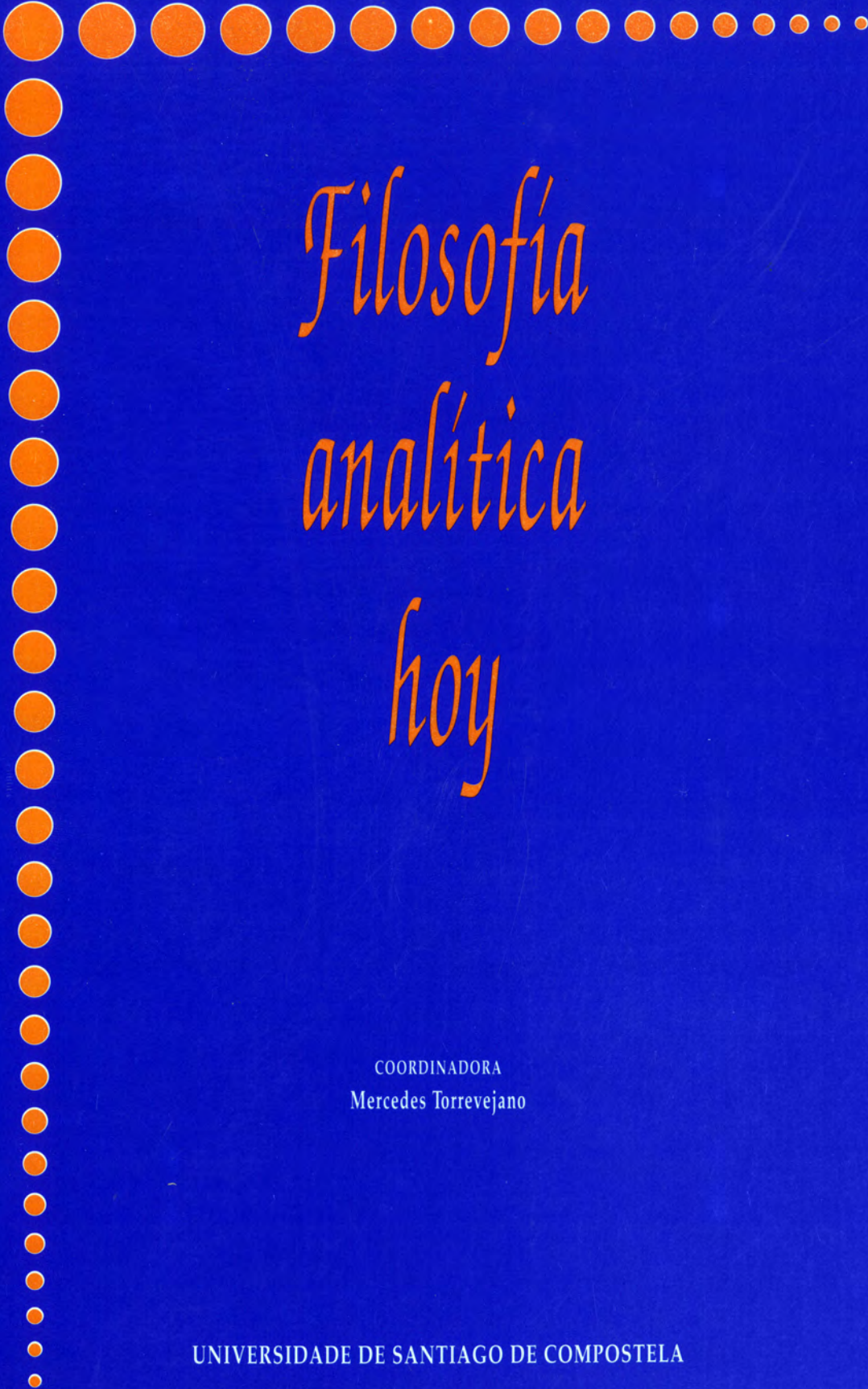


A decorative border of orange circles is located on the left and top edges of the cover. The circles on the left are of varying sizes and are arranged in a vertical line. The circles on the top are of varying sizes and are arranged in a horizontal line.

Filosofía analítica hoy

COORDINADORA
Mercedes Torrevejano

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

FILOSOFÍA ANALÍTICA HOY

Encuentro de tradiciones



**CURSOS E CONGRESOS DA
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
Nº 67**



FILOSOFÍA ANALÍTICA HOY

Encuentro de tradiciones

Coordinadora

MERCEDES TORREVEJANO

1991

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

**FILOSOFIA analítica hoy : Encuentro de tradiciones / Coordinadora: Mercedes Torre-
vejano. — Santiago de Compostela : Universidade. Servicio de Publicacións e Intercam-
bio Científico, 1991. — 284 p. ; 24 cm. — (Cursos e Congresos da Universidade de
Santiago de Compostela ; 67). — Índice: p. 5. — D.L.: C. 582-1991. — ISBN 84-7191-
722-X.**

**I. Torrevejano, Mercedes, coord.; II. Universidade de Santiago de Compostela. Servicio
de Publicacións e Intercambio Científico, ed.**

140



© Universidade de
Santiago de Compostela, 1991

EDITA: Servicio de Publicacións
e Intercambio Científico
Casa Gradín
Campus Universitario

IMPRIME: Imprenta Universitaria
Pavillón de Servicios
Campus Universitario

Dep. Leg.: 582-1991
ISBN: 84-7191-722-X

INDICE

Presentación	7
Después del análisis: significado, comprensión e intencionalidad	9
<i>Juan José Acero</i>	
La filosofía analítica y su concepción del método de la filosofía	27
<i>Ernesto Tugendhat</i>	
Sobre el pretendido realismo básico del Tractatus	35
<i>Alfonso G. Suárez</i>	
El juego de saber y el lenguaje de la certeza	49
<i>Angel Alvarez Gómez</i>	
La diferencia entre explicar e interpretar según Davidson	83
<i>Manuel García-Carpintero</i>	
Panorama actual de la filosofía analítica de la mente: funcionalismo y experiencia	121
<i>Vicente Sanfélix Vidarte</i>	
Escritura y "phone": el desplazamiento derridiano de la tradición analítica	155
<i>Jorge Pérez de Tudela Velasco</i>	
Putnam, o las perplejidades de la filosofía	177
<i>Mercedes Torrevejano</i>	
El "realismo" en la tradición analítica. El caso de Hilary Putnam	193
<i>J.M. Sagüillo</i>	
Antología de texto de Hilary Putnam	217
<i>Luis Villegas</i>	



PRESENTACION

El presente libro ha surgido de la conjunción de unas cuantas circunstancias de muy diversa índole, todas ellas felices, a mi parecer.

En primer lugar, la preocupación de un grupo de profesores de esta Universidad, por realizar, a lo largo del curso 1988-89 —tras una alentadora experiencia en el curso anterior— una serie de encuentros a modo de Seminario de Postgraduados, dedicados al diálogo y la reflexión filosófica en torno a algún tema actual y del interés de todos; un tema que nos permitiera examinar la tradición analítica en alguno de sus puntos de encuentro con otras tradiciones filosóficas. El plural interés cristalizó en la propuesta de lectura de algunos textos de Putnam.

Por otra parte, el gran colectivo de Profesores de Filosofía del Curso de Orientación Universitaria de Galicia, coordinados por quien suscribe, se propuso, a lo largo de ese mismo curso, realizar una revisión crítica y actualizada del tópico "filosofía analítica" tal como se integraba a la sazón en la programación del COU. Ello nos condujo a la realización de un gran ciclo de temas sobre la Filosofía analítica que, en tres etapas, se ocupó sucesivamente de estudiar: los orígenes del movimiento filosófico así llamado (Orense); examen monográfico de una de sus figuras clásicas y paradigmáticas: Wittgenstein (La Coruña), por último: la situación presente de la filosofía analítica y su encuentro con otras tradiciones filosóficas. Esta tercera etapa del trabajo desembocó —aunándose con la primera circunstancia reseñada— en la celebración en Santiago de Compostela, en mayo de 1989, de un *curso-extraordinario* donde se concentraron un buen número de estudiosos y especialistas de la tradición analítica de nuestro país, "presididos", en cierto, sentido, por el eminente profesor Ernst Tugendhat, de la Freie Universität de Berlín (ciudad hoy felizmente integrada y definitivamente identificable como única, desde cualquiera de las flechas de la rosa de los vientos).

En el libro que ahora ofrecemos se refleja básicamente el curso celebrado en Santiago; pero asimismo, aparecen reflejos de las etapas de estudio llevadas a cabo en Orense y La Coruña —por lo que hace al programa de trabajo del profesorado de COU—; naturalmente se refleja también el trabajo del Seminario de Postgraduados realizado en el seno de la Sección de Filosofía de esta Universidad, como ya hemos indicado.

El conjunto de la tarea, con todo lo que implica de organización y esfuerzo económico no habría sido posible sin la colaboración, siempre grata, generosa y encomiable de diversas entidades y personas. Para todos, mi reconocimiento:

- La Consellería de Educación y la Dirección General de Enseñanzas Medias de la Xunta de Galicia. Sus entonces titulares respectivos, Aniceto Núñez y Emilio García Bonome, se mostraron alentadores y magnánimos. (Tal vez su condición profesional de filósofos les hizo fácil seguir el impulso...)
- El Vicerrectorado de Organización Académica de esta Universidad, responsable de la Coordinación general del COU en toda Galicia, no les fue a la zaga. El celo de su titular el Prof. López Rivadulla, en esta tarea, sólo sería comparable a su inmensa e inigualable comprensión.
- Los Departamentos de la Sección ("Filosofía y Antropología Social"; "Lógica y Filosofía de la Ciencia"; "Filosofía del Derecho, Moral y Política") prestaron ayuda económica y toda la infraestructura necesaria.
- El *Seminario Permanente* de Filosofía del COU aportó generosamente ideas, tiempo y trabajo. A los compañeros Ramón Areas de Orense y a José Luis Castro y Herminia Astray de La Coruña se debe todo el éxito de las sesiones que en uno y otro lugar se celebraron. Agustín Carracedo merece nuestra gratitud por los trabajos de ordenador, de diseño y edición de los programas.
- Los compañeros de la Universidad de Santiago, muy especialmente los que se dedican a la Lógica, Filosofía de la Ciencia, Filosofía del Lenguaje, a temas ontológicos, a Teoría del Convencimiento, que animaron sesión a sesión, el diálogo del Seminario sobre los textos de Putnam, saben que mi presencia en esta página es cuestión de unas leyes físicas; o si se quiere, como en las reuniones, de orden: me han dado la palabra.
- Antes y después, como estímulo definitivo y el más valioso, están aquellos en razón de los cuales nuestro trabajo es profesión socialmente asumida e institucionalmente consagrada: los estudiantes. A aquellos que han elegido la Filosofía en la Universidad, más o menos sabedores de que se adscriben con ello a una tarea escurridiza y frustrante, aunque seductora ("a ella se vuelve una y otra vez como a una amante" decía Kant, y perdonémosle el femenino de la expresión); tarea que al mismo tiempo persigue al seducido, "como la ira de los dioses" ...; a todos ellos, repito, se debe también este libro.

Compostela, Julio de 1990
Mercedes TORREVEJANO

DESPUÉS DEL ANALISIS: SIGNIFICADO, COMPRENSION E INTENCIONALIDAD

Juan José Acero

Después del análisis

Cuando estamos a punto de entrar en la última década del siglo, ¿puede hablarse de la Filosofía Analítica como de un movimiento o una escuela filosófica en activo? No; no se puede. Como tal escuela o conjunto de escuelas, la Filosofía Analítica tuvo en John Austin a su último exponente indisputado, y Austin falleció en 1960. Opino esto por las razones que paso ya a exponer. La Filosofía Analítica abarca un abanico de escuelas o tendencias filosóficas que comparten entre sí

(i) una concepción de la naturaleza y origen de los problemas filosóficos.

Según esa naturaleza, y muy a grandes rasgos, la complejidad del lenguaje humano y la circunstancia de que nuestro conocimiento del lenguaje sea en gran medida conocimiento implícito da frecuentemente pie a equívocos y errores en nuestra comprensión de él. De esas ideas o imágenes equivocadas acerca del lenguaje, demostraron los filósofos analíticos, surgen muy distintos problemas filosóficos. Puesto que el análisis del lenguaje resulta, así, de importancia decisiva para la filosofía, caracteriza también a la Filosofía Analítica

(ii) la consideración de la teoría del significado como disciplina filosófica central,

desplazando con ello a la epistemología -según es usual opinar. Naturalmente, puesto que filósofos analíticos los hubo de orientaciones diversas, no puede esperarse que existiera unanimidad en cuanto a cómo entender la llamada 'teoría del significado'. Para Frege -según la conocida exposición de Dummett¹-, la teoría del significado era simplemente Lógica y, por lo tanto, algo bien distinto de lo que uno infiere de los textos de Austin. Para Wittgenstein, la teoría del significado no era tal teoría: lo que nosotros llamamos así pertenece al ámbito de lo que no puede decirse o de lo que resulta imposible hacer.

¹ Véase Frege. *The Philosophy of Language* (London: Duckworth, 1973), especialmente, cap. 19.

Resulta de aquí que la filosofía tiene mucho que ver con la aplicación de esa teoría o adecuada imagen de la naturaleza y funcionamiento de los mecanismos lingüísticos. Mejor que como una disciplina cuyo objetivo es la obtención de verdades, la filosofía habría de concebirse como una actividad que nos permite rectificar equívocos o errores a que da lugar una errónea comprensión de los recovecos de nuestro lenguaje. Por lo tanto, el tercer rasgo distintivo a considerar puede recogerse en la tesis de que

(iii) la filosofía no es una ciencia, sino una actividad.

Este rasgo es especialmente interesante. Nos permite reconocer como analíticos a los (ampliamente denostados) filósofos (neo-positivistas sin tener que ver en el rechazo de la metafísica una condición necesaria de la pertenencia al movimiento global. Y también, desde luego, nos permite incorporar a la lista a muchos otros filósofos contrarios a uno o más apartado del credo vienés. Esto es esencialmente como debería ser, en mi opinión. Así mismo, (iii) es una condición notablemente fuerte: si se considera que la reflexión del filósofo cristaliza (o ha de cristalizar) en teorías o en propuestas teóricamente sustantivas, entonces se corre el riesgo de transgredirla. (Todo depende luego de los matices).

Ahora bien, ¿cuántos de entre los que alguna vez se incluyeron (o fueron incluidos) en las filas de alguna facción de la Filosofía Analítica se adhieren hoy a estas *tres* tesis señaladas? Y si los hubiera, ¿cuántos de entre los discípulos directos o indirectos de los grandes filósofos analíticos lo hacen hoy? Yo no conozco a ningún filósofo ni a ninguna escuela filosófica que en la actualidad esté desarrollando un programa que reivindique las tesis (i) - (iii).

Hay, claro está, Filosofía Analítica en un sentido muy diferente. Y de las diversas opciones que abre este sentido -a saber, combinaciones de algunas de las tesis (i), (ii) y (iii), sólo quiero considerar aquí una de ellas. Hay quienes se adhieren a (ii) sin hacerlo a (i). Para estos autores, el dominio en el que el filósofo desarrolla su actividad es el dominio de los conceptos (o los significados), y para poder desarrollar esa actividad se precisa de un modelo de la naturaleza y del funcionamiento del lenguaje. Si el modelo en cuestión lo es de la comprensión del lenguaje por parte de sus usuarios, entonces el caso de Michael Dummett caería dentro de este apartado. Sin embargo, lo más frecuente resulta, entonces, que (iii) sea indigerible para los que así opinan. Una teoría del significado es una *teoría* del significado: un conjunto de principios que pretenden establecer un conjunto de verdades acerca del significado de las expresiones de una lengua. Filosofía Analítica, podría generalizarse entonces, es toda aquella que persigue resolver mejor algunos temas que los grandes filósofos analíticos sacaron a la palestra (y quizás en términos próximos a como aquéllos lo hicieron). Por ejemplo, por esta regla de tres, Quine, Davidson, Kripke y Montague! serían filósofos analíticos. (Hao Wang, por ejemplo, como es sabido, considera a Quine un filósofo analítico). Esto, sin embargo, es absurdo; o así me lo parece a mí, al menos. (Quine niega incluso la posibilidad de la teoría del significado; y Davidson y Kripke, así como muchos otros autores que han trabajado en uno u otro aspecto del análisis

semántico, han dado un lugar central en su pensamiento a varios de los temas de la metafísica clásica²).

Sinceramente, no creo que este criterio sea muy útil a la hora de poner orden clasificatorio en el multifario panorama de la filosofía contemporánea, aunque tampoco deseo empeñarme en combatirlo. Más acertado me parece hablar a este respecto de *tradición* que de tendencias o escuelas. La tradición analítica tiene su origen, claro está, en diferentes ideas, expuestas en diferentes momentos y por diferentes escuelas, de Frege, Russell, Wittgenstein, Carnap, Ryle, Hempel, Austin y otros. Esas ideas están en el origen del actual estado de la cuestión en filosofía de la ciencia, del lenguaje y de la mente. Sin embargo, me parece innegable que su aceptación va de la mano del rechazo mayoritario de las tesis (i) y (iii).

Por lo tanto, en lo que sigue voy a suponer sin otro argumento que nos encontramos en una etapa post-analítica. Defenderé, además, que existen interesantes puntos de confluencia entre doctrinas centrales dentro de la tradición analítica y doctrinas también centrales en otras tradiciones filosóficas contemporáneas, especialmente en las tradiciones hermenéutica y fenomenológica. A mi modo de ver, éste es un aspecto interesante de la filosofía contemporánea que, por su complejidad, necesitaría de un estudio mucho más detenido y más comprensivo que el que recibe en este breve trabajo. De entre las diferentes manifestaciones en que esa confluencia es manifiesta, yo me voy a ocupar tan solo en las próximas páginas de ciertas confluencias habidas a propósito de las relaciones entre lenguaje, pensamiento y realidad. Más concretamente, de acuerdos tácitos que afectan a la naturaleza del significado, la comprensión (o interpretación) del lenguaje y la intencionalidad de la mente.

Dos avisos antes de entrar en materia. Primero, en todos estos casos, no es mi deseo dar la impresión de que la confluencia de intereses, temas y doctrinas es mayor de la que realmente es. Por el contrario, creo que las diferencias no sólo son muy numerosas, sino también muy sustantivas. Mi objetivo será, sin embargo, subrayar los puntos en común y el material escogido para ello, así como su tratamiento, se justifican por ese objetivo. Y segundo, mi referencia a la tradición analítica y a otras tradiciones no debe servir nunca para dar la impresión de que en uno u otro caso me estoy refiriendo a autores con ideas o sistemas filosóficos muy similares. Por lo que respecta a la tradición analítica, en una panorámica global fiel a la realidad todo lo que veríamos sería un variado mosaico de posiciones en las que las diferencias privan sobre los puntos en común.

Lenguaje y pensamiento, ¿qué es primero?

El primero de los acuerdos tácitos se plantea de un lado en la defensa -que hace Apel, entre otros- de un concepto hermenéutico-trascendental del lenguaje frente a un concepto

² En opinión de E. Tugendhat, esto no tiene nada de extraño, pues es precisamente el dominio del análisis semántico, y en particular el de la semántica formal, el apropiado en el que afrontar con éxito las cuestiones centrales de la ontología tradicional. Véase su *Traditional and Analytical Philosophy. Lectures on the Philosophy of Language* (Cambridge University Press, 1982), Parte I, cap. 3 («Ontology and Semantics») y 4 («Has Formal Semantics a Fundamental Question?»).

solipsista del mismo³. Al igual que Heidegger, primero, y Gadamer posteriormente, Apel es también de la opinión de que los grandes filósofos griegos de la antigüedad clásica fueron responsables de una devaluación de las funciones cognitivas del lenguaje. Ellos hicieron del lenguaje esencialmente un instrumento, un sistema de expresión del pensamiento, que así apareció como una realidad o un dominio autónomo. El lenguaje es, entonces, «mera expresión secundaria o instrumento ... del pensamiento» (TF, II, 320). Ésta es la concepción del lenguaje como un sistema de signos. Y decir que una expresión es un *signo* es, en boca de estos autores, una crítica, no la declaración de un hecho.

La concepción instrumental del lenguaje se generalizó en los siglos XVII y XVIII con la irrupción en la escena filosófica de la tesis racionalista de que la mente tiene acceso directo a los contenidos de sus propios estados mentales y de que ella no puede engañarse a este respecto. Y se consumó con la doctrina de John Locke de que, en su significación primaria, las palabras están en lugar de las ideas que se encuentran en la mente de los hombres. El *solipsismo metodológico*, que a juicio de Apel representa la posición oficial de la Filosofía Analítica, en lo que se refiere a las relaciones entre lenguaje y pensamiento, tiene sus raíces en estas doctrinas. Según el solipsismo metodológico, la mente humana es un complejo dispositivo que trabaja con representaciones. ('Ideas' fue el término que los empiristas clásicos acuñaron para lo mismo). Así, por ejemplo, recordar es traer a la unidad de control que ejerce la función consciente la consiguiente representación desde el lugar de la memoria donde está almacenada; y razonar es llevar a cabo operaciones *formales* sobre representaciones. A su vez, las representaciones tienen su origen en las relaciones causales entre el organismo y el medio externo (e interno). O, como sostiene la Hipótesis del Innatismo del Lenguaje del Pensamiento, en esas relaciones *más* un sistema de representación innato⁴.

El solipsismo metodológico adolece de los inconvenientes de la concepción instrumentalista del lenguaje. En efecto, para un solipsista, el significado de una palabra es precisamente el *concepto* (o idea) asociado a esa palabra; y el significado de una oración declarativa es el *pensamiento* -un complejo de ideas- que esa oración expresa. Ahora bien, mientras que una palabra o una oración tiene significado, conceptos y pensamientos tienen *contenido*. Y éste es el que es en virtud del lugar que ocupa en el sistema de percepciones, estados mentales y conducta del *individuo* en cuestión. 'Rojo' -el término- significa lo que significa porque ROJO -el concepto- tiene el contenido que tiene. Y el contenido de este concepto depende de su rol en las percepciones del individuo, así como en ciertas pautas inferenciales: si *x* es ROJO, entonces *x* no es VERDE, por ejemplo; y si *x* es VERDE, entonces *x* no es ROJO, pongamos por caso. En otros casos, el contenido del concepto se agota en sus relaciones inferenciales. Eso ocurre con el concepto Y. Si el pensamiento ESTAMOS EN PRIMAVERA Y LLUEVE es verdadero, entonces es verdadero el pensamiento de que

³ Éste es un tema recurrente dentro de los escritos de Apel. Yo me baso especialmente en diversos ensayos de *La transformación de la filosofía* [TF en lo que sigue], Madrid: Taurus, 1985, vol. 2.

⁴ Sobre esta hipótesis, véase J. Fodor: *El lenguaje del pensamiento*, Madrid: Alianza Editorial, 1985; «The Present Status of the Innateness Controversy», en *RePresentations*, Brighton, Sussex: Harvester Press, 1981; *Psychosemantics*, Cambridge, Mass.: The M.I.T. Press, 1987, especialmente, Apéndice II.

ESTAMOS EN PRIMAVERA y también es verdadero el pensamiento de que LLUEVE⁵. Es más, la intencionalidad del lenguaje, es decir, su capacidad para referir una expresión a un objeto, o para describir una situación una oración o una preferencia, es un fenómeno derivado. La intencionalidad del lenguaje deriva de la intencionalidad del pensamiento.

Lo importante del caso es (i) la diferencia entre *significado* de expresiones y *contenido* de conceptos (y pensamientos), (ii) el hecho de que el análisis del contenido es *absolutamente* independiente del significado de las expresiones del lenguaje y (iii) el viejo principio cartesiano de que la mente tiene acceso directo e incorregible a sus propios estados mentales. Todo eso junto hace posible una imagen del uso lingüístico en que un individuo escruta dentro de su mente el contenido de uno de sus estados mentales, identifica el pensamiento que desea expresar y a continuación echa mano de la expresión (palabra, oración, etc.) que mejor lo expresa. Pero entonces, cuando ésa es la perspectiva adoptada, el lenguaje se nos muestra como un mero instrumento.

Bien, podría decirse ahora, ¿es que el lenguaje es otra cosa -algo más, quizás- que un instrumento? Apel así opina. El lenguaje, nos dice «es una magnitud trascendental en el sentido kantiano; más exactamente: es una condición de posibilidad y validez del acuerdo y del auto-acuerdo y, con ello, a la vez del pensamiento conceptual, del conocimiento objetivo y del obrar con sentido» (TF, II, 318). Estas tesis constituyen el concepto *hermenéutico-trascendental* del lenguaje.

El primero de los tres puntos de confluencia a los que he aludido más arriba es precisamente el rechazo del solipsismo metodológico. Apel critica esta doctrina porque, de acuerdo con ella, el lenguaje se muestra «sólo como un *instrumento* del pensamiento» (TF, II, 329) y considera que la moderna filosofía del lenguaje es culpable de cometer este error. Pero, ¿qué moderna filosofía del lenguaje? Una parte central de la filosofía del lenguaje que se ha elaborado en las dos últimas décadas -la que se conoce como Nueva Teoría de la Referencia- es igualmente crítica hacia el solipsismo metodológico. Y crítica en dos sentidos: por ser *solipsista* y por ser *individualista*.

Putnam, por ejemplo, es (o debería ser) conocido por su tesis de que el significado de ciertos términos, los términos de especie natural -'tigre', 'jade', 'olmo', 'limón', 'oro', etc.-, significan lo que significan en virtud de otros factores que las consiguientes ideas que los hablantes puedan tener en su cabeza. El estado mental del hablante que emplea uno de estos términos no determina de qué está hablando, a qué porción del mundo se refiere. Yo puedo ser incapaz de distinguir los olmos de las hayas, pero si digo: 'Hay olmos en Castilla' no estoy diciendo que haya hayas en Castilla. Y, de otro lado, como prueba el famoso ejemplo de la Tierra Gemela, dos personas pueden estar en el mismo estado psicológico y, sin embargo, hablar de cosas diferentes al emplear la misma forma verbal -'agua', por ejemplo. La moraleja es manifiesta: «Los significados no están en la cabeza⁶.» No lo están, porque lo que hace que al emplear un término de especie natural me refiera a algo no depende de lo que esté en la

⁵ En el momento de impartir la conferencia, a las 12 horas del día 10 de Abril de 1989, llovía.

⁶ «The Meaning of 'Meaning'», en *Philosophical papers*, vol. II: *Mind, Language and Reality*, Cambridge University Press, 1975; *Representation and Reality*, Cambridge, Mass.: The M.I.T. Press, 1988, cap. II.

cabeza del hablante, sino de la existencia de expertos sobre el tema y de la posibilidad, que la organización social debe permitir, de que en caso de dudas sobre si algo se le aplica o no un término tal esos expertos puedan emitir un dictamen.

La hipótesis de la División del Trabajo Lingüístico -como Putnam denomina a esta última idea- es, por lo tanto, incompatible con la posición solipsista. Si hay expresiones cuyos criterios de aplicación por parte del grupo social depende de una 'cooperación estructurada' entre estos hablantes y los correspondientes expertos, entonces sucede que cuál sea el contenido de los conceptos y pensamientos de los hablantes -por ejemplo, qué pienso cuando pienso que HAY OLMOS EN CASTILLA- es cosa que depende de acuerdos previos; de acuerdos que sólo pueden adoptarse haciendo uso del lenguaje. Si aceptar esto es aceptar que el lenguaje tiene una dimensión trascendental, entonces Apel y Putnam siguen caminos que conducen al mismo lugar.

Lo mismo, creo, puede decirse de algunas ideas que T. Burge ha venido exponiendo desde finales de la pasada década⁷. Estoy sugiriendo que interpretemos las críticas de Burge al individualismo en semántica y filosofía de la mente como una crítica al solipsismo. Como en el caso de Putnam, Burge rechaza también que lo que las palabras de uno significan dependa exclusivamente de lo que el acceso a los contenidos de sus propios estados mentales dicte. Por el contrario, «depende en parte de las ventajas sociales [que se siguen] de mantener cánones de comunicación y de atribución mentalista comúnmente establecidos⁸.»

El argumento de Burge es conocido, pero bien merece la pena volver a exponerlo aquí. Según el Diccionario de la Real Academia de la Lengua, el término 'artritis' significa inflamación de las articulaciones. Consideremos, entonces, el caso de María. Es una hablante estándar del castellano y además, para su desgracia, siente dolores en el muslo. No me pregunten cómo, pero María cree que tiene artritis. Naturalmente, su creencia es falsa. Sus dolores son musculares o de otro tipo, pero no los causa una inflamación de ninguna articulación. María no tiene artritis. Lo que diríamos en este caso es que María no dispone del adecuado concepto ARTRITIS, pese a manifestar sufrir esa dolencia. Con el término 'artritis' María asocia un concepto distinto. Pero ahora introduzcamos una pequeña (¿o no es pequeña?) variante en el ejemplo. Todo es como acabo de decir, pero María vive en un mundo en que el término 'artritis' expresa un concepto en el que se incluyen no sólo las inflamaciones de las articulaciones, sino otros procesos, reumatoides, por ejemplo, que también producen dolores. La pregunta acerca de cómo es la creencia de María *en esta nueva situación*, si verdadera o falsa, no recibe la misma respuesta que antes. ¡Pero, entonces, el contenido de nuestros estados mentales, de nuestras creencias, y el significado de algunas de nuestras palabras, como 'artritis', o de algunas de nuestras oraciones, como 'Tengo artritis' (dicha por María), depende del medio lingüístico y cultural en el que se encuentran! Si negamos los factores sociales en las atribuciones mentalistas y en el análisis semántico, nuestra imagen de unas y otro quedarán muy seriamente desdibujadas.

⁷ Especialmente en «Individualism and the Mental», en P. A. French, Th. E. Uehling & H. K. Wettstein (eds.): *Midwest Studies in Philosophy*, vol. IV: *Studies in Metaphysics*, Minneapolis: University of Minnesota Press, 1979. El tema recurre en muchos otros más recientes ensayos de este autor.

⁸ Cf. *Op. cit.*, pág. 116.

Traducción e interpretación: la importancia del acuerdo

Cambio ahora de tema y me fijo en un año que ha resultado ser importante para el devenir de la filosofía contemporánea: 1960. Es el año en que Quine publica el más importante de sus libros, *Palabra y objeto*⁹, realmente un hito para la teoría del significado, el análisis lógico del lenguaje y la epistemología recientes. Sin embargo, 1960 no es un año menos importante para la tradición filosófica que va de Heidegger a Derrida o Ricoeur. Es el año en que se publica *Verdad y método*¹⁰, de Gadamer, la obra de referencia para la filosofía hermenéutica. Obvio es decirlo una vez más, entre la obra de Quine y la de Gadamer hay diferencias manifiestas, y no es mi propósito ni negarlas ni disimularlas. La filosofía del lenguaje de Quine está estrechamente vinculada al tema «[d]el desarrollo y la estructura del aparato referencial» del lenguaje (PO, 11), mientras que la de Gadamer tiene como punto de referencia la objetividad de la obra de arte. Quine no cree en la idea de una filosofía primera y, por lo tanto, ve con escepticismo la posibilidad de separar ciencia y filosofía, filosofía del lenguaje y lingüística. Gadamer, por su parte, es muy explícito en lo que respecta a sus relaciones con la lingüística. Ésta no es para él un punto de referencia: «Nos interesamos ... por lo contrario de lo que intenta investigar la ciencia del lenguaje» (VM, 483) -asegura. Estas y otras importantes discrepancias están ahí¹¹. Pero también existen interesantes puntos comunes que quiero aquí subrayar. Y no sólo entre Quine y Gadamer, sino entre diversos cualificados representantes de las tradiciones analítica y hermenéutica.

El principal punto de acuerdo se produce en el tema del estudio del lenguaje y la cultura de otras comunidades lingüísticas y culturales que la propia. En la tradición analítica los problemas de este estudio se abordan desde el punto de vista de la teoría de la traducción o de la interpretación *radical*, es decir, «la traducción del lenguaje de un pueblo al que se llega por vez primera» (PO, 41). Muchas y muy importantes páginas de Quine, Davidson, Wallace, Grandy y varios otros autores han analizado este tema. Ahora bien, Apel se ha apercibido de que la problemática de la traducción radical tiene su equivalente en la

⁹ PO en lo que sigue. (Edición castellana en Barcelona: Labor, 1968.)

¹⁰ VM en lo que sigue. (Edición castellana en Salamanca: Ediciones Sígueme, 1977.)

¹¹ La lucidez de Tomás Calvo, en un seminario impartido conjuntamente sobre el tema de la metáfora en 1987, en la Universidad de Granada, me permitió apercibirme de que un serio obstáculo para la mutua comprensión de estas dos tradiciones estriba en la concepción del signo respectivamente adoptada. En la tradición analítica, la teoría del signo que ha servido de constante punto de referencia ha sido la fregeana. En la tradición hermenéutica, especialmente en autores como Ricoeur o Derrida, no ha sido Frege, sino Saussure la autoridad aceptada. Aquí el signo se considera integrado por significante y significado, y en ninguno de ambos intervienen factores de índole referencial. Éste es, sin duda, uno de los inconvenientes más formidables para la mutua comprensión. A quienes han sido educados en la idea de que la referencia es la relación semántica fundamental y para quienes han seguido los avatares de la campaña de Quine, y de sus discípulos, contra la semántica intensional, la idea saussuriana del signo les parecerá centrada en cuestiones que serían, en el mejor de los casos, marginales y, muy frecuentemente, culpable de cometer errores fundamentales. Dos ensayos de Ricoeur pueden considerarse una contribución de interés precisamente por la luz que arroja sobre estos extremos. Me refiero a «La estructura, la palabra, el acontecimiento» y «El problema del doble sentido», cuya versión castellana aparece en *Hermenéutica y estructuralismo*, Buenos Aires: Ediciones Megápolis, 1975.

tradición hermenéutica¹². El problema central no lo constituye aquí exclusivamente el acceso a una lengua desconocida, sino la interpretación de textos, en general; y no exclusivamente textos -que en el inicio de esta tradición filosófica son de carácter bíblico o jurídico-, sino cualesquiera productos culturales. Las diferencias no son, sin embargo, decisivas. En efecto, según Gadamer, la interpretación lingüística es la forma general de la interpretación (VM, 478).

Con esto por delante, he aquí los aspectos más dignos de mención de la confluencia de temas y de doctrinas que se ha producido entre las tradiciones analítica y hermenéutica. El primero de ellos ha sido citado por autores tan distintos como Follesdal y Apel¹³, y yo quiero mencionarlo aquí de pasada. En *Palabra y objeto*, en sus observaciones sobre la traducción de la lógica del nativo de la jungla y en la traducción de lo que llama oraciones estimulativamente analíticas ('*stimulus-analytic*'), Quine hizo ver la necesidad de operar con un Principio de Caridad. En su opinión, cuanto más absurdas y/o exóticas sean las creencias que el traductor se ve forzado a imputar al nativo, tanto mayor habrá de ser su desconfianza hacia su traducción (PO, 82). (Pero desconfianza no significa necesariamente error). Y siguiendo la misma línea argumentativa, Davidson ha hecho uso de una máxima que exige que un manual de traducción deba confeccionarse de forma que maximice el acuerdo: «Un amplio acuerdo es el único posible trasfondo contra el que pueden interpretarse las disputas y los errores¹⁴.»

La posición de Gadamer no es, creo, distinta. El problema de la interpretación es esencialmente un problema de lograr un acuerdo sobre un asunto por medio del medio del lenguaje. No hay interpretación sin el supuesto de que el lenguaje del texto o los rasgos de la obra artística cuyo significado queremos entender es nuestro mismo lenguaje; o que ese significado es expresable con *nuestras* propias palabras. Por ello, un proceso de interpretación, dice Gadamer, «tendrá que elaborar un lenguaje común» (VM, 466).

El lenguaje como medio universal y la autonomía del sentido

Más interesante todavía que el citado punto de acuerdo es el siguiente. Tanto para Gadamer y otros hermeneutas como para una gran mayoría de quienes del otro lado se ocupan de los problemas del análisis semántico una teoría de la interpretación como una teoría del significado ha de ser una teoría de la *comprensión*. Esto, que puede resultar indiscutible cuando tengamos entre manos un texto o una obra de arte, no es obvio en absoluto para la teoría del significado. ¿Por qué ha de consistir una teoría del significado de las

¹² Véase «Intentions, Conventions, and Reference to Things: Dimensions of Understanding in Hermeneutics and in Analytic Philosophy of Language», que es la contribución de Apel a la Conferencia Internacional sobre *Significado y Comprensión* que tuvo lugar en 1979 en el castillo de Cerisy-la-Salle, en Normadía. Las actas de esta conferencia fueron editadas por H. Parret y J. Bouveresse, con el título de *Meaning and Understanding*, Berlin: Walter de Gruyter, 1981.

¹³ K. O. Apel, *op. cit.*, pág. 81; D. Follesdal: «Hermeneutics and the Hypothetical-Deductive Method», *Dialectica*, vol. 33 (1979), pág. 337.

¹⁴ *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), pág. 153.

expresiones de una lengua *L* en una teoría de la comprensión de *L*? En la semántica filosófica de los últimos veinte años se ha preferido casi de forma mayoritaria esta opción. Pero no es la única ni mucho menos. Adoptarla ha supuesto que se ensayaran y que se aceptaran determinadas líneas argumentales. Y, además, todavía hoy tenemos a destacados autores, como R. Thomason o como S. Soames, que o no comparten, o que aceptan en una medida muy restringida, semejante compromiso¹⁵.

Así, por ejemplo, una bien conocida línea argumental de Dummett arranca de la premisa de que el significado de una expresión determina y viene exhaustivamente determinado por el uso que hagan de ella quienes la empleen¹⁶. Ahora bien, no habría posibilidad para un aprendiz del lenguaje de adquirir el significado de algunas expresiones si no pudiera identificar o reconocer que se dan las circunstancias en que el uso de tales expresiones resulta apropiado. Pero entonces, cuando uno ha aprendido a reconocer semejante tipo de circunstancias o situaciones, uno comprende el significado de la expresión.

El camino de la interpretación a la comprensión es recto y corto; y el camino del significado a la comprensión tiene sus vericuetos, pero llega a donde ha de llegar, en opinión de muchos filósofos, lógicos y lingüistas. ¿Qué decir ahora de la dirección inversa? Nos hallamos, es mi impresión, ante la situación simétrica de la que antes encontrábamos. Si sabemos lo que los hablantes captan cuando comprenden las expresiones que usan u oyen usar, sabemos lo que esas expresiones significan. Así que la cosa parece natural para la teoría de la traducción. Pero ¿es igual de simple para la hermenéutica de textos u otros productos culturales? ¿Es la comprensión sólo y exclusivamente interpretación? Una de las tesis más interesantes de *Verdad y método* es que sí, que «[l]a forma de realización de la comprensión es la interpretación» (VM, 467. La cursiva es de Gadamer). El capítulo duodécimo de esta obra («El lenguaje como medio de la experiencia hermenéutica») expone una defensa de esta tesis que la hermenéutica de Schleiermacher hace obligada. Esa defensa se basa, si yo entiendo bien a Gadamer, en estas dos premisas:

(I) Una concepción del lenguaje como Medio Universal,

y

(II) Un principio semántico de Autonomía del Sentido.

La primera de estas premisas es clara, repetida y enfáticamente enunciada por Gadamer: «El lenguaje es el medio universal en el que se realiza la comprensión misma» (VM, 467. La cursiva es de Gadamer). Tanto se trate de un texto como de una obra de arte, comprender implica, dice Gadamer, *dialogar* o conversar con el uno o con la otra. Y eso quiere decir pergeñar hipótesis interpretativas que formularemos por medio de *nuestras* palabras o de *nuestros* conceptos.

¹⁵ R. Thomason: «Introducción» a *Ensayos de filosofía formal*, de R. Montague, Madrid: Alianza Editorial, 1976; S. Soames: «Semantics and Psychology», en J. Katz, ed.: *The Philosophy of Linguistics*, Oxford University Press, 1985.

¹⁶ *Truth and Other Enigmas* (Cambridge, Mass.: Harvard University Press, 1978), pág. 216).

La segunda premisa, el principio de autonomía del sentido, le permite a Gadamer rechazar que los problemas de interpretación textual sean problemas que atiendan a cuestiones de orden subjetivo, especialmente, de las cuestiones relativas a cuáles eran las intenciones que animaron al autor del texto o la obra a hacer lo que hizo y cómo lo hizo. Por ello, el problema de la interpretación no es el de «una simple reproducción del proceso psíquico original» (VM, 464). Por el contrario, el texto, la obra producida tienen uno o más sentidos propios. De ahí que el texto escrito sea el material que más adecuadamente da lugar al problema de la interpretación: «En la escritura, el sentido de lo hablado está ahí por sí mismo, enteramente libre de todos los momentos emocionales de la expresión y de la comunicación» (VM, 471).

Las dos premisas actúan en tándem de la siguiente forma: Por el principio (I), nuestro diálogo con un texto se hará en nuestros propios términos lingüísticos. Es con ellos que, por (II), hemos de despejar las incógnitas que el texto plantee, dejando a un lado los problemas de índole psicológica. Esos problemas son reales y no se resuelven lanzando rayos noéticos que vuelvan transparente lo que no lo es. Textos y obras de arte dan lugar, como sostiene -con razón, a mi modo de ver- Follesdal, al típico trabajo del método hipotético-deductivo: conjeturas, refutaciones y, cuando somos afortunados, confirmaciones también. Gadamer dice lo mismo pero con otras palabras: los problemas de la hermenéutica son problemas de «extrañeza y de superación de la misma» (VM, 465). La comprensión es interpretación porque nuestro diálogo con el texto o con la obra de arte consiste en *esa* resolución de problemas.

Pues bien, tanto (I) como (II) han tenido una amplia aceptación dentro de la tradición analítica de las últimas décadas. La autonomía del significado es justamente un estandarte que han defendido los opositores al programa nominalista de Grice, programa consistente en la definición de la noción de significado en términos de la noción psicológica de intención (del hablante). Davidson, por ejemplo, lo ha incorporado, incluso bajo la misma denominación, en su teoría de la interpretación radical¹⁷. En lo que respecta a la premisa (I), Hintikka ha venido sosteniendo, desde hace aproximadamente una década, que se trata un *leit-motiv* central en el pensamiento de filósofos analíticos como Frege o Wittgenstein -éste en su doble trayectoria-, y de filósofos post-analíticos como Quine. Pero es también -en una hipótesis reciente de un discípulo de Hintikka, M. Kusch- una doctrina central en el pensamiento de Heidegger y de Gadamer¹⁸. El caso de este segundo filósofo merece destacarse.

Volvamos a (i). Si el lenguaje es un medio universal, como cree Gadamer, entonces no existe una posición de ventaja, al margen de todo esquema conceptual o lingüístico, desde la cual acceder más fácilmente al significado del texto o de la obra de arte. Este principio

¹⁷ Véase *Inquiries Into Truth and Interpretation* (Oxford: Clarendon Press, 1984), págs. 113 y s., 127, 164 y s. y 174 y s.

¹⁸ Véase, por ejemplo, M. B. Hintikka & J. Hintikka: *Investigating Wittgenstein*, Oxford: Basil Blackwell, 1986, cap. 1; M. Kusch: «Husserl and Heidegger on Meaning», *Synthese*, vol. 77 (1988) 99 - 127. Las interesantes ideas de Kusch sobre la filosofía del lenguaje de Gadamer se exponen en «*Language is the Universal Medium... -Gadamer's Philosophy of Language*», Oulu: Oulun Yliopisto Historian Laitos, 1987, así como en un libro mucho más extenso que Kluwer ha anunciado para 1989.

nos compromete con la tesis de que cualquier revisión de este sistema se hace desde dentro y descansando siempre en el supuesto de que hay una única interpretación de este lenguaje, un solo y definido conjunto de relaciones entre nuestras palabras y el mundo. No estemos, sin embargo, necesariamente condenados a valernos siempre del mismo sistema conceptual, tal y como nuestro lenguaje lo expresa y refleja. Pero también es cierto que el horizonte de las propias ideas del intérprete, es decir, el sistema de conceptos, creencias y valores del intérprete, así como la tradición a la que pertenece, siempre están presentes en la interpretación. La premisa (I), como ha visto Kusch, va inexorablemente de la mano de una posición metafísica *internalista*:

«No se puede querer mirar desde arriba el mundo lingüístico ...
pues no existe ningún lugar fuera de la experiencia lingüística del
mundo desde el cual éste pudiera convertirse por sí mismo en objeto»
(VM, 543).

Tampoco puede entenderse desde fuera la obra de arte desde fuera de todo horizonte. El internalismo es, así pues, un principio básico de la interpretación, según Gadamer. Pero igualmente es un aspecto central del pensamiento de filósofos post-analíticos tan destacados como Quine o como Putnam.

Indeterminación de la interpretación

Si no coincidencia en las doctrinas mismas, el compartir premisas como las dos que acabo de exponer sí que explica que filósofos analíticos y hermenéuticos extraigan consecuencias que guardan entre sí un notable paralelismo. Quine y Davidson, entre otros, son conocidos por su defensa de la tesis de que la traducción radical está empíricamente indeterminada. Es decir, que puede haber dos diferentes manuales de traducción del lenguaje de una comunidad de la jungla, manuales que sean los unos incompatibles con los otros, pero todos ellos en una consonancia igualmente buena con las disposiciones a la conducta lingüística de los miembros de esa comunidad. «En esta misma medida,» dice Quine, «no puede haber fundamento para decir cuál de dos traducciones, manifiestamente diferentes, de oraciones sueltas es la correcta¹⁹.» Si *O* es una oración de esa lengua para la que distintos manuales ofrecen traducciones manifiestamente divergentes, entonces la pregunta '¿Qué significa realmente *O*?' no tiene respuesta objetiva.

La tesis de la indeterminación de la traducción es una de las doctrinas de la filosofía del lenguaje de esta segunda década de siglo que más se ha debatido. Ahora bien, en la tradición hermenéutica hay una doctrina paralela, una doctrina que también aquí ha dado lugar a una polémica repetidamente recogida. Así, frente a lo que después objetó Hugo Betti, también la traducción está histórica o tradicionalmente indeterminada. Toda interpretación es relativa «a la situación hermenéutica a la que pertenece». Por ello, de forma análoga a la

¹⁹ *La relatividad ontológica y otros ensayos*, pág. 106.

indeterminación objetiva del significado de una expresión por la conducta lingüística de sus usuarios, «no puede haber una interpretación correcta ‘en sí’» (VM, 477). Una expresión significa lo que el correspondiente manual de traducción dice que significa; una obra de arte se interpreta en función del horizonte del intérprete.

Naturalmente, indeterminación no significa arbitrariedad. Dos o más manuales de traducción pueden considerarse correctos a la vez, y la pregunta ‘¿Cuál de los dos *realmente*...?’ estará fuera de lugar. Paralelamente, la indeterminación histórica de la interpretación no abre paso al relativismo (si esto implica arbitrariedad). Betti sostuvo que era así, pero la razón que Gadamer expone es convincente. La razón es básicamente (I): «a toda interpretación le es esencialmente inherente un carácter lingüístico» (VM, 478), lo cual significa que toda interpretación es por fuerza relativa a un determinado horizonte. Sin embargo, (I) no es suficiente para que Gadamer pueda probar sus tesis de la indeterminación. (I) debe ir de la mano de la premisa de que existe una pluralidad de horizontes distintos, como efectivamente sostiene Gadamer²⁰. Si uno sostiene que no es posible salirse de su propio horizonte, entonces siempre queda una cuestión por responder: ‘¿Cómo interpretar el producto cultural *x* con respecto al horizonte *H*?’ Mi opinión es que *esta* cuestión es perfectamente objetiva y que no hay otra.

Conocimiento: teorías vs. preguntas y respuestas

Si en este dominio filosófico las tradiciones analítica y hermenéutica presentan una comunidad de intereses e incluso de doctrinas, ¿por qué se los pasa por alto? Quiero finalizar esta parte del presente trabajo señalando que una razón para ello se encuentra en una seria y profunda diferencia en cuanto al modelo de conocimiento que en uno u otro caso se asume. El desarrollo del conocimiento consiste para la tradición analítica en filosofía (y en filosofía de la ciencia) en la elaboración y el perfeccionamiento de teorías, donde las teorías -hablando muy en general- se entienden como entidades lingüísticas o extralingüísticas de naturaleza compleja. Ahora bien, cuando leemos a Gadamer, a Heidegger o a otros autores de allende las fronteras analíticas, tendemos a suponer sin más que sea ésa precisamente la idea de desarrollo del conocimiento de la que ellos hagan uso. Consiguientemente, se tiene la tendencia a considerar metafóricas cuando menos las afirmaciones de Gadamer sobre la *conversación hermenéutica*. Esto es, sin embargo, un error. Gadamer y los demás están expresándose literalmente cuando obran así. Lo hacen literalmente porque su modelo de conocimiento o de investigación es un modelo *dialógico* o *interrogativo*: la investigación y el desarrollo del conocimiento es un proceso de preguntas y de respuestas, y la lógica del conocimiento es la lógica de las preguntas y respuestas. Ciertamente, los modelos dialógicos no han sido los más populares en epistemología y filosofía de la ciencia, pero han estado ahí en los últimos años. La obra de Laudan *El progreso y sus problemas* es un clásico

²⁰ Esta idea me fue señalada por Tomás Calvo, que no es, sin embargo, responsable de ninguna afirmación que se haga en este trabajo.

de esta orientación y el modelo interrogativo de investigación que Hintikka ha venido estudiando a lo largo de esta década es otro caso de lo mismo. Es particularmente directa y digna de mención aquí la reivindicación del internalismo que Hintikka hace desde su modelo²¹.

Intencionalidad e intensionalidad

Las relaciones entre la tradición analítica y otras tradiciones filosóficas contemporáneas permanece, salvo en una excepción, en la oscuridad. (En lo que antecede, he tratado de arrojar luz sobre el tema.) La excepción la tenemos en las relaciones entre Filosofía Analítica y Fenomenología. Sobre ellas, y de forma muy especial, sobre las relaciones entre Frege y Husserl existe un número de publicaciones de gran calidad. En particular, los esfuerzos hechos en esta dirección por D. Follesdal y J. N. Mohanty, aunque discrepantes sobre cuestiones fundamentales específicas, muestran cuán estrechos son los vínculos que existen entre los pensamientos de estos autores y de otros que luego les han seguido²². No es, sin embargo, mi objetivo ocuparme ahora de estas relaciones. Quiero llamar la atención sobre la profunda afinidad existente entre el análisis husserliano de la intencionalidad de la mente y una cierta rama de la semántica lógica: la semántica de los mundos posibles.

Esta afinidad muestra hasta qué punto resulta equivocada una extendida opinión que corre en círculos no-analíticos. Así, Apel y Habermas -para no citar a colegas más próximos- parecen pensar que la Filosofía Analítica tuvo una evolución lineal en la que, finalmente, autores como Wittgenstein y Austin lograron invertir la dirección que seguía el estudio del lenguaje; una dirección que le había sido dada por autores como Frege y Russell. La Filosofía Analítica se desarrolló, según Apel, «de la *semántica lógica del lenguaje ideal*, como su punto original de partida, hacia la *semántica* y la *pragmática* del

²¹ La versión castellana de la importante obra de Laudan, publicada originalmente en 1977, es de Encuentro Ediciones, Madrid, 1986. En lo que respecta a las ideas de Hintikka, véase «On the Logic of an Interrogative Model of Scientific Inquiry», *Synthese*, vol. 47 (1981) 69 - 83, especialmente pág. 79, donde puede percibirse el punto de vista internalista aludido. Otras publicaciones del mismo autor sobre este tema u otros muy próximos son «The Logic of Information-Seeking Dialogues: A Model», en W. Becker & W. Essler, eds.: *Konzepte der Dialektik*, Frankfurt a. M.: Vittorio Klostermann, 1981; (junto con M. B. Hintikka) «Sherlock Holmes Confronts Modern Logic», en E. Barth, ed.: *Theory of Argumentation*, Amsterdam: Benjamins, 1981.

²² D. Follesdal: *Husserl und Frege*, W. Nygaard, Oslo: I. Kommissjon Hos H. Aschehoug and Co., 1958; «Husserl's Notion of Noemá», *Journal of Philosophy*, vol. LXVI (1969) 680 - 687; «An Introduction to Phenomenology for Analytic Philosophers», en R. E. Olson & A. M. Paul, eds.: *Contemporary Philosophy in Scandinavia*, Baltimore: The John Hopkins Press, 1972; «Phenomenology», en E. C. Carterette y M. P. Friedman, eds.: *Handbook of Perception*, vol. I, New York: Academic Press Inc., 1974. J. N. Mohanty: «Husserl and Frege: A New Look at Their Relationship», *Southwestern Journal of Philosophy*, vol. V (1974). Reimpreso en H. Dreyfus, ed.: *Husserl, Intentionality and Cognitive Science* (Cambridge, Mass: The M.I.T. Press, 1983); *Edmund Husserl's Theory of Meaning*, The Hague: Martinus Nijhoff, 1976; *Husserl & Frege*, Bloomington: Indiana University Press, 1982. Una importante colección de ensayos al respecto es la editada por H. Dreyfus, *op. cit.* Cf. también, para aspectos puntuales, E. Tugendhat, *op. cit.*, *passim*. Una obra interesante sobre las relaciones entre la Filosofía Analítica y la Fenomenología, que contiene además una bibliografía breve, pero muy pertinente, es H. A. Durfee, ed.: *Analytic Philosophy and Phenomenology*, The Hague: Martinus Nijhoff, 1976.

*lenguaje natural*²³.» Ahora bien, esto no es cierto por dos razones. Primera, la dirección de desarrollo no es la que Apel dice. La línea de análisis Frege-Russell-Carnap-Church-Quine no se interrumpe con Austin. Durante los sesenta y los setenta, esa línea se prolonga, entre otros, con Kripke, Hintikka, Davidson, Montague, D. Lewis, B. Partee, N. Belnap jr., Stalnaker, D. Kaplan, y a ella se adhieren, esta vez en los ochenta, H. Kamp, Barwise, Perry, B. Richards y un larguísimo etcétera. No hay un cambio unánime de paradigma. (Seguramente, porque nunca ha habido un único paradigma.) Lo que acontece entre 1892 y la década de los cincuenta de este siglo es la delimitación de un nuevo dominio de problemas y, consiguientemente, la apertura de una nueva línea de trabajo. Pero la línea o líneas existentes, esas que tuvieron su origen en los escritos semánticos de Frege y Russell, o en obras como el *Tractatus*, de Wittgenstein, o como *Meaning and Necessity*, de Carnap, han seguido y siguen fortaleciéndose (como lo demuestra el elevado número de trabajos que sobre el análisis lógico del lenguaje y sus aplicaciones filosóficas se publican de forma regular en revistas como el *Journal of Philosophical Logic*, *Synthese* o *Logique et Analyse*, o en series como la G.R.A.S.S., los Midwest Studies in Philosophy o las que publican conocidas editoriales inglesas, americanas y holandesas). La tesis de Apel, tan extendida entre nosotros, adolece de un mal que sufren (casi) todos los manuales: la excesiva simplificación.

Y segundo, y filosóficamente más interesante. Lo que la obra de estos autores demuestra, y lo que Montague llevó a una formulación explícita es que lenguas naturales y lenguajes formalizados pueden estudiarse dentro de un marco teórico único. Para Montague ése era el marco de la *teoría de modelos* de una lógica intensional de predicados de segundo orden. «En mi opinión,» sostuvo Montague, «no hay diferencia teórica importante entre las lenguas naturales y los lenguajes artificiales de los lógicos; de hecho, considero que es posible abarcar la sintaxis y la semántica de ambos tipos de lenguajes en una única teoría matemáticamente precisa²⁴.» Es este marco de la teoría de modelos el que ha dado lugar a una mejor comprensión de múltiples fenómenos semánticos: cuantificación, referencia pronominal, condicionales, deixis, verbos de actitud proposicional, de percepción, proposiciones, adjetivación no-intersectiva y varios más. Hoy en día este marco está en discusión y tiene serios competidores, como la Semántica de Situaciones o la Teoría de la Representación del Discurso. Pero nadie que trabaje en este campo duda de la necesidad del uso de métodos formales para mejor conocer parcelas muy centrales de las lenguas naturales. Yo creo que muy pocos de ellos pondrían en tela de juicio que una forma idónea de arrojar luz sobre esas parcelas consiste en construir lenguajes formalizados que reproduzcan los fenómenos con que uno topa cuando se enfrenta a una lengua natural -una propuesta cuya primera formulación explícita se halla en el *Tractatus* de Wittgenstein.

²³ «Intentions, Conventions, and Reference to Things», *loc. cit.*, pág. 81. En su *Festschrift für Perreet*, Apel sostiene que fue Austin quien introdujo el cambio de paradigma. Habermas cita aprobatoriamente esta tesis en su *Teoría de la acción comunicativa* (Madrid: Taurus, 1987), vol. I, pág. 357.

²⁴ Así se abre su ensayo «Gramática universal», en *Ensayos de filosofía formal*, *loc. cit.*, pág.

Pues bien, es precisamente dentro de este marco donde, como han demostrado D. W. Smith y R. McIntyre²⁵, se ha producido una confluencia sorprendente de la semántica de mundos posibles y la teoría de la intencionalidad de Husserl. Dedicaré el resto de estas páginas a presentar un esbozo de ese notable paralelismo.

El tema general de la fenomenología de Husserl es la intencionalidad de la mente humana. Por intencionalidad se entiende ese rasgo de la conciencia por el que ésta parece apuntar o señalar *algo*, un cierto tipo de objeto. La visión es visión *de* algo; la creencia es siempre creencia *en* algo o creencia *de* que algo es el caso, y así sucesivamente. 'Conciencia' es aquí un término cobertor útil por el que nos referirnos a lo que Russell llamó las actitudes proposicionales: creencia, conocimiento, percepción, esperanza, temor, etc.; es decir, los diversos tipos de fenómenos mentales. Brentano pensó que lo que caracteriza a un fenómeno mental como *mental* es su relación con un tipo de objeto *sui generis*, pero en seguida se hicieron manifiestas las dificultades de un punto de vista así. Husserl hizo frente a esas dificultades proponiendo una teoría de la *mediación*. Según esa teoría, aunque en todo acto mental el sujeto apunta a algo, un objeto, no es imprescindible que ese objeto exista. Sí que es necesario que el acto se halle vinculado a una entidad abstracta, consistente en un conjunto de propiedades o rasgos. Si ese conjunto de propiedades determinan unívocamente un objeto, éste es el objeto al que el acto apunta. Esa entidad que desempeña una función intermedia o de mediación es el *contenido intencional* del acto, el *noema* del acto. Un noema, y especialmente su *sentido*, determina un objeto porque las propiedades de que consta (en su conjunto) determinan unívocamente ese objeto. Resulta realmente decisivo en este análisis que no confundamos objeto real y contenido intencional. Son entidades distintas, de las cuales la primera está unívocamente determinada por la segunda. La conciencia no es conciencia del objeto intencional, sino del objeto real, cuando éste existe. El objeto al que la conciencia apunta, «*el objeto intencional de la representación es el mismo que su objeto real y -dado el caso- que su objeto exterior, y es un contrasentido distinguir entre ambos*²⁶.»

Es sabido que para Husserl se accede a los noemas -para ser exactos, los noemata- por medio de un tipo especial de análisis al que se conoce como análisis fenomenológico; un análisis en el cual se practica la *reducción fenomenológico-transcendental*. Este análisis se centra en un determinado *nivel de descripción* de los actos mentales. Lo característico de ese nivel de descripción es que en él suspendemos toda referencia a, o ignoramos, tanto los objetos reales de esos actos como los nexos que vinculan al sujeto de dicho acto con su origen causal y las circunstancias contextuales en que tiene (o tuvo) lugar. Así, por ejemplo, cuando investigamos el noema de mi apercibimiento (visual) de que Juan lleva una corbata de seda, hacemos abstracción de si Juan llevaba efectivamente una corbata de seda, de si yo estaba en situación de reconocer que Juan llevaba una corbata de seda, de si las condiciones perceptivas ambientales eran las adecuadas, y así sucesivamente. Todo lo que importa es

²⁵ Husserl and Intentionality [HI en lo que sigue] Dordrecht: D. Reidel, 1982.

²⁶ Investigaciones lógicas (Madrid: Revista de Occidente, 1967), V, Apéndice a la #11, pág. 227 del vol. II.

el contenido y la naturaleza del noema. Si se desea decirlo así, el objetivo del análisis es qué significa ver un cierto estado de cosas. Es decir, en virtud de qué mi percepción es percepción de que Juan lleva una corbata de seda y no, por ejemplo, de que los árboles están comenzando a florecer.

Las connotaciones semánticas de este análisis de la intencionalidad de la conciencia resultan obvias cuando la distinción fregeana entre signo, sentido y referencia y las doctrinas que la acompañan son familiares. El mismo Husserl admitió expresamente tal paralelismo cuando escribió -en la que es la cita favorita de Husserl para Follesdal- que «el noema es una generalización de la noción de significado en el dominio de todos los actos²⁷.» Esto no es, sin embargo, suficiente para propugnar que dicha teoría sea para el análisis de los actos (estados) mentales lo que la semántica de los mundos posibles es para el análisis de oraciones como

(1) Veo que Juan lleva hoy una corbata de seda.

El modo de analizar estas oraciones característico de semejante enfoque -que se debe a J. Hintikka²⁸- postula que sostener que (1) es una oración verdadera equivale a afirmar que

(2) Juan lleva hoy una corbata de seda

es verdadera en todos los mundos posibles que resultan *compatibles con lo que Juan ve en éste*, el real. Es decir, que especificar el contenido de mi acto mental consiste en distinguir esos mundos en los que Juan lleva hoy una corbata de seda de *esos otros* mundos posibles en los que tal cosa no es el caso. Separar unas alternativas perceptivas al mundo real de las otras es justamente decir qué es lo que veo (ahora).

Smith y McIntyre hacen dos cosas que, a mi modo de ver, son muy importantes. Argumentan de forma detenida (en *HI*, caps. V y VI) que el análisis que Husserl hizo del noema en términos de la noción de *horizonte* del objeto del acto reproduce casi literalmente, en lo que a la letra se refiere, pero muy fielmente, en lo que concierne al espíritu, los términos del análisis semántico. Husserl llamó horizonte del objeto de una percepción visual al conjunto de las posibilidades que el sentido del noema del acto deja abiertas. Y llamó horizonte del acto al conjunto de todos los posibles actos del mismo tipo que tienen el mismo sentido que el primero, pero que añaden más propiedades o determinaciones del objeto que

²⁷ *Ideen 3* (The Hague: Martinus Nijhoff, 1952), pág. 89.

²⁸ «On the Logic of Perception», en *Models for Modalities*, Dordrecht: D. Reidel, 1969; «Information, Causality, and the Logic of Perception», en *The Intentions of Intentionality and New Models for Modalities* Dordrecht: D. Reidel, 1975. J. N. Mohanty («Intentionality and Possible Worlds», en H. Dreyfus, ed. *Husserl, Intentionality and Cognitive Science; Husserl & Frege*, pág. 83) ha argüido contra la sugerencia de interpretar los noemata en términos del concepto de mundo posible que, al no ser epistémica la noción de mundo posible, no explica qué es lo que se capta cuando uno entiende una expresión. Esta objeción de Mohanty sería válida, si los mundos posibles se invocaran para analizar conceptos modales aléticos, pero no cuando el análisis lo es de actitudes psicológicas (como la creencia, o la percepción). Aquí, los mundos posibles implicados son aquellos compatibles con lo que el sujeto *cree* o *percibe*, por lo que la noción de mundo posible utilizada sí que es epistémica.

éste. El horizonte del acto de mi ejemplo lo constituiría la diversidad (*Mannigfaltigkeit*) de percepciones de que Juan lleva una corbata de seda y *algo más*, posiblemente distinto en cada caso: que lleva sombrero de fieltro, que lleva camisa blanca, etc. Para cada uno de estos actos, el consiguiente objeto intencional será un estado de cosas más restrictivo: que Juan lleva una corbata de seda y que lleva sombrero de fieltro; o que Juan lleva una corbata de seda y que lleva camisa blanca. Y así sucesivamente. Obviamente, cuando hemos especificado todas las posibles formas de determinar más el objeto del acto hemos especificado todos los mundos posibles en que Juan lleva una corbata de seda que son compatibles con lo que yo veo en éste.

Pero, segundo, Smith y McIntyre muestran (*HI*, cap. IV, Parte 3, y cap. VIII) cuáles son los límites del método de la reducción fenomenológica al poner de manifiesto cómo la 'puesta entre paréntesis' de los vínculos causales coloca a Husserl ante serios obstáculos a la hora de explicar cómo un noema puede prescribir un determinado objeto. Si estoy viendo el rostro de Ana, la estoy viendo a *ella*, aunque María sea una hermana gemela de Ana con un rostro indistinguible del de la primera. ¿Cómo es posible semejante diagnóstico, si seguramente -la hipótesis es perfectamente razonable- los noemata de los rostros de Ana y María serían idénticos? Esta dificultad es absolutamente paralela a una que se ha debatido muy extensamente en la reciente teoría semántica. La Nueva Teoría de la Referencia ha sido elaborada, a lo largo de la pasada década, como respuesta a una cuestión análoga planteada en el análisis semántico de los términos singulares (nombres propios, expresiones deícticas y demostrativas, descripciones definidas) y para los términos de especie natural (cf. más arriba). El paralelismo puede trazarse del siguiente modo: Primero, así como, según Husserl, todo acto mental tiene (necesariamente) un contenido intencional y (sólo contingentemente) un objeto real, toda expresión de una lengua tiene, según Frege y otros autores, (necesariamente) un sentido y (contingentemente) un referente extralingüístico. En segundo lugar, del mismo modo que el contenido intencional determina (a menudo) un objeto real, también el sentido determina (a menudo) un referente. Y, en tercer lugar, noema y sentido hacen esto en la medida en que, como conjuntos de determinaciones o propiedades que son, hay una entidad que las posee en exclusiva. El problema, descubierto por Kripke y por Putnam, para la teoría semántica, y extendido por Smith y McIntyre para la fenomenología husserliana, es que esta elucidación de las relaciones entre la expresión y su referente no puede ser correcta: tenemos para la segunda el problema de rostros idénticos de personas distintas; tenemos para la primera que ni el estado mental del hablante (Putnam) ni ningún conjunto de condiciones necesarias y suficientes basta para explicar por qué un término singular '*t*' refiere a un objeto *x*. No es éste el momento de volver a repetir los argumentos de Kripke. Están ahí y han hecho de *Naming and Necessity* un lugar clásico de la reciente filosofía del lenguaje²⁹.

²⁹ La versión castellana, *El nombrar y la necesidad*, ha sido editada en México por la U.N.A.M., 1982. Esta obra recoge conferencias inicialmente dadas en 1970. Una panorámica reciente sobre estos temas se contiene en las colaboraciones de G. Bealer & U. Mönningh, N. Salmon y G. Forbes al volumen IV del *Handbook of Philosophical Logic*, titulado *Topics in the Philosophy of Language*, ed. por D. Gabbay & F. Guenther, Dordrecht: D. Reidel, 1989.

Es éste un ejemplo, muy importante a mi juicio, de cómo puede llegarse a obstáculos que, si no los mismos, son traducibles entre sí, desde tradiciones filosóficas diversas; y de cómo cabe aprender a superarlos desde un lado de la barrera viendo cómo se han resuelto en la otra. Un ejemplo, por ello, que puede animarnos a pensar que en filosofía nuestra particular formación no tiene por qué impedir la discusión y la colaboración con quien parte de métodos y supuestos muy distintos. Pero también es cierto que no impedir no es lo mismo que facilitar.

Juan José Acero
Septiembre, 1989



LA FILOSOFIA ANALITICA Y SU CONCEPCION DEL METODO DE LA FILOSOFIA

Ernesto Tugendhat

Quiero presentar un esbozo sobre lo que pienso que es el método analítico en la filosofía. No es una concepción aceptada, sino un intento propio, basado en parte sobre ideas de Wittgenstein. Seguramente corresponde a mucho, pero no a todo lo que se está haciendo en la filosofía analítica. En general los filósofos analíticos no reflexionan sobre lo que están haciendo. En efecto, también casi ningunos de los filósofos no-analíticos reflexionan sobre lo que están haciendo. Existen muchos estudios filosóficos sobre los métodos de la ciencia, pero casi nada sobre el método de la filosofía misma.

Si hacemos una investigación sobre el método de la filosofía, debemos primero ponernos de acuerdo sobre lo que *es* filosofía, es decir, ¿qué es lo que todos estamos haciendo cuando filosofamos?. Ya esto parece ser una cuestión de demasiada amplitud y quizás demasiado disputada. Pero pienso que se pueden enumerar dos rasgos que son más o menos característicos de casi todo lo que se ha llamado filosofía. Uno es que tiene algo que ver con la totalidad -con la totalidad no de las cosas, sino con la totalidad de nuestra vida, o como también podríamos llamarlo más simplemente: con nuestra vida en general. Aún si se tematizan en la filosofía aspectos especiales, parecen siempre ser aspectos especiales de nuestra vida como totalidad.

El segundo rasgo característico de todo lo que se ha llamado filosofía es que consiste en una clarificación de conceptos. Ahora no se trata en la filosofía de cualesquiera conceptos, sino específicamente de aquellos conceptos que tienen algo que ver con nuestra vida en general. Así que los dos rasgos que acabo de mencionar -de que se trata de nuestra vida en general y que se trata de una clarificación de conceptos- están entrelazados entre sí. No quiero pretender que toda filosofía ya está definida sólo por estos dos rasgos, pero sí quiero afirmar que aunque muchas concepciones de filosofía contienen más que estos dos rasgos, estos dos rasgos forman un núcleo que se encuentra en todas las concepciones de filosofía. Pensemos por ejemplo en Sócrates y Platón. Las preguntas que ellos pusieron delante de la gente eran preguntas como «qué significa ser bueno, qué entendemos bajo saber, qué significa identidad». Ahora estos conceptos, lo bueno y lo malo, el saber, la identidad, o también tiempo, verdad, justicia, voluntad, responsabilidad, etc., etc., son precisamente conceptos sin los cuales no podemos concebir nuestra vida, son por lo tanto conceptos que tienen algo que ver con nuestra vida en general. O tomemos la Lógica de Hegel. Lo que Hegel hizo en su Lógica no era otra cosa que una manera muy especial de poner en orden y así clarificar los conceptos básicos de nuestra vida. O tomemos aún a Heidegger. Menciono a Heidegger porque en Heidegger el hecho de que se trata de los conceptos

básicos de nuestra vida no está tan claramente en la superficie de lo que él mismo dice. Según él el tema de la filosofía es el ser. Pero, ¿qué entiende Heidegger bajo «ser»? El ser en Heidegger es el horizonte de todo nuestro entendimiento. Ahora bien, nuestro entendimiento en general es idéntico con lo que yo llamé nuestra vida en general. En realidad es quizás mejor hablar de entendimiento que de vida, pues la filosofía naturalmente no trata de nuestra vida biológica, sino de la vida precisamente en el sentido de nuestro entendimiento.

Para aclarar mejor de qué conceptos se trata y cómo estos conceptos se delimitan respecto de los demás conceptos, podemos recurrir a otro concepto que parece ser característico de toda filosofía. Me refiero a la oposición entre lo apriori y lo empírico. Generalmente cuando se hace la distinción entre lo apriori y lo empírico se trata de enunciados o sea de juicios apriori o empíricos. Pero en el mismo Kant, para el cual el problema de los juicios apriori tenía una importancia tan grande, en un sitio algo recóndito -en la parte sobre el método de la filosofía al final de la *Crítica de la Razón Pura*- se encuentra una referencia a conceptos apriori dados. Por ejemplo, espacio, tiempo, causalidad, número, etc., son para Kant conceptos apriori dados, y ahora podemos decir que estos conceptos son precisamente los que presuponemos en nuestro entendimiento en general. Kant delimita los conceptos apriori dados de conceptos empíricos como por ejemplo el concepto de lo rojo o el concepto de oro. Conceptos como rojo, oro, gato, ajedrez (para dar algunos ejemplos muy diferenciados entre ellos) no son conceptos apriori dados, no pertenecen a nuestro entendimiento en general, podemos muy bien pensar que podríamos vivir sin ellos. Esto significa que los conceptos apriori dados parecen ser conceptos indispensables para nosotros. Con eso llegamos a otro concepto que muchos filósofos han usado para aquello de lo que trata la filosofía. Si estos conceptos son indispensables, son necesarios. No se trata de juicios necesarios, sino de conceptos necesarios, y esto parece tener el sentido de que son necesarios para nuestro entendimiento. Esto nos lleva a la siguiente pregunta: si se trata de conceptos necesarios y apriori dados, ¿tendría esto por consecuencia que también los enunciados de la filosofía es decir los juicios que la filosofía hace sobre estos conceptos son apriori? Pero se sabe que juicios apriori tienen que tener el sentido o de apriori sintéticos o de apriori analíticos. Dejo aquí al lado la opinión muy especial de Kant de que hay juicios sintéticos apriori, pues para creer en la posibilidad de ellos hay que hacer toda una serie de presuposiciones que son muy específicas de Kant.

Hay que tener en cuenta para ello su concepción de idealismo trascendental. Además, aún en Kant, los juicios apriori sintéticos no son juicios en que se clarifican conceptos; y seguramente la totalidad de las tres críticas de Kant no consiste en una serie de juicios apriori sintéticos. ¿Deberíamos entonces decir que los juicios a los cuales la filosofía quiere llegar son juicios analíticos? Pero si son analíticos, serían tautológicos, y parece obvio que ninguna filosofía jamás consista en tautologías. Si entonces los juicios de la filosofía no son ni apriori sintéticos ni apriori analíticos, la única alternativa que parece permanecer es que los juicios de la filosofía son empíricos. Esto es una tesis que en efecto voy a defender al final de mi ponencia. Si por el momento, Vds. me permiten asumir esto como ya concluido, esto significaría que la filosofía trata de conceptos apriori dados pero consiste en juicios empíricos. Esto permite una neta delimitación entre los siguientes tres campos de investigación: primero, ciencias empíricas; segundo, lógica y matemáticas, y tercero, filosofía.

Las ciencias empíricas formulan juicios empíricos haciendo uso de conceptos empíricos, lógica y matemática hacen juicios apriori analíticos pero no tratan de (sino usan) conceptos apriori dados, la filosofía por el contrario trata de conceptos apriori dados pero en juicios empíricos.

Si lo que acabo de decir más o menos corresponde a lo que hacemos en la filosofía, entonces la cuestión del método tiene que tener el sentido de cuál es el método con que podemos clarificar conceptos y, en especial, conceptos apriori dados.

Antes de contestar a esta pregunta, tengo que hacer un paso intermedio. Me refiero a una famosa cita en las Confesiones de San Agustín, donde dice: «¿Qué es el tiempo? Si nadie me pregunta lo sé, pero si se me pregunta, no lo sé». Esta afirmación de San Agustín, en nuestros tiempos reafirmada por Wittgenstein, puede servir muy bien para caracterizar la diferencia entre conceptos apriori dados y conceptos empíricos. De los conceptos empíricos es siempre posible, en contraste con los conceptos apriori dados, que alguien los desconozca; pero, si en cambio, los conoce, entonces también puede explicarlos. Tomemos como ejemplo un color un poco menos conocido que el rojo, el bermellón; yo personalmente no lo sabría reconocer, pero si alguien lo conoce, me lo podría mostrar. El caso del oro es más complicado pero en nuestro contexto es similar. La pregunta «qué es oro?» no se puede contestar con simplemente mostrarlo. Para responder a esta pregunta hay que hacer investigaciones empíricas sistemáticas, por la razón de que en el caso del oro se trata de una estructura. Pero en nuestro contexto se puede decir lo mismo que en el caso del bermellón. Se puede desconocer qué es oro, pero si alguien tiene la explicación científica y así lo conoce, claramente puede contestar la pregunta qué es oro. El caso del tiempo y todos los otros conceptos apriori dados es completamente diferente. Por una parte, no se puede concebir que alguien no entienda qué significa la palabra tiempo, -no se puede concebir precisamente porque es un concepto apriori dado, es decir un concepto que es indispensable para nuestro pensamiento. Sólomente niños pequeños no entienden esta palabra ni las palabras que son esenciales para su entendimiento, es decir «antes» y «después». Lo mismo ocurre con todos los conceptos apriori dados. Que son apriori dados no significa que nacemos con ellos, sino que no hay entendimiento adulto que no sepa perfectamente usar estas palabras, por ejemplo «antes» y «después», aunque como le ocurrió a San Agustín en general no sepamos cómo explicarlas. Con esto he dado una explicación más precisa de lo que San Agustín quiso decir cuando dijo algo, paradójicamente, que sí sabe cuando no se pregunta y no sabe cuando se pregunta. Saber aquí tiene dos sentidos. Que «sí lo sabe cuando no se pregunta» significa que sabe perfectamente usar esta palabra, es decir, la entiende, y que «no sabe cuando se le pregunta» significa que no sabe explicarla. Con los conceptos apriori dados tenemos esta distinción -entre entender y saber explicar- que no tenemos con los conceptos empíricos. El primer filósofo que se dió cuenta de esta situación fue Sócrates. Sócrates o el joven Platón inauguraron lo que para toda filosofía desde ellos quedó claro: que la filosofía, en contraste con las ciencias, es una investigación reflexiva. Es reflexiva porque siempre ya entendemos estos conceptos y no tenemos que hacer, sobre ellos ninguna investigación empírica. Pero ahora viene la gran pregunta: ¿en qué consiste esta reflexión filosófica? Esta es precisamente la pregunta sobre el método de la filosofía. Para esta pregunta existen en la historia de la filosofía, según lo que yo alcanzo a ver, tres respuestas.

La respuesta de Platón era que la reflexión consiste en una intuición interna o intelectual. Esta respuesta fue dominante en gran parte de la historia aunque muchos filósofos como por ejemplo Heidegger simplemente la tomaron como algo presupuesto. El filósofo que más agudamente reflexionó sobre este método de intuición de esencias fue Husserl. La segunda respuesta, ya también anticipada en Platón pero desarrollada más profundamente en el idealismo alemán es que el método de la Filosofía es la dialéctica. Hay otro sentido de dialéctica que en mi opinión es un buen sentido, pero la dialéctica de Fichte y de Hegel es en mi opinión algo completamente errado, pues era un intento de una quasi deducción de los conceptos apriori a partir de uno o de dos conceptos fundamentales. Pero en verdad sólo existe la deducción lógica, que no es una deducción de conceptos sino de juicios, y aplicar la idea de deducción a los conceptos no llega a tener sentido. La tercera respuesta es la de la filosofía analítica y consiste en entender que la reflexión filosófica consiste en demostrar cómo se usan las palabras correspondientes; y demostrar cómo se usa una palabra significa precisamente explicar esta palabra.

Esta afirmación de que la clarificación de conceptos consiste en la explicación del uso de las palabras se encuentra casi siempre con una fuerte resistencia. En primer lugar se objeta que no nos interesa el mero uso de las palabras sino que lo que queremos es investigar las cosas. Pero ¿qué cosas? En el caso de conceptos empíricos esta objeción sería válida. Por ejemplo en el caso del oro sería tonto simplemente reflexionar sobre el uso de la palabra; no, ahí es necesario hacer investigaciones empíricas, y esto precisamente significa la investigación de las cosas. Recurrir a las cosas significa recurrir a la experiencia y especialmente a la experiencia sistemática de las ciencias. Pero ¿qué sentido tiene el postulado de recurrir a las cosas cuando se trata de conceptos apriori dados? Podría solamente tener el sentido de recurrir a una experiencia intelectual, y esto es precisamente lo que afirma Husserl cuando dijo que hay algo como una intuición de esencias, un ver intelectual. Sólo que, primero, nunca se ha logrado demostrar en qué consiste este ver intelectual y cómo alguien lo puede llevar a cabo, y segundo, aun si existiese una tal intuición, ella no tendría ningún valor para el entendimiento mutuo o intersubjetivo. Y como la filosofía no es una investigación solitaria sino que cada cual quiere demostrar algo a todos los demás, el método intuitivo aun si existiese sería completamente inútil para la labor filosófica intersubjetiva. La única manera como se puede clarificar un concepto intersubjetivamente, excepto si es un concepto científico, es por la explicación del uso de la palabra. La clarificación de conceptos es, por lo tanto, una clarificación de sentidos de palabras.

Pero entonces -y esto es la segunda objeción- ¿no se reduce la filosofía a la lingüística? Esta pregunta es muy importante porque la manera filosófica de explicación de palabras se puede aclarar precisamente cuando uno la distingue de la explicación lingüística. Hay una distinción fundamental entre semántica filosófica y semántica lingüística, aunque existen filósofos analíticos como Donald Davidson que no la reconocen. Esta distinción consiste en que la semántica filosófica es precisamente reflexiva, mientras que la semántica lingüística no lo es. La semántica lingüística se caracteriza esencialmente por el uso de un metalenguaje. Se explican los sentidos de las palabras de un idioma ajeno usando las palabras del propio idioma. Esto significa que se presupone en la lingüística que ya se

entienden otras palabras correspondientes a las palabras que se explican. Pero esto no es una explicación del entendimiento mismo de esas palabras. Por el contrario, la tarea de la explicación filosófica consiste en explicar una palabra sin usar una palabra correspondiente; lo que tenemos que hacer, pues, es darnos cuenta de cómo explicaríamos una palabra que corresponde a un concepto apriori dado a alguien que no la conoce aún ni conoce una palabra correspondiente en otro idioma. Esta es la razón por qué en Wittgenstein la pregunta acerca de cómo explicaríamos tales palabras, por ejemplo las palabras «antes» y «después», a un niño, es tan importante. La respuesta a la pregunta de San Agustín es: sólo cuando nos damos cuenta de cómo explicaríamos la palabra a alguien que no la entiende aún, nos damos cuenta de cómo entendemos nosotros mismos la palabra. Explicar una palabra en este sentido, es hacer explícito lo que ya entendemos.

Pero hay una segunda diferencia en relación a la lingüística, a saber, que muchas veces tenemos que inventar en la filosofía nuevas palabras, palabras terminológicas. Heidegger seguramente ha exagerado esta necesidad, pero en el caso de Aristóteles se puede muy bien estudiar la necesidad de inventar palabras. Ahora bien, esto mismo se podría usar como un argumento de que la filosofía no puede consistir en la mera explicación de las palabras ya existentes. Esto es cierto, pero no es un argumento contra el método analítico, porque los conceptos apriori dados algunas veces no se encuentran en palabras, sino en estructuras del lenguaje. Por ejemplo las palabras «sujeto» y «predicado», tan fundamentales para la filosofía aristotélica como también para la filosofía analítica moderna, son invenciones filosóficas, pero lo son precisamente para explicar estructuras lingüísticas dadas.

Tal vez desde afuera puede parecer muy fácil el método de la explicación, pero en verdad es una tarea muy ardua. La razón es que raros son los casos en que la pregunta por una explicación se puede contestar por una simple definición. Un caso de este tipo fácil era la respuesta a la cual llegó Platón en su diálogo Teeteto sobre qué significa saber. Aunque la respuesta queda como puramente tentativa al final del diálogo, esta respuesta quedó como respuesta standard para toda la historia de la filosofía hasta Russell. Según Platón como también según Russell, se dice de una persona que sabe una proposición X, cuando cumple tres condiciones; primero, la persona cree que X, segundo X es verdad y tercero la persona tiene una justificación suficiente para X. Aún en este caso fácil Platón tuvo que escribir todo un largo diálogo para llegar a esta conclusión. Pero en general la situación es más difícil; no se puede llegar a una definición, y esto por dos diferentes razones. La primera es que, en general, nuestros conceptos apriori dados no son aislados, sino que forman lo que se podría llamar una red de conceptos; en términos de la filosofía analítica esto significa que no podemos explicar el sentido de una palabra por una definición, sino que tenemos que encontrar el sentido de esta palabra en su conexión con los sentidos de otras palabras. Pongamos como ejemplo el concepto de moral. Uno tiene en un tal caso que averiguar muy cuidadosamente, cuáles son las otras palabras esencialmente importantes para entender la palabra «moral» y cómo se explican; pensemos en palabras como «bueno» y «malo», «deber», «encomio» y «reprobación». Esta conexión con otras palabras es el caso más típico para casi todo lo que se está haciendo en la labor filosófica, pero hay otra razón de por qué una definición no es accesible. Tal es el caso de las palabras que son tan fundamentales que no se pueden explicar verbalmente. Piensen por ejemplo en la palabra «y» o en la palabra

«no». Un caso de este tipo son también las palabras fundamentales para entender el tiempo, en especial las ya mencionadas palabras «antes» y «después», como también «futuro», «presente» y «pasado».

¿Cómo se podría explicar estas palabras a alguien que no las conoce todavía, como por ejemplo a un niño? Wittgenstein lo ha demostrado en su Libro Marrón, de una manera que me parece ejemplar, pero que no tengo tiempo de reproducir aquí. Ahora supongamos que los dos casos -el de las palabras que no se pueden explicar verbalmente y el caso en que la palabra está conectada con otras- se combinan entre sí. Sospecho que un caso así es el del concepto de identidad. Nadie ha explicado todavía la expresión «x es idéntico con y».

Debo añadir que una gran parte de la labor analítica es crítica -crítica contra la filosofía tradicional. Es un malentendido muy difundido que la filosofía analítica quiere acortar o disminuir la dimensión filosófica. Aunque esto es cierto en el caso de muchos filósofos analíticos, esto no tiene nada que ver con el sentido de la filosofía analítica. La crítica analítica contra la filosofía tradicional es una crítica contra la ingenuidad en el uso de las palabras. La polémica se dirige contra el mero uso de las palabras sin darse cuenta cuidadosamente de su sentido o mejor dicho, sentidos, como ya lo hizo ejemplarmente Aristóteles, a quien se le puede llamar el primer filósofo analítico.

Quiero defender ahora la asersión que hice antes de que los juicios a que la filosofía aspira cuando clasifica los conceptos apriori dados son empíricos.

Esto podría parecer que está en contraposición con la distinción que se hace entre la filosofía y las ciencias empíricas. Pero el sentido de esta contraposición está completamente contenido en el hecho de que la filosofía es una investigación reflexiva y no se refiere a hechos dados en la experiencia. Que es reflexiva significa que en el filosofar reflexionamos sobre nuestro entendimiento, pero nuestro entendimiento no es una conciencia pura sino que es un hecho empírico. La medida de verdad o adecuación de una clarificación de un concepto apriori dado es que la clarificación sea conforme al entendimiento que nosotros *de hecho* (y esto significa empíricamente) tenemos. Si tomamos una vez más como ejemplo el caso fácil de la explicación del concepto de saber, la pregunta acerca de si la explicación que Platón dió es correcta o no, significa si corresponde adecuadamente a un rasgo que de hecho tenemos en nuestro entendimiento. La investigación filosófica es por lo tanto tan empírica como la ciencia y la diferencia consiste simplemente en la dirección que la investigación toma. La investigación filosófica va hacia adentro; se puede también decir, menos metafóricamente, que se efectúa en primera persona, pero primera persona del plural -se trata de *nuestro* entendimiento- mientras que la investigación científica va hacia afuera, en el sentido normal de experiencia. Digamos, finalmente, que cuando se trata de una ciencia lingüística, se efectúa en tercera persona.

Esta manera de ver las cosas permite también ver el problema del relativismo filosófico de una nueva manera. Muchas veces se oye la pregunta: ¿qué garantía podemos tener de que, por ejemplo, todo lenguaje humano tenga la estructura de sujeto y predicado, que esto sea una estructura realmente universal?. Se ha alegado que el chino no tiene esta estructura. Esto se debía a una confusión entre estructura predicativa sintáctica y estructura predicativa semántica. Pero supongamos que un día llegáramos a encontrar un idioma que realmente no contuviese una estructura predicativa en su semántica. En este caso simplemente se

ampliaría nuestra concepción de entendimiento, así como se amplió nuestra concepción de espacio cuando se descubrieron los espacios no-euclídeos. No hay nada asombroso en eso excepto para quien pensara que un concepto apriori dado tiene, como para Kant, el sentido que es dado a una conciencia pura y por lo tanto debe permanecer el mismo para todos los humanos y para todo entendimiento futuro. Dicho de otra manera, que un concepto es apriori dado no implica la afirmación de que es universal. Si concebimos los conceptos apriori dados, como conceptos dados en la reflexión para nosotros, la universalidad se limita a este «nosotros» y el ámbito de este nosotros se puede ampliar. Con esto el relativismo adopta un sentido que no tiene nada de temible. Al contrario, tanto como toda investigación científica es provisoria, toda investigación filosófica lo es también. No hay que dar pruebas de que lo que se encuentra en la filosofía es universal, porque los enunciados filosóficos bien entendidos no pretenden este carácter universal.

Y ante el caso de la amplificación del entendimiento al encontrar otros idiomas u otras maneras de ver el mundo, otras maneras de entender, ahí tenemos el punto donde la filosofía analítica, correctamente entendida, se encuentra con el método hermenéutico.

Cuando llega a tener un sentido hermenéutico, el método analítico pasa de la primera a la segunda persona. La segunda persona, siempre en el plural, pertenece también al método reflexivo y se distingue netamente de un pensamiento en tercera persona. La distinción es la siguiente: cuando encontramos otro idioma, si lo tratamos en tercera persona, entonces lo traducimos al nuestro, usamos nuestro idioma como metalenguaje y hablamos en tercera persona de «ellos». Cuando en cambio tratamos al otro idioma en segunda persona, hablamos no de ellos sino con ellos y les pedimos a ellos que nos expliquen sus palabras, y esta explicación tiene el mismo sentido que ha tenido desde el principio en la primera persona del plural. Ahora somos nosotros los que no entienden todavía las palabras de los otros ni tenemos palabras correspondientes en nuestro idioma, y si logramos que ellos nos expliquen sus palabras, las llegamos a entender y así se amplifica el horizonte de nuestro propio entendimiento.

Ernesto Tugendhat
Septiembre, 1989



SOBRE EL PRETENDIDO REALISMO BASICO DEL TRACTATUS

Alfonso G. Suárez

1. En su reciente libro *The False Prison*¹ el Profesor David Pears sostiene que el sistema del *Tractatus* es básicamente realista: «Wittgenstein -afirma Pears- vió la estructura subyacente de la realidad como una retícula de estados de cosas posibles, con los objetos en los puntos nodales, y son las naturalezas de los diferentes tipos de objetos las que determinan el modo en que se articula la retícula (...) esta retícula impone una constricción a todos los lenguajes fácticos: éstos sólo pueden describir la realidad en la medida en que se conformen a ella en su propia estructura interna» (p.6). Además la retícula «no es en modo alguno dependiente de ninguna contribución hecha por nuestras mentes» (p.9). Los objetos no son datos sensoriales² ni «entidades invocadas para amoldarse a una teoría semántica»³, sino fenómenos en el sentido kantiano, objetos del «mundo tal como lo encontramos». Son, por ello, entidades especificables independientemente y su identificación discurre por las vías clásicas usuales: o bien por descripción definida o por ostensión.

Pears afirma que «en el *Tractatus* el comienzo del lenguaje es la nominación de los objetos» (p.9). Este realismo básico es un «realismo acrítico» porque «los objetos son los *partenaires* dominantes en sus relaciones con los nombres y nada se dice sobre contribución alguna hecha por nuestras mentes en este punto» (p. 29). Es cierto que somos nosotros quienes imponemos nombres a los objetos, pero «una vez que un nombre le ha sido impuesto a un objeto, la naturaleza del objeto se hace cargo de la conducta lógica del nombre y la controla, haciendo que tenga sentido en algunos contextos oracionales pero no en otros» (p. 88).

Finalmente, según Pears, Wittgenstein «no suscribe el solipsismo en el *Tractatus*, porque eso le comprometería a tratarlo como una teoría susceptible de ser verdadera o falsa. Meramente concede que el solipsista ha logrado hacerse con un buen punto de un tipo que no puede enunciarse en el discurso fáctico sino sólo mostrarse (...) Tampoco, naturalmente es el primer Wittgenstein un idealista (...) Es, claro está, un realista» (p.188).

¹ D. Pears, *The False Prison. A Study of the Development of Wittgenstein's Philosophy*, Oxford: Clarendon Press; Vol. I, 1987; Vol. II, 1988. (En adelante las referencias a páginas sin más indicación serán a las de esta obra).

² Como pretenden M.B. y J. Hintikka en *Investigating Wittgenstein*, Oxford, Blackwell, 1986, cap. 3.

³ H. Ishiguro, «Use and Reference of Names», en P. Winch (ed.), *Studies in the Philosophy of Wittgenstein*, Londres: Routledge and Kegan Paul, 1969, pp. 20-50; la cita es de la p 21 (Hay traducción de L. Mirilas: *Estudios sobre la filosofía de Wittgenstein*, Buenos Aires: Eudeba, 1971). La misma interpretación es defendida por B. McGuinness en «The So-called Realism of the *Tractatus*», en I. Block (ed.), *Perspectives on the Philosophy of Wittgenstein*, Oxford, Blackwell, 1981, pp. 60-73.

En el presente trabajo trataré de mostrar que esta interpretación, aunque aparentemente cuadra con ciertas proclamas clave en el *Tractatus* y permite explicar elegantemente el cambio de la primera a la última filosofía de Wittgenstein como un abandono del realismo, sin embargo no logra dar cuenta adecuadamente del principio contextual formulada en 3.3 y fuerza a Pears a una lectura incorrecta de la sección 5.6, en la que se aborda el tema del solipsismo.

2. De acuerdo con la interpretación de Pears, el contacto entre nombre y objeto es previo al uso del nombre en proposiciones y es la identidad del objeto la que determina el uso del nombre. Pero esto parece entrar en conflicto con el principio contextual:

Sólo la proposición tiene sentido; sólo en el contexto de la proposición tiene un nombre referencia⁴.

Como es bien sabido, Wittgenstein tomó este principio de Frege y siguió prestándole adhesión en las *Investigaciones filosóficas*. No obstante, Moore recuerda que en sus lecciones de 1930-33 dijo que el principio en cuestión «es verdadero o falso», según como ustedes lo entiendan⁵. En el *Tractatus* ha de entenderse, al menos parcialmente, como una reacción contra la concepción russelliana de los nombres lógicamente propios. Para Russell, los nombres contrastan con los «símbolos incompletos» en que tienen significado por sí mismos, siendo este significado el objeto que denotan. Son por tanto irrelevantes los contextos proposicionales en que pueden aparecer. Pears caricaturiza atinadamente esta posición diciendo que, para Russell, «la imposición de nombres a objetos es exactamente como fijar etiquetas a equipajes, una operación que queda completada en el momento en que se ata el nudo» (p.102).

Pues bien, él pretende distanciarse de esta posición sin renunciar a la atribución de un realismo básico al *Tractatus*. Para ello propone una interpretación del principio contextual que habría de conciliar el carácter dominante del objeto en la relación nominativa con el hecho de que un nombre carece de referencia cuando se lo toma aisladamente. Según la interpretación de 3.3. propuesta por Pears⁶,

a) habría un acto inicial, de naturaleza similar al contemplado por Russell, de imposición del nombre al objeto; pero, aunque este acto es necesario para que el nombre lo represente, no es suficiente, pues

b) el contacto entre el nombre y el objeto queda en adelante condicionado al uso correcto del nombre en oraciones, siendo la condición «que las oraciones en las que aparece el nombre deben presentar posibilidades reales para la cosa nombrada» (p.75).

⁴ L. Wittgenstein, *Tractatus logico-philosophicus* (en adelante citado sólo mediante su numeración decimal), texto alemán y trad. inglesa de C.K. Ogden, Londres, Routledge and Kegan Paul, 1922; texto alemán y trad. de D.F. Pears y B.F. McGuinness, Routledge and Kegan Paul, 1961; 3.3. (Hay trad. de E. Tierno Galván, Madrid, Revista de Occidente, 1957; revisada en Madrid, Alianza Editorial, 1973; nueva trad. de J. Muñoz e I. Reguera, Alianza Editorial, 1988).

⁵ «Wittgenstein's Lectures in 1930-33», en G.E. Moore, *Philosophical Papers*, Londres, Allen & Unwin, 1959, p. 261 (Trad. de C. Solís, *Defensa del sentido común y otros ensayos*, Barcelona, Orbis, 1983).

⁶ Cf. pp. 75, 102-3, 109, 131.

Interpretado así, 3.3. a través de la condición mencionada, *matiza* el contacto nombre-cosa por imposición directa, pero no lo *reemplaza* por algo diferente.

Pears defiende esta propuesta de dos maneras: apoyándola mediante evidencia textual y tratando de desacreditar la interpretación rival de Ishiguro y McGuinness.

3. Debe quedar claro que, salvo que Pears se muestra impreciso sobre la naturaleza del acto inicial de imposición del nombre al objeto, también en este modelo la *imposición* queda completada en el momento en que concluye esta ceremonia. Así pues, por lo que respecta a (a) no hay diferencia sustancial con el modelo russelliano. Toda la diferencia reside en (b), la condición que *matiza* el etiquetado inicial.

La salida es sin duda ingeniosa. Pero, independientemente de sus méritos propios, ¿es ésta la posición de Wittgenstein? ¿Qué apoyos textuales puede aportar Pears para atribuírsela? Según él, «hay fuerte evidencia en favor de esta interpretación en los *Diarios*» (p.111). El pasaje en que se apoya reza:

Un nombre representa una cosa, otro otra cosa y ellos mismos están conectados; así el todo representa -como una figura viviente- el estado de cosas.

La conexión lógica ha de ser, por supuesto, posible entre las cosas representadas, y tal será siempre el caso si las cosas están realmente representadas⁷.

Creo que Pears malinterpreta este texto. Ciertamente el segundo párrafo establece una condición que, aparentemente, coincide con la condición que, en el modelo que Pears le atribuye a Wittgenstein, *matiza* la asignación inicial del nombre al objeto. Pero la condición que se expone aquí *no es una condición impuesta a la relación nominativa* -la relación en virtud de la cual un nombre se refiere a (*bedeutet*) un objeto, lo representa (*vertretet*) o nombra (*nennt*)- sino *una condición impuesta a la relación representativa (darstellende Relation) o figurativa (abbildende Beziehung)*. Esta última no es una relación entre nombres y objetos, sino entre la proposición («el todo») y el estado de cosas. Y lo que Wittgenstein está afirmando es que, si la conexión lógica entre los nombres no es posible entre las cosas representadas, la proposición no presentará un estado de cosas posible, no determinará un lugar en el espacio lógico. Que en este pasaje se están discutiendo las condiciones bajo las cuales la proposición presenta (*darstellt*) un estado de cosas, y no las condiciones bajo las cuales un nombre se refiere a (*Bedeutet*) un objeto, queda claro si echamos un vistazo a su contexto. La pregunta que acaba de plantearse no es «¿Cómo se refiere un nombre a un objeto?» sino «¿cómo determina la proposición el lugar lógico? ¿Cómo representa la figura un estado de cosas?»⁸.

⁷ L. Wittgenstein, *Notebooks 1914-1916*, texto alemán y trad. inglesa de G.H. von Wright y G.E.M. Anscombe, Oxford, Blackwell, 1961 y 1979, 4.11.14, (Trad. de J. Muñoz e I. Reguera, *Diario filosófico, 1914-1916*, Barcelona, Ariel, 1982).

⁸ *Ibid.*

En segundo lugar, aunque se tratara aquí de establecer una condición matizadora de la relación nominativa, y no de la relación figurativa, el pasaje sólo apoyaría la interpretación de Pears si la cláusula condicional fuese entendida como una condición suficiente. Pero hay razones para pensar que Wittgenstein hace aquí un uso laxo de «si» que expresa más bien una condición necesaria. Esta lectura de la cláusula condicional parece apoyada por la anotación del día siguiente, en donde leemos:

Para que una proposición represente un estado de cosas sólo resulta, pues, *necesario* que sus partes constitutivas representen las del estado de cosas y que aquéllas estén en una conexión posible para éstas.⁹

En cualquier caso, este pasaje demuestra una vez más que en esta secuencia de anotaciones Wittgenstein se ocupa de las condiciones que deben cumplirse «para que una proposición represente un estado de cosas» y no de las que serían exigibles para que un nombre esté por un objeto.

En tercer lugar, el pasaje que Pears cita es de octubre de 1914 y encontramos que en junio de 1915 Wittgenstein todavía no había llegado a aceptar el principio contextual. Así en una anotación de 14.6.15 se pregunta: «Pero ¿qué ocurre con la referencia del nombre fuera del contexto proposicional?» Y en la anotación del día siguiente su respuesta es un *rechazo* del principio contextual:

Está de todo punto claro que puedo adscribir a este reloj que tengo delante, y que marcha ante mis ojos, un nombre, y que este nombre tendrá referencia fuera de toda proposición (...)

Así pues, si en aquel momento Wittgenstein aún no había aceptado el principio contextual, mal puede leerse el pasaje que Pears cita como una explicación de la matización que ese principio impondría al contacto nombre-cosa por imposición directa.

Si tenemos, por otra parte, en cuenta que, por lo que hace a las consideraciones generales que Pears cita en favor de su interpretación¹⁰, él mismo reconoce que no son decisivas, hemos de concluir que la interpretación propuesta tiene, a lo sumo, una muy débil evidencia textual favorable. Veamos, entonces, el intento que Pears hace de desacreditar la interpretación rival.

4. Bajo la interpretación defendida por Ishiguro y McGuinness, la identidad del objeto referido por un nombre depende de las proposiciones en que aparece el nombre. Aunque Ishiguro cree que la referencia de un nombre depende no sólo de qué contextos proposicionales tienen *sentido* cuando se inserta el nombre en ellos, sino también de qué contextos proposicionales arrojan *verdad* cuando se inserta el nombre en ellos, McGuinness se conforma con lo primero.

Pears objeta que incluso la versión débil que ambos comparten «parece ir contra el principio de representación. Si la ‘posibilidad de las proposiciones se basa en el principio

⁹ *Ibid.*, 5.11.14; los subrayados son míos.

¹⁰ *Cf.* pp. 103-4.

de que los objetos tienen signos como sus representantes' (4.0312), ¿cómo puede depender la identidad de los objetos representados por estos signos de los sentidos de las proposiciones en que aparecen?» Así «según esta interpretación, no se podría establecer el sentido de una proposición antes de las referencias de sus nombres ni sus referencias antes de su sentido, y así el lenguaje nunca podría despegar del suelo» (p. 100). La dificultad que Pears encuentra con respecto a esta versión del principio contextual es paralela a la que Max Black encontró con respecto a la afirmación de Wittgenstein de que las elucidaciones que explican las referencias de los signos primitivos sólo pueden entenderse si se conocen ya las referencias de esos signos¹¹. Pero Wittgenstein no está diciendo que la comprensión de los nombres es previa a la comprensión de las proposiciones y a la vez que la comprensión de las proposiciones es previa a la comprensión de los nombres. Esto sería palmariamente circular. La captación de las referencias de los signos primitivos y la comprensión de las elucidaciones forman un todo inseparable porque la identidad de los objetos referidos por los nombres no es independiente del sentido de las proposiciones en que éstos aparecen. Esto no obsta para que, a fin de dar cuenta de la creatividad lingüística, reconozcamos que, como Dummett lo expresa¹², en el orden del *reconocimiento* el significado del nombre es primario. Entendemos la oración al entender las palabras que la componen y su estructura. Pero no podemos entender las palabras componentes independientemente de su aparición en proposiciones. No hay circularidad aquí, del mismo modo que no la hay en el hecho de que sólo podemos entender la oración «Es pobre pero honrada» entendiendo el significado de sus componentes y a la vez no podemos entender el significado del componente «pero» sin entender su aparición en contextos oracionales. Trátense *todas* las palabras como sincategoremáticas, siguiendo la recomendación de Quine¹³, y el misterio se disipa.

Pears reconoce que Ishiguro y McGuinness tienen una respuesta a su objeción:

La respuesta es que, aunque el análisis completo de una oración fáctica consiste en nombres en combinación inmediata mutua, la obra efectiva del análisis se termina en el nivel penúltimo, donde no hay nombres sino sólo cuantificadores y variables (...) toda la tarea del análisis se expresa en quíns y el paso final, que traduce el resultado al russelliano, es puramente decorativo (...) los nombres del *Tractatus* son nombres postizos (*dummy names*), porque no sirven para distinguir un objeto de otro, sino sólo para designar objetos que ya han

¹¹ 3.263. Dice Black que la «inquietante» respuesta que Wittgenstein dió a la pregunta de cómo puede comunicarse el significado de los nombres «es que es imposible explicar el significado de un nombre explícitamente: el único modo de transmitir el significado es usar el nombre en una proposición, presuponiendo con ello que el significado ya se comprende. Bajo este punto de vista, el logro de una referencia común por parte de hablante y oyente se vuelve misterioso» (*A Companion to Wittgenstein's Tractatus*, Ithaca, N.Y., Cornell University Press, 1964, pp. 114-5).

¹² M. Dummett, *Frege: Philosophy of Language*, Londres, Duckworth, 1973, pp. 3-7.

¹³ W.V. Quine, «Five Milestones of Empiricism» en *Theories and Things*, Cambridge, Mass., Harvard University Press, 1981.

sido distinguidos entre sí por los contextos proposicionales, proporcionados por el nivel penúltimo de análisis, en los que pueden insertarse sus nombres. Estos contextos proposicionales contienen expresiones predicativas que pueden entenderse sin familiaridad con los particulares designados por nombres propios russellianos y esto hace posible que el lenguaje despegue del suelo (pp. 100-1).

Pero, replica Pears, esta respuesta suscita a su vez la cuestión de cómo, en opinión de Wittgenstein, obtendrían entonces *sus* significados las expresiones predicativas que constituyen los contextos proposicionales. «¿Se supone que mantenemos constantes las referencias de los nombres propios mientras barajamos las expresiones predicativas para hallar en qué contextos proposicionales tienen *ellas* sentido? Una vez más parece que el lenguaje no podría despegar del suelo» (pp. 108-9).

Como el propio Pears reconoce, es difícil aventurarse en este terreno dada la obscuridad de las opiniones de Wittgenstein acerca de las expresiones predicativas. Tentativamente me atrevo a sugerir que una posible solución está en el hecho de que en realidad el nivel penúltimo en el que se termina la tarea afectiva del análisis *no contendría expresiones predicativas*. 3.1431-2 y 4.22 parecen apoyar la idea de que cualquier expresión funcional puede ser reemplazada por una cierta disposición u ordenación de sus argumentos. Así en vez de « $\phi(x,y)$ » podríamos escribir « x y y » (donde la colocación de las variables a esa distancia indica que entre los argumentos se da una cierta relación). Tampoco habría problemas con expresiones cuantificadas. Por ejemplo, en vez de « $(\exists y)fy$ » podríamos escribir « $(\exists y)\lambda a$ » (donde el hecho de que « y » se escribe invertida indica que al menos un objeto tiene una cierta propiedad)¹⁴. La tarea del análisis se expresaría entonces no en quinés, sino en ultra-quinés, que contiene cuantificadores y variables, pero carece de expresiones predicativas.

Hay otra «insuperable objeción» (p.113) que Pears hace a la interpretación de Ishiguro y McGuinness. Según él, «ambos dicen que Wittgenstein usa un criterio de identidad que hace idénticos a cualesquiera dos *objetos* en el *Tractatus* si son indiscernibles» (p. 105; el subrayado es mío). Pero, es bien sabido que Wittgenstein rechaza la identidad de los indiscernibles en 5.5302. De donde se seguiría que la interpretación de Ishiguro y McGuinness es errónea.

Aquí Pears debe de estar dando a entender que el principio de la identidad de los indiscernibles se sigue de su aparente contrapartida en el modo formal. En efecto, lo que Ishiguro y McGuinness le atribuyen a Wittgenstein es la proposición inversa del llamado principio de sustitutividad de los idénticos. Llamemos a esta proposición PIS. Podríamos formularla así:

¹⁴ Véanse sobre este punto I.M. Copi, «Objects, Properties, and Relations in the *Tractatus*», *Mind* 67 (1958): 145-65 y G.E.M. Anscombe, *An Introduction to Wittgenstein's Tractatus*, Londres, Hutchinson, 1959, cap. 7 (Trad. de M. Pérez Rivas, *Introducción al Tractatus de Wittgenstein*, Buenos Aires, El Ateneo, 1977).

(PIS) si dos nombres *a* y *b* son mutuamente sustituibles en toda proposición verdadera, *salva veritate*, son correferenciales.

Es este principio, y no la identidad de los indiscernibles, el que Ishiguro y McGuinness le atribuyen a Wittgenstein. De hecho mientras que en la versión de PIS de Ishiguro que Pears cita se habla de *nombres* y en la de McGuinness de *signos*, el principio de la identidad de los indiscernibles trata de *objetos*. Podríamos formularlo así:

(Id.Ind.) Si dos objetos *x* y *y* comparten todas sus propiedades, son idénticos.

Ahora bien, la aceptación del primero no nos compromete a la aceptación del segundo. Aunque consideremos falso a Id. Ind., hemos de convenir en que PIS es trivialmente verdadero. Pues si el antecedente de PIS se da, i.e. si *a* y *b* son mutuamente sustituibles en toda proposición verdadera *salva veritate*, serán sustituibles en «*a=a*», que expresa una proposición verdadera. Reemplazando entonces *a* por *b* a la derecha del signo de identidad, obtendremos «*a=b*» *salva veritate*. Pero si «*a=b*» expresa una proposición verdadera, *a* y *b* serán correferenciales. Por tanto PIS es verdadero. En conclusión, puesto que PIS es trivialmente verdadero e Id.Ind. falso, el primero no entraña el segundo.

5. Pears pretende que en el *Tractatus* Wittgenstein no abraza el solipsismo sino que ofrece un argumento en su contra. Esto no cuadra muy bien con lo que dice 5.62:

Pues lo que el solipsismo *pretende* es totalmente correcto, sólo que no puede *decirse*, sino que se muestra.

Que el mundo es *mi* mundo se muestra en que los límites *del* lenguaje (del único lenguaje que yo entiendo) significan los límites de *mi* mundo.

Pues bien, si lo que el solipsismo pretende es «totalmente correcto», ¿cómo podría Wittgenstein querer ofrecer un argumento en su contra? Pears, como otros muchos comentadores, se acoge en este punto a la confesa *inexpresabilidad* del solipsismo: si el propio Wittgenstein reconoce que la doctrina solipsista no puede ser dicha, ¿cómo podríamos atribuírsela? Pero Pears peca aquí contra un principio de exégesis que, en otro lugar de su libro, él mismo establece diciendo: «La ontología del comienzo no es algo de lo que se suponga que podemos prescindir porque es un intento de decir cosas que sólo pueden mostrarse. Por el contrario, aquí, *como en cualquier otro lugar del Tractatus*, la estricta imposibilidad de formular una tesis en el lenguaje fáctico es, si acaso, un signo de su importancia» (p. 112; el subrayado es mío). Si Pears hubiera sido coherente, habría aplicado también este principio a la sección 5.6. Si dejáramos de atribuirle a Wittgenstein una postura por el mero hecho de que él afirma que es inexpresable, tendríamos que decir que no suscribe prácticamente ninguna proposición del *Tractatus*. Tal es la irónica situación en que lo coloca la doctrina del mostrar. Si lo que el solipsismo pretende es totalmente correcto, el presunto argumento habría de dirigirse, si acaso, contra el intento de decirlo o contra el acto de decirlo, no contra el *contenido* de ese pretendido *acto*.

Mas, ¿por qué es infiel Pears a su propio principio de exégesis?. La respuesta es, creo yo, que al atribuirle a Wittgenstein un realismo básico se ve forzado a negar que fuera

solipsista. Sin embargo, en sus escritos anteriores Pears había adoptado la postura que aquí estamos defendiendo. Así en su conversación con Brian Magee sobre las dos filosofías de Wittgenstein presenta el solipsismo como ejemplo de verdad inexpressable, como las verdades de la religión, de la moralidad y de la propia filosofía¹⁵. Y en su artículo «Wittgenstein's Treatment of Solipsism in the *Tractatus*»¹⁶, afirma que ciertamente Wittgenstein fue solipsista, aunque su solipsismo no fue restrictivo, sino lingüístico o transcendental.

Pears señala que el contexto de la discusión del solipsismo en el *Tractatus* es la investigación del modo en que está limitada la realidad empírica. De acuerdo con 5.5561, «la realidad empírica está limitada por la totalidad de los objetos. El límite se muestra también en la totalidad de las proposiciones elementales». Así hay un límite *general* del lenguaje: sólo puede contener funciones de verdad de proposiciones elementales en las que se mencionan objetos simples. En los 5.55 se considera un intento fallido de poner además un límite *específico* a las formas posibles de proposiciones elementales. La conclusión que se alcanza es que la lógica no puede decirnos de antemano qué formas diversas de proposiciones elementales hay. Pears pretende entonces que los 5.6 discutirían otro intento fallido de imponer un límite específico -egocéntrico esta vez- al discurso fáctico. Wittgenstein no se ocuparía de un solipsismo de base fenomenológica que pretende usar mis datos sensoriales como fundamento para trazar los límites del lenguaje, sino de un solipsismo basado en el ego que pretende trazarlos usando mi yo como punto de referencia. El alcance del discurso significativo estaría determinado por el ámbito de mi familiaridad con objetos. O en el modo material: la realidad empírica, que está limitada por la totalidad de los objetos, quedaría restringida a los objetos presentados a mi ego.

Me parece que en este enfoque hay ya una distorsión inicial de la posición de Wittgenstein. De nuevo ocurre aquí que Wittgenstein se ve colocado en la irónica situación de tener que decir lo que, según la doctrina del mostrar, no puede decirse. Y, sin embargo, lo dice *de hecho*: «Los límites de mi lenguaje significan los límites de mi mundo» (5.6), «El mundo y la vida son una sola cosa» (5.621), «Yo soy mi mundo. (El microcosmos)» (5.63), «El yo entra en filosofía por el hecho de que 'el mundo es mi mundo'» (5.641)¹⁷. El carácter *aseverativo* de estas proclamas cuadra muy mal con la tesis de Pears de que Wittgenstein está considerando un intento fallido de imponer un límite egocéntrico al discurso fáctico. Tal es la fuerza asertórica de estas proclamas solipsistas que Pears, al comentar 5.621, cae en contradicción con sus previas afirmaciones acerca del realismo básico del *Tractatus*. En efecto, tras comentar que el mundo que es identificado con la vida es el mundo de los fenómenos, añade:

¹⁵ Cf. B. Magee, *Modern British Philosophy*, Londres, Secker & Warburg, 1971, «Conversation with David Pears: The Two Philosophies of Wittgenstein», p. 33.

¹⁶ *Crítica* 6 (1972): 57-84; reimpresso en S. Shanker (ed.), *Ludwig Wittgenstein: Critical Assessments*, vol. I, Londres, Croom Helm, 1986, pp. 170-84, cf. pp. 173, 178, 181-184 (Citado en adelante con las siglas WTST).

¹⁷ Y no se olviden las *aserciones* solipsistas aún más violentas que Wittgenstein hace en los *Diarios*. Especialmente agresivas son las de 2.9.16: «Qué me importa la historia» ¡Mi mundo es el primero y el último! Quiero informar acerca del mundo como yo lo encontré (...) Soy yo quien tengo que enjuiciar el mundo, que medir las cosas».

Si parece extraño que Wittgenstein identifique este mundo con la vida, vale la pena recordar que percibir e imaginar son cosas que hacemos nosotros. Sin embargo, su observación va aún más allá. No quiere decir que hagamos esas cosas contra el trasfondo de un mundo independiente: quiere decir que ellas *constituyen* el mundo (p.174).

Pero esto contradice su anterior afirmación de que el realismo básico del *Tractatus* «es un realismo acrítico, porque la cuestión de si contribuimos algo a la constitución de ese mundo ni siquiera se plantea» (p. 29).

6. Un aspecto discutible en la interpretación de Pears es su asunción de que los objetos, cuya totalidad limita la realidad empírica, aunque no necesariamente datos sensoriales, son objetos adecuados para la familiaridad, fenómenos, objetos del «mundo tal como lo encontramos» (pp. 93-9, 107-8, 154-6). De nuevo hay aquí un choque frontal con la interpretación rival de Ishiguro y McGuinness. En cualquier caso, las pruebas textuales que Pears aporta en favor de su propuesta distan de ser convincentes. En diversos lugares se apoya en 5.631, que reza:

Si yo escribiera un libro titulado *El mundo tal como yo lo encontré*, tendría que informar en él sobre mi cuerpo (...).

Pears utiliza este pasaje para descartar la interpretación fenomenista suscrita por Jaakko y Merrill Hintikka y para establecer la tesis de que los objetos son fenómenos en el sentido kantiano, no datos sensoriales mentales. Y la razón sería que, puesto que el informe del solipsista tendría que mencionar su cuerpo, su mundo no es un mundo de datos sensoriales. Esto no es convincente. El pasaje en cuestión no puede ser usado como soporte de una u otra interpretación de la noción de objeto porque lo que el solipsista menciona en su informe *no* son objetos simples, pues su cuerpo es obviamente un complejo.

Débil es también el apoyo que le prestan a su propuesta proclamas del *Tractatus* como las siguientes:

Si conozco un objeto, conozco también todas las posibilidades de su aparición en estados de cosas (2.0123).

Es obvio que un mundo imaginado, por muy diferente que sea del real, debe tener algo -una forma- en común con el real (2.022).

Pears comenta: «en este pasaje hay referencias a la posibilidad de nuestra familiaridad con (los objetos), y un mundo alternativo al nuestro es llamado ‘un mundo imaginado’. Estas indicaciones sugieren que los mundos posibles en todos los cuales existen los objetos son meramente los resultados de variar la disposición de los objetos que subyacen al mundo tal como lo encontramos y que por ello ‘limitan la realidad empírica’» (pp. 107-8). Ahora bien, esta conclusión sólo se seguiría si el conocimiento mencionado en 2.0123 fuese conocimiento por familiaridad y si un mundo imaginado fuese necesariamente un mundo

de objetos susceptibles de familiaridad. Pero, en primer lugar, no hay por qué tomar 'kennen' en 2.0123 («Wenn ich den Gegenstand kenne») como conocimiento por familiaridad, entendido como conocimiento de objetos dados a los cinco sentidos. Puede tomarse como conocimiento en el sentido en que puedo decir que conozco un teorema, un poema o una lengua. Y, en segundo lugar, 2.022 implica que cualquier mundo imaginado admisible siempre será el resultado de variar las disposiciones de los objetos, que tienen que ser comunes a todos los mundos posibles. Pero de ahí no se infiere que haya de tratarse de objetos del mundo fenoménico. Un mundo imaginado no tiene por qué ser un mundo de «objetos adecuados para la familiaridad»¹⁸. Por el contrario, del requisito de que los objetos pertenezcan a todo mundo posible más bien se infiere que no se trata de objetos espacio-temporales.

7. El argumento en contra del solipsismo que Pears le atribuye a Wittgenstein es un dilema que pretende arrinconar a su oponente entre los cuernos de la autorrefutación o de la vacuidad. La premisa mayor del dilema es: «O bien este ego al que se presentan los únicos objetos que existen es identificado independientemente de esos objetos o de lo contrario su identificación no es independiente de ellos» (pp. 228-9). La primera alternativa constituye el cuerno autorrefutador, pues «si la identificación del ego es independiente de los objetos que se le presentan, su afirmación será autorrefutatoria, dado que el único modo en que puede lograr una identificación independiente es a través de su cuerpo, que es un cuerpo entre otros en el mundo físico» (p. 229). Pero la segunda alternativa lleva a la vacuidad, «porque, cuando señala 'los únicos objetos que existen', él permanecerá identificado meramente como el sujeto al que se presentan esos objetos. Como dice después en *El cuaderno azul*, si el solipsista acepta este cuerno del dilema, será como alguien que construye un reloj con la esfera unida a la manecilla de modo que ambas den vueltas juntas» (*ibid.*).

Como argumento contra el solipsismo empírico éste es sin duda un buen argumento. Lo que ya es discutible es que represente la posición que Wittgenstein adoptó en el *Tractatus*. Más bien sucede que Pears le atribuye argumentos que utilizó en sus escritos posteriores a 1929¹⁹, contra el uso que el solipsista hace del pronombre «yo». Una vez más son las exigencias de su atribución al *Tractatus* de un realismo básico las que lo fuerzan a postular aquí una continuidad en el pensamiento de Wittgenstein que no es admitida en la mayoría de los restantes temas, incluso en temas estrechamente vinculados con éste, como la preocupación por el velo impenetrable de los datos sensoriales.

¹⁸ WTST, p. 174.

¹⁹ Así en *El cuaderno azul*, cuando el solipsista afirma «Cuando cualquier cosa es vista (realmente vista), soy siempre yo quien la ve», Wittgenstein le replica «¿Siempre quién?» (*The Blue and Brown Books*, Oxford, Blackwell, 1958, p. 61; trad. de F. Gracia Guillén, *Los cuadernos azul y marrón*, Madrid, Tecnos, 1968). Y en «Notes for Lectures on 'Private Experience' and 'Sense Data'», cuando el solipsista afirma «Yo tengo algo que nadie tiene», Wittgenstein le responde: «¿Pero quién es yo?» (*The Philosophical Review*, 77 (1968), p. 283; trad. de J. Lascrain y E. Villanueva, «Notas para las conferencias sobre 'Experiencia privada' y 'Datos sensibles'», en E. Villanueva (ed.), *El argumento del lenguaje privado*, México, UNAM, 1979).

Hay una inconstancia, por parte de Pears, acerca de cuál sea la consecuencia de adoptar la alternativa de que la identificación del sujeto no es independiente de los objetos de los que se percibe. Así, se nos dice en un lugar que la vaciedad del solipsismo consiste en que, cuando se lo lleva a sus últimas consecuencias, coincide con el realismo (p. 184). Pero anteriormente se nos había dicho que la consecuencia de dejar flotar libremente al ego es el panpsiquismo: «Si no hay ningún criterio empírico para su individuación, el ego se esparcirá sobre todas las cosas. Pero esto significa que se extenderá a las cosas inanimadas» (pp. 170-1). En otro lugar la alternativa es presentada así: el solipsista proclama que los únicos objetos que existen son aquellos de los que él se percibe; pero sólo puede identificarse a sí mismo como el sujeto que se percibe de ellos. Mas si ésta es la condición inevitable del solipsista, de lo que se trata es de circularidad y no de vacuidad. En cualquier caso, las consecuencias que se presentan son mutuamente incompatibles: *o bien* el solipsismo conduce al realismo y es vacío *o* conduce al panpsiquismo con lo cual su carácter vacío desaparece *o* es palmariamente circular.

Sea como fuere, el dilema, puesto en los términos de Pears, *no es exhaustivo*, porque en ambos cuernos se da por sentado que el yo es un objeto, que el yo pertenece al mundo -bien porque se lo identifica a través de mi cuerpo o bien porque se lo hace coincidir con «el sujeto pensante, representante» (5.631). Pero Wittgenstein está contemplando una *tercera* posibilidad: la que «yo» y «mi» en «Yo soy mi mundo», «El mundo es mi mundo», etc., se refieran, no al yo empírico, sino al yo trascendental, al sujeto metafísico. De éste se dice que «no pertenece al mundo, sino que es un límite del mundo» (5.632). Creo que Pears no toma suficientemente en serio esta noción. Pero es necesario hacerlo, pues sólo así podemos dar cuenta coherentemente de toda una serie de aspectos, a primera vista dispersos, del *Tractatus*: la inexpresabilidad del solipsismo, la doctrina de la proyección, los comentarios sobre la voluntad y la ética, etc. En particular la inexpresabilidad del solipsismo queda clara. Que el mundo es mi mundo no puede decirse porque el yo al que se refiere la palabra «mi» no es un objeto. Pero todo lo que puede decirse en el lenguaje puede decirse por medio de funciones de verdad de proposiciones elementales. Y éstas son combinaciones de nombres, que a su vez son representantes de objetos. Pero el yo no puede ser el referente de un nombre porque no es un objeto. Por tanto, la verdad del solipsismo no puede ser dicha.

8. La inconstancia de Pears acerca de las consecuencias de adoptar el segundo cuerno del dilema va acompañada de una fluctuación en la especificación del sujeto. Esto queda patente si comparamos dos especificaciones distintas del sujeto que se nos ofrecen, una en el libro y otra en el artículo «Wittgenstein's Treatment of Solipsism in the *Tractatus*». En el artículo, la consecuencia es presentada de esta manera:

Los objetos básicos son los objetos de «mi mundo» (5.6). Ahora bien, esta especificación suya sólo será auténticamente restrictiva si va acompañada de una identificación de mí mismo. Pero en el *Tractatus* el yo al que apela el solipsista es siempre «el yo filosófico» (5.641) o ego que «no pertenece al mundo» (5.632). Así ciertamente

la tesis de Wittgenstein debe ser que el solipsismo no es auténticamente restrictivo, porque usa un punto de referencia inútil²⁰.

Es cierto, como hemos señalado antes, que el ego al que apela el solipsista es el ego metafísico. Y la consecuencia que se extrae -que el solipsismo no es auténticamente restrictivo- es aceptable si se la toma *cum granu salis*. En efecto, en 5.64 Wittgenstein afirma que «el solipsismo, llevado a sus últimas consecuencias, coincide con el puro realismo. El yo del solipsismo se encoge en un punto inextenso y queda la realidad coordinada con él». Pears se apoya en esta afirmación para sustentar su lectura realista. Pero esto sería poner totalmente el énfasis en un aspecto de la postura de Wittgenstein y relegar el otro aspecto. Si hemos de enfatizar justamente ambos respectos, debemos decir que el solipsismo transcendental del *Tractatus* es a la vez un realismo empírico, a la manera en que Kant señaló que lo era su idealismo transcendental. Parte de lo implicado en esta aserción es lo que Wittgenstein desarrolló en su segundo período cuando advirtió que el desacuerdo entre solipsistas, idealistas y realistas no es un desacuerdo sobre ninguna cuestión práctica: «¿Me compadecerá más un realista que un idealista o un solipsista?»²¹. Es desacuerdo sobre la notación correcta por medio de la cual pueden representarse los hechos. El solipsista desea una nueva notación que subraya determinados rasgos y disimule otros. Su error consiste en interpretar una nueva forma de visión como la visión de un nuevo objeto, en interpretar un movimiento gramatical como un fenómeno cuasi-físico²². Ahora bien, mediante una nueva notación nada puede cambiarse; los hechos de geografía siguen siendo los mismos²³. «¡Pero el solipsista tampoco quiere obtener ninguna ventaja práctica al defender su concepción!»²⁴. Su error consiste en creer que hay alguna justificación en los hechos para la nueva notación.

En contraste con esto, en *The False Prison* Pears expone el cuerno vacuo del dilema diciendo que si el solipsista postula su ego

como el punto focal tras el campo de su conciencia, como lo que llamó después «el ojo geométrico» (...) este procedimiento a priori vaciará su tesis (...) El ego «geométrico» no es descubrible por introspección ni de ninguna otra manera y no puede servir como punto de referencia identificable. Wittgenstein hace buen uso de esta idea de Schopenhauer (p. 157).

²⁰ WTST, p. 178.

²¹ *The Blue Book*, p. 48; y cf. p. 59.

²² L. Wittgenstein, *Philosophische Untersuchungen/ Philosophical Investigations*, texto alemán y trad. inglesa de G.E.M. Anscombe, Oxford, Blackwell, 1953, §401 (Trad. de A. García Suárez y U. Moulines, *Investigaciones filosóficas*, Barcelona, Crítica, 1988).

²³ *The Blue Book*, p. 57.

²⁴ *Philosophische Untersuchungen*, § 401.

Aquí la especificación del ego es distinta. El ego que se postula no es el ego trascendental sino el sujeto que se percibe de los objetos existentes. Es el ego empírico, como demuestra el hecho de que esa entidad es demolida haciendo uso de la idea de Schopenhauer (y anteriormente utilizada por Hume y Kant) de que no es descubrible en el mundo. Si se tratara del yo filosófico, la apelación al argumento de la no-encontrabilidad sería irrelevante, pues ese argumento es ineficaz contra el yo filosófico, ya que éste no forma parte del mundo.

En conclusión, si el ego del que se trata en el dilema es el ego empírico, el dilema no es exhaustivo, pero si es el ego trascendental, entonces la conclusión debe ser que el solipsismo trascendental es un realismo empírico.

Alfonso G. Suárez
Septiembre, 1989



EL JUEGO DE SABER Y EL LENGUAJE DE LA CERTEZA

Angel Alvarez Gómez

«Man sieht eben nicht, wie sehr spezialisiert der
Gebrauch von 'Ich weis' ist».

(Wittgenstein: UG, 11).

La obra de Wittgenstein ha venido provocando reacciones diversas, pero a todos ha recordado algunas condiciones que acompañan necesariamente a una reflexión filosófica: la primera, la dificultad, si no la práctica imposibilidad, de mantener posiciones por mucho tiempo una vez pasada la época de los grandes sistemas, de ceder a la tentación de considerarlas definitivas; o, lo que es lo mismo, la necesidad de practicar la reflexión sobre la decisión tomada. La segunda, esa especie de desprendimiento o mirada lejana sobre los problemas que acosan al pensamiento de siempre; como si siendo lo inevitable, alimentara su relevancia desde fuera. Si unimos a ello la habilidad para mostrar el halo místico que rodea a los problemas más claros, o el fondo insondable del campo del lenguaje, podemos considerar que tenemos datos para explicarnos el efecto irresistible de sus escritos. Es todo ello más fundamental que la apariencia de heterodoxia frente a todo lo que pudiera mostrar reclamos de firmeza, aunque esto cuente más en la generalización del atractivo.

Me propongo hacer algunas reflexiones sobre las anotaciones que Wittgenstein traía entre manos hasta sus últimos días de vida *Sobre la certeza*¹.

Esas consideraciones dejadas en hojas sueltas y aparentemente desordenadas se refieren a un tema individualizado provocado por la lectura de los escritos de Moore: «Proof of external World» y «Defence of Common Sense»² donde se busca el cumplimiento de la pretensión de saber con seguridad que una serie de proposiciones son verdaderas.

Se trata en Moore y en Wittgenstein de proposiciones tan elementales como: «Aquí hay una mano y aquí otra», «La Tierra existía desde mucho antes de mi nacimiento», «Nunca

¹ *Über Gewissheit* (G.E.M. Anscombe y G.H. von Wright comp. Oxford, Basil-Blackell, 1969. Utilizo la edición bilingüe preparada por J.U. Prades y V. Raga. Barcelona, Gedisa, 1988). La obra ha aparecido ya en las: WITTGENSTEIN, L.: *Werkausgabe* in 8. Bänden. Frankfurt, Surkhamp, 1986 ss.

Me referiré a la obra en el texto con UG seguido del número que corresponda. Otro tanto haré con las *Philosophische Untersuchungen* (= PhU) de las que utilizó la edición bilingüe alemán-castellano de Basil-Blackwell, Londres 1958 y UNAM, México, 1986, editada por Crítica, Barcelona 1988 con trad. de A. García Suárez y U. Moulines.

² MOORE, G.E.: *Defensa del sentido común y otros ensayos*. Barcelona, 1983 (Trad., C. Solís con Introd. de J. Muguerza). Para la relación de Wittgenstein con Moore, cfr: DE MARTELAERE, P.: «Wittgenstein critique de Moore. De la certitude sans savoir». *Rev. Phil. de Louvain*, 1986 (84), 208-227; WHITE, A.: «Common sense. Moore and Wittgenstein». *Rev. Int. Philos.*, 1986 (40), 313-330.

me he alejado mucho de la superficie de la Tierra». La prueba de verdades tan «sencillas» genera una obra indeterminada de hecho, interminable propiamente y, por encima, ligada a la «defensa del sentido común». ¿Acaso no sabemos con seguridad cosas tan elementales? ¿No tiene la filosofía asuntos más complicados que tratar? O, si es la filosofía la que ha complicado las cosas, ¿Quién las ha de devolver a la naturalidad que les corresponde?.

A lo primero es preciso advertir que la investigación no pone en duda la convicción de tales verdades, sino que quiere saber el cómo y el por qué de la misma, «Cuáles son las características de ese estar convencido» (UG, 246). A lo segundo digamos al menos dos cosas: Primera, que lo generalmente aceptado como supuesto de toda posibilidad de problemas encierra en sí mismo la complejidad que viene a destaparse en la generación de esos problemas; de otro modo no se podría entender por qué los hay. Pero no se confunda complejidad con duda.

Segunda, que hay ciertamente en el lenguaje filosófico «proposiciones a las que volvemos una y otra vez como hechizados» (UG, 31). Las ganas de eliminarlas se traducen en la regla de prescindir al menos «de las proposiciones que no nos permiten avanzar» (UG, 33). Pero habrá que saber qué clase de avances es propio del análisis filosófico.

Sobre la certeza está provocado por Moore y recibe la herencia del mismo Wittgenstein, de Kant, de Hume, de Descartes. Es una vez más el empeño en la lucha contra el escéptico desde la aceptación del escéptico. Pero en la forma de recoger el guante nuevamente lanzado. Por eso la figura del escepticismo es y no es la misma y el desplazamiento de sus posiciones obliga a la invención de nuevos soportes técnicos y, en definitiva, a nuevas concepciones sobre el común problema del «saber»: Solo en el campo del lenguaje y en la teoría de los juegos tiene sentido considerar los márgenes permisibles de la duda. Del hecho de que del parecer no se pueda concluir el ser, tampoco se justifica que tenga sin más sentido la duda. La posibilidad de asegurarse o no «pertenece al juego del lenguaje. Es uno de sus rasgos esenciales». (UG, 2 Y 3).

Resulta, pues, obligado advertir que la clarificación de cualesquiera problema de que se ocupa el filósofo, como el de la pretensión de saber con seguridad, tiene su lugar propio en el lenguaje; es nuestro único recurso. En esos términos, la filosofía de todos los tiempos se ha comportado del mismo modo: Siempre ha sido análisis conceptual. Muy esquemáticamente cabría decir que lo que diferencia a unas filosofías de otras es la forma de entender el camino del análisis y, en consecuencia de ello, la diversidad de concepciones y teorías sobre el problema³.

La forma de Wittgenstein procede del círculo de los usos conocidos a las complicaciones buscadas. Dicho de otro modo, la progresión se enrosca en círculo y, si la obra no fuera interminable, mostraría la totalidad del lenguaje⁴. Pero el carácter de interminable pertenece en propiedad al problema; no es una cuestión meramente de hecho. Sería oportuno al respecto recordar el aviso de Neurath: «Somos como navegantes que tienen que reconstruir

³ Cfr.: WHITE, A.: «Análisis conceptual», en *La lechuza de Minerva* (Ch. J. Bontempo y S. J. Odell comp.), Madrid, Cátedra, 1979, pp. 113-126.

⁴ Cfr.: PEARS, D. Fr.: *The false prison. A Study of development of Wittgenstein's philosophy*, vol., I-II. Oxford, Clarendon Press, 1987.

su nave en pleno mar, sin jamás poder dismantelarla en un dique de carena y reconstruirla con los mejores materiales»⁵, con lo cual podemos observar que la sensación que al filósofo producen los problemas ha cambiado. De Bacon a Kant la confianza en la tarea de la ciencia asienta sobre aquél: «neque infinitum et ultra mortale fingant»⁶ y la osadía de Hume frente a la prudencia y seguridad de Kant participan de diverso modo de la misma convicción de que es posible examinar los recursos antes de aplicarlos⁷.

Ahora, por el contrario, no podemos reparar la nave antes de navegar, ni podemos mandar al piloto a la escuela antes de que se vea gobernando el barco; ni cabe pensar que alguna vez el barco esté definitivamente reconstruido antes de botarlo. El lugar del lenguaje y sus juegos coincide con la idea más apropiada de que el diseño está integrado en la realización y que uno y otro vienen condicionados por la obra ya realizada y por el aprendizaje que nos permite usarla. Así es como el análisis no puede por menos de enroscarse y tener carácter interminable y la filosofía ofrecer el aspecto de un laberinto en el que, más que a salir, hay que aprender a orientarse.

Esta es la situación: Queremos saber con seguridad que unas cuantas proposiciones simples relativas a nosotros, al mundo y a nuestra presencia en él, son verdaderas. El lugar para saberlo, no sabemos si el ideal, pero sí al que siempre hemos de recurrir, es el lenguaje. El lenguaje es en realidad un juego de lenguajes. Pero el concepto de juego es un concepto de «bordes borrosos» (PHU, I, 71). Las dificultades se hacen patentes en las condiciones del problema y se confirman al seguir el texto de Wittgenstein.

Un recurso útil podría ser hacer del problema del saber una serie de cortes en rompecabezas o una rotación de círculos siempre dentro del lenguaje; lo permite el juego del lenguaje y el texto de Wittgenstein.

Las líneas de los cortes parece lógico que nos respondan a lo previsible. Ello nos permite saber que nada ocurre porque no coincidamos con Wittgenstein, a condición de que descubramos al fin los fundamentos que nos permiten saber con seguridad⁸. Lo que importa es que «el lenguaje está en orden como está». Por tanto que el ideal coincide con la realidad, aun en las proposiciones más vagas, porque no habría sentido si no fuera el orden perfecto (PHU, I, 98).

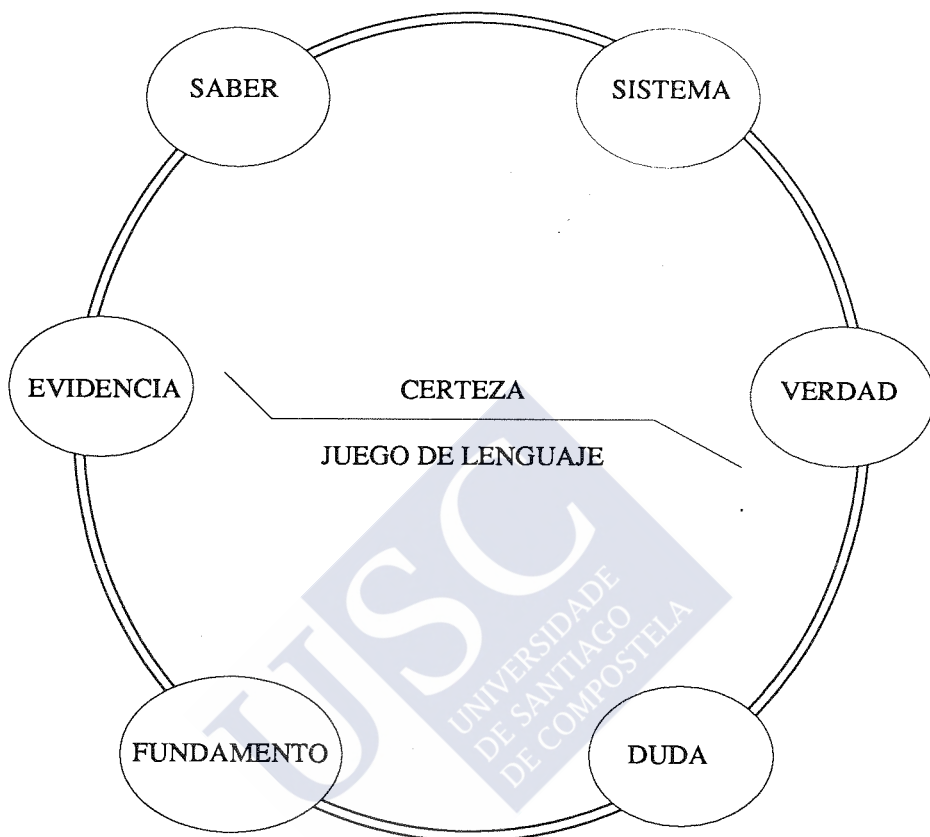
Así podría figurarse el problema:

⁵ NEURATH, O.: «Proposiciones protocolares», en AYER, A.J. (comp): *El positivismo lógico*. México, 1965, p. 206.

⁶ BACON, Fr.: *Instauratio magna*, praefatio. KANT, I.: *Kritik der reinen Vernunft*, B II.

⁷ HUME, D.: *Treatise of human nature*, I^a, IV^a, VII^a (pp. 263-264 ed. Selby-Bigge /p.415 trad F. Duque). KANT, I.: *Prolegomena...*, A, 17. Cfr: MANNISON, D.: «Hume and Wittgenstein: Criteria vs skepticism», *Hume Stud.*, 1987, 138-165.

⁸ Cfr: SCRIVANO, F.: «Sullo stile di Wittgenstein», *Riv. Est.*, 1987 (27), n° 25, 105-114.



1. APROXIMACION PROBLEMATICA A LA CERTEZA.

Fundamentar la pretensión de saber con seguridad que determinadas proposiciones son verdaderas exige:

1. Desarrollar constantes aclaraciones terminológicas.
2. Definir algunos ejes conceptuales.
3. Establecer la relación del saber con otras constelaciones.

Se adelantan dos advertencias: La primera, que la confusión puede empezar en no discernir ya lo que son meras aclaraciones de la identificación de los ejes conceptuales de cada constelación, dado que uno y otro ocurre en el juego del lenguaje. Dicho con más propiedad, que «los conceptos fluyen unos en otros» (PHU, I, 544)⁹. Pero no conseguirlo significaría que en el juego epistemológico no existen relevancias, lo cual daría al traste con la pretensión final.

⁹ Para orientación de la obligada referencia a las *Investigaciones*, cfr: BAKER, G.P. -HACKER, P.M.S.: *An analytical commentary on Wittgenstein's Philosophical Investigations*, Univ. of Chicago Press, 1985.

La segunda, que «saber con seguridad» se asocia con «certeza» pero «un estado de cosas no se sigue de la propia certeza. La certeza es, *por así decirlo*, un tono en el que se constata cómo son las cosas, pero del tono no se sigue que uno esté justificado» (UG, 30).

Preguntamos, pues, ¿cómo funciona el juego de «saber»? O, ¿qué papel corresponde a cada una de las piezas y cuáles son las reglas a las que nos atenemos?.

1.1. DEMARCAACION EN EL JUEGO DEL LENGUAJE

«Llamaré 'juego de lenguaje' al todo formado por el lenguaje y las acciones con las que está entretejido» (PHU, I, 7)¹⁰. Más que intentar conocer la esencia del juego -pretensión a la que no se podría satisfacer (PHU, I, 65)-, importa recordar que «hay *innumerables* géneros: innumerables géneros de empleo de lo que llamamos 'signos', 'palabras', 'oraciones'. Y esta multiplicidad no es algo fijo, dado de una vez por todas; sino que nuevos tipos de lenguaje, nuevos juegos del lenguaje, como podemos decir, nacen y otros envejecen y se olvidan». (PHU, I, 23)¹¹.

Que todos son juegos quiere decir que tienen ciertos parecidos, lo que no llega más allá de cierto «aire de familia», pues es así como se superponen y entrecruzan los diversos parecidos que se dan entre los miembros de una familia: estatura, facciones, color de los ojos, andares, temperamento, etc. etc.». (PHU, I, 66 Y 67).

Es difuso el concepto de juego -borroso-, pero no quiere decir que incontrolable en el uso. Podemos describir ejemplos de juegos, mostrar cómo se puede por analogía construir otros, discernir lo que sí y lo que no llamaríamos juego y cosas por el estilo (PHU, I, 75). Pero no podemos trazar una figura nítida porque ello equivaldría a eliminar el carácter borroso: «Si alguien traza un límite nítido yo podría no reconocerlo como el que siempre quise trazar también o el que he trazado mentalmente... Se puede, pues, decir: su concepto no es el mismo que el mío, sino uno emparentado con él» (PHU, I, 76 y 77).

Es decir: 1º. No podemos conocer la esencia del lenguaje; es más, pretenderlo sería algo así como dejarse hechizar por el super-orden de los super-conceptos, cuando la palabra 'lenguaje' debe tener un empleo tan bajo como las palabras 'mesa', 'lámpara', 'puerta'. (PHU, I, 97).

2º. Lenguaje y acciones están de tal manera entramados que no podemos saber lo que dice y hace uno sin mirar a lo que hace y cómo lo hace. La reflexión acerca de lo que nos permite estar seguros de ciertas verdades, no es separable de lo que hacemos cuando las afirmamos.

3º. No es posible mantener formas privilegiadas en los juegos, ni, por consiguiente, en el juego de lenguaje. El privilegio solo existe después de definir una preferencia especial

¹⁰ Acerca del alcance de los juegos de lenguaje, cfr: *Sprachspiel und Methode. Zum Stand der Wittgenstein Diskussion* (hrsg. von D. Birnbacher u.A. Burkhard). Berlín-New York, de Gruyter, 1985. BROSE, K.: «The limits and possibilities of the language-game. On the systematic structure of Wittgenstein's Philosophical Investigations» *Ratio*, 1985 (27) 121-132. CAVANCHICK, S.M.: «Juegos de lenguaje, significado y sentido» *Rev. Latinoam. de Filos.*, 1988 (14) 35-52.

¹¹ Acerca de la comprensión y esencia del significado y la transformación de la relevancia hacia el uso de las palabras, cfr: KAMBARTEL, Fr.: *Experiencia y estructura*, cap. V. (Buenos Aires, Sur, 1972, Garzón Valdés trad. pp. 196-213).

por alguna de las semejanzas: los entretenidos, los inteligentes, los de mayor habilidad, etc. No hay, pues, una función primaria; luego, el lenguaje expositivo, descriptivo, declarativo, no goza de ella.

4º. El carácter «borroso» del juego afecta de múltiple modo a nuestro tema del 'saber'. En primer lugar nosotros comparamos la estatura y el color de los ojos de los que son parientes, no de todos los hombres, y encontramos el parecido plasmándose sobre el parentesco; pero este es algo previamente sabido, o al menos conjeturado. En segundo lugar, el carácter «borroso» es el fundamento del acuerdo y del desacuerdo, de la nitidez compartida y no compartida. En sus extremos, del solipsismo y del entendimiento mutuo; además, en una figura siempre móvil, expuesta a continua transformación.

5º. A pesar de ello, «en los casos normales nos es claramente prescrito el uso de una palabra... Cuanto más anormal es el caso, más dudoso se vuelve lo que debemos decir entonces». Si se llegara al extremo de que regla y excepción tuvieran la misma frecuencia, entonces el juego habría perdido su quid (PHU, I, 142)¹². Esto quiere decir que: tenemos que prescindir de las proposiciones que no nos permiten avanzar; y que las explicaciones que justifican el uso de una palabra «deben tener un fin». Explicaciones y límite de las mismas pertenecen a la relación de regla y excepción (UG, 34).

6º. Que algunas proposiciones son verdaderas tiene que ver con la vinculación de verdad (y falsedad) y proposición. Pues bien, «llamamos una proposición a aquello a lo que aplicamos en *nuestro lenguaje* el cálculo de las funciones de verdad». Parece que eso significa que «lo que *engrana* en el concepto de verdad» (como en una rueda dentada), eso es una proposición». Pero lo que es una proposición está determinado en un sentido por las reglas de formación de las oraciones y, en otro sentido, por el uso del signo en el juego de lenguaje; además, el uso de 'verdadero' y 'falso' puede ser parte constitutiva del juego. Entonces *pertenecen* para nosotros a la proposición, pero no se *ajustan* a ella (PHU, I, 136). Ello obliga a buscar al tiempo que las razones que cumplan la pretensión de saber, la trama del juego al que pertenecen. Es esa trama la que establece lo que se puede y lo que no se puede plantear, decir, aseverar, dudar.

1.2. COMPLICACIONES CONCEPTUALES

La pretensión de seguridad está muy cerca del tono de certeza. Es esto lo que hay que precisar y, en el momento que se toca el asunto, saltan a la vista numerosas complicaciones.

1ª. Para estar seguros de la verdad de una proposición necesitamos probarla o derivarla de otras proposiciones. En el segundo caso, puede ocurrir que éstas no sean más seguras que aquellas (UG, 1). Una forma de lo primero puede consistir en la ostensión directa y sus consecuencias; el juego de lenguaje permite el paso de la inseguridad a la seguridad «mirando más de cerca». (UG, 3).

2ª. Se presentan con frecuencia proposiciones que, admitiéndose sin discusión por todos, son tales que requieren ser reducidas a otras más claras. Otras veces, son tan evidentes

¹² Cfr: BOUVERESSE, J.: *La force de la règle. Wittgenstein et l'invention de la nécessité*. París, Minuit, 1987; KREBS, V.: «Objectivity and Meaning. Wittgenstein on following rules» *Philos. Invest.*; 1986 (6) 177-186; WERHANE, P.H.: «The constitutive nature of rules» *South. J. Philos.*, 1987, (25) 239-254.

que nos sorprende la pregunta por su justificación. En: «Sé que soy un hombre», la intención es tan ambigua que necesito apoyarla en algún punto. Por ejemplo: «Sé que tengo un cerebro». En: «La Tierra existía mucho antes de que yo naciera», no interesa que se sepa, sino cómo llega a saberse. Ambos son ejemplos de verdades elementales. Pero la reducción de la primera nos coloca ante la perplejidad: No dudo y, sin embargo, nunca he visto que tenga un cerebro; no tengo de ello experiencias y no me está prohibido imaginar que mi cráneo está vacío (UG, 4). La segunda es una verdad que todos sabemos, no porque hayamos practicado determinadas operaciones de pensamiento o contemos con archivos de la historia pasada (UG, 84 y 85). Tocar los límites de la experiencia como fuente del conocimiento requiere apuntar al fondo que sustenta toda experiencia y recorrer las vías del pensamiento revela que éstas se extienden sobre algún fundamento previo. Desembocamos así en que «la verdad de algunas proposiciones empíricas depende de nuestro sistema de referencia» (UG, 83) y en que «la prueba suficiente de un enunciado... pertenece a la descripción del juego del lenguaje» (UG, 82. Cfr: PHU, I, 669). Lógica, experiencia, fundamento y sistema quedan así requeridos.

3ª. Hay complicaciones en la indistinción entre 'saber' y 'sé...'. Moore asimila 'saber' a 'creer', 'conjeturar', 'dudar', etc., mientras que 'sé...' no puede constituir error. De ese modo, «sería posible inferir de una declaración la verdad de una aserción» y entonces no cabrían otras disposiciones ni, lo que es más grave, la misma posibilidad de error (UG, 21). Detrás de ese uso incorrecto, pues, se manifiesta «un estado mental peculiar y muy importante» (UG, 6). Quizá la consecuencia más importante tenga que ver con la pretensión de que lo enunciado en primera persona pueda ser afirmado también en primera persona del plural; en el paso de 'se...' a 'sabemos'.

4ª. La seguridad extiende sus raíces desde la vida misma, sin que se haya de cumplir la formalidad expresa de asegurarse (UG, 7 y 9). 'Saber' y 'estar seguro' no se diferencian «excepto cuando 'sé...' quiere decir: No puedo equivocarme» (UG, 8). La excepción apunta a la diferencia entre 'verdad' y 'seguridad' que remonta hasta la distinción griega entre 'alétheia' y 'aspháleia'. Es difícil desenredarse de ese juego y la renuncia a la seguridad en aras de la verdad viene persistentemente a caer en nuevas formas de seguridad. A Nietzsche no le costaría demasiado esfuerzo desenmascararlo¹³.

Pero este asunto presenta dos aspectos diferentes: uno, la tendencia universal (natural?) a la seguridad; otro, que no sea necesario asegurarse de aquello de lo que se está seguro. Ambos aspectos solicitan la perspicacia del filósofo para saber por qué se puede vivir seguro de lo que no se sabe cómo es seguro.

Una solución está en separar la ciencia de la vida, la investigación de la utilidad. Es una solución que se parece bastante al tratamiento que Descartes da a la experiencia sensible¹⁴,

¹³ NIETZSCHE, Fr.: *Die Fröhliche Wissenschaft*, Vº, 355 (*Werke*, hrg. von G. Colli u. M. Montinari. Berlín-New York, De Gruyter, 1973, V/2, p. 275-277).

¹⁴ Cfr: DESCARTES, R.: *Meditaciones*, VIª. Una síntesis de la relación de Wittgenstein con Descartes puede verse en: LUTTERFELS, W.: «Wittgensteins Rehabilitierung des Cartesianischen Paradigmas», en *Philosophie des Geistes, Philosophie der Psychologie*. (Arkten des 9. Int. Wittgenstein's Symposium, 19-26 August, 1984, hrg. von R. Chisholm, J. Mack, J. Blackmore and A. Hubner). Wien-Hölder-Tempsky, 1985, p. 417-419. (Las Actas contienen comunicaciones de sumo interés acerca de relevantes aspectos de la filosofía del autor).

pero responde a una cuestión que el pensamiento griego recoge de más ancestrales tradiciones míticas y religiosas.

En la medida que se es consciente de la provisionalidad de esa solución, es preciso remontar las condiciones que la imponen, lo que significa inmediatamente que el núcleo de la cuestión no es experiencia sí o no.

Este remonte es posible si: 1º. Se tienen presentes determinadas distinciones; 2º. Se responde a determinadas exigencias; 3º. Se descubre el asiento último irreductible de toda seguridad¹⁵.

1º. Sabemos de la ilegitimidad de inferir el 'ser así' de la declaración: 'Sé que es así' de otra persona; de la legitimidad de 'ser así' de mi declaración: 'Sé que es así'; como que 'allí hay una mano' de la proposición: 'Sabe que etc.' por último, que de la declaración: 'Sé...' de otro, no se sigue que lo sepa.

Un primer grupo de distinciones se refiere a las relaciones de inferencia entre 'saber que x es así', dicho por mí o por otra persona y 'ser así'; incluso a la diferencia entre declarar 'sé...' y saberlo (UG, 13). El segundo apunta a la garantía, como un hecho, de lo que se describe como estado de cosas (UG, 12), es decir, a la advertencia de que 'sé...' puede significar «creo saber».

2º. Las exigencias tienen que ver con la necesidad de romper el círculo de la subjetividad; esa es la función de la demostración. Pero, ¿cómo conseguirlo si es siempre el sujeto quien tiene que romper el círculo de sí mismo?. La demostración consiste en establecer la correspondencia entre la objetividad y la seguridad, eliminando cualquier forma de ésta que no se sienta como infalible (UG, 14, 16).

Bien pudiera decirse que lo que caracteriza a la Teoría del Conocimiento moderna es haber perdido el escrúpulo de la pureza de la verdad incontaminada de las condiciones subjetivas (de certeza, seguridad, etc). El concepto de objetividad se desplaza del lugar de la mirada límpida al «entre» de la relación sujeto-objeto que constituye todo conocimiento. Es así como la objetividad ya no es separable de las condiciones subjetivas; un problema que obtiene su primera orientación en la exigencia de 'claridad y distinción' para nuestras ideas, aspecto nuevo del problema del criterio no elegido arbitrariamente, sino impuesto con la misma fuerza con que se nos impone la verdad de la propia existencia¹⁶. Otra versión de lo mismo la encontramos en Kant, lo que determina que en adelante el problema de la claridad (= garantía de objetividad) haya de plantearse mejor en términos de «cómo dar claridad a nuestras ideas»¹⁷.

3º. La infalibilidad «ha de poder ser establecida de un modo *objetivo*» (UG 16). En este punto tenemos que tocar el fondo o hundirnos para siempre. La situación que se plantea es similar a la del Descartes de la segunda meditación¹⁸. La salida en este caso consiste en saltar por encima del curso ordinario del pensamiento; de ahí que todas las inferencias hayan de

¹⁵ Cfr. CARAWAY, C.: «Criteria and conceptual change in Wittgenstein's later philosophy», *Metaphilosophy*, 1986 (17) 162-171.

¹⁶ DESCARTES, R.: *Meditationes...*, III, AT. VII, 35, 7-11; *Discours de la Méthode*, IVª, AT. VI, 33, 16-24.

¹⁷ PEIRCE, Ch. S.: *How to make our Ideas Clear*, Collected Papers (comp. by Ch. Hartshorne and P. Weiss. Cambridge Mass. The Belknap Press of Harvard Univ. 1934, Vol. V, p. 248-271).

¹⁸ AT. VII, 23, 24-24,9.

resultar, al menos inicialmente, chocantes. (¿No lo es la «res cogitans» frente a la «res extensa»?). Lo inevitable en ese modo de proceder es que el estado de seguridad del sujeto determine el estado de cosas, que la verdad asiente en la certeza. La cuestión en adelante será saber si se está de acuerdo con el carácter de ficción que es propio de la ciencia así construida. (Ni que decir tiene que a este respecto, el paso del 'yo' al Dios garante de objetividad nada cambia de lo esencial).

Wittgenstein entiende que no es posible saltar el círculo de la subjetividad y mucho menos replegándose hacia dentro. O se está ya en la objetividad o no hay solución. Que se está en la objetividad se sabe de dos maneras:

Una, mostrando que todo pensamiento se mueve en el círculo de las cosas, sean éstas reales o ilusorias. La segunda, recordando que el pensamiento tiene su asiento en el lenguaje y éste es siempre público¹⁹. Cuando se entiende el juego de lenguaje, se conocen las razones que avalan las afirmaciones y «se ha de poder imaginar *cómo* puede saberse una cosa semejante» (UG, 18). Conseguimos hacer pie y recobrar impulso para salir a superficie en la comprensión de un juego que delimita lo que hay que decir y lo que apenas necesita ser mostrado, no por el procedimiento de la ostensión, ni por el más sofisticado de la observación científica acompañada de prueba teórica. «La duda solo tiene lugar en un juego de lenguaje» (UG, 24); esto es, la seguridad de las proposiciones de la vida ordinaria se justifica en que el juego ha establecido las circunstancias en las que la duda, tanto como el error, son posibles (UG, 25, 32).

La proximidad de la seguridad al tono de la certeza nos hace saltar por encima de la fascinación de la experiencia y de la conciencia -las dos formas tradicionales de resolver la cuestión- para situarnos en el fundamento último del juego de lenguaje en el que se producen las aseveraciones. La solución toca al problema que preocupa centralmente al Wittgenstein de la primera y de la última época: la perplejidad de la filosofía y el sinsentido de la metafísica. Experiencia y conciencia son ambas respuestas desde la confusión de la gramática superficial y profunda, de uso empírico y aplicación lógica (Cfr: PhU, I, 664, 221, 422, 520; Zettel, 458).

2. LA FUNCION GRAMATICAL

El afán de saber con seguridad encierra la bipolaridad del problema de la verdad: objetividad y validez son inseparables; pero la validez comporta exigencias de justificación y «del tono no se sigue que uno esté justificado» (UG, 30). Se puede entender que tales exigencias quedan cumplidas en el darse de los objetos a la intuición, siempre que rompa con la ingenuidad y se sitúe en la correlación de naturalezas y nociones simples. Se puede perfilar más el asunto si alcanzamos a descubrir las condiciones de posibilidad de toda experiencia; o, lo que viene a importar lo mismo, poniendo en juego lo que es propio de toda subjetividad. Por si surgen desconfianzas acerca de uno y otro, se pueden tantear esfuerzos sin otra garantía que la resultante del análisis intersubjetivo en las mejores condiciones de rigor hasta el momento conocidas.

¹⁹ PhU, I, 275. Cfr: GARCIA SUAREZ, A.: «La lógica de la experiencia, Madrid, Tecnos, 1976; SAUVE, D.: «Le problème du 'langage privé' et la conception Wittgensteinienne du langage», *Dialogue*, 1988 (27), 417-449.

Pero los problemas filosóficos tienen esa naturaleza endiablada que al mismo Wittgenstein le hacía sentirse en el infierno²⁰. Podemos saber lo que es la claridad y sufrir alucinaciones, conocer las reglas y no saber aplicarlas, o saber cómo aplicarlas en condiciones normales y no poder describir las circunstancias; con lo que venimos a quedar en que «la práctica de usar una regla muestra también qué es un error en su utilización» (UG, 29; Cfr: UG, 26-28). La seguridad, que parecía venir avanzando hacia determinaciones más rigurosas, retorna al lugar de la garantía práctica en cada caso.

Estamos tentados a entender que la elaboración de las condiciones de seguridad ha fracasado y que quedamos entregados a los avatares de la inseguridad. Pero Wittgenstein entiende otra cosa: que «cada proposición porta el método de su propia verificación». Del método se sentía únicamente Wittgenstein seguro en medio de todas las complejidades; es justamente su naturaleza la que ha de ponerse de manifiesto²¹.

En los dominios del lenguaje sabemos que estamos en el ámbito de todas las posibilidades y también de todas las travesuras. A él pertenecen todo tipo de proposiciones, la tentación persistente a confundir las de un tipo con las del otro, y, en consecuencia los «intentos frustrados» (UG, 36) de decir, demostrar, explicar o dar instrucciones sobre lo que no compete. Hay proposiciones que enuncian:

1) Posibilidades/imposibilidades físicas, tales como: «solo una persona puede dormir en una cama de 20 centímetros de ancho».

2) Posibilidades/imposibilidades gramaticales, tales como: «la varilla tiene longitud» o «los solitarios los juega uno solo».

3) Posibilidades con aspecto experiencial que expresan realmente barreras gramaticales, como: «nadie puede sentir mi dolor», o «las sensaciones son privadas». (PhU, I, 248)²².

La capacidad de representación pictórica de las proposiciones de tipo 3) son las responsables del engaño metafísico que «destruye la distinción entre la investigación factual y la conceptual». (Zettel, 458). Por eso es necesario empezar por cambiar la formulación de las expresiones de modo tal que sugieran, no las posibilidades/imposibilidades experienciales, sino las gramaticales; tales como: «No se puede...» por: «no hay nada de ese tipo en este juego»; «no puedo exhibir mi sensación», por: «en el uso de la palabra 'sensación' no se da el poder exhibir todo lo que se siente»; etc. (Z. 134).

Son confusiones que surgen «Cuando el lenguaje se va de vacaciones» (PhU, I, 38), «Cuando el lenguaje marcha en el vacío, no cuando trabaja» (PhU, I, 132). De esta índole son los problemas metafísicos, tales como «si existen objetos físicos». (UG, 35-36).

La exigencia, pues, de justificación nos pone ante la más dura problemática de la filosofía y nos sume en la perplejidad del lenguaje que tiene por sí mismo un comportamiento normal y otro anormal. ¿Cómo, podemos preguntarnos, evitar éste sin destruir aquél?. ¿Son los problemas metafísicos cuestiones para individuos ociosos o virus que se repro-

²⁰ FANN, K.T. (ed): *Wittgenstein: The Man and his Philosophy. An Antology*. Delta Book, New York, 1967, p. 52.

²¹ HILMY, S. St.: *The later Wittgenstein. The emergence of a new Philosophical Method*. Oxford, Basil Blackwell, 1987.

²² Cfr: FANN, K.T.: *El concepto de Filosofía en Wittgenstein*, Madrid, Tecnos, 1975, p.113.

ducen y mutan perteneciendo a la entraña de la vida misma?. ¿Está el lenguaje enfermo de metafísica?. Sí es así, ¿Cómo se puede curar?. A esas cuestiones podemos responder aclarando:

1º. Nos encontramos una vez más ante aquél «destino singular» («besondere Schicksal») ²³ que impulsaba a la investigación Kantiana trasladado ahora al lugar natural del lenguaje.

2º. No podemos separar la conducta sana de la enferma, aunque consiguiéramos curar todas las enfermedades, por la simple razón de que no podemos separar en el conocimiento la intervención del sujeto lingüístico.

3º. A la filosofía corresponde elucidar la gramática, los *preparativos* para poder usar el lenguaje, pero no le corresponde la aplicación; del mismo modo que forma parte de los preparativos del ajedrez aprender las posiciones de las piezas y las reglas de sus movimientos, pero ellas no constituyen aún jugadas ²⁴. Pero solo puede cumplir esa función alejando la tentación de decir las jugadas que corresponden a la ciencia natural o a las declaraciones de experiencia. El orden de relevancia se ha trasladado; parece destruir todo lo importante y renunciar a grandes aspiraciones, «pero son sólo castillos en el aire los que destruimos y dejamos libre la base del lenguaje sobre la que se asientan». (PhU, 118).

4º. Tocamos los límites de las cuestiones en cuanto abordamos las cuestiones; así le ocurre a la filosofía. El peligro reside en que la preocupación por desarrollar las posibilidades haga olvidarse de las fronteras. El problema de los límites acompaña a la filosofía desde su infancia, especialmente desde que se afronta con la conciencia de que no contamos con otra especie de ayudas que nos socorran desde fuera. De ahí que «nada más útil pueda investigarse que qué sea el conocimiento humano y hasta donde se extienda». La misma idea convertida en objeto central en Kant y a través de él en el Wittgenstein de siempre; basta recordar la advertencia que prologa el *Tractatus*: «Este libro quiere, pues, trazar uno límites al pensamiento, o, mejor, no al pensamiento, sino a la expresión de los pensamientos» ²⁵.

Tocar el límite en filosofía obliga a reflexionar no solo acerca de los distintos tipos (límites de hecho/derecho, de razón/experiencia, de la razón pura/histórica, etc.), sino del para qué del señalar el límite mismo: Hay límites que se trazan para que nadie salga o entre; otros para alentar el riesgo de saltarlos; otros para avisar de dónde termina una propiedad y dónde empieza otra (PhU, I, 499). Esa advertencia es necesaria cuando se dismantelan las confusiones anteriormente mencionadas de la metafísica y queda el olor azufrado de la tormenta recién pasada. A la filosofía compete elucidar los significados para saber en todo caso a qué atenerse acerca del sentido de las proposiciones. Algo así como la «lógica de nuestro lenguaje» (TrLPh, Vorwort); o la esencia, no como «algo que esté ya abiertamente

²³ KANT, I.: KrV, A VII.

²⁴ De nuevo resultaría fácil el paralelismo con la función de la filosofía transcendental Kantiana: KANT, I.: *Welche sind die wirkliche Fortschritte*, AK. 7, p. 260. (Trad. F. Duque, Madrid, Tecnos, 1987, p.9). Una sucinta exposición de las relaciones de la filosofía crítica en YOUNG, J.: «Wittgenstein, Kant, Schopenhauer and critical Philosophy», *Theoria*, 1984 (50), 73-105.

²⁵ DESCARTES, R.: *Regulae ad directionem ingenii*, VIII^a, AT., X, 397-398 y passim. KANT, I.: KrV, A XI-XII, A 793/B 767 y ss. El *Tractatus* y las PhU tratan de lo mismo; pero no conviene pasar por alto las diferencias internas que perfilan las -así llamadas- filosofías de Wittgenstein: La primera, que en el *Tractatus* se busca el límite del lenguaje, mientras que para las PhU existen los límites; la segunda, que en el Tr. el límite se descubre, mientras que en las PhU se ortoga. Cfr: PhU, I, 119 y Zettel, 467.

de manifiesto y que se vuelva visible sinópticamente mediante una ordenación, sino algo que yace *bajo* la superficie». (PhU, I, 92).

5º. Es enfermiza la propensión a confundir los usos empírico y lógico de las proposiciones; es enfermiza la metafísica porque, teniendo por objeto la elucidación de los conceptos, se deja caer en la tentación de los usos y aspira a convertirse en algo como ciencia natural. La curación no puede consistir en dar carpetazo al análisis gramatical, sino en hacerlo en profundidad, como profundas e importantes son las inquietudes que provocan los constantes malentendidos. (PhU, I, 111).

Así es como se llega a abordar el problema del realismo, o el escepticismo del idealista ante cuestiones tan aparentemente peregrinas como: «Si existen objetos físicos». El primero afirma que el objeto tiene en sí mismo un ser real; para el segundo el conocimiento constituye al objeto de modo que éste no tiene otro ser que el ser en el conocimiento. Mientras la cuestión y la polémica se centre en la respuesta a la dificultad, el caso no puede ser resuelto. Es preciso advertir que la misma expresión del problema -y no menos la solución- «todavía podría ser una expresión completamente incorrecta». (UG, 37).

Al tratamiento clásico del problema le faltan dos ingredientes: el primero, la distinción entre conocimiento lógico y conocimiento empírico; el segundo, la delimitación de la competencia conceptual en relación al juego de que se trate. (Cfr: UG, 58 - 65).

3. RESTOS DE UN MODELO MATEMATICO DE PENSAR

El problema de la distinción entre investigación factual y conceptual en el dominio del lenguaje tiene la posibilidad de aclararse recurriendo al modo matemático. Observamos así una constante que iniciada en la temprana época griega, reaparece en cuanto que las inquietudes filosóficas reciben nuevo asiento o alcanzan nueva formulación. Las «largas cadenas de razones» de que se sirven los geómetras pueden generalizarse -sin confundirlas con la tarea de las ciencias matemáticas particulares- con solo tomar en consideración las diversas relaciones que guardan entre sí los objetos y hacerlo procediendo con orden desde las cosas más simples a las más complejas²⁶. Ese era el modelo cartesiano al que la ciencia moderna rinde tributo. Para hacerlo posible hubo de acabar primero con el hechizo de la naturaleza omitiendo la heterogeneidad de las esencias de los seres corporales²⁷.

Podría decirse que a esa posibilidad responde aún, a través de Kant, el *Tractatus*. Pero enfrentado desde las *Investigaciones* a la necesidad de aceptar el lenguaje en toda su multiplicidad y complejidad, a la exigencia de trazar los límites en lugar de descubrir el límite, a «establecer un orden en nuestro conocimiento del uso del lenguaje: un orden para una finalidad determinada; uno de los muchos órdenes posibles, no el orden» (PhU, I, 132); urgido, en fin, por la necesidad de distinguir la diversidad de las palancas de su uso, de pasar de la consideración de las formas a la de las funciones de las palabras (PhU, I, 11 y 12), la apelación al modo de pensar/hacer matemático adquiere otro signo²⁸.

²⁶ DESCARTES: DM, IIª, AT, VI, 19, 6-20, 24.

²⁷ DESCARTES: *Regulae*..., XIVª.

²⁸ Una consideración más detenida exigiría un comentario a las *Bemerkungen über die Grundlagen der Mathematik-Remarks on the Foundations of Mathematics* (Oxford, Basil Blackwell, 1978, 3ª. Cfr: KAMBARTEL: l.c. nota (11).

1º. La misma historia interna de la matemática impulsa a la transformación de una mentalidad que cede a la tentación de anclarse en una concepción fija, rigurosa y única del lenguaje: «Hay *innumerables* géneros diferentes de empleo de todo lo que llamamos signos, palabras, oraciones... (Una *figura aproximada* de ello pueden dárnosla los cambios en la matemática)». (PhU, I, 23).

2º. La filosofía, si en algún modo ha de seguir el hacer matemático, no será nunca resolviendo sus problemas. Las cuestiones matemáticas las tiene que resolver la matemática; a la filosofía corresponde «hacer visible sinópticamente el estado de la matemática que nos inquieta, el estado *anterior* a la solución de la contradicción», porque las cosas no marchan como habíamos previsto (PhU, I, 124-125). Eso y la necesidad de *tratar* lo que el matemático está inclinado a decir sobre la objetividad y realidad de los hechos matemáticos, es lo que a la filosofía le queda de su vieja función fundamentadora (PhU, I, 254). Compárese con el espíritu que anima la afirmación cartesiana de que: «estas (demostraciones metafísicas) son más ciertas que aquellas (matemáticas)»²⁹.

3º. La matemática pertenece a los dominios de lo imaginable y de lo no imaginable. Como tal, rompe las barreras de la representación para contener la realidad; pero rompe también las barreras del ser para expresar lo que debe ser; y rompe la concepción de lo pictórico para saltar sobre los límites. Lo mismo que las pruebas matemáticas nos llevan a «revisar lo que considerábamos el dominio de lo imaginable» (PhU, I, 517, 518 y 519), la filosofía puede llevarnos a descubrir que nos equivocamos cuando creemos entender una pregunta, sobre todo cuando no nos acordamos cómo ni en qué ocasiones se emplean las oraciones. Justamente ese es el riesgo que con mayor frecuencia corre y en el que persistentemente cae por tener que cumplir con la función de conjurarlo (PhU, I, 513-515). El sinsentido acecha en las lindes de todo hacer lingüístico con sentido, como acecha al borde de los problemas matemáticos y de los quehaceres manuales.

4º. La matemática no deja lugar a discusión acerca del entramado de reglas que rigen el juego. No se llega a las manos por las operaciones, sino por el modo de hacerlas. La filosofía ha estimulado siempre el ansia de conocer la verdad, al parecer arraigado en la propia naturaleza humana. El problema está en saber cómo se sastiface esa necesidad y cómo se justifica. A la constelación de la concordancia ('homóiosis', 'adaequatio', 'Ubereinstimmung') ha pertenecido durante siglos la respuesta a esas cuestiones. Concordancia del entendimiento y la cosa, o de la filosofía con la realidad y la experiencia; pero, en definitiva, de todos los sujetos dotados de entendimiento o capaces de reflexión filosófica. No es difícil terminar afirmando que «la concordancia de los hombres decide lo que es verdadero y lo que es falso». (PhU, I, 241).

Pero no es lo mismo concordancia a la que se llega que concordancia en la que se está o a la que se pertenece. No es lo mismo concordancia del pensamiento con las cosas, o de las cosas con su 'eidos', que concordancia de los hombres en el lenguaje. Solo esta última versión corresponde propiamente, después de todos los corrimientos habidos en la Teoría del Conocimiento³⁰.

²⁹ Descartes et Burman, AT, V., 177 en comentario a VII, 35, 30-36, 8.

³⁰ Cfr: SKARICA, M.: «Enunciación y verdad según Wittgenstein», *Philosophica* (Valp). 1986-1987 (9-10) 213-241.

Una mirada superficial y precipitada se interesa por la concordancia de las opiniones y el asombro que produce que ésta se dé, suscita la fantasía de una especie de rayo noético que atraviesa desde el pensamiento a las cosas por medio del lenguaje. Pero no es el lenguaje un medio, sino la base sobre la que se asientan los pensamientos acerca de las cosas. Por eso la concordancia está como determinada por una «forma de vida» y alcanza tanto a las definiciones como -por extraño que parezca- a los juicios.

¿Qué significa esto?. Que tenemos que, en primer lugar, distinguir entre describir los métodos de medida, medir y formular los resultados de hacer mediciones; en segundo lugar, que la garantía del «medir» se encuentra en la constancia de los resultados de las mediciones (PhU, I, 242). Se llama así la atención tanto sobre la diferencia de requisitos formales y aplicaciones, como sobre la circularidad entre los requisitos y los resultados de la actividad.

Parecería que estamos condenados a entendernos, o, por lo menos, que todos mentimos de la misma manera, lo que nos daría el método para detectar en todo caso y destruir la falsedad. Sin embargo, la historia del problema no resulta tan prometedora y Wittgenstein lo sabe. Lo mismo que se ha trasladado el centro de atención de la concordancia, se traslada el centro de la dificultad. Siendo el lenguaje la base, en él reside el problema.

La capacidad aleccionadora de las matemáticas tiene, como se ve, múltiples entradas de las que en *Sobre la certeza* interesan de manera especial: De una parte, la diferencia entre lo que es objeto de cálculo y lo que lo es de experiencia; de otra, el modo de expresar el resultado del cálculo y la experiencia del sujeto calculador. En ambos casos, a la espera de esclarecer cuál es el límite de la duda, bien sea acerca del contenido o del testimonio y siempre en un orden de relevancia que marca el interés paradigmático: «Lo que es interesante es cómo *utilizamos* las proposiciones matemáticas» (UG, 38), no lo que nos pasa cuando calculamos.

Otro tanto hay que decir del resto de las proposiciones: «Un significado de una palabra es una forma de utilizarla» (UG, 61). Lo que quiere decir que cuando se sabe que el centro del problema reside en el lenguaje, no se puede seguir mirando por detrás, como ha estado tentada a hacer, por mala conciencia, la Teoría del Conocimiento. Sobran los procesos psicológicos, las seguridades transcendentales y, en lo sometido a la regla, las mismas constataciones; como sobra, del lado de los contenidos, la pretensión de *esencias* y, del lado de las reglas, la regla de las reglas: «Así se calcula; calcular es *eso*» (lo que se hace cuando se calcula). (UG, 47, 39, 42-46).

Del mismo modo que el cálculo es una forma de calcular que hemos aprendido en la escuela, el significado «es lo que aprendemos cuando una palabra se incorpora a nuestro lenguaje por primera vez» (UG, 61). Por eso se corresponden 'significado' y 'regla' y los significados mismos no son otra cosa que las 'funciones' (UG, 64); las funciones están ligadas a los 'juegos'. Así, pues, cuando éstos cambian, también lo hacen los conceptos y significados (UG, 65).

Como no hay regla de la regla, tampoco hay posibilidad de ir por detrás de su aplicación. Ese es el punto donde se detiene la justificación; de nada sirve -y solo contribuye a confundir- que afirmemos que *sabemos* que « $a + b$ es igual que $b + a$ » (UG, 113). En el fondo nos queda la sospecha de que no nos tenemos por qué fiar. En sí misma es una sospecha razonable. Tanto como la que en un tono dramático relataba Descartes cuando se empeñaba

en sumergirse en el profundo remolino de la duda; o la que con tono más bromista servía a Locke para socavar la sustancia; o la que tentaba a Hume después de la desastrosa travesía que le condujo al naufragio escéptico³¹. Se trata de la sospecha que nace de la falta del cero epistemológico, ante la que parecen no quedar más alternativas que la desesperación escéptica o la irresponsabilidad de establecer algún fundamento. Cuando la filosofía hace eso, destapa la caja de los truenos, pues ya no es posible justificar en modo alguno la confianza de unos sujetos para con otros, ni en sus testimonios, ni en sus acciones. La historia del problema es suficientemente aleccionadora: Cuando no se ha caído en el enredo de los estados de ánimo, se ha apelado a instancias externas, como si pudiéramos salirnos de donde ya siempre estamos.

«Pero recuerda: Incluso mi consideración de que el calculo está firmemente establecido no es más que una decisión con una finalidad práctica» (UG, 49).

Tres notas nos advierten para en adelante: La primera que no hay algo así como la roca firme, «fundamentum inconcussum veritatis», sino la barca a merced de las olas, pero que flota y navega; puede sonar a ridículamente poético, pero la filosofía sobrenada abismos.

La segunda, que la confianza en que metidos en ese artefacto pueda hacerse la travesía se alimenta de la decisión tomada de hacerla. (No es tan fácil como podría parecer abandonar la tentación humeana para optar por las seguridades Kantianas. Acaso hayamos aprendido a ver mejor el barco agujereado y las deficiencias del capitán avisado)³². Habrán de pagarse inevitablemente por ello las consecuencias: No se podrá ir a más velocidad que la que el barco pueda desarrollar, ni transportar mayor variedad de pasajeros, ni hacer frente a otros avatares que los que se puedan aguantar; ni tomarse vacaciones. Esto quiere decir que cuando algo se ha establecido como seguro no tiene sentido preguntar si es seguro; los riesgos ya acompañan a las decisiones.

La tercera, que la decisión ha sido tomada guiados por el afán y la prueba de conseguir fines prácticos. Podría decirse que, no siendo ciega, se ha desarrollado por tanteos y queda expuesta a las necesidades de desarrollo y transformación.

Todo ello es lo que ha venido a quedar cuando se ha desinflado el globo de la relación entre intelecto y voluntad que tan duramente atormentó a la reflexión filosófica en el pasado y en torno a cuyo eje giró especialmente el pensamiento moderno. Cuando el globo se desinfla se viene a tierra; a la tierra de las necesidades reales como gozne, ante las cuales se desvanece el prejuicio de la pureza cristalina de la lógica (PhU, I, 108). Nos interesamos por las oraciones y proposiciones en el mismo sentido que tienen en la vida ordinaria. Esto es «algo que me suena a pragmatismo. Aquí caigo en el enredo de una especie de concepción del mundo (Weltanschauung)» (UG, 422).

El significado como función, la función dentro del juego que se juega en vistas a la satisfacción de necesidades reales. Como en todo pragmatismo, la verdad juega a la par según la concepción de la coherencia y de la correspondencia: aquella, en el sistema de nuestras creencias y ésta, en la adecuación a nuestras expectativas según necesidades. Así se puede presentar como una concepción mediadora a la par que simplificadora.

³¹ DESCARTES, R.: *Meditaciones*, II^a, AT., VII, 23, 22-24, 13; LOCKE, J.: *An Essay Concerning human Understanding*, II^a, XXIII, 2; HUME, D.: *Treatise of human Nature*, I, IV, VII (Selby-Bigge, 263-264).

³² Cfr. HUME, D.: l.c.; KANT, I.: «*Prolegomena*, A 17-18.

4. FORMAS DE SEGURIDAD

Desde esos supuestos ya no es difícil aclarar en qué relación están entre sí las diferentes formas de seguridad; el mismo juego establece qué cálculos son absolutamente dignos de confianza y qué otros no están todavía completamente establecidos (UG, 48).

La primera forma recuerda la necesidad de corregir las expresiones como 'puedo/no puedo' por expresiones como 'cabe/no cabe'. A ella pertenecen las proposiciones de la lógica que describen situaciones conceptuales (lingüísticas) y los cálculos matemáticos: «La proposición matemática ha sido oficialmente sellada con la etiqueta de la incontestabilidad. Es decir: discutid sobre otras cosas; *ésta* se mantiene bien firme, es el eje en torno al cual puede girar vuestra disputa» (UG, 655). Se entiende que las proposiciones que pertenecen a ese orden constituyen el armazón. ¿Nos empeñaremos en probarlas empíricamente cuando son la posibilidad de toda proposición empírica?. ¿Cómo no hemos de caer en el enredo si nos empeñamos en justificar la gramática profunda como si de gramática superficial se tratara?.

Surgen en este punto dos dificultades de indudable transcendencia:

1ª. Las proposiciones matemáticas se obtienen a partir de una serie de actuaciones que no se diferencian en absoluto del resto de las acciones de nuestra vida y que se ven expuestas a olvidos, faltas de atención e ilusiones (UG, 651). Si puedo sufrir alucinaciones cuando digo que: «Me llamo X», ¿no podría ocurrir lo mismo con cada una de las proposiciones de la tabla de multiplicar?. ¿Con qué derecho contraponer la seguridad de las matemáticas a la relativa inseguridad de las proposiciones empíricas?.

2ª. ¿Cómo garantizar que no se han de cambiar las proposiciones matemáticas? (UG, 652). Pero entonces, ¿qué hay de su carácter incontestable?.

Por ambas dificultades parecen hacerse vías de agua. Necesitamos acaso recorrer otras formas de seguridad antes de ver si se pueden taponar.

Del mismo modo que existe la tendencia a dotar de empiricidad a las proposiciones lógicas, se pretende dar valor gramatical a proposiciones temporales avaladas por la fuerza de una evidencia incontestable. ¿Qué clase de seguridad corresponde a proposiciones como: «Aquí hay una mano?», o: «Me llamo X». ¿Acaso no son tan indiscutibles como que « $12 \times 12 = 144$ »? En caso de que no sea así, ¿sobre qué principios podremos asentar la afirmación de la realidad?.

La palabra 'principio' puede tomarse en varios sentidos: No es lo mismo buscar una 'noción común', que pueda servir para probar la existencia de todos los entes que se conozcan, que buscar un 'ser' cuya existencia nos sea más conocida que la de otros, de modo que nos pueda servir de principio para conocerlos. Esta exigencia cartesiana³³ sastifecha por el principio de contradicción y por el 'cogito' respectivamente, marcó el rumbo de la filosofía moderna. Pero, si es verdad que el principio de contradicción queda encerrado en los límites de la lógica, el principio de que 'nuestra alma existe' pretende gozar de idéntica firmeza.

Este es el punto que para Wittgenstein hay que pensar con más calma. Primero apartándose de formulaciones de cuño mentalista como la que precede; lo que da

³³ A Clerselier, junio o julio, 1646, AT, IV, 444, 4-445, 8.

obligadamente al traste con la solución de la duda para Descartes. Aserciones como «Aquí hay una mano» (es decir, la *mí*a), serían del tipo preferido por Wittgenstein. Pero justamente esas habían caído ya ante la fuerza arrolladora de la duda³⁴; luego han de recibir su fuerza incontestable de otra fuente.

Segundo, porque son proposiciones relativas a cuerpos, como: «Existe un planeta a tal distancia del sol»; y pese a «no haber una frontera bien delimitada entre ambas» la primera «no puede ser catalogada como una hipótesis» (UG, 52).

La salida cartesiana dejaba en suspenso el mundo (en el que va incluido el propio cuerpo). La pretensión de Wittgenstein corre el riesgo de declarar incorrectos todos nuestros cálculos. Porque, ¿cómo fijar la frontera entre nuestras proposiciones hipotéticas y las que no lo son si todas se refieren a la única clase de objetos (los cuerpos)? Si es manifiesta la posibilidad de errar acerca de los planetas, también habrá esa posibilidad cuando hablo de mi cuerpo.

La naturaleza del principio de realidad quedaba en Descartes incontaminada de corporalidad por un salto a otro género en el que no cabe error. En el único género que queda para Wittgenstein muere esa posibilidad parece que de modo irremediable. La única diferencia de unas proposiciones a otras está en que cuanto más nos acercamos a nosotros mismos, más improbable se vuelve el error; es tanto como decir que todo dependerá de lo minuciosa que sea la observación. Pero el empirismo y últimamente el verificacionismo ya enseñaron con profusión sus debilidades como para que nadie -mucho menos Wittgenstein- pretenda acogerse a él.

La alternativa queda, pues, en escepticismo o decisionismo. Pero la decisión no es relativa a determinadas experiencias *qua* tales, sino la que impone el juego del lenguaje que establece también la posibilidad misma del error y de la duda. Ejemplo: se puede dudar de si algo existe porque sabemos de algo que existe. La misma cinta transportadora del lenguaje trae las condiciones de su aplicación con uno u otro tipo de seguridad y lo que no cabe como aplicación; de no ser así, «sería concebible que nos equivocáramos en *todos* los enunciados sobre objetos físicos, que todos los enunciados que hiciéramos fueran incorrectos». Pero, en ese caso, no sabríamos tampoco qué significa (a donde se aplica) 'corrección'. Luego, el error, «cuando llegamos a cierto punto ya no es ni siquiera concebible, la duda pierde gradualmente su sentido... y todo lo que describe el juego de lenguaje pertenece a la lógica» (UG, 54, 56).

Con lo dicho se pueden esclarecer otros problemas: uno relativo a la gramática de las proposiciones; otro, a los distintos grados de seguridad.

«Qué clase de objeto es algo lo dice la gramática» (PhU, I, 373). Esto nos permite distinguir la imagen que está en primer plano de su trasfondo difícil de reconocer (PhU, I, 422). Es en ese juego de fondo donde nacen las confusiones, atraídos a la par por la estructura de las operaciones y por la aparente proximidad de las aplicaciones: capacidades gramaticales las asimilamos a capacidades físicas y barreras gramaticales son contempladas como imposibilidades físicas. Por ese camino podemos encontrar una nueva versión del principio de realidad que, fracasado el modelo cartesiano, habíamos perdido.

³⁴ DESCARTES, R.: *Meditaciones*, IIª, AT., VII, 24, 14-25, 5 y 26, 1-27, 8.

La proposición: «Sé que aquí hay una mano, no solo lo supongo», es una proposición gramatical y no solo temporal, comparable a esta otra: «Los solitarios los juega uno solo» (UG, 57; PhU, I, 248). ¿Cómo solemos actuar?.

De una parte, asimilamos la seguridad de la primera a la de cualesquiera otra situación experiencial del mismo género que, por ejemplo: «Hay un cuadro rojo en la pared»; de otra, convertimos el sentido conceptual de la segunda en sentido factual (como si pensáramos: «Nunca he visto a más de una persona haciendo solitarios»); por último equiparamos el sentido factual de todas las proposiciones: las relativas a mi experiencia, a la experiencia de los cuerpos externos y a la de las oraciones conceptuales. La consecuencia es la confusión que obliga a recurrir al antídoto de una duda imparable porque ya nada se distingue y manda el género de seguridad más débil. Sin embargo, pronto nos curamos de espantos y volvemos a nuestras seguridades indiscutibles sin que, por cierto, los motivos de duda se hayan difuminado.

Es en el fondo la vieja imagen de incoherencia del escéptico que, no pudiendo aceptar ningún criterio de verdad, lo vemos entregado a los más que sospechosos de la costumbre, o la naturaleza, o las impresiones³⁵. Wittgenstein parece haber encontrado la verdadera razón de ese modo de actuar, corrigiendo sobre la marcha la defensa, en apariencia peregrina e indiscriminada, del obrar conforme a la naturaleza de Hume.

Aquello a lo que la naturaleza nos arrastra de modo incontrovertible tiene que ver con las proposiciones gramaticales, no con las temporales. Por tanto, podemos inducir afirmaciones generales relativas a lo primero, no a lo segundo. La consecuencia: «De modo que hay objetos físicos» a partir de «Aquí hay una mano», no permite que: «De modo que hay colores» se pretenda concluir de: «Sé que veo rojo». Esta última es una proposición temporal, aunque tenga una estructura sintáctica y aparentes componentes experienciales del mismo género que la relativa a la mano.

Hay, pues, seguridades relativas a reglas gramaticales, como la del juego del solitario; hay otras relativas a experiencias que se *tienen*, como la de quien tiene un dolor que nadie más puede sentir y que son incontrovertibles (PhU, I, 246). Hay, además, todo el cúmulo de las experiencias comunes, que son tales porque todos las observan y más bien nadie las tiene, en torno a cuya seguridad caben inagotables sospechas.

Pero entonces necesitamos reconsiderar nuestro uso del 'sé...'. De momento para discernir cuándo es importante que acompañe a las proposiciones, a aquellas cuya *verdad* pertenece a nuestro sistema de referencias (UG, 83). Hablar de tal sistema comporta más que afirmaciones relativas al propio cuerpo. Entenderlo así equivaldría a haber caído una vez más en las redes de la experiencia. Pero ya hemos visto que: «Aquí hay una mano» tiene una seguridad gramatical; del mismo género todas aquellas que para nada dependen de investigaciones específicas o que admitamos porque se hayan seguido determinadas vías de pensamiento. Más bien hemos de pensar que se trata de proposiciones que, por estar al margen de toda duda, conforman como «los márgenes del camino que recorre la investigación» (UG, 88).

³⁵ SEXTO EMPIRICO: *Hypothiposis pirrónicas*, I, 11. Para algunos aspectos de la relación con el escepticismo, cfr: MALCOM, N.: «Wittgenstein's Scepticism in *On Certainty*», *Inquiry*, 1988 (31) 277-293; BOTWINICK, A.: «Wittgenstein and Scepticism. And Essay in the unitary of Wittgenstein's thought», *Philos. Res Arch.*, 1986-1987 (12) 163-176; LEWIS, A.: «Wittgenstein and rule-scepticism» *Philos. Quarterly*, 1988 (38) 280-304.

Nos las habemos, pues, con expresiones en las que 'sé...' es un conocimiento lógico que comporta que «no hay en tal caso nada como una duda». Pero siendo esto así, hay que decir: primero, que ninguna importancia tiene la expresión ni el pronombre que la introduce ('yo sé...', 'nosotros sabemos...') (UG, 58). No es el sujeto quien instaaura la verdad por arte de algún poder oculto, sino que se descubre a sí mismo como perteneciendo a un ámbito en el que una experiencia locucionaria puede tener sentido. No es el sujeto el que instaaura la verdad por medio del lenguaje, sino que es el lenguaje el que instaaura al sujeto en la verdad.

Segundo, que está fuera de lugar la pretensión de probar el realismo a partir de los que son conocimientos lógicos referenciales (UG, 59). Ese ha sido el error de Moore y, en general, del realismo tradicional; cuando menos habría de hablarse de un realismo lingüístico.

Tercero, que no podemos mantener la confusión entre «la imagen que nos hacemos del conocimiento y la que propiamente sirve de fundamento» (UG, 90). Porque se tiene la tendencia a aplicar a 'saber' un significado primitivo análogo a 'ver' y nos lo representamos como la percepción de un proceso externo que ocurre por medio de rayos visuales que lo proyectan en el ojo de la conciencia tal como es. De esa confusión nace la pretensión de hablar como si supiéramos lo que sucede en el mundo externo, más allá de los meros datos sensoriales. Wittgenstein está tocando con la ironía habitual en el nervio de esa especie de entusiasmo generalizado por resituar el problema del conocimiento en los dominios de la percepción, como si con ello ya se resolviera el asunto.

Cuarto, la firmeza de cuanto sirve de fundamento nos puede sorprender cuando descubramos que no tenemos *razones* adecuadas para nuestras convicciones. Nos quedamos con la imagen de una estatua, más o menos brillante, que tiene los pies de barro. Creemos que «la Tierra existe desde hace millones de años» porque siempre se ha dicho que era de ese modo. Si indagamos por qué fué así (dicho) y no que «existe desde cinco minutos antes de yo nacer», es posible que no haya en el fondo más que razones de *simplicidad* o de *simetría*. En tal caso, «nos limitamos a decir: *así debe ser*» (UG, 92).

La adopción del punto de vista cultural se ha llevado aquí uno de sus mejores bocados, aunque se lo debe al fracaso del modelo basado en la seguridad de la ciencia física al que Wittgenstein empezó por rendir tributo³⁶.

Situados en esa perspectiva, ¿qué pasaría si alguien viniera y nos convenciera de otro punto de vista?. Nada más que esto: que nos veríamos obligados a considerar el mundo de otra manera. No puede pasar nada más porque aquello no era «más que el trasfondo que me viene dado y sobre el que distingo entre lo verdadero y lo falso» (UG, 94). El pragmatismo que poco más arriba hemos visto confesado, envuelve un relativismo antropológico desde el que resulta factible eludir los problemas de fundamentación. Las cosas son como son y son como tienen que ser mientras nuestra concepción del mundo no cambie.

³⁶ Cfr. PUTNAM, H.: *Reason, Truth and History*, Cambridge Univ. Press, 1982, cap. Vº, final.

5. SOLIDIFICACION MITICA COMO SUSTENTO DE CONCEPTOS

Existen proposiciones que describen el mundo y a las que remitimos, como a un fondo seguro, todas las demás; «pertenecen a una suerte de mitología» (UG, 95). El mito fué la primera palabra; y el mito es la antesala del símbolo, como el símbolo es la antesala de la abstracción y del concepto -advértía nuestro poeta. Tienen los mitos el carácter de lo sagrado y verdadero y, como cuentan la fundación del mundo (o de la institución, o del trabajo), constituyen el paradigma de todo acto significativo; además, por contener la referencia «originaria», en él se alcanza el dominio y la manipulación de las cosas que explica³⁷. De una forma precipitada acaso habíamos dado por válida la imagen positiva de una evolución hasta el pleno dominio del concepto. La historia de la filosofía y el curso de nuestra civilización no serían otra cosa que el correr de ese proceso.

No deberíamos, sin embargo, perder de vista que, bien avanzada esa historia, se hayan hecho dramáticas llamadas en favor del concepto³⁸. La llamada toma de conciencia de haber quedado presos en las redes de la representación. Y esta es una preocupación que se tiene que arrastrar hasta Wittgenstein por cuanto que sobre ella, como gozne, ha girado la concepción moderna del conocimiento. La diferencia ahora está en la constricción, no para avanzar al concepto, sino para volver al mito, en cuanto fundamento de todo hacer racional.

El mito fue la primera palabra y tiene una función semejante a las reglas de un juego que se aprende, con explicitación de las reglas o no, jugando. Así aprenden los niños el lenguaje.

Sí ahora preguntáramos ¿por qué son incuestionables las proposiciones gramaticales?, diríamos: porque el mito así las mantiene. Son la parte solidificada del lenguaje que funciona como un canal sobre el que fluyen las proposiciones empíricas. Para que suene a mayor escándalo y aclaremos alguna pregunta que nos ha salido al paso (cómo las proposiciones matemáticas, determinándose de la misma manera que las demás, son incontrovertibles y están garantizadas?), podemos afirmar que la matemática y la lógica son nuestro almacén mitológico, que rellenamos de aventuras una vez que los perfiles de los personajes y de las grandes acciones han quedado diseñados.

Como tal almacén, parece que haya de ser indiscutible y, por ende, intocable; es decir, que la lógica no pueda cambiar. Más, ¿acaso no cambian los mitos?. Si nos servimos de la comparación, ¿no habremos de asumir todas las condiciones que impone?.

Los mitos ciertamente cambian, aunque también parece que siempre toquen los mismos problemas y apunten para ellos clarificaciones similares (familiares?). No es que Wittgenstein se haya dejado embrujar por el lenguaje cuando pretendía deshacer el hechizo: le vemos reconociendo que el lenguaje cambia y, a pesar de todo, algo hay como que siempre se mantiene.

Lo primero nos sorprende ya en los mismo términos que explicitan la comparación: «Esta relación cambia con el tiempo, de modo que las proposiciones que fluyen se solidifican y las sólidas se fluidifican» (UG, 96).

³⁷ Cfr: SCHULTE, J.: «World-picture and Mythology», *Inquiry*, 1988 (31) 323-334; FELLMANN: «Von der Sprache des Mythos zum Mythos der Sprache», *Phenom. Fors.*, 1988 (21) 114-132.

³⁸ ¿Acaso no es paradigmática la actitud de Hegel? Cfr: *Phänomenologie des Geistes*, Vorrede, (Ed. Moldenhauer-Michel, Frankfurt, 1970, vol. 3, p. 58).

Lo segundo nos deja perplejos: «El lecho del río de los pensamientos puede desplazarse» (UG, 97). Lo paradójico está en que, aunque se desplace, siempre el lecho del río se mantiene, es diferente del agua que fluye. Aquí parece reaparecer el hechizo y reencontrarnos en la red^{38bis}.

Podemos comportarnos ante ese hecho de dos maneras: Una, saltando por encima del valor limitado de los ejemplos, que son tan ilustrativos como entorpecedores; en definitiva, «andaderas del juicio»³⁹. Entiéndase que esa actitud tiene importancia porque en el fondo quiere escapar de la fuerza que constriñe al mito para acceder a la fuerza constrictiva del concepto. Recuérdese que Dédalo pudo escapar del laberinto que había construido fabricándose también las alas y por ello se convierte en un modelo a imitar por el hombre moderno⁴⁰.

Pero también podemos comportarnos sumergiéndonos en la realidad de la que emanan las cuestiones que nos planteamos; en suma, por tentador que pueda ser volar, moverse en el laberinto sin el riesgo de ser devorados. En ese caso tenemos que, distinguiendo dos tipos de embrujo, descubrir las posibilidades del lenguaje. En otros términos, el hechizo acompaña al lenguaje como los pasadizos que se cruzan y retuercen constituyen el laberinto; otra cosa es el mugir del monstruo que lo habita, contra el que siempre tiene que batirse la filosofía. Gramática profunda y gramática superficial vuelven constantemente a reclamar nuestra atención.

Wittgenstein, consciente de la pretensión que recorre toda la modernidad, que asienta sobre la intención de ser «maestros y poseedores de la naturaleza», vuelve la mirada al fondo mismo que alimenta esa pretensión, lo que significa entregarse al hechizo sin caer en el embrujamiento. Por eso sigue valiendo ahora que «los límites del lenguaje significan los límites de *mi* mundo» (TrLPh, 5.62).

Aún puede Wittgenstein sorprendernos más: Todo el mundo distingue entre el agua que fluye y el lecho sobre el que discurre; la sorpresa está en esto: «que no haya una distinción precisa entre una cosa y la otra» (UG, 97). Y ahora que lo pensamos nos enteramos de que es así: que cuando el agua se agita arrastra envuelta la parte del lecho que consigue arrancar. ¿Hasta dónde podemos llegar?. No puede haber río que fluya sin lecho sólido. La corriente es capaz de remover la parte arenosa del lecho y depositarla a capricho en puntos diversos; pero el márgen del río es en parte de roca que no está sometida a alteración o lo está solo a cambios imperceptibles (UG, 99).

Redescubrimos así la profunda verdad que inspiró a Heráclito⁴¹, la naturaleza de nuestras seguridades y su movilidad dentro siempre de los márgenes. Por eso, «por más que

^{38bis} Cfr. RORTY, R.: *La filosofía y el espejo de la naturaleza*, Madrid, Cátedra, 1983 (trad. J. Fernández Zulaica). SAINTZ-FLEUR, J.P.: *Logiques de la representation. Essai d'épistémologie Wittgensteinienne*. Louvain-la-Neuve, Academia, 1988; CURBELO, I.: «Wittgenstein y el concepto de representación del mundo en *On Certainty*» *Diálogos*, 1988 (23) nº 52, 109-126.

³⁹ KANT, I.: KrV, A 134/B 173.

⁴⁰ Me gustaría poder demostrar esto: cuando la recuperación de los clásicos generaliza la noción de «laberinto», lo interesante no es la fácil aplicación del término a los razonamientos escolásticos -que es lo que más se divulga- sino el poder de dominar el laberinto del mundo por haberlo construido, y en esa historia andan tanto Bacon, como Descartes y Leibniz.

⁴¹ Diels-Kranz: 5, 12 (1^o), 126, 64, etc.

sea cierto que la misma proposición puede considerarse, a veces, como una proposición que ha de ser controlada por la experiencia y, otras veces, regla de control», se equivocaría quien dijera que la lógica es una ciencia empírica (UG, 98). De otra parte, se nos alcanza que el pensamiento de Wittgenstein sea a la par un pensamiento de la inmediatez de las formas de vida y de la intemporalidad de lo que cambia a ritmo para nosotros imperceptible.

Ahora podemos entender mejor las «pruebas» del mundo externo de Moore. Se trata de proposiciones empíricas que desempeñan un papel lógico bien particular; por ello son verdades de tal naturaleza que «si él las sabe, todos las sabemos» (UG, 136, 100).

Otro tanto se puede decir del riesgo del error: «Para que un hombre se equivoque ha de juzgar ya de acuerdo con la humanidad» (UG, 156). Lo primero en lo que pensamos es en la debilidad de los testimonios, que se mide en el contraste con la clase de testimonios incontestables. Así los filósofos han discutido sobre el valor relativo de la percepción, de la memoria o de la tradición; y siempre figura al fondo alguna especie de conocimiento sobre el que no cabe duda alguna, alguna especie de intuición o de elaboración racional definitiva. Incluso cuando no nos atrevemos a afirmar ese fondo, podemos mantener nuestras convicciones mientras no se presenten pruebas en contra de mayor fuerza. Pueden, pues, ocurrir desmoronamientos, pero siempre serán relativos. «Por ejemplo, estoy completamente seguro de la fecha en que ocurrió la batalla, pero, si encuentro una fecha distinta en un libro de historia que sea digno de crédito, cambiaré de opinión sin perder, por ello, la confianza en todos los juicios» (UG, 66).

Tanto aquella tradición que apela a una fuente última de conocimiento seguro, como ésta que ignora (o niega) tal fuente, pero no se siente perturbada por desmoronamientos, apuntan a un fondo apenas alterable que solo en los últimos tiempos se ha podido identificar. Ante la vieja polémica entre dogmatismo y escepticismo ya Hume había intuido alguna forma de coincidencia, tanto en las dificultades como en las seguridades, aunque «por hábito, por capricho o por inclinación» unos insistan más en lo uno y otros en lo otro⁴².

¿De dónde viene que haya dificultades tanto con respecto a los sentidos como con respecto a cualquier ciencia que son absolutamente irresolubles con un método regular?. ¿De dónde que nos encontremos en la imperiosa necesidad de pensar, creer, razonar sobre toda clase de temas, e incluso de dar nuestro asentimiento con confianza y seguridad?.

Merodeamos en torno al error mientras pretendemos liquidar definitivamente las dificultades o, cuando no conseguimos esto, queremos acabar con nuestras seguridades. La seguridad está impregnada de dificultad y ésta solo crece del suelo nutricio de aquella. Podemos cometer errores justo porque hay supuestos acerca de los que no cabe error. La lógica del lenguaje establece también dónde podemos empezar a dudar y hasta dónde nos podemos equivocar.

Pero equivocarse «allí donde consideramos que un error es imposible y donde, de hecho, no nos hemos encontrado con ninguno» (UG, 67; cfr: 155) es, más que de error, asunto de perturbación mental, que ha de ser tratado de otro modo. Una vez más, siguiendo los pasos de Descartes, hemos tenido que recluir a la locura para evitar que se adueñe del mundo.

Cuando sentimos el cerco de la posibilidad de errar acudimos precipitadamente a la comprobación; y lo hacemos indiscriminadamente, como si todas nuestras convicciones fueran de la misma naturaleza. Pero la comprobación es un procedimiento para el que la

lógica tiene establecidos los límites de su valor y aplicación. Una vez más hemos de recordar la sorpresa que nos causa la idea humeana de que cuanto más revisamos nuestros cálculos más aumentan nuestras dudas, acostumbrados a la concepción contraria del cálculo de probabilidades⁴³.

Sabemos que nuestras cuentas no se convierten en más seguras porque las revisemos veinte veces, o lo hagan veinte personas diferentes (UG, 77). Desde el momento que nos instalamos en la falibilidad natural de nuestro juicio, las reglas de la lógica sufren una disminución continua hasta la completa extinción de toda creencia y evidencia. Pero sabemos que vamos a seguir actuando de acuerdo con lo que, a falta de mejor conocimiento, llamamos la «fuerza imperiosa de la naturaleza».

Pues bien, es ese fondo de la 'naturaleza', cuya ciencia quiso construir Hume, el que ahora se nos revela como constituyendo la lógica del lenguaje⁴⁴. Por ello necesitamos delimitar errores de perturbaciones y abandonar la idea de que una refinada estrategia de comprobaciones pueda acabar con todo motivo de duda o de seguridad. De otro modo quedaría destruida por completo la razón humana.

No podemos asimilar error y perturbación mental. Los límites entre uno y otro sólo los conocemos cuando hemos dejado atrás los planteamientos dogmáticos y escépticos que giran en torno a la ultimidad de un fundamento racional para la seguridad de nuestros conocimientos. Un error ocurre dentro de un sistema que sabe perfectamente quien se equivoca. La perturbación ocurre cuando nos salimos del juego, como quien alocadamente se colocara fuera del mundo. Colocarse fuera del juego es tanto como echar a un lado el lenguaje como portador de la experiencia de la historia que corrige las alucinaciones de los sujetos y construye el terreno medio donde ocurren las coincidencias y se hace posible la comunicación y el entendimiento que nos asegura de un trato adecuado con las cosas.

De las perturbaciones podemos encontrar las causas; «un *error* no solo tiene una causa, sino también una razón» (UG, 74). ¿Podemos observar con mayor nitidez la modernidad del pensamiento de Wittgenstein manteniendo la transformación de la 'causa' en 'ratio'? ¿Hay una forma más clara de afirmar la superación de esa misma modernidad haciendo nacer la 'ratio' de la gramática? A partir de ese momento ya no somos responsables de algunas dificultades que han salido al paso y que forman parte del elenco de problemas de la Teoría del conocimiento clásica, agudizados en la época del positivismo. Establecidas las reglas del cálculo, ¿qué sentido tiene empeñarse en repetir cada una de las operaciones realizadas una y otra vez para comprobar que no hay error?.

Por ese señuelo quedó atraído Hume y lo que él dice responde al modo habitual de actuar el sentido común. Pero, establecidas las reglas y aplicadas, la seguridad discurre por sí misma. Esto quiere decir también, independientemente del sujeto que calcula, que acaso

⁴² HUME, D.: *Dialogues concerning human Religion*, XII (En *Philosophical Works*; ed by Th. H. Green and Th. H. Grose, vol. II, p. 459, nota 1).

⁴³ HUME, D.: *Treatise...* I, IV, I (p. 182-183 Selby-Bigge/313-315 Duque).

⁴⁴ Sobre la diversidad de los errores y cómo se resuelven, cfr: MALCOM, N.: «Moore y el lenguaje común», en CHAPPELL, V.C.: *El lenguaje común*, Madrid Tecnos, 1971 (Trad. J.R. Capella), p.26-27.

esté condenado como tal sujeto, a no poder vencer nunca el azoramiento de la inseguridad y la sospecha⁴⁵.

La parte que en este asunto toca a la filosofía está en: primero, mirar por detrás de tales dudas y hacerse cargo de que tanta razón hay para dudar de nuestros cálculos como para no dudar (UG, 78); y, segundo, deshacer el embrujo, provocado por el lenguaje, que nos lleva a pretender dar razón hasta de aquello que marca los márgenes de la razón y hacerlo bajo la responsabilidad de sujetos individuales y empíricos. No, «lo que ha de ser considerado como prueba suficiente de un enunciado, pertenece a la lógica. Pertenece a la descripción del juego de lenguaje» (UG, 82).

Hay en el fondo de los pensamientos de Wittgenstein en torno al principio de razón resonancias que hacen aconsejable una discusión para el contraste con los textos de Leibniz que inaguran formalmente el mandato de ese principio; podría quedar de manifiesto el fundamento sobre el que asienta en uno y otro pensamiento, el alcance que le corresponda como «corolario» del gran principio y la clase de lógica a la que sirva⁴⁶.

Así como no podemos eliminar la perturbación mental, sí podemos eliminar en muchos casos la posibilidad de error. El ámbito donde aquella se produce no lo podemos dominar; tiene causas, pero no razones. Pero sí podemos dominar la ocurrencia de errores dominando las razones (UG, 650). Se deben a la no aplicación debida de las reglas que acompañan el juego. Lo demás es rebasar las posibilidades mismas que el juego instituye; por ejemplo, si no dudamos de la efectución acorde con las reglas de ciertos cálculos, pero dudamos del cálculo mismo (en cuanto a juego). Sería una situación similar a la de quien no duda que haya jugado correctamente las piezas del ajedrez, sino que teme que las reglas del ajedrez no sean las correctas.

Esa es una forma de mirar por detrás que no está permitida porque equivale a salirse del mundo. Y no vale con decir que la cosa puede ser clara en matemáticas pero que ya no lo es tanto en cuestiones de experiencia. Por ejemplo, ¿me puedo equivocar si digo que nunca he estado en la Luna?. (UG, 662). A poco que considere el modo de comportarme, descubro que la posibilidad de error queda abierta en lo que son condiciones diferentes de juego, tales como haber sido transportado allí por medios desconocidos mientras dormía. Cuando yo digo que 'no me puedo equivocar...', lo hago como jugador que sabe qué jugadas caben en ese juego; lo que queda fuera no es afectado por mi aserción (UG, 663).

Así llegamos a poder decir que nuestro juego de seguridades asienta sobre un fondo de inseguridad permanente que no cuenta con otra garantía que la de la vida misma. Los límites del error han sido trazados en conformidad con una experiencia pasada de casos en número significativo. La lógica del lenguaje no ha sido construída arbitrariamente, sino respondiendo a los hechos. Ahora es la corriente de la vida la que con su fuerza ataca los márgenes del río cambiando su disposición, o arranca alguna parte del lecho que previamente se había consolidado. Ahí queda explícito el dominio del principio de experiencia; la experiencia

⁴⁵ Desde ahí se iluminan perspectivas relativas a la pérdida del sujeto en la historia del pensamiento moderno y, más concretamente, al problema de la «identidad». Cfr: HUME, D.: *Treatise...*, I, IV, VI; véase: ARREGUI, J.V.: «Yo y persona. El problema del sujeto en Wittgenstein», *An. Filos.*, 1985 (18), nº1, 113-133.

⁴⁶ LEIBNIZ, G.W.: A Amauld, junio, 1686, Gerhardt, II, IX^a, p.56.

originaria entreverada de lógica que se convierte en almacén del pensamiento. Por eso está antes de toda racionalidad o irracionalidad, como lo está el juego que sobre ella se construye y ambos son, «por decirlo así, algo imprevisible... Están ahí -como nuestra vida» (UG, 559, 558).

Que el juego sea algo imprevisible puede inducir a universalizar la cautela a todas las proposiciones, de manera que bajemos de la seguridad del 'sé...' a lo movedizo del 'creo...'. Cuando descubrimos la falsedad de una creencia podemos corregirla sin que el desmoramiento nos deje a la intemperie; pero si son los supuestos mismos los que resultan falsos, parece que lo que se lleva por delante es el cimiento mismo sobre el que descansan todos los juicios (UG, 492). Pero, ¿acaso eso puede ocurrir?. Y, ¿qué haríamos si llegáramos a tener esa impresión?

Las proposiciones gramaticales son como el cuerpo encallecido de la experiencia que puede sustentar los tejidos más débiles de los hechos que nos acontecen. Para que los supuestos resultaran falsos habría de producirse un cataclismo sobre cuya posibilidad nunca discurre la vida, aunque la contemple en el límite. Por eso, la reacción sería decir que «nunca habíamos pensado en ello». Porque entretenerse con ese pensamiento mientras el cataclismo no se produce entorpece la actividad de pensar en conformidad con la lógica de la experiencia. Y, para pensarlo después de que ocurra... ya habrá tiempo.

No se trata de inmunizarse contra el problema auxiliándose de la ciencia, sino reconocer que «de determinadas proposiciones no se puede dudar sin renunciar a todo juicio» y esa no es una noción empírica, ni pertenece a la psicología, sino que tiene el carácter de una regla (UG, 494). Una revisión del juicio que afecte a los supuestos mismos, equivale a la aniquilación de todo patrón de medida, sin el que no es posible emitir juicio alguno.

6. EVIDENCIA Y VERDAD

Las proposiciones gramaticales de carácter matemático y de soporte experiencial delimitan el campo dentro del cual se juega y sus reglas.

Constituyen el «fundamento de todas las operaciones con el pensamiento» (UG, 401). La eliminación que hemos visto del error enlaza con la elucidación del papel de la *evidencia*. También este es un concepto que atrae sobre sí una vieja y compleja búsqueda, sobre la que ha pesado y aún pesa lo más grave del problema del conocimiento⁴⁷. También en este caso se habría caído en el enredo de querer mirar el mundo por detrás, respondiendo a la ilusión ancestral de encontrarse con las cosas cara a cara. Mejor, de sorprenderlas en su intimidad. Como si pudiéramos llegar a resarcirnos del asombro inicial que suscitan nuestras preguntas llegando hasta la desnudez de las cosas mismas.

Desnudez por desnudez, como exigía Descartes: una mente pura y atenta, liberada del ir y venir de los sentidos y de los cantos de la imaginación para una idea transparente y aislada de toda contaminación⁴⁸. Y, en medio de ese juego, la cuestión de si intuición sensible o también intuición intelectual, rozando siempre los límites de la razón⁴⁹ hasta venir

⁴⁷ Véase: CHISHOLM, R.: *Teoría del conocimiento* Madrid, Tecnos, 1982, (Trad.: V. Peris) especialmente caps 2 y 4.

⁴⁸ DESCARTES, R.: *Regulae...*, III^a, AT. X, 368, 13-17; *Principia*, I^a, 45^a.

⁴⁹ Véase al respecto: GILSON, E.: *Index Scholastico-cartesien*, nrs: 84, 85, 87. París, Vrin, 1979.

a dar en la rebaja humeana plasmada en el reconocimiento de un realismo ontológico y un fenomenismo gnoseológico. En ese grado del problema, el paso a través del tándem Hume-Kant fué definitivo. Tanto como para reorientar el asunto en la renuncia a la desnudez, tanto de la mente como de la idea de la cosa. La recomposición posterior pasa por las posibilidades de la fenomenología husserliana y de la significativa sugerencia de Peirce para «dar claridad a nuestras ideas». Al final nos encontramos inevitablemente instalados en los dominios del lenguaje. Entonces es cuando sabemos lo que quiere decir 'evidencia'.

La posibilidad de conocimiento asienta, en efecto, sobre alguna evidencia porque, de lo contrario, no podemos confiar ni siquiera en la evidencia presente y, si eso ocurre, ya no podemos confiar en nada. «Suponiendo que no fuera verdad que la Tierra existía desde mucho antes de mi nacimiento, ¿cómo habríamos de imaginar el descubrimiento de ese error?» (UG, 301).

La evidencia está ligada al significado y el significado al uso. Querer volver el uso del revés para generar nuevos problemas es hurgar en lo que ya no forma parte de nuestra vida. Ese es el sentido de la rueda que gira alocadamente⁵⁰. No puedo decir (= no cabe) que «no sé si aquí hay una mano» cuando sé lo que significa la palabra 'mano', y no solo para mí y ahora, porque yo estoy comprometido con el uso que se hace de la palabra en el juego (UG, 306).

¿De dónde viene el afán de ir más allá del uso, de dónde el impulso de la rueda alocada?. Ya ha quedado manifiesto que no es cosa de pretender aquél «fundamentum inconcussum veritatis» del comienzo de la modernidad y que, a pesar de ello, la cuestión del fundamento no puede sin más ser despachada; como ha quedado claro cuánta es la fragilidad y movilidad que afecta a la lógica del lenguaje. ¿No tendremos que terminar reconociendo que nos movemos -y zarandeamos- del fundamento gramatical al ansia de poder dar razón del mismo, olvidando en el azoramiento que todo dar razón ocurre ya siempre dentro de aquél fundamento?. No hay fundamento para el fundamento y podremos dar miles de razones de nuestra manera de actuar, «pero ninguna será tan segura como la que debería de fundamentar» (UG, 307).

Al tiempo que necesitamos desbaratar la pretensión abusiva de dar razones, se diluyen otros aspectos del problema de la evidencia: Uno de ellos tiene que ver con la relación que la evidencia dice a la verdad; el otro con la imagen que nos hacemos de la evidencia.

Asociamos la evidencia a «lo que es verdadero con toda certeza» (UG, 197), con lo que estamos suponiendo que podemos poseer tal tipo de conocimientos y, cuando inquirimos más allá buscando qué significa 'verdadero' desembocamos en el tópico del estar «de acuerdo con lo hechos» (UG, 199). Hasta aquí no debería de haber complicaciones, si no fuera porque de inmediato aparece la imagen de una especie de rayo noético que va del pensamiento a las cosas y que resulta ser el garante del acuerdo. Ante la instantaneidad del rayo, a nadie se le ocurre pensar en el significado del 'acuerdo'. Como cuando contemplamos el paisaje no se nos ocurre detenernos a considerar la naturaleza del medio en el que viajan los rayos de luz, así nos forjamos la idea de la evidencia «como si fuera una especie de *ver* por nuestra parte» (UG, 204).

Pero es justamente en ese «entre» (el pensamiento y los hechos, el sujeto y el objeto, la idea y la realidad, etc.) donde el asunto del conocimiento se nos despista y la rueda de la

⁵⁰ MALCOM; N.: l.c., p. 19-20;

filosofía empieza a girar en el vacío. Pues bien, Wittgenstein ha querido señalar, de una parte, lo erróneo de un planteamiento que focaliza la atención hacia el predominio de un papel, sea del sujeto o sea del objeto; y, de otra, desmarcarse de aquellos enfoques erróneos acerca de la naturaleza del «entre» que previamente han detectado adecuadamente el lugar del problema. Ni una interpretación ontológica en cuanto cooptenencia al ser de la idea y la cosa; ni un acuerdo sobre la base de la cooptenencia a la idea del sujeto y la sustancia⁵¹. El acierto en haber resituado el problema, en tales casos, no queda compensado ante la desmesura de la interpretación.

La aquí llamada desmesura se corresponde con la pretensión metafísica que, en último término, funciona como ya advirtiera Kant: cuestiones inevitables a las que buscamos respuesta apresurada. Un análisis más riguroso nos permitiría sobre esa base descubrir que apelamos a principios trascendentes cuando no podemos soportar mantenernos en los límites de las condiciones transcendentales. Con la diferencia de que la unidad del sujeto ahora salta hecha pedazos ante la diversidad de juegos que son los que definen las condiciones.

Resulta difícil aceptar que, en suma, todo pende del hilo de una decisión y acaso nos quedamos decepcionados al constatar que el entendimiento sigue a la voluntad. Otro viejo problema nunca suficientemente desenredado, pero definitivamente claro para toda la modernidad.

Aparentemente es un hilo demasiado débil y se vería lógica la necesidad de algún fundamento más firme. Ello es porque la historia compleja del problema del entendimiento y voluntad nos ha hecho inclinarnos a la simplificación que opone racionalidad a arbitrariedad en términos duros y las actitudes duras de la filosofía ya sabemos que ocultan cegueras. Pero «hemos de considerar en primer lugar el papel de la decisión a favor y en contra de una proposición» (UG, 198).

La parte de verdad que encierra la simplificación es que no podemos proporcionar el fundamento que corresponde a la decisión a favor o en contra de la verdad. El abuso de la simplificación está en llegar a entender que no sea posible decidir a favor o en contra.

Nos decidimos por una proposición cuando todo habla a su favor y nada en contra. El fondo de arbitrariedad queda compensado por la seguridad de una experiencia habitual que no tendríamos por qué alejar demasiado de los términos humanos. La cuestión siguiente es si podemos darle el carácter de «objetivamente segura» (UG, 203).

Ni el mismo Wittgenstein, después de todas las llamadas a la modestia de la teoría de los juegos, recomendaría que en cada caso acompañáramos la muletilla de: «... en nuestro juego de lenguaje». La confusión surge del lenguaje por cuanto que lo que podemos llamar objetivamente seguro, lo entendemos como de *acuerdo sin restricciones* como el mundo de los hechos. Pasamos sin darnos cuenta de lo que significa 'estar de acuerdo' a hacer de ello aplicación.

⁵¹ Cfr: HEIDEGGER, M.: *Identität und Differenz* (Pfullingen, Neske, 1957/Ed. Bilingüe a cargo de A. Leyte y H. Córtes, Barcelona, Antropos, 1988). HEGEL, G.W.Fr.: *Phänomenologie des Geistes*, Vorrede, ed. cit. p. 22-23).

Si llamamos 'acuerdo' a las evidencias que hablan a favor de nuestra proposición en el juego de lenguaje, dejamos atrás la representación de la intuición como un *ver* que nos pone ante la intimidad de la cosa y nos quedamos con lo que *hacemos* al justificar nuestra decisión por o contra la verdad de algo (UG, 204). Eso significa el límite de justificación de la evidencia. No al modo clásico, en el que constaba la escasez de objetos de intuición y la dificultad para alcanzarla, sino la limitación a un significado que no se sale nunca del lenguaje -como el ser vivo no se puede salir del mundo- para ir y volver de las palabras a las operaciones en las que cobrarían sentido los significados.

¿No nos encontramos con que el brillo del problema de la verdad también se ha debilitado? ¿Y no es ello mismo efecto del desembrujo que Wittgenstein ha visto necesario para todos los problemas?. Solo nos cabe preguntar por esa especie de destino ciego en el que caen la filosofía y el lenguaje, condenada aquella a destejer lo que éste continuamente teje.

7. LOGICA Y FUNDAMENTO

Tejido con los problemas anteriores aparece el de la lógica y el fundamento. En realidad, resulta inevitable la referencia desde que se empieza a hablar de experiencia, de error, de duda o de evidencia.

La sorpresa primera de la que por hábito histórico del problema nos cuesta sobreponernos, es aquella de que el fundamento del pensamiento forman parte tanto proposiciones lógicas como empíricas. Estaríamos tentados a seguir viendo ahí la distinción entre juicios analíticos y sintéticos. Pero es justamente esa tradición la que hay que romper, porque a su cargo cuentan suficientes confusiones.

Lo más claro es decir que «no hay ningún límite claro entre proposiciones metodológicas y proposiciones en el seno de un método» (UG, 318). Es lo mismo que decir que no hay límite claro entre proposiciones lógicas y empíricas (UG, 319). Tales afirmaciones resultan relativamente escandalosas, pero pertenecen al sentido mismo que se ilumina desde dentro del juego. Del mismo modo que estamos acostumbrados a entender la razón como fundamento y no como viniendo del fundamento, así entendemos que la claridad es algo que funda y no es fundado. Pero claridad/confusión son términos (conceptos) que hemos aprendido a usar en circunstancias determinadas de manera correcta. Y ya no es que el límite no sea preciso entre proposiciones empíricas y matemáticas; el mismo concepto de 'proposición' no está bien delimitado (UG, 320). Pero lo está suficientemente. En este sentido, Wittgenstein ha podido de nuevo recurrir a la tradición para volverse contra la interpretación usual que de ella se ha hecho. A fin de cuentas, el escándalo que provocan estas últimas aserciones no debe ser mayor que el que suscitaban las consideraciones sobre la suficiencia de la enumeración mientras se estaba proyectando una ciencia consistente en conocimientos ciertos y evidentes⁵².

No aprendemos primero lo que es una proposición en todo caso y luego entramos en la clasificación de diversos géneros. Eso es lo que nos gustaría hacer y lo que inconscientes de la naturaleza del lenguaje/juego vienen haciendo los filósofos hasta el autor del

⁵² DESCARTES: *Regulae...*, VII^a.

Tractatus. En consideraciones de esa naturaleza abundan las *Investigaciones*⁵³. Pero tenemos que llegar a «mirar la proposición como un instrumento y su sentido como su empleo» (PhU, I, 421).

La sorpresa segunda es mayor y deja perplejo al mismo Wittgenstein: «Toda proposición empírica puede transformarse en un postulado -y entonces se convierte en una norma de descripción» (UG, 321). Como el lecho del río que se forma de proposiciones que se han solidificado y que además se desplazan.

Que la distinción de los tipos de proposiciones se difumine lo entendemos mejor cuando observamos cómo funciona el lenguaje (no simplemente cuando lo usamos) que cuando complicamos demasiado los problemas filosóficos. Todo el mundo se comporta en la convicción de que no es razonable quien «se niega a creer que esta montaña está en su sitio desde tiempo inmemorial» (UG, 322).

La distinción desaparece llevándose consigo tanto el privilegio incontestable de las unas como la fragilidad de las otras. En primer lugar, no se quiere decir que ahora todo sea tan inseguro e imprevisible como lo empírico, sino tan incontestable como lo lógico. En segundo lugar, esa firmeza que se ha ampliado hasta lo empírico, contamina a lo lógico de la movilidad que acompaña por necesidad a todo juego.

De ello se desprenden dos consecuencias que han de ser vistas como terreno ganado. La primera, que nos libramos de la responsabilidad estrictamente subjetiva de nuestras aserciones de fundamento empírico y podemos hacer desaparecer el 'yo sé...', con todos los enredos psicológicos que le acompañan. Queda entonces solo el *saber* que poseería cualquier persona razonable en la misma situación sin asomo alguno de duda (UG, 325). O quien todavía las tenga, habrá de ser tomado como quién se sale del mundo o no es capaz de aprender a jugar (UG, 329).

La segunda ventaja reside en que, al tiempo que pertenece a la lógica de nuestras investigaciones científicas el que *en la práctica* ciertas cosas no puedan ser puestas en duda, *podemos* investigarlo todo. Por paradójico que a primera vista parezca, lo segundo es posible por lo primero: «Si quiero que la puerta se abra, los goznes han de estar bien firmes» (UG, 343).

La vieja referencia a los razonamientos sobre la cantidad y el número, o a los razonamientos experimentales sobre cuestiones de hecho o de existencia, lanzada por Hume con intenciones incendiarias y que se convirtiera en declaración programática del positivismo⁵⁴, se recupera ahora con carácter de fundamento en ambos lados. Quiere decir que se ha superado el dualismo del principio lógico y el principio real y que el «juego de la física es tan seguro como el de la aritmética».

Pero no olvidemos que «esto puede ser malentendido» (UG, 447). No estamos ante una observación psicológica, sino lógica. ¿Por qué, sin embargo, desconfiamos aún? ¿No será que seguimos imaginando una situación de privilegio para la lógica frente a la menesterosidad de lo empírico? ¿Tiene la lógica un origen diferente a lo demás?

⁵³ Cfr: PhU, I, 92-93, 134-138, 317, 522, 527, etc. Véase: LEYVRAZ, J.P.: «Le langage quotidienne comme fondement de notre liaison au monde dans la pensée de Wittgenstein» *Stud. Philos.*, 1981, (40) 17-25.

⁵⁴ HUME, D.: *Enquiry*, XII, III.

La génesis de la lógica hemos de verla en el contexto de la necesidad de supervivencia del ser humano. Pertenece a la situación de un ser en estado primitivo y no ha surgido de un razonamiento (UG, 475). No solo no tiene origen diferente a la experiencia, sino que nace al calor de los tanteos de ésta. ¿Cómo podría, pues, tener estatuto diferente?. ¿Qué sentido cabría para la distinción entre proposiciones lógicas y proposiciones empíricas?.

Pero si descubrimos esto, ¿podemos seguir pretendiendo describir la lógica? (UG, 501). Y si lo hacemos, ¿sirve para algo?. No se puede aprender la lógica al margen de la vida y describir sus operaciones no nos hace más eficaces. Tampoco podemos seguir pretendiendo que haya entrado definitivamente en el seguro camino de la ciencia mientras otras partes seguirían llamando a la puerta. Hay aquí una advertencia más contra la manía del pensamiento abstracto que induce al malentendido. Y seguro que no andamos descaminados si entendemos que es la «escisión» entre lógica y experiencia la que sostiene el empeño clarificador de las últimas obras de Wittgenstein, justamente por haber caído también en el hechizo de la superación.

Por lo que se ve, el problema de la lógica y la experiencia viene a ser el problema del fundamento. Sabemos que «algo se nos debe enseñar como fundamento» (UG, 449). Es esta una cuestión que está a la base de la trayectoria de la Teoría del conocimiento clásica; como el que su abandono representa que los problemas del conocimiento se sitúen a la misma altura que otros problemas de índole científica que le son próximos, tales como los de la lógica o la psicología⁵⁵.

Pero el abandono de aquél enfoque ¿significa sin más el abandono de la preocupación?. Wittgenstein dice que algo se nos debe *enseñar como* fundamento. Significa: Primero, que puede que no lo sea; y, segundo, que para nosotros funciona como tal. Parece como si de nuevo dejáramos abierta la puerta a la separación entre fundamento real y fundamento epistemológico, por mucho que el nombre de fundamento real nos suene ya a cosa pasada. Siendo ello así y habiéndose trasladado -como la corriente del río traslada sus orillas- el asiento del ficcionalismo del territorio de las ideas de la razón a los dominios del entendimiento científico (y común), parece que todas nuestras convicciones quedan suspendidas en el aire; o, lo que viene a ser lo mismo, sostenidas por el problemático soporte de una decisión. (En todo caso, una decisión consecuente con la lógica de la experiencia).

De inmediato se nos ocurre preguntar si no sería mejor escapar del hechizo del 'fundamento'. Pero entonces, ¿por dónde podríamos empezar?.

Por muchas veces que se diga, nos sigue sonando a extraño que el lenguaje contenga todas nuestras posibilidades y todas nuestras trampas; juntas más íntimamente que el trigo y la cizaña. No podemos desterrar la palabra 'fundamento' porque en ella se resuelve el problema del 'comienzo'. Necesitamos algo por lo que empezar y nos atormenta -al menos desde Aristóteles- la idea de que haya que ir aún más atrás: «Es difícil comenzar por el principio. Y no pretender retroceder más allá» (UG, 471).

Podemos con cuanto precede entender, sin embargo, el por qué de esa inquietud y la necesidad de aquella resolución. Lo primero, porque cuanto ha solidificado en el lenguaje

⁵⁵ Cfr: QUINE, W von O.: *La relatividad ontológica*, 3: «Naturalización de la epistemología». Madrid, Tecnos, 1974 (Trad: M. Garrido y J.Ll. Blasco). pp. 93-120.

siempre puede fluidificar; lo segundo porque, aún bajo sospecha, la vida sigue funcionando. Es de nuevo la metáfora del barco que se repara mientras navega y no se hunde; acaso también la forma definitiva de salir al paso de tanta filosofía empeñada en la institucionalización de la «sospecha». De tener algo que ver con esto, sería el complemento aún no alcanzado hasta el presente de la conjura cartesiana contra la locura.

Wittgenstein es consciente de que el barco no se puede varar en el dique hasta haberlo restaurado, como lo será Quine de que «no podemos eliminar los aderezos conceptuales sentencia tras sentencia y dejar una descripción del mundo objetivo»⁵⁶. Es también lo que desde otro contexto denunciara Hegel como inseparabilidad de método y contenido⁵⁷.

Todo ello va contra la ingenuamente desmedida pretensión, la misma de la que arranca la concepción fundamentalista, de «limpiar el terreno» de prejuicios y prevenciones. «Cuando el niño aprende el lenguaje, aprende al mismo tiempo lo que es preciso investigar y no investigar... De la misma manera que al escribir, aprendemos un determinado modelo de letras que luego cambiamos, aprendemos en primer lugar que las cosas permanecen, como una norma que luego se somete a variaciones. Este juego se acredita a sí mismo» (UG, 472, 473, 474).

Podemos intentar seguir mirando por detrás; o salirnos del mundo. No será sino la manifestación de la dificultad de encontrar el *comienzo* de una vez por todas. Como podemos preguntarnos no solo por la causa, sino también por la razón de que juguemos al juego. La diferencia que se nos ofrece está en que Descartes no puede descansar hasta que encuentre esa razón (fundamento), mientras que para Wittgenstein se trata de una pregunta con la ambigüedad de detectar el juego y no poder ser satisfecha.

Vamos de un planteamiento fundamentalista de la Teoría del Conocimiento al reconocimiento de la existencia de un fundamento que como tal nos es enseñado⁵⁸. Falta que intentemos saber: De una parte, qué diríamos si hubiéramos de cambiar de opinión en las cosas fundamentales; de otra parte, qué prueba tenemos de aquellas proposiciones que son fundamentales; por último, cómo podemos estar seguros de que no volvemos a caer en las redes del subjetivismo.

Lo primero nos lleva a pensar en una posibilidad límite; la del cataclismo. Es lo mismo que si el río perdiera un día el lecho, el agua ya no podría correr. Como tal posibilidad, no hay nada que decir, porque no hay que cambiar la opinión sobre lo fundamental; justamente porque es fundamental. El lenguaje habría perdido todas las claves de referencia y lo que dijéramos sería tan chocante como que todas las cosas del mundo se volvieran del revés. Pero es que la misma fantasía del cataclismo sólo se entiende desde la vigencia del orden del mundo (del lenguaje). Por tanto, el mundo está en orden y, si lo dejara de estar, ya en nada se podría confiar; la vida ya no sería posible. Wittgenstein no puede por menos de haber pensado en la paradoja del pirrónico (lo de la incoherencia o falta de responsabilidad que con más frecuencia se ha dicho es menos grave): quien llega a dudar de lo fundamental

⁵⁶ QUINE, W. von O.: *Palabra y objeto*. Barcelona, 1968, p. 19. Lo que Quine sigue diciendo difícilmente podría ser compartido por Wittgenstein.

⁵⁷ HEGEL, G.W.Fr.: *Phänomenologie...*, Vorrede (ed. Moldenhauer-Michel, III, p. 47).

⁵⁸ Cfr: TARCA, L.: *Il linguaggio sub specie aeterni. La filosofia di Wittgenstein come attività razionale ed esperienza mistica*. Padova, 1986.

encuentra razones para dudar de las razones de aquella duda; por lo tanto, lo mejor que puede hacer es mantener las antiguas creencias (UG, 512-516). Así Pirrón, conociendo la debilidad de las razones que argumentan a partir de las impresiones, viene a recalcar en el mejor criterio de lo que manda la naturaleza. Pirrón no sabía lo que esconde esa oscura «naturaleza»; y la filosofía que se propone investigarlo está al servicio de la mejor posibilidad de vivir, no de hacerlo imposible; aunque sabemos que no siempre presta esa atención.

Lo segundo atañe al alcance de la arbitrariedad. ¿Por qué otorgamos carácter gramatical a proposiciones de forma empírica?. Una concesión de tal carácter por parte de los sujetos individuales carecería de valor para ellos mismos. En su propio contexto Kant había apuntado la necesidad de buscar el camino que va de la «persuasión» (*Überredung*) a la «convicción» (*Überzeugung*)⁵⁹. Es necesario que conste la validez para todos los sujetos sobre la base de la confirmación debida. La primera condición rompe con los límites estrechos de las propias creencias y, al tiempo, con las inseguridades que la aqueja: yo «sé con seguridad que tal y cual» y sé que todos los demás lo saben. Pero, además, yo lo sé porque así lo he aprendido. Ese es el soporte de nuestras afirmaciones categóricas. Cualquier duda al respecto chocaría con el grueso de nuestras confirmaciones, pero no por otra razón que porque en esa dirección van precisamente las confirmaciones. De tales fundamentos está construida nuestra concepción de la naturaleza, de tal manera que el cambio de convicciones traería aparejado el cambio en la concepción del mundo. ¿Diremos todavía que hay que sospechar? ¿Y para qué?.

La condición de confirmación debida parece exigir el rigor de una experimentación controlada; otro proceder sería poco serio. Pero, «la ardilla no llega por medio de la inducción a la conclusión de que el próximo invierno también va a necesitar reservas. Tampoco nosotros necesitamos un principio inductivo para justificar nuestras acciones y predicciones» (UG, 287). Hemos recibido una herencia ancestral que se nos ha adherido como si de esquemas intuitivos se tratara y no tenemos razón alguna para dudar de ella. Gracias a que llegamos a esa claridad podemos quedarnos tranquilos en los dominios de la «creencia», porque la misma distinción entre ‘saber’ y ‘creer’ se ha desplazado de donde la situó la Teoría del Conocimiento clásica. De nuevo cambiando de contexto habríamos de decir que «el opinar es demasiado poco, el saber es excesivo»⁶⁰. No podemos ya superar a Hume (UG, 414), aunque sepamos explicarlo mejor: «todo eso ¿lo se?. Lo creo... Estoy firmemente convencido de que los demás creen, creen que saben, que todas esas cosas son así» (UG, 288 y 289; cfr: UG, 289-297).

Podemos concluir que cuando hablamos de la «fundamentación empírica» mentamos un arraigo en la experiencia que no depende de la confirmación de la expectativa en cada caso, sino que se debe a que «formamos parte de una comunidad unida por la ciencia y la educación» (UG, 298).

Ahora ya estamos en condiciones de responder a quien pregunte por cómo escapar del subjetivismo. Lo inmediato es que no es un «yo» quien sabe. Lo siguiente es que ‘saber’ se

⁵⁹ KANT, I.: KrV, A 820-21/B 848-49.

⁶⁰ KANT, I.: KrV, A 823/B 851.

ajusta al juego de lenguaje, incluso es bueno imaginar un juego en el que no exista nuestro concepto de saber. Además, distinguimos el alcance de expresiones como 'sé que tengo dolor' y 'estoy seguro de que él...'. Mientras que la última nos da la seguridad subjetiva, la primera sirve para marcar la diferencia entre quien lo sabe y quien no lo sabe por falta de propia experiencia. Por último reconocemos la existencia de un dominio común -el del cálculo- en el que 'sé...' es siempre una prueba (UG, 560-563)⁶¹.

CONCLUSION

La perplejidad acecha a quien se asoma a la obra de Wittgenstein, le acosa en la lectura y se vuelve insoportable cuando uno piensa haber tocado fondo. Me temo que ello ocurre por la condición misma del filosofar.

Por otra parte, una obra como *Sobre la certeza*, tan interminada como interminable, tiene la dificultad suplementaria de tener que terminar de pensar en el orden de una articulación suficiente las sugerencias sueltas sobre problemas o aspectos de problemas que muestran, en reiteración de ocasiones, variantes diferentes que se dejan a poca costa entender como meras repeticiones.

Wittgenstein ha resituado, expoleando por las tesis de Moore, los problemas de la Teoría del Conocimiento en el lenguaje y sus juegos, con intención de aclarar qué pasa con la 'certeza'.

Quién actúa desde dentro de un juego se encuentra de súbito con las condiciones que éste impone, que se corresponden con la aclaración de cuestiones como las del 'fundamento', la relación de 'lógica' y 'experiencia' y la 'verdad'. Solo que no se alcanzan tan prometedores resultados sin haber pasado por la cruz del análisis conceptual obligado a conjurar la magia de palabras como: 'saber' y 'seguridad', 'creencia' y 'certeza', 'duda' y 'demostración', 'error' y 'perturbación'. Lo más representativo es que haya que hacerlo a base de devolverlas su carácter mítico, con lo que, a primera vista, parece que retornáramos a la situación epistemológica primigenia. El sentido de la vuelta tiene el de reconocimiento de aquello en lo que siempre se estuvo y que fué olvidado.

La trama interna de esa advertencia acusa directamente a la filosofía por haber multiplicado el embrujamiento del lenguaje en lugar de conjurarlo; de ese modo se ha vuelto abstrusa y, como tal, debe desaparecer para regenerarse en la forma purificadora de toda suerte de hechizos. Muerte y regeneración han de ocurrir a la par, y las notas *sobre la certeza* son de ello buena prueba⁶². De todos los rostros de la filosofía sabemos que nos toca atender al del conocimiento, porque de él se originan los demás.

Pero el afán de conocer no es sino una necesidad vital que ha de ser satisfecha. Conocer es una exigencia pragmática de la naturaleza que no se basta con que nos vayan saliendo las cosas bien, sino que eso sólo puede ocurrir cuando discurra sobre un fondo de seguridad

⁶¹ Cfr. VOLTOLINI, A.: «Dal solipsismo come verità ineffabile al truismi come certezza inespressa», *An. Scuola Sup. di Pisa*, 1985 (15) 1035-177.

⁶² Al respecto véanse: KAHL, St. A.: «Wittgenstein ordinary language an the critique of metaphysics», en *Die Aufgaben der Philosophie in der Gegenwart* (Akten des 10. Int. Wittg. Symposiums 18-25 August 1985). Wien, Hölder-Pichler-Tempsky, 1986, p. 297-300. (una vez más tienen interés la generalidad de las comunicaciones). COOK, J.W.: «The Metaphysics of Winttgenstein's On Certainty» *Philos. Invest.*; 1985 (8) 81-119.

inamovible. A falta de una dotación ciega de la naturaleza, en el curso del tiempo hemos construido, guiados por la experiencia, los canales por los cuales fluyen ordenados los acontecimientos y, como en el curso del río, podemos prever a dónde desembocan.

Eso es lo que podemos llamar una «concepción del mundo», que establece tanto las reglas de la lógica como los márgenes de la racionalidad y lo que sobre ésta se construye. El núcleo de ésta presentación de Wittgenstein reside en la transformación de la gramática de las proposiciones que ya había desarrollado en las *Investigaciones filosóficas*.

La Teoría del Conocimiento se nos ha dado la vuelta. En el texto de Wittgenstein hay resonancias constantes de los creadores de la teoría clásica; a lo largo de este estudio hemos hecho aparecer algunas de ellas como muestra. La naturaleza de esa inversión nos deja por fuerza la impresión de una seguridad asentada sobre la inseguridad de una decisión. Ahí parece residir su mayor fuerza, donde más transparece su debilidad. Tanto más cuanto que se evitan los enredos psicologistas y la amenaza subjetivista queda superada en la pertenencia de todos a una misma ciencia y educación. Otro tanto habría que decir del insuperable escepticismo o de la alternancia realismo/idealismo⁶³.

Sabemos, en fin, que nuestros conocimientos acerca de nosotros mismos o del mundo externo son objetivamente ciertos porque comúnmente aceptados por la fuerza constrictiva de la lógica de la experiencia. El problema del conocimiento no es anterior, sino posterior a esas certezas, como lo es el problema de la razón.

Sólo que al final de cada recodo nos asaltan las inquietudes de siempre. Por mucho que se nos recomiende empeñar todo el esfuerzo filosófico en espantar esa suerte de fantasmas. A Wittgenstein mismo le acosan de tal forma que su obra se enreda y retuerce de forma interminable. Sus claves, sin embargo, están claras.

Esto es lo que después del recorrido necesitado de tanto esfuerzo -no es más fácil el segundo Wittgenstein que el primero- nos da cierta tranquilidad y satisfacción; porque «aunque solo raramente hayamos dado en el blanco, podemos haber reconocido los objetivos que siempre ha tenido presentes» (UG, 387).

Angel Alvarez Gómez.
Septiembre, 1989

⁶³ Por muchos caminos se hace presente la influencia de Wittgenstein; cfr: *Wittgenstein and contemporary Philosophy* (ed.: Br. F. McGuinness y A. Gargani) Pisa, ETS, 1985.

LA DIFERENCIA ENTRE EXPLICAR E INTERPRETAR, SEGUN DAVIDSON*

Manuel García-Carpintero

Departamento de Lógica, Historia y Filosofía de la Ciencia
Universidad de Barcelona

El tema de estas jornadas es la relación entre la filosofía analítica y otras tradiciones filosóficas contemporáneas. Lo escaso de mis conocimientos sobre las aportaciones de esas otras tradiciones filosóficas del presente hace que me sea imposible contribuir a ellas con una comparación bien elaborada entre el modo en que se abordan los problemas filosóficos -o algunos problemas filosóficos concretos- en la tradición analítica y el modo en que se abordan en otros ámbitos del pensamiento filosófico contemporáneo. Me ha parecido una alternativa adecuada examinar una propuesta filosófica analítica que, por lo que sé, resulta ser una solución similar a la ofrecida por otras «escuelas» a problemas también similares. El conocedor de esas otras tradiciones podrá así comparar el modo en que se hace filosofía en la tradición analítica con el modo en que se abordan los problemas en otros ámbitos de esta actividad intelectual.

Explicación y Compresión

Un viejo tema de disputa filosófica es la cuestión de si existen diferencias conceptuales entre las «ciencias naturales», como la física, la química o la biología, y las «ciencias humanas», como la psicología, la sociología, la historia o la lingüística. Una tradición proveniente del siglo anterior sostiene que ese es el caso; que, mientras en las ciencias naturales se ofrecen *explicaciones*, en las ciencias humanas se ofrecen *interpretaciones*; que las unas aportan *explicación*, (*Erklärung*) y las otras *compresión* (*Verständnis*). Naturalmente, por el momento sólo contamos con nombres, con un cierto poder descriptivo, de las cualidades distintivas de ambos tipos de empresas intelectuales. Necesitamos una buena explicación filosófica de la naturaleza de las supuestas diferencias.

Es cierto que los nombres apuntan a una diferencia que, en principio, indudablemente existe entre la información suministrada por la física y la información suministrada por la historia -por tomar un ejemplo-. La dinámica newtoniana nos explica el errante movimiento observable de Marte, su aparente ir y venir sobre el fondo de las «estrellas fijas». Para ello, la teoría «sitúa» el movimiento observable de Marte en el marco de un cierto modelo dinámico (un modelo que podemos visualizar fácilmente imaginando un sistema de bolas de distinto tamaño girando unas en torno a otras sobre un fondo relativamente muy lejano

* Este artículo forma parte del proyecto de investigación PB87-0834-CO3-O3- -subvencionado por la DGICYT- 'Cuantificadores generalizados, anáfora e información condicional'.

-atadas a sus circunvoluciones por una cierta fuerza que satisface los postulados de la teoría-, e imaginando cómo sería el movimiento aparente de una cierta bola visto desde la superficie de otra)¹. Por otra parte, una cierta hipótesis histórica puede explicarnos igualmente la decantación progresiva de la mayoría del pueblo alemán hacia las tesis del Partido Nacionalsocialista a lo largo de los años veinte y los primeros años treinta, en términos de las consecuencias económicas y morales del pago de las sumas impuestas por los aliados a Alemania en el Tratado de Versalles en concepto de reparaciones de daños producidos por la I Guerra Mundial. Ambas teorías nos proporcionan información, información que nos hace inteligibles los fenómenos en cuestión. Pero puede observarse una cierta diferencia: mientras que la segunda hipótesis consigue este resultado *dándoles sentido* a las acciones de ciertos sujetos, no cabe en absoluto decir lo mismo de la primera.

Hay aquí una diferencia intuitiva, ciertamente; y una diferencia similar puede observarse contrastando cualquier explicación de las «ciencias naturales» con muchas explicaciones de las «ciencias humanas». Podríamos intentar una primera formulación de la distinción: generalmente, las explicaciones de las ciencias humanas involucran el *dar sentido* a las conductas de ciertos sujetos, sujetos capaces de *actuar*. Nada similar parece ocurrir en las ciencias naturales. Y con esto podemos apreciar mejor cuál es nuestro problema, si queremos sostener que existe una diferencia conceptualmente importante entre explicación y comprensión. Debemos justificar que esta práctica de dar sentido o interpretar un acaecimiento es sustantivamente diferente a la práctica de explicar un fenómeno natural. Porque una diferencia intuitiva puede ser una diferencia tan sólo aparente, una diferencia tras la cual no haya nada sustantivo -salvo quizás algún prejuicio característico de nuestra especie-.

Existe un buen número de intentos desencaminados de teorizar la distinción intuitiva que acabamos de describir. Resulta provechoso examinar alguno de ellos.

La generalidad de las explicaciones de la «ciencia natural»

Está, en primer lugar, quien encuentra una diferencia en cuanto a la generalidad. Las explicaciones de las ciencias naturales presuponen (o contienen) leyes completamente generales, leyes no restringidas a ámbitos específicos. Sin embargo, sólo en un sentido muy laxo de 'ley' puede decirse lo mismo de las explicaciones características de las ciencias humanas. Voy por la calle y me encuentro a un sujeto con uno de sus zapatos en la mano, agitándolo frenéticamente. A primera vista, encuentro muy rara su conducta; pero pronto doy una explicación: el sujeto *pretende quitarse una molesta piedra del zapato*. Esta es una buena explicación del fenómeno; pero difícilmente puede decirse que involucra una ley de tipo 'todo sujeto que pretende quitarse una molesta piedra del zapato lo agita frenéticamente con una mano hasta que la piedra sale'. A lo sumo, involucra la ley '*ceteris paribus*, todo sujeto que pretende quitarse una molesta piedra del zapato lo agita frenéticamente con una mano hasta que la piedra sale'. Esto es, 'todo sujeto que, *en condiciones parejas a éstas*,

¹ Presupongo que en esta descripción el análisis de las teorías científicas de lo que se denomina hoy la «concepción semántica». Véase Bas Van Fraassen, *The Scientific Image*, Clarendon Press, Oxford, 1980, cap. 3; y Ronald Giere, *Explaining Science*, The University of Chicago Press, Chicago, 1988, cap. 3.

pretende quitarse una molesta piedra del zapato, lo agita frenéticamente con una mano hasta que la piedra sale'. La cláusula *ceteris paribus* contiene implícita una referencia a una situación particular; y lo peor del caso es que no podríamos eliminarla, reemplazándola por una descripción general de las situaciones en las que la ley tiene validez. (El sujeto no debe creer que sus pies despiden un olor insoportable, olor que se hará patente al quitarse el zapato, y que hay presente alguien a quien le ha de molestar tal olor; pero, en ese caso, no debe tampoco desear librarse de la presencia del molesto sujeto que hay ante él utilizando para ello el hedor insoportable que producen sus pies librados de la opresión del zapato. El sujeto no debe creer que tiene un agujero de considerables proporciones en el calcetín, agujero que será apreciado sin duda con reprobación por algún posible paseante. El sujeto no debe estar dominado por alguna urgencia corporal irrefrenable. Etc., etc., etc..)

Aunque el ejemplo del párrafo anterior no constituía una explicación de alguna de las ciencias humanas, sí ejemplifica bien una dificultad de que adolecen las explicaciones que estas disciplinas suministran, y que no encontramos (al menos en el mismo grado) en las explicaciones características de la física. Pues (como quizás se apreciará mejor más adelante) ese ejemplo, por humilde que sea, constituye un paradigma de *interpretación*, de explicación consistente en *dar sentido* a un acaecimiento -en verlo como la *acción* de un ser racional-. No hay ninguna diferencia esencial entre la explicación proporcionada en el ejemplo y la explicación del asesinato de Escobedo en términos de ciertos juicios de Felipe II; y lo que hace características a las explicaciones psicológicas, históricas, sociológicas o económicas es que presuponen explicaciones del tipo ejemplificado.

Así pues, tenemos aquí una diferencia. Pero no es una buena estrategia recurrir a ella para justificar la distinción entre *interpretar* y *explicar*, porque se trata de una diferencia gradual. Hay todo un continuo de distinciones en este respecto, y no necesariamente están agrupadas en la cola de ese continuo las explicaciones de las ciencias humanas. Las explicaciones más cercanas al ideal de máxima generalidad sugerido en el párrafo anterior son las de la física fundamental; pero ya estas mismas explicaciones, cuando no constituyen caracterizaciones de modelos ideales -en los que «no existe rozamiento»-, sino que se ofrecen como hipótesis explicativas de acaecimiento concretos, contienen cláusulas *ceteris paribus* difícilmente eliminables. (El lector sólo debe pensar en la explicación que proporciona la física de que el avión de papel, lanzado desde donde ha sido lanzado, haya ido a caer precisamente donde ha caído; o en la explicación que da la meteorología de las lluvias torrenciales caídas en los últimos días sobre el norte de España. Tales explicaciones contienen cláusulas *ceteris paribus* tan difícilmente eliminables como las del ejemplo.) Y no cabe duda alguna de que muchas de las explicaciones de la geología o la meteorología están, en el continuo antes mencionado, mucho más alejadas del ideal que algunas explicaciones psicológicas, económicas o sociológicas.

Diferencias en cuanto a la aplicación del concepto de *causa*

Una segunda propuesta filosófica busca la distinción en el papel de la relación causal en la explicación científica. Así, por ejemplo, en el contexto de una larga discusión destinada a ejemplificar las diferencias entre explicación y comprensión con ayuda del psicoanálisis, Jürgen Habermas dice lo siguiente:

Mientras en la manipulación técnica de la naturaleza hacemos trabajar en nuestro interés a la naturaleza en razón de nuestro conocimiento de las relaciones causales, la comprensión analítica aprehende la causalidad del inconsciente como tal; la terapia no se basa, como la «causal» en sentido estricto de la medicina somática, en una utilización de las conexiones causales conocidas, debe más bien su eficacia a la abolición de las conexiones causales mismas².

Aunque quizás con mi interpretación no hago justicia a las ricas connotaciones del término alemán en que Habermas hace descansar la parte fundamental de su elucidación de las diferencias entre «el sentido estricto» de la causalidad y la «causalidad del inconsciente» (*Aufhebung*), lo que parece estarse diciendo en este texto es lo siguiente: Los psicoanalistas sostienen que existe una conexión explicativa entre la represión de determinadas emociones, juicios o sentimientos, -su consiguiente alojamiento en el inconsciente-, y ciertos síntomas neuróticos, histéricos, ciertos actos fallidos o el contenido de ciertos sueños. La conexión en cuestión, según Freud, es la misma conexión causal que vincula, pongamos por caso, la no ingestión de vitamina C y el resfriado común, o la ingestión de determinadas setas y determinado tipo de envenenamiento. Pero Habermas encuentra una diferencia entre unas y otras relaciones causales: la relación causal entre la represión del pensamiento del padre visitando un prostíbulo y la parálisis neurótica del brazo se puede «disolver» o «eliminar» (haciendo consciente el pensamiento reprimido), mientras que la relación causal entre la ingestión de setas y el envenenamiento no se puede abolir de ningún modo: es un «hecho bruto».

Si eso es lo que Habermas dice, está cayendo en una importante confusión. Una cosa son las relaciones causales entre sucesos concretos, particulares e irrepetibles. Esas no se pueden «disolver», por la misma razón que no se puede cambiar el pasado. Ayer perdí la única copia que tenía de la primera versión de este artículo, y la causa fue que, por error, tecleé 'Delete' seguido del nombre y la dirección del archivo que contenía esa primera versión. Me gustaría deshacer el entuerto, pero eso no hay quien lo cambie. Del mismo modo, si la causa de que Anna tuviera ayer el brazo paralizado es el pensamiento reprimido de su padre en un prostíbulo, no hay quien pueda deshacer tan poderosa ligazón.

Cosas muy distintas son las leyes causales. Las leyes causales no vinculan acaecimientos particulares, sino propiedades o características generales. Por ejemplo, existe una ley causal que vincula los sucesos que consisten en la ingestión de setas venenosas, en general, con los sucesos que consisten en el envenenamiento del individuo que las ha ingerido; existe una ley causal que vincula los acaecimientos que consisten en la ingestión de ácido acetilsalicílico, en general, con la desaparición del dolor de cabeza en el sujeto de la ingestión; y existe una ley causal que vincula la falta de vitamina C, en general, con el resfriado común. Estas leyes tampoco se pueden disolver; son hechos, hechos entre los hechos del mundo (sólo que hechos generales, regularidades o quizás algo más importante)³. Pero su conocimiento puede

² Jürgen Habermas, *Conocimiento e interés*, pág. 269-270, Taurus, Madrid, 1982.

³ La opinión aquí expresada sobre el debatido asunto filosófico de la causalidad se justifica un poco más adelante.

ayudarnos a evitar que se produzcan ciertos sucesos (particulares), que de otro modo podrían haberse producido, o a hacer que se produzcan ciertos sucesos (igualmente particulares), que de otro modo no se hubiesen producido. Si yo, conociendo la ley causal que vincula el ácido acetilsalicílico con la desaparición del dolor de cabeza, y padeciendo un dolor de cabeza de considerables proporciones, me tomo una aspirina, puedo conseguir que desaparezca mi dolor de cabeza. Con ello no «disuelvo» la relación causal que vinculaba el suceso particular anterior consistente en la no presencia de ácido acetilsalicílico y el suceso particular anterior consistente en mi dolor de cabeza, ni menos aún llevo a cabo la «abolición» de la ley causal que vincula el ácido acetilsalicílico con la desaparición del dolor de cabeza; antes bien, gracias a que esa ley causal sigue estando en pleno vigor, produciendo la presencia de ácido acetilsalicílico en mi organismo llevo a cabo la desaparición de mi dolor de cabeza. (Una historia similar podría narrar sobre el modo en que evito mi envenenamiento por la ingestión de *Amanita Phalloides* sin «disolver» la ley causal que vincula ingerir setas de ese tipo y sufrir un cierto envenenamiento, sino apoyándome en mi conocimiento de ella.)

Pues bien, el psicoanalista (si la historia etiológica que nos cuenta es cierta) no hace otra cosa. Apoyándose en su conocimiento de la ley causal que vincula los acaecimientos consistentes en que está presente la represión del pensamiento de Anna de su padre en un prostíbulo, con los acaecimientos consistentes en la parálisis del brazo de Anna, el psicoanalista lleva a cabo la eliminación del síntoma mediante la eliminación de la represión. No por ello disuelve la relación causal singular entre el acaecimiento anterior consistente en la represión del pensamiento y el síntoma neurótico, ni menos aún la ley causal que vincula (al menos en el caso de Anna) la represión del pensamiento con la parálisis del brazo, sino que apoyándose en ella impide acaecimientos futuros consistentes en la parálisis del brazo impidiendo acaecimientos futuros consistentes en la represión del pensamiento. La única posible diferencia entre eliminar el síntoma eliminando la represión, y eliminar el dolor de cabeza eliminando la falta de ácido acetilsalicílico, está en que la ley causal involucrada parece más general en el último caso que en el anterior. Pero de esa cuestión ya habíamos hablado antes⁴.

Otros autores, éstos de la tradición analítica, han sostenido también que existe una diferencia conceptual entre explicar y comprender, y han tratado de encontrar la diferencia en el mismo lugar que Habermas: en una diferencia en cuanto al sentido de las locuciones causales cuando se predicán de sucesos naturales y cuando involucran acaecimientos mentales, tales como juicios, pensamientos, intenciones, propósitos, etc., cosa ésta que característicamente sucede en las explicaciones del tipo que venimos llamando hasta aquí 'interpretación' o 'compresión'. Es característico de una interpretación que se diga que un determinado acaecimiento «natural» (el que un organismo agite frenéticamente un zapato) ha sido producido por «estados mentales», estados del tipo de los antes mencionados, de algún individuo. G.E.M. Anscombe y G.H. Von Wright, por ejemplo, han sostenido que

⁴No presupongo en el párrafo anterior que existan leyes causales del tipo de las que los psicoanalistas postulan. Cf. el excelente estudio crítico de Adolf Grünbaum, *The Foundations of Psychoanalysis*, University of California Press, Berkeley, 1984. La crítica a Habermas está tomada de ese mismo trabajo.

existen diferencias importantes entre el «producir» que puede darse entre, pongamos por caso, un propósito concreto (el propósito de hacer salir la piedra que tengo en el zapato agiténdolo frenéticamente) y, digamos, los movimietos corporales resultantes (el frenético movimiento de mi mano empuñando el zapato) y el producir que puede darse entre dos acaecimientos del mundo físico concretos (la caída del rayo y la destrucción del roble). Diferencias conceptuales tan importantes, como para reservar el término 'causa' para estas últimas producciones y evitarlo en las primeras.

Anscombe sostiene que, mientras que en el caso físico, causa y efecto son «existencias distintas», no ocurre lo mismo en el caso, digamos, «mental»⁵. Pero es parte fundamental de nuestro concepto de causa que causa y efecto deben ser acaecimientos diferentes. Von Wright sostiene que, mientras en el caso físico es posible encontrar evidencia de que se ha dado el efecto independiente del darse de la causa, y viceversa, no ocurre lo mismo en el caso «mental»⁶. Pero es parte fundamental de nuestro concepto de causa que pueda hallarse evidencia de que se ha dado la causa independientemente de la evidencia que muestre que se ha dado el efecto, y viceversa.

Sin duda los argumentos bosquejados habrán producido al lector no familiarizado con ellos una impresión de paradoja. Por un lado, parece que «algo» hay en ellos; por otro, parece que la aparente verdad de sus premisas se apoya en algo que no acaba de estar claro. A este respecto, Davidson ha contribuido a iluminar enormemente estos problemas, y los problemas conceptuales relacionados con ellos (que, puede decirse, son lo problemas fundamentales de la filosofía). La obra de Donald Davidson (junto a la de W.V.O. Quine, en la que en buena medida está inspirada) puede situarse, en el marco del conjunto de temas de que me ocupo en este trabajo, en la misma línea de Habermas, Anscombe y Von Wright. Davidson opina que existe una diferencia conceptual importante entre las explicaciones de la ciencia natural y las explicaciones de la psicología (en general, de las disciplinas que usan entre los elementos de su batería conceptual nociones tales como *creencia*, *deseo*, *intención*, *pensamiento*, *propósito*, etc.); y la diferencia concierne al papel de la casualidad en esas explicaciones.

Davidson sostiene, en contra de Anscombe o Von Wright, que cuando decimos que una intención concreta y particular produjo un determinado acaecimiento queremos decir exactamente lo mismo que cuando decimos que el golpe de una determinada bola de billar contra otra produjo el movimiento de la segunda. Interpretar una acción particular no se diferencia en nada de explicar un acaecimiento físico no intencionalmente producido, como la caída de un rayo: en ambos casos, se trata de colocar el suceso en la trama de relaciones

⁵ La muerte de Sócrates no puede causar el enviudamiento de Xantipa, por cuanto muerte y enviudamiento son un mismo suceso. De acuerdo con Anscombe, mi intención de elevar mi brazo izquierdo no puede ser causa de mi elevación del brazo izquierdo, porque uno y otro son un mismo acaecimiento. Cf. G. E. M. Anscombe, *Intention*, Cornell University Press, Ithaca, 1976.

⁶ Según Von Wright, aunque mi intención de elevar el brazo izquierdo y mi elevación del brazo izquierdo sean sucesos diferentes, la única evidencia posible de que mi elevación del brazo izquierdo sea «intencional», es decir, no sea un movimiento reflejo o una conducta hipnóticamente teledirigida, sino una verdadera *acción* mía, está en que yo tenga efectivamente la intención de elevar el brazo izquierdo; y la única evidencia posible de que tengo tal intención está en la elevación de mi brazo izquierdo. Cf. G. H. Von Wright, *Explicación y comprensión*, Alianza, Madrid, 1979.

causales entre eventos particulares que constituye, en palabras de Hume, «el cemento del universo».

No sólo hay argumentos directos que muestran que ello es así, y razones para descalificar los argumentos de Anscombe y Von Wright⁷, sino que aceptándolo podemos también comprender cabalmente otros muchos conceptos que fácilmente producen paradojas y perplejidades; por ejemplo, los conceptos de *incontinencia*, o *debilidad de la voluntad*, y en general los conceptos aplicables a acciones y estados que caen bajo el epígrafe genérico de la *irracionalidad*: *pensamiento desiderativo*, *autoengaño*, etc..

Pues bien, pese a ello, Davidson (y Quine independientemente) sostiene que existe una diferencia sustantiva entre las explicaciones que involucran causalmente lo mental (las interpretaciones) y las que no lo hacen. La diferencia no atañe a la explicación causal de acaecimientos particulares, sino a la naturaleza de las leyes causales presupuestas. En cierto sentido que habremos de estudiar, no hay, según Davidson, leyes causales que involucren lo mental. Y el problema no deriva, al menos directamente, de la escasa generalidad de las presuntas leyes causales de la psicología o de la historia (aunque su escasa generalidad está relacionada con la cuestión). El problema es mucho más profundo (es «de naturaleza conceptual»).

Yo mismo no creo en la validez de las tesis últimas de Davidson; de modo general, no creo que exista una diferencia sustantiva entre *explicación* y *compresión*. Sin embargo, estoy seguro de dos cosas. Una es que las consideraciones de Davidson han clarificado enormemente todos estos problemas; por decirlo con alguna grandilocuencia, nos han ayudado a ubicar el lugar de la razón en un mundo de acaecimientos causalmente interrelacionados. Otra, que los argumentos de Davidson indican con claridad el precio que hay que estar dispuesto a pagar para negar que exista alguna diferencia conceptual entre la psicología y la biología, y también que es una cuestión todavía abierta que ese precio pueda pagarse. Lo que persigo con este trabajo es transmitir algo de la clarificación proporcionada por la obra de Davidson, y a la vez establecer una cierta tesis crítica. Una tesis que será, pues, condicional: si se dan ciertas circunstancias (circunstancias que han de satisfacer exigentes requisitos, por lo que no sólo es un hecho contingente que se den, sino que ni siquiera es obvio que hayan de darse), entonces no hay diferencia conceptual importante entre, pongamos por caso, la psicología y la biología o la geología.

El argumento de Davidson depende de una clarificación previa de ciertas nociones, fundamentalmente las nociones necesarias para construir una buena teoría filosófica de las relaciones causales. Es preciso explicar qué son las relaciones causales entre acaecimientos singulares, y dar cuenta de la naturaleza de las entidades relacionadas por ellas, lo que hasta

⁷ La exposición más satisfactoria de ambas cosas es el artículo del propio Davidson, 'Actions, Reasons and Causes', recopilado (como la mayoría de los artículos que mencionaré después) en *Essays on Actions and Events*, Clarendon Press, Oxford, rep., 1982. Este artículo contribuyó decisivamente a cambiar las opiniones de la mayoría de los filósofos analíticos sobre la cuestión que aquí discutimos. Se puede comprobar que, mientras en los escritos anteriores a 'Actions, Reasons and Causes' -en buena medida por influencia del segundo Wittgenstein- se da por supuesto que la «carga de la prueba» está en quien piensa que la conexión entre intenciones y acciones es la causal, los artículos posteriores dan por supuesto justamente lo contrario.

aquí he venido denominando ora «eventos», ora ‘acaecimientos’, ora ‘sucesos’. Es preciso elucidar la naturaleza de las leyes causales. Es preciso enumerar algunos hechos fundamentales sobre las características semánticas de los enunciados en que establecemos relaciones causales. Y es preciso explicar qué establece la diferencia (si algo la establece) entre los acaecimientos mentales y los «puramente físicos». Por tanto, buena parte de las páginas que siguen están dedicadas a estas cuestiones.

No voy a seguir siempre al propio Davidson a la hora de introducir esas nociones; la razón es que no pretendo sólo exponer sus ideas, sino explicarlas críticamente, y mis diferencias con Davidson en cuanto a la cuestión central de este trabajo dependen de mis diferencias con él en cuanto al análisis de las nociones antes mencionadas. Sin embargo, si no en los argumentos ni en el marco conceptual (ontológico) presupuesto, la discusión sobre la causalidad y la mente coincidirá en las conclusiones más importantes con Davidson.

Así, sostendré que en su sentido epistémicamente primario, la relación causal se da entre acaecimientos⁸, y que los acaecimientos son entidades particulares que nuestra ontología debe suponer⁹. Sostendré que lo que identifica a los acaecimientos son los procesos causales en los que intervienen¹⁰. Mantendré el carácter nómico de la causalidad, esto es, que las relaciones causales entre acaecimientos particulares presuponen leyes de dependencia causal entre propiedades¹¹. Sostendré que los enunciados en que se aseveran relaciones causales entre particulares son extensionales, mientras que las leyes causales presentan un cierto tipo de intensionalidad (en el sentido tradicional del término); y sostendré que unos y otros enunciados poseen un alto grado de «intensionalidad», de origen no semántico sino pragmático¹². Por último, sostendré que una acción es un acaecimiento causado por una combinación de estados mentales, estados mentales que constituyen las razones del agente para producir el acaecimiento¹³. Este elevado grado de coincidencia en las conclusiones es lo que me permite representarme, en el grueso de este trabajo, como meramente transmitiendo, interpretada, la comprensión que Davidson nos ha proporcionado de estos temas.

Causalidad singular, causalidad general y leyes causales

En su sentido más básico, la relación causal se da entre acaecimientos. La batalla de Los Arapiles y el vuelo en *Dodge* de Carrero son acaecimientos, pero también lo son la segunda proyección de ‘Avanti!’ en el cine Verdi ayer y el embotellamiento que padecí en la

⁸ Davidson sostuvo esto en ‘Causal Relations’, en la recopilación antes citada *Essays on Actions and Events*.

⁹ Cf. ‘The Logical Form of Action Sentences’ y ‘Events as Particulars’, recopilados también en la obra citada antes.

¹⁰ Cf. ‘Criticism’, Comment and Defence a ‘The Logical Form...’ y ‘The Individuation of Events’, *ibíd.*, y ‘Reply to Quine on Events’, en E. Lepore y B. Maclaughlin (eds), *Perspectives on the Philosophy of Donald Davidson: Actions and Events*, Basil Blackwell, 1985, donde abandona su posición ante los sutiles argumentos de Quine en ‘Events and Reification’ en el mismo volumen. (En mi opinión, es preciso resistirse a aceptar las conclusiones de esos argumentos).

¹¹ Cf. ‘Mental Events’, en la recopilación ‘Actions and Events’.

¹² La intensionalidad, o «relatividad a una descripción», de las leyes causales es un tema constante en los escritos de Davidson. Cf., por ejemplo, ‘Causal Relations’.

¹³ Esta es la tesis defendida por Davidson en su citado clásico ‘Actions, Reasons and Causes’; cf. también ‘Agency’ y ‘Intending’, los tres en la recopilación antes citada.

autopista B-30 el dieciocho de diciembre de 1989 durante cuatro horas. Un acaecimiento es algo que ocurre en algún lugar durante algún tiempo. Empleamos la palabra 'acaecimiento' para acaecimientos especialmente relevantes, lo que hace que mi uso resulte un tanto extraño. Pero cualquier otra palabra que pudiera parecer adecuada en principio ('evento', 'suceso'.) plantea problemas similares; estipulo, pues, este uso particular. Los acaecimientos, como los objetos físicos, son particulares localmente ubicados, irrepetibles y contingentes. Un acaecimiento consiste en la ejemplificación de una propiedad por un particular, o en la ejemplificación de una relación n -ádica por n particulares, o en una compleja suma de relaciones como estas. Los acaecimientos están relacionados entre sí, por relaciones espaciotemporales (y por relaciones causales). Dos acaecimientos pueden ser simultáneos, o pueden solaparse parcialmente en el tiempo, o uno de los dos puede preceder completamente al otro en el tiempo.

En mi uso de la expresión, no hay acaecimientos posibles (el embotellamiento que podría haber habido, aunque no lo hubo, ayer en la A-17), como tampoco hay objetos físicos posibles (el hijo de James Dean). Pero el mismo lenguaje (el mismo pensamiento) que utilizamos para referirnos a los particulares y a los acaecimientos, para describirlos o para expresar hechos sobre ellos, en virtud de las reglas semánticas que nos permiten hacer todas esas cosas, caracteriza necesariamente diferentes mundos o circunstancias posibles, algunos reales y otros diferentes al real. Hablando laxamente, podemos decir que los mundos o circunstancias diferentes al real están «habitados» por «otros» particulares y acaecimientos (si es que este modo de hablar no ha de confundirnos).

Aunque los acaecimientos son particulares, en ocasiones nos referimos de un modo general a los acaecimientos que comparten ciertos rasgos; decimos, por ejemplo que los particulares involucrados (aún siendo diferentes) tienen una cierta propiedad, o están en una cierta relación. Hablamos de que ayer hubo un embotellamiento, o de la cuarta *proyección de 'Avanti!' en el Verdi* ayer. Me referiré a estos acaecimientos generales con las expresiones 'tipos de acaecimiento' y 'acaecimiento-tipo'.

Sean 'c' y 'e' variables sustituibles por nombres habituales de acaecimientos (o variables cuyos valores son acaecimientos). ¿Qué podría servir como análisis de «c causa e»? De acuerdo con algunos filósofos (David Lewis y John Mackie son los más destacados), la causalidad singular es primitiva, y la verdad de una aserción causal singular no implica la verdad de ninguna aserción causal general -aserciones que expresan relaciones causales entre tipos de acaecimiento, no entre acaecimientos singulares-. Cuando decimos que c causó e queremos decir que, si c no hubiera ocurrido, no habría ocurrido e¹⁴. (Un enunciado de este tipo es un «contrafáctico», un enunciado en que describimos lo que hubiera ocurrido si los hechos hubiesen sido otros de los que en realidad fueron.) Ulteriormente, Lewis y Mackie (por caminos diferentes) tratan de explicar el contenido de un contrafáctico de ese tipo sin apelar a regularidades, es decir, sin presuponer la existencia de relaciones causales generales. (En mi opinión, Mackie tiene en esto mucho menos éxito que Lewis. El camino

¹⁴ Cf. David Lewis: 'Causation', 'Counterfactuals and Comparative Possibility' y 'Events', con sus correspondientes «postscripts», en *Philosophical Papers*, vol. 2, Oxford University Press, Oxford, 1986; y John Mackie: *The Cement of the Universe*, Clarendon Press, Oxford, Pap. rep., 1980, cap. 2.

elegido por este último consiste en tomar como primitiva la noción de *mundo posible* y la relación de ser *x un mundo posible más similar al mundo a que el mundo y*. A grandes rasgos, «si *c* no hubiera ocurrido, no habría ocurrido *e*», en el análisis de Lewis, es verdadero si y solamente si, para cualquier mundo posible en el que *c* no se da y sí se da *e*, hay un mundo posible más similar al mundo real, en el que se da *c* y también se da *e*.)

Por otro lado, de acuerdo con otro grupo de filósofos, cuando decimos que un acaecimiento es causa de otro, presuponemos que el acaecimiento causa, *c*, es un caso particular de un tipo de acaecimiento, *C*; el acaecimiento efecto, *e*, es un caso particular de un tipo de acaecimientos, *E*; y que los acaecimientos de tipo *E* dependen causalmente de los acaecimientos de tipo *C*. Es decir, las relaciones de dependencia causal entre acaecimientos particulares presuponen lógicamente relaciones de dependencia causal entre tipos de acaecimientos. Hay relaciones causales entre acaecimientos singulares porque hay relaciones causales entre tipos de acaecimientos que los acaecimientos singulares ejemplifican; y el 'porque' anterior es el de la consecuencia lógica (o analítica, si se prefiere).

Este segundo grupo de filósofos, aquellos que piensan que una aserción causal singular implica la verdad de una aserción causal general, se dividen a su vez en dos grupos, digamos los «humeanos» y los «aristotélicos». Todo depende de la respuesta que se dé a esta pregunta: ¿En qué consiste la relación de dependencia causal entre tipos de acaecimientos? ¿Cuál es la semántica de una aserción causal general? Esta relación debe consistir, como mínimo, en una regularidad (cuando se da *C*, típicamente se da *E*). Pero debe tratarse de una regularidad compleja, porque de otro modo habríamos de retirar el carácter de leyes causales a lo que los científicos consideran haber establecido como tales. Así, las regularidades causales no necesariamente deben consistir en que siempre que se dé *C* se dé *E* -en el que se dé un acaecimiento de tipo *C* sea suficiente para que se dé un acaecimiento de tipo *E*. (Que a un ser humano le corten la cabeza es causa de que muera, y es suficiente para ello también; pero fumar es causa del cáncer de pulmón, aunque no es suficiente que alguien fume para que contraiga cáncer de pulmón). Y el mismo ejemplo muestra que una ley causal no debe implicar que el acaecimiento-tipo *C* sea necesario para el acaecimiento-tipo *E*; que nunca se dé *E* si no se da *C*: no es necesario que alguien haya fumado para que desarrolle un cáncer de pulmón.

La mayoría de las formulaciones de leyes causales ordinarias contienen implícitas cláusulas *ceteris paribus*. Un factor causal del cáncer de pulmón de Juan es el que durante cincuenta años hubiera fumado más de veinte cigarrillos diarios. No porque todos los que fumen así desarrollen cánceres de pulmón, ni porque sólo desarrollen cáncer de pulmón los que fuman así; sino porque fumar así, *en condiciones parejas en lo restante* a aquellas que se dan en el caso de Juan, basta para desarrollar un cáncer de pulmón. Y al decir esto, sabemos que no somos capaces de sustituir la cláusula *ceteris paribus* por una especificación general precisa de esas condiciones.

Así pues, si *C* causa *E*, no es preciso que *C* sea suficiente para *E* ni que *C* sea necesario para *E*. Pero, si fumar es causa del cáncer de pulmón, entonces fumar, junto con otros factores posiblemente desconocidos (digamos *x*)-presentes en el caso de Juan- es suficiente para contraer un cáncer de pulmón. Y esos factores, por sí solos, no son suficientes para producir el cáncer de pulmón. De modo que *C* es causa de *E* sólo si *C* es parte necesaria de

una condición suficiente para E; sólo si siempre que se da C en circunstancias x se da también E; Y nunca se da E cuando sólo se dan las circunstancias X y no se da C (abreviemos esto así: sólo si C es una condición NS de E)¹⁵.

Una aserción causal general es, como mínimo, una regularidad NS del tipo descrito en el párrafo anterior. Pero no puede ser sólo eso. Las regularidades NS son normalmente reversibles (fumar, en circunstancias X, basta para el cáncer de pulmón; pero también basta el cáncