción, queremos conocer las razones del agente, de manera que podamos ver por nosotros mismos qué había en esa acción que la hizo atractiva para el agente. Pero sería una tontería suponer que existen leyes estrictas que estipulan que siempre que un agente tiene ciertas razones, ese agente llevará a cabo una determinada acción.

Las propiedades normativas y las propiedades causales de los conceptos mentales están relacionadas. Si hubiéramos de abandonar el aspecto normativo de las explicaciones psicológicas, no servirían ya para los propósitos para los que sirven.

Tienen un interés tan vivo por las razones para actuar del agente y para sus cambios en las creencias que estamos dispuestos a aceptar el compromiso que suponen explicaciones que no se pueden ajustar perfectamente con las leyes de la física. La física, por otra parte, tiene como propósito aspirar a leyes tan completas y precisas como podamos hacerlas —una aspiración distinta—. El elemento causal de los conceptos mentales ayuda a compensar su falta de precisión; es parte del concepto de acción intencional el hecho de que sea causada y explicada mediante creencias y deseos; y es parte del concepto de creencia o de deseo que tiende a causar, y por tanto a explicar, acciones de determinados tipos.

Buena parte de lo que he dicho acerca de lo que distingue a los conceptos mentales de los conceptos de una física desarrollada podría decirse para distinguir los conceptos de muchas de las ciencias especiales, como la biología, la geología y la meteorología. De modo que incluso si estoy en lo cierto en que el carácter normativo y causal de los conceptos mentales los separa definicional y nomológicamente de los conceptos de una física desarrollada, podrá parecer que hay algún hecho más básico o fundamental que explica esa separación. Creo que sí existe.

En la mayoría de los casos el conocimiento de los contenidos de nuestras propias mentes debe ser trivial. La razón es que, dejando aparte casos especiales, el problema de la interpretación no puede darse. Cuando se me pregunta por los contenidos proposicionales de mi mente, he de utilizar mis propias oraciones. Usualmente la respuesta es absurdamente obvia: mi oración «La nieve es blanca», como mi pensamiento de que la nieve es blanca, es verdadera si y sólo si la nieve es blanca. Mi conocimiento de los contenidos de otra mente es posible, según he argumentado, únicamente en el contexto de una concepción del mundo compartida y correcta en líneas generales. Pero ese conocimiento difiere del conocimiento que tengo de mi propia mente, puesto que es necesariamente indirecto, al depender, entre otras cosas, de las correlaciones observadas entre la conducta hablada y el resto del comportamiento de una persona y de los acaecimientos en el entorno común.

La diferencia fundamental entre mi conocimiento de otra mente y del mundo físico compartido se origina en una fuente distinta. La comunicación y el conocimiento de las otras mentes que ella presupone es la base de nuestro concepto de objetividad, de nuestro reconocimiento de la distinción entre las creencias falsas y verdaderas. No hay ningún modo de salirse de esta norma o medida para comprobar si estamos en lo cierto, no más de lo que podemos comprobar si la barra de platino-iridio que se encuentra en la Oficina Internacional de Pesos y Medidas de Sèvres (Francia) pesa un kilogramo. (Esta comparación era válida cuando esa barra de Sèvres proporcionaba la definición de kilogramo.) Podemos, por descontado, recurrir a una tercera parte o a una cuarta para ampliar y asegurar las medidas o normas interpersonales de lo que es real, pero esto no conduce a algo intrínsecamente distinto, sino sólo a más de lo mismo.

He hablado antes de una analogía entre cómo se asignan los números para registrar las relaciones entre objetos con respecto a la temperatura o el peso y cómo usamos nuestras propias oraciones para identificar los contenidos de los pensamientos y las preferencias de los demás. Pero esta analogía es imperfecta; la naturaleza de la escala que se utiliza difiere en los dos casos. Para llegar a un acuerdo acerca de las propiedades de los números y el tipo de estructuras de la naturaleza que nos permiten representar numéricamente esas estructuras dependemos de nuestras interacciones lingüísticas con los demás. No podemos del mismo modo ponernos de acuerdo sobre la estructura de las oraciones o los pensamientos que usamos para explorar los pensamientos y significados de los demás, pues el intento de alcanzar un acuerdo así simplemente nos remite al proceso mismo de interpretación del cual depende todo acuerdo.

Sugiero que es aquí donde llegamos a las fuentes últimas de la diferencia entre la comprensión de las mentes y la comprensión del mundo como algo físico. Una comunidad de mentes está en la base del conocimiento y proporciona la medida de todas las cosas. No tiene ningún sentido cuestionar la adecuación de esta medida o buscar una norma o medida ulterior.

Hemos tratado extensamente de la inevitabilidad del aspecto objetivo de todo pensamiento. ¿Qué es lo que queda del aspecto subjetivo? Está claro que no hemos suprimido la diferencia entre autoconocimiento y conocimiento de las otras mentes; el primero sigue siendo directo y el segundo indirecto. Hemos remitido la objetividad a las intersecciones de puntos de vista —para cada persona, la relación entre sus propias reacciones al mundo y las reacciones de los demás—. Estas diferencias son reales. Nuestros pensamientos son «internos» y «subjetivos» debido a que sabemos cuáles son de un modo en que nadie más puede saberlo. Pero aunque la posesión de un pensamiento es necesariamente individual, su contenido no lo es. Los pensamientos que formamos y que nos ocupan se ubican conceptualmente en el mundo que habitamos y sabemos que habitamos junto con otros. También nuestros pensamientos acerca de nuestros propios estados mentales ocupan el mismo espacio conceptual y se ubican en el mismo mapa público.

La concepción filosófica de la objetividad está cargada con una historia y un conjunto de supuestos acerca de la naturaleza de la mente y el significado que separan el significado de una preferencia o el contenido de un pensamiento de cuestiones relativas a su realidad externa, «mi» mundo del mundo tal como les parece a otros. Esta concepción tan popular mantiene que lo subjetivo es prioritario con respecto a lo objetivo y que existe un mundo subjetivo anterior al conocimiento de la realidad externa. Es evidente que la panorámica del pensamiento y el significado que he esbozado aquí no deja espacio alguno para tal prioridad, puesto que predica el autoconocimiento acerca del conocimiento de las otras mentes y del mundo. Lo objetivo y lo intersubjetivo son, por tanto, esenciales a cualquier cosa que podamos llamar subjetividad y constituyen el contexto en el cual ésta toma forma. Collingwood lo dijo sucintamente:

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 mí⁵.

Puede parecer que si compartir una concepción general del mundo es una condición del pensamiento, las diferencias en cuanto al carácter intelectual e imaginativo entre las mentes y las culturas se perderá de vista. Si he dado esta impresión es porque he querido concentrarme en lo que me parece que es primario, y por tanto tiende a no ser notado: el necesario grado de comunidad esencial a la comprensión de otra persona y la medida en que tal comprensión proporciona el fundamento del concepto de verdad y realidad del cual depende todo pensamiento. Pero no quisiera sugerir que no podamos entender a aquellos con los cuales diferimos sobre vastas regiones de la opinión física y moral. Sucede también que la comprensión es asunto de grado: los demás pueden saber cosas que nosotros no sabemos ni quizá podamos saber. Lo que es seguro es que la claridad y efectividad de nuestros conceptos crece con el desarrollo de nuestra comprensión de los demás. No existen límites definidos respecto a lo lejos que puede llevarnos o nos llevará el diálogo.

A algunos filósofos les preocupa que si todo nuestro conocimiento, o al menos nuestro conocimiento proposicional, es objetivo perderemos contacto con un aspecto esencial de la realidad: nuestras perspectivas personales, privadas. Creo que esta preocupación carece de fundamento. Si estoy en lo cierto, nuestro conocimiento proposicional no tiene su base en lo impersonal sino en lo interpersonal. Así, cuando miramos al mundo natural que compartimos con otros, no perdemos contacto con nosotros mismos, sino que más bien reconocemos nuestra pertenencia a una sociedad de mentes. Si yo no supiera lo que otros piensan, no tendría pensamientos propios y por tanto no sabría qué es lo que pienso. Si no supiera qué es lo que pienso, carecería de la capacidad de calibrar los

⁵ R. G. Collingwood, *The Principles of Art*, 248.

pensamientos de los demás. Calibrar o estimar los pensamientos de los demás requiere que viva en el mismo mundo que ellos y que comparta muchas de sus reacciones a sus rasgos más sobresalientes, incluyendo sus valores. De modo que no hay peligro de que cuando vemos objetivamente el mundo perdamos contacto con nosotros mismos. Los tres tipos de conocimiento forman un trípode: si se perdiera alguna de las patas, ninguna de sus partes se mantendría en pie.

Lista de contenidos de los volúmenes de ensayos de Donald Davidson

VOLUMEN 1

Essays on Actions and Events

Intention and Action

1. Actions, Reasons, and Causes (1963)
2. How is Weakness of the Will Possible? (1969)
3. Agency (1971)
4. Freedom to Act (1973)
5. Intending (1978)

Event and Cause

6. The Logical Form of Action Sentences (1967)
7. Causal Relations (1967)
8. The Individuation of Events (1969)
9. Events as Particulars (1970)
10. Eternal *vs.* Ephemeral Events (1971)

Philosophy of Psychology

11. Mental Events (1970)
12. Psychology as Philosophy (1974)
13. The Material Mind (1973)
14. Hempel on Explaining Action (1976)
15. Hume's Cognitive Theory of Pride (1976)

Appendices

- A. Adverbs of Action (1985)
- B. Reply to Quine on Events (1985)

VOLUMEN 2

Inquiries into Truth and Interpretation

Truth and Meaning

- 1. Theories of Meaning and Learnable Languages (1965)
- 2. Truth and Meaning (1976)
- 3. True to the Facts (1969)
- 4. Semantics for Natural Languages (1970)
- 5. In Defence of Convention T (1973)

Applications

- 6. Quotation (1979)
- 7. On Saying That (1968)
- 8. Moods and Performances (1979)

Radical Interpretation

- 9. Radical Interpretation (1973)
- 10. Belief and the Basis of Meaning (1974)
- 11. Thought and Talk (1975)
- 12. Reply to Foster (1976)

Language and Reality

- 13. On the Very Idea of a Conceptual Scheme (1974)
- 14. The Method of Truth in Metaphysics (1977)
- 15. Reality Without Reference (1977)
- 16. The Inscrutability of Reference (1979)

Limits of the Literal

- 17. What Metaphors Mean (1978)
- 18. Communication and Convention (1982)

Appendix to Essay 10

- Replies to Quine and Lewis (1974)

VOLUMEN 3

Subjective, Intersubjective, Objective

Subjective

1. First Person Authority (1984)
2. Knowing One's Own Mind (1987)
3. The Mith of the Subjective (1988)
4. What is Present to the Mind (1989)
5. Indeterminism and Antirealism (1997)
6. The Irreducibility of the Concept of the Self (1998)

Intersubjective

7. Rational Animals (1982)
8. The Second Person (1992)
9. The Emergence of Thought (1997)

Objective

10. A Coherence Theory (1983)
Afterthoughts (1987)
11. Empirical Content (1982)
12. Epistemology and Truth (1988)
13. Epistemology Externalized (1990)
14. Three Varieties of Knowledge (1991)

VOLUMEN 4

Problems of Rationality

Rationality

1. The Problem of Objectivity (1995)
2. Expressing Evaluations (1984)
3. The Objectivity of Value (1995)
4. Judging Interpersonal Values (1996)
5. Objectivity and Practical Reason (1999)

Problems and Proposals

6. Turing's Test (1990)
7. Representation and Interpretation (1990)
8. Problems of Explaining Action (1987)

9. Could There be a Theory of Rationality? (1995)
10. What Thought Requires
11. The Unified Theory of Thought and Action (1980)

Irrationality

12. Paradoxes of Irrationality (1974)
13. Incoherence and Irrationality (1985)
14. Deception and Division (1986)
15. Who is Fooled? (1997)

VOLUMEN 5

Truth, Language and History

Truth

1. Truth Rehabilitated (2000)
2. The Folly of Trying to Define Truth (1996)
Replies to Rorty, Stroud, and McDowell
3. Meaning, Truth, and Evidence (1990)
4. Pursuit of the Concept of Truth (1995)
5. What is Quine's View of Truth? (1994)

Language

6. A Nice Derangement of Epitaphs (1986)
7. The Social Aspect of Language (1994)
8. Seeing Through Language (1997)
9. Joyce and Humpty Dumpty (1989)
10. The Third Man (1992)
11. Locating Literary Language (1993)

Historical Thoughts

12. Plato's Philosopher (1985)
13. The Socratic Concept of Truth (1992)
14. Dialectic and Dialogue (1994)
15. Gadamer and Plato's *Philebus* (1997)
16. Aristotle's Action
17. Method and Metaphysics (1993)

Referencias bibliográficas

- AGASSI, Joseph, *Science in Flux*, Dordrecht, Reidel, 1975.
- ALSTON, William, «Varieties of Privileged Access», *American Philosophical Quarterly*, 9 (1971).
- AVRAMIDES, Anita, «Davidson and the New Sceptical Problem», en U. Zeglen (ed.), *Donald Davidson: Truth, Meaning and Knowledge*, Londres, Routledge, 1999.
- AYER, A. J., *The Problem of Knowledge*, Londres, Macmillan, 1956.
- (ed.), *Logical Positivism*, Nueva York, Free Press, 1959.
- «Privacy», en *The Concept of a Person and Other Essays*, Nueva York, Macmillan, 1963.
- BENNETT, Jonathan, *Linguistic Behavior*, Cambridge, Cambridge University Press, 1976.
- BLACK, Max, «The semantic Definition of Truth», *Analysis*, 8 (1948).
- BRENTANO, Franz, *Psychology from an Empirical Standpoint*, ed. O. Kraus, trad. A. Rancurello, D. B. Terrell, y L. L. McAlister, Londres, Humanities Press, 1973.
- BURGE, Tyler, «Individualism and the Mental», en P. A. French, T. E. Uehling y H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, Minneapolis, University of Minnesota Press, 1979.
- «Two Thought Experiments Reviewed», *Notre Dame Journal of Formal Logic*, 23 (1982).
- «Other Bodies», en A. Woodfield (ed.), *Thought and Object*, Oxford, Oxford University Press, 1982.
- «Individualism and Psychology», *Philosophical Review*, 95 (1986).
- «Cartesian Error and the Objectivity of Perception», en P. Petit y J. McDowell (eds.), *Subject, Thought, and Context*, Oxford, Oxford University Press, 1986.
- «Intellectual Norms and Foundations of Mind», *Journal of Philosophy*, 83 (1986).

- «Perceptual Individualism and Authoritative Self-Knowledge», en R. Grimm y D. Merrill (eds.), *Contents of Thought*, Tucson, University of Arizona Press, 1988.
- «Individualism and Self-Knowledge», *Journal of Philosophy*, 85 (1988).
- CARNAP, Rudolf, «Psychology in Physical Language», *Erkenntnis*, 3 (1932-3).
- CHAN, Wai Pang, PRETE, Frederick y H. DICKINSON, Michael, «Visual Input to the Efferent Control System of a Fly's "Gyroscope"», *Science*, 280 (10 de abril de 1998).
- CHOMSKY, Noam, *Knowledge of Language: Its Nature, Origin and Use*, Nueva York, Praeger, 1986.
- *Language and Problems of Knowledge*, Cambridge, Mass., MIT Press, 1988.
- CHURCHLAND, Paul, *Scientific Realism and the Plasticity of Mind*, Cambridge, Cambridge University Press, 1979.
- COLLINGWOOD, R. G., *The Principles of Art*, Oxford, Oxford University Press, 1938.
- DAVIDSON, Donald, *Essays on Actions and Events*, Oxford, Oxford University Press, 1980.
- *Inquiries into Truth and Interpretation*, Oxford, Oxford University Press, 1980.
- «A Nice Derangement of Epitaphs», en R. Grandy y R. Warner (eds.), *Philosophical Grounds of Rationality*, Oxford, Oxford University Press, 1986.
- «Meaning, Truth and Evidence», en R. Barrett y R. Gibson (eds.), *Perspectives on Quine*, Oxford, Blackwell, 1990.
- «The Structure and Content of Truth», *Journal of Philosophy*, 87 (1990).
- «Three Varieties of Knowledge», en A. P. Griffiths (ed.), *A. J. Ayer Memorial Essays*, Cambridge, Cambridge University Press, 1991.
- «Pursuit of the Concept of Truth», en P. Leonardi y M. Santambrogio (eds.), *On Quine*, Cambridge, Cambridge University Press, 1995).
- «Replies», *Critica*, 30 (1998).
- DENNETT, Daniel, «Beyond Belief», en A. Woodfield (ed.), *Thought and Object*, Oxford, Oxford University Press, 1982.
- «Real Patterns», *Journal of Philosophy*, 88 (1991).
- DUMMETT, Michael, *Truth and Other Enigmas*, Londres, Duckworth, 1978.
- *The Interpretation of Frege's Philosophy*, Londres, Duckworth, 1981.
- EVANS, Gareth, *The Varieties of Reference*, Oxford, Oxford University Press, 1982.

- FODOR, Jerry, «Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology», *Behavioral and Brain Sciences*, 3 (1980).
- «Cognitive Science and the Twin Earth Problem», *Notre Dame Journal of Formal Logic*, 23 (1982).
- *Psychosemantics*, Cambridge, Mass., MIT Press, 1987.
- *A Theory of Content*, Cambridge, Mass., MIT Press, 1990.
- FØLLESDAL, Dagfinn, «Knowledge, Identity, and Existence», *Theoria*, 33 (1967).
- GOODAMAN, Nelson, *Fact, Fiction, and Forecast*, 4.^a ed., Cambridge, Mass., Harvard University Press, 1983.
- GUHRAUER, G., *Gottfried Wilhelm Freiherr von Leibniz: Eine Biographie*, Hildesheim, Gb. Olms, 1966.
- HAHN, Lewis E. (ed.), *The Philosophy of Donald Davidson*, Library of Living Philosophers, Chicago, Open Court, 1999.
- y SCHILPP, P. A. (eds.), *The Philosophy of W. V. Quine*, Library of Living Philosophers, Chicago, Open Court, 1986.
- HEMPEL, Carl, «On the Logical Positivist's Theory of Truth», *Analysis*, 2 (1935).
- «The Empiricist Criterion of Meaning», *Revue Internationale de Philosophie*, 4 (1950).
- HENRICH, Dieter (ed.), *Kant oder Hegel?*, Stuttgart, Klett-Kotta, 1983.
- JEFFREY, Richard, *The Logic Decision*, Chicago, University of Chicago Press, 1965.
- KAPLAN, David, «Quantifying In», *Synthese*, 19 (1968).
- KRIPKE, Saul, «Naming and Necessity», en D. Davidson y G. Harman (eds.), *Semantics of Natural Language*, Dordrecht, Reidel, 1972.
- *Wittgenstein on Rules and Private Language*, Cambridge, Blackwell, 1982.
- LEPORE, Ernie (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Cambridge, Blackwell, 1986.
- LEWIS, C. I., *Mind and the World Order*, Nueva York, Scribner's, 1960.
- LEWIS, David, «Languages and Language», en K. Gunderson (ed.), *Language, Mind and Knowledge*, Minneapolis, University of Minnesota Press, 1975.
- MALACHOWSKI, Alan (ed.), *Reading Rorty*, Oxford, Blackwell, 1990.
- MALCOLM, Norman, «Thoughtless Brutes», *Proceedings and Addresses of the American Philosophical Association*, 46 (1972-3).
- MARGALIT, Avishai (ed.), *Meaning and Use*, Dordrecht, Reidel, 1979.
- MATES, Benson, *The Philosophy of Leibniz: Metaphysics and Language*, Oxford, Oxford University Press, 1986.
- NAGEL, Thomas, «The Objective Self», en C. Genet y S. Shoemaker (eds.), *Mind and Knowledge*, Oxford, Oxford University Press, 1983.

- NEURATH, Otto, «Protocol Sentences», *Erkenntnis*, 3 (1932-2).
- PUTNAM, Hilary, «The Meaning of "Meaning"», en *Philosophical Papers*, ii: *Mind, Language, and Reality*, Cambridge, Cambridge University Press, 1975. [Traducción al castellano en L. Ml. Valdés Villanueva (ed.), *La búsqueda del significado*, Madrid, Tecnos, 1991.]
- *Meaning and the Moral Sciences*, Londres, Routledge & Kegan Paul, 1978.
- *Realism and Reason*, Cambridge, Cambridge University Press, 1983.
- QUINE, W. V., «Quantifiers and Propositional Attitudes», *Journal of Philosophy*, 53 (1956).
- *Word and Object*, Cambridge, Mass., MIT Press, 1960. [Traducción al castellano, *Palabra y objeto*, Barcelona, Labor, 1968.]
- *Ontological Relativity and Other Essays*, Nueva York, Columbia University Press, 1969.
- *The Roots of Reference*, La Salle, Ill., Open Court, 1974.
- «The Nature of Natural Knowledge», en S. Guttenplan (ed.), *Mind and Language*, Oxford, Oxford University Press, 1975.
- «On Empirically Equivalent Systems of the World», *Erkenntnis*, 9 (1975).
- «Intensions Revisited», en *Theories and Things*, Cambridge, Mass., Harvard University Press, 1981.
- *Theories and Things*, Cambridge, Mass., Harvard University Press, 1981.
- «Reply to Stroud», en P. A. French, T. E. Uehling y H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, Minneapolis, University of Minnesota Press, 1981.
- «Events and Reification», en E. Lepore y B. McLaughlin (eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Cambridge, Blackwell, 1985.
- *Pursuit of Truth*, ed. rev., Cambridge, Mass., Harvard University Press, 1992.
- *From Stimulus to Science*, Cambridge, Mass., Harvard University Press, 1995.
- y GOODMAN, Nelson, «Steps toward a Constructive Nominalism», *Journal of Symbolic Logic*, 12 (1947), 97-122.
- RAMSEY, Frank, «Truth and Probability», en *Philosophical Papers: F. P. Ramsey*, ed. D. H. Mellor, Cambridge, Cambridge University Press, 1990.
- RORTY, Richard, «Incorrigibility as the Mark of the Mental», *Journal of Philosophy*, 67 (1970).
- *Philosophy and the Mirror of Nature*, Princeton, Princeton University Press, 1979.
- *Consequences of Pragmatism*, Minneapolis, University of Minnesota Press, 1982.

- «Pragmatism, Davidson and Truth», en E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford, Blackwell, 1986.
- RYLE, Gilbert, *The Concept of Mind*, Nueva York, Barnes & Noble, 1949. [Traducción al castellano, *El concepto de lo mental*, Buenos Aires, Paidós, 1967.]
- SCHIFFER, Stephen, *Remnants of Meaning*, Cambridge, Mass., MIT Press, 1987.
- SCHLICK, Moritz, «The Turning Point in Philosophy», *Erkenntnis*, 1 (1930-1).
- «Positivism and Realism», *Erkenntnis*, 3 (1932-3).
- «The Foundation of Knowledge», *Erkenntnis*, 4 (1934).
- «Meaning and Verification», *Philosophical Review*, 4 (1936).
- SEARLE, John, *Intentionality*, Cambridge, Cambridge University Press, 1983.
- «Indeterminacy and the First Person», *Journal of Philosophy*, 84 (1987).
- SELLARS, Wilfred, «Empiricism and the Philosophy of Mind», en H. Feigl y M. Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, Minneapolis, University of Minnesota Press, 1956.
- SHOEMAKER, Sidney, *Self-Knowledge and Self-Identity*, Ithaca, NY, Cornell University Press, 1963.
- STICH, Stephen, «Autonomous Psychology and the Belief-Desire Thesis», *Monist*, 61 (1978).
- *From Folk Psychology to Cognitive Science*, Cambridge, Mass., MIT Press, 1983.
- STRAWSON, Peter, *Individuals*, Londres, Methuen, 1959.
- «Truth», en *Logico-Linguistic Papers*, Londres, Methuen, 1971.
- TARSKI, Alfred, «The Concept of Truth in Formalized Languages», en J. H. Woodger (ed.), *Logic, Semantics, Metamathematics*, Oxford, Oxford University Press, 1956.
- WALLAS, Graham, *The Art of Thought*, Nueva York, Harcourt Brace, 1926.
- WEISS, Donald, «Professor Malcolm on Animal Intelligence», *Philosophical Review*, 84 (1975).
- WITTGENSTEIN, Ludwig, *Philosophical Investigations*, Nueva York, Macmillan, 1953. [Traducción al castellano, *Investigaciones filosóficas*, Barcelona, Crítica, 1988.]
- WOODFIELD, Andrew (ed.), *Thought and Object*, Oxford University Press, 1982.
- WRIGHT, Crispin, «Kripke's Account of the Argument against Private Language», *Journal of Philosophy*, 81 (1984), 759-78.

Índice analítico

- acceso privilegiado, 28, 30
acción intencional, 146, 180-1
actitudes proposicionales, 47, 48, 53, 54-5, 148, 207, 287-8
animales y, 97-8, 149-50, 177-8
y autoridad de primera persona, 25-6, 31, 63, 64, 65
y ciencia, 111-3
contenido de, 61
e indeterminismo, 111, 112, 115
e intersubjetividad, 175
naturaleza de, 101-2
y pensamientos, 146, 150
poseedores de, 141, 150-1, 153
racionalidad y, 141, 142-3
y significado, 80
subjetividad de, 102-3
véanse también creencias; deseos; intenciones; estados mentales
adscripciones de actitudes proposicionales:
asimetría entre, 25, 28-9, 30-1, 32, 33-6, 44-5, 282
primera persona, 26-9, 30, 31, 33
tercera persona, 31, 32
véase también actitudes proposicionales;
Agassi, J., 30
Alston, W., 34
American Philosophical Association, 216
analítico-sintético, 75, 122, 203-4, 242, 268
antiindividualismo, 270, 272-7
antirrealismo, 257-8
Dummett y, 117, 257
y el indeterminismo, 110-29
y el realismo, 111, 247, 255, 263
antisubjetivismo, 81-2
aprendizaje del lenguaje, 77-9, 84-5, 131, 134-5, 157, 164, 173, 278
asentimiento provocado, 207
aseverabilidad justificada, 259
autoridad de primera persona, 25-40, 44, 48, 50, 53, 54, 61, 71, 87, 102, 116, 279
alcance de, 51
y antisubjetivismo, 84
y contenidos mentales, 66
y estados mentales, 71
existencia de, 52
y externismo, 66, 271-2, 274
e indeterminación, 124
Ayer, A. J., 29n, 226, 291
Bennett, J., 149n
Black, M., 247n
Boyd, R., 245

- Brentano, F., 109, 115
 Burge, T., 48, 52, 53, 55-61, 64-5,
 70, 84, 270, 272-6, 278
- Carnap, R., 98, 117, 223, 224, 232,
 235, 240, 242, 243
- Chomsky, N., 118, 166n
- Church, A., 254
- Churchland, P., 100n
- ciencias naturales, 112, 133, 176
véase también física
- cientifismo, 112-3
- Círculo de Viena, 225, 228, 234
- Collingwood, R. G., 298, 299n
- comunicación, 38-40, 58n, 112, 163
- condiciones para, 90, 154-5,
 166-7, 213
- y objetividad, 286-7, 288-91, 297
- y pensamiento, 286-7
- conceptos, 42, 130, 176-9
- causales, 113, 295
- contenidos de, 88, 112, 270
- de uno mismo, 130-8
- mentales, 283-6
- conducta, 32, 88, 103, 180
- y conocimiento, 280, 281
- y creencias, 179
- e intensidad, 126
- conductismo, 52, 110, 147
- conocimiento empírico, 222, 225-6,
 227-8, 244
- conocimiento, 27, 53
- autoconocimiento, 41-71, 137-8,
 265, 270-1, 280-1, 283, 291,
 292, 296
- y creencia, 216, 234, 285
- del mundo externo, 75-6, 280-1,
 282, 291, 292
- de otras mentes, 41-2, 280-1,
 282, 291, 292, 286
- empírico, 82, 88, 222, 225-6,
 228-9, 230
- escepticismo sobre, 80, 205-6,
 281-4
- fundamentos de, 83, 235
- perceptual, 273-4
- proposicional, 288-300
- teoría de, 80, 81, 234
- teorías coherentistas del, 193-
 221, 230, 236
- tres variedades de, 133-4, 280-
 300
- contenido empírico, 222-44
- Convención T (Tarski), 196
- creencias, 26, 45-6, 52, 105, 126-7,
 241, 291-2
- animales y, 144-5
- condiciones para, 150-3
- y conocimiento, 195-6, 199,
 215-6, 234, 285-6
- contenido proposicional de,
 101, 177-8, 184-5
- contenidos de, 101, 177-8, 184-5
- y deseos, 179-80
- falsa, posibilidad de, 267-70,
 291-2
- justificación de, 202-3
- perceptual, 80, 261-2, 280
- y sensaciones, 200-1
- y significado, 207-11, 214-5
- y sorpresa, 152-3
- y verdad, 146, 194-5, 196, 205-
 6, 214-5, 242, 243-4, 245,
 246, 255-6, 259-60, 285
- cuantificación, 121, 220
- decisión, teoría de la, 179, 195
- Dennett, D., 51, 52n, 53, 70, 103,
 115, 125, 126, 127-8
- Descartes, R., 47
- deseos, 26, 27, 45, 127
- Dewey, J., 253
- dicotomía esquema-contenido, 74-6,
 78
- dudas, 27, 80-1
- Duhem, P., 30
- Dummett, M., 93n, 109, 117, 203,
 245, 247n, 259-60, 263

- empirismo, 33, 82, 110, 201n, 224, 246, 266
- epistemología, 18-20, 82, 83, 203, 234
- externización de, 265-79
- naturalizada, 76, 220, 265-6
- error, 26, 38, 45, 60, 80-1, 83, 95, 136, 213-4, 274
- escepticismo, 80-1, 203, 205, 208, 220-1, 275-6
- sobre otras mentes, 31-3, 36, 82-3, 200, 284
- estados mentales, 46-8, 68-9, 73, 84-5, 88-9, 90, 100, 102-8, 110, 282
- y la autoridad de primera persona, 71, 108
- contenidos de, 47-8, 87, 278-9
- inconscientes, 30
- restringidos, 86-7
- estados psicológicos, 47, 53, 86
- restringidos, 47-8, 53, 61, 86
- véase también* externismo
- estímulos:
- respuesta a, 153-4, 170-5, 290-1
- triangulación de, 128, 185
- Evans, G., 54n, 70, 94
- evidencias, 225-31, 232-3, 242-3, 261
- externismo, 52, 61, 68, 79-80, 90, 270
- y la autoridad de primera persona, 270-1, 274-5
- perceptual, 272, 273-6
- social, 272-3
- física, 54, 64, 102-3, 112-5, 128, 176, 181, 293-7
- contenido empírico y, 222-44
- Fodor, J., 50, 52, 53, 69, 70, 87, 103, 107, 115
- Føllesdal, D., 94
- Frege, G., 51, 91, 254
- Freud, S., 30
- Goodman, N., 226, 245
- Grice, H. P., 163
- Hempel, C., 223, 224-5, 226-7, 230, 236, 237, 239-40, 241, 253
- Henrich, D., 130
- holismo, 83, 107, 203
- de lo mental, 144-6, 176-7, 180-1
- Hombre del pantano, 46-7
- Hume, D., 70, 92, 199
- identidad, teorías, 63-5
- indeterminación, 119, 121, 293
- y antirrealismo, 110-29
- y actitudes proposicionales, 112
- y la autoridad de primera persona, 124-5
- de la interpretación, 112, 115-6, 293
- de la traducción, 115-6, 118, 188, 204, 293
- tesis de, 115-6, 124-5
- inescrutabilidad de la referencia:
- tesis de, 121-3
- intenciones, 27, 163, 166, 168-9, 179-81
- e interpretación, 162-3
- y significado, 59, 163
- intensionalidad, 94, 147
- interacción social:
- y aparición del pensamiento, 184-5
- interpretación, 37-40, 59, 123, 129, 134-8, 160-2, 166, 211-2, 244, 269
- indeterminación de, 112, 115-6, 293
- intención y, 161-2
- radical, 208, 212
- intersubjetividad:
- y actitudes proposicionales, 174
- del lenguaje, 242-3
- y objetividad, 128, 138, 154
- véase también* triangulación

- irracionalidad, 146-7
véase también racionalidad
- justificación de las creencias, 199-202, 212
- Kant, I., 74
- Kaplan, D., 70, 94, 103
- Kripke, S., 94, 161, 163n, 164, 166, 167n, 169, 175, 213n, 286n
- lenguajes, 34, 188-90, 263
 - intersubjetividad de, 242
 - metalenguaje y, 124, 248, 251
 - naturaleza abstracta de, 156-8
 - y pensamiento, 148-50, 184, 285
 - privado, 166-9, 175, 286-7
 - productividad, 189-90
 - y reglas, 163-5
 - y teoría semántica, 185-6
 - y verdad, 247, 248, 259-60
- Lewis, C. I., 74, 254
- Lewis, D., 158n
- Malcolm, N., 143, 148, 151-2
- Malinowski, B., 30
- materialismo, 110
- medición, metrización:
 - fundamental, 186-8, 190
 - numérica, 99-100, 106-7, 117-8
- mentes:
 - conocimiento de la propia, 41-71, 137-8
 - conocimiento de otras, 281-3, 290-1, 295
 - contenidos de, 72-90, 137-8, 280-1, 282, 290, 291, 295
 - véase también* estados mentales
- monismo, 65, 114
- monismo anómalo, 65, 114
- Motherwell, R., 41, 44
- Nagel, T., 131
- Neurath, O., 19, 222-3, 224, 225, 232-3, 235-6, 238, 240-3, 253
- objetividad, 129, 134, 183, 238, 277-8, 298
 - y comunicación, 286-7, 288-9, 287
 - e intersubjetividad, 128, 138
- objetos abstractos, 91-2, 156-8
- objetos internos, 69, 108
 - relación con los objetos externos, 109
- objetos mentales, 70-1, 91-109
- observabilidad, suposición de la, 147
- oraciones de creencia, 69, 94, 96-7, 116-7, 119-20, 122-3
- oraciones observacionales, 201n, 204, 210, 212, 224, 226, 231, 232, 237-8, 257
- oraciones ocasionales, 212-3
- oraciones protocolares, 222-9, 230, 231-2, 233-6, 237-44
- oraciones teóricas, 257
- oraciones, 128-9, 136, 159
 - interpretación y, 209, 237
 - verdad y, 250-2, 253-5, 256-7, 258-63
- ostensión, 131-4, 277-8
- Parménides, 110
- Peirce, C. S., 245
- pensamientos, 44, 85, 90, 184
 - aparición de, 176-90
 - y comunicación, 286-7
 - contenidos de, 50, 52-4
 - y creencia, 145
 - y lenguaje, 148-50, 185, 285-6
 - y objetos mentales, 69-70, 91-109
- percepciones, 19, 26
 - y sensación, 82, 199-200, 228-9
- Platón, 110, 219
- positivismo lógico:
 - y contenido empírico de las oraciones, 222-3, 234-5, 237
- predicciones, 126-7, 262
- Principio de Coherencia, 288

- Principio de correspondencia, 288
 principios de caridad, 208-9, 287
 preferencias, 34, 104-5, 106, 119-20, 123, 158
 contenido proposicional de, 287
 y creencias, 105-6
 verdad de, 111
- Problema de la Tierra Gemela, 46, 48-9, 50, 52, 59, 61-2, 101-2
- psicología, 53-4, 87-8, 112-3, 240
- psicología del sentido común, 53-4, 103
- Putnam, H., 19, 45-9, 52-3, 59-63, 65, 69-70, 78, 84, 87, 101, 197, 213n, 216, 219, 245, 258, 259, 272
- Quine, W. V. O., 69, 76, 94n, 95, 98, 107, 112, 114-6, 118-9, 120, 121, 125, 160, 201, 203-4, 207-10, 212-4, 220, 241, 245, 257-8, 266, 288
- racionalidad, 17, 135, 155, 178
 y actitudes proposicionales, 141, 142-5
 y animales, 17, 141-55
- realismo, 113, 197, 253-4, 258, 264
 y antirrealismo, 112, 246, 254, 264
 interno, 197, 245, 258-9
 metafísico, 197, 259
 transcendental, 246
- referencia, 117
 de re y *de dicto*, 94
 y oraciones-T, 247-8
 teoría de la inescrutabilidad, 121-2
- relativismo conceptual, 72-4, 83
- representación, 81-2, 90
- Rorty, R., 19, 32, 198, 216, 217, 218, 219, 220, 253
- Russell, B., 45, 94-5
- Ryle, G., 28-9, 41, 44
- satisfacción, 121, 196
 relación de Tarski de, 119n, 157, 190, 247-8
- Schiffer, S., 97n
- Schlick, M., 19, 222-6, 230-9, 241-2
- Searle, J., 47-8, 50, 106, 271n
- Sellars, W., 19, 32n
- semántica composicional, 97n
- sensaciones, 19, 26, 33
 y creencias, 198-201, 204-5
 y percepciones, 82, 199-200, 228-9
- Shoemaker, S., 34
- significado, 45-51, 56-8, 71, 124, 164, 203-5
 y creencia, 203-5, 206-8, 208-9
 e intención, 59, 163
 Putnam sobre, 45, 46-9
 y verdad, 193-4, 204, 259-60
- significado, teorías del, 203-4
- solipsismo metodológico, 53, 103, 223
- sorpresa:
 y creencia, 151-2
- Stich, S., 46n, 53, 70, 89n, 103
- Strawson, P., 31-2, 254
- subjetividad, 102, 298-9
 mito de, 72-90
- subjetivismo, 256
- Tarski, A., 119n, 157, 187, 190, 218, 224
- teorías coherentistas, 187, 217, 223, 228, 236
 del conocimiento, 193-221, 230, 235
 Neurath y, 230-4, 241-2
 y la verdad, 193-221, 230, 235, 245, 255
- teorías de la correspondencia, 111, 119n, 193
 sobre la verdad, 82, 196-7, 216-7, 246-8

- teoría semántica, 186-90
 teoría verificacionista, 223
 teorías de la verdad de estilo tarskiano, 119n, 217, 246-8, 249-50, 259
 traducción:
 indeterminación de, 115-6, 118, 188, 204, 283
 radical, 208, 212
 triangulación, 18, 20, 128, 132-4, 154, 173-5, 184-5, 277-8, 280-1
 y aparición del pensamiento, 183-6

 verdad, 155, 219, 224, 258-9
 concepción objetiva de, 183, 256
 concepto de, 111, 128, 155, 253, 263
 definiciones de, 247-9
 epistemología y, 245-64
 inmanencia de, 246, 257
 y lenguaje, 247, 248
 Putnam sobre, 258-9
 y significado, 193-4, 204, 259-60
 teorías coherentistas de, 193-221, 230, 235, 245, 255
 teorías de la correspondencia, 82, 196-7, 216-7, 246-8
 teorías del estilo de Tarski de, 118n, 217, 246-8, 248-50, 259
 verificacionismo, 110, 146-7, 204

 Wallas, G., 41, 44
 Weiss, D., 152
 Wittgenstein, L., 17, 42-3, 128, 156, 161, 164-5, 168, 184, 285
 Woodfield, A., 50, 63, 65, 84, 271n
 Wright, C., 163n, 245, 259

 Yablo, S., 93
 Yo:
 concepto de, 130-8

NOTA FINAL

Le recordamos que este libro ha sido prestado gratuitamente para uso exclusivamente educacional bajo condición de ser destruido una vez leído. Si es así, destrúyalo en forma inmediata.

Súmese como voluntario o donante, para promover el crecimiento y la difusión de la Biblioteca



Para otras publicaciones visite
www.lecturasinegoismo.com
Referencia: 3644

DONALD Davidson es uno de los filósofos norteamericanos vivos cuya influencia filosófica desde la década de 1960 ha sido más profunda y amplia, siendo reconocido por la colectividad filosófica como un clásico contemporáneo.

Este volumen ofrece una selección de sus trabajos producidos desde 1980 a 1990 sobre el conocimiento, la mente y el lenguaje. Todos tenemos conocimiento de nuestras propias mentes (subjetivo), conocimiento del contenido de otras mentes (intersubjetivo) y conocimiento del espacio que compartimos (objetivo). Davidson analiza la naturaleza y el status de cada uno de estos tres tipos de conocimiento y las conexiones y diferencias entre ellos. En su recorrido nos habla también sobre la verdad, la racionalidad humana y las relaciones entre el lenguaje, el pensamiento y el mundo.

0111086

ISBN 84-376-2085-6



9 788437 620855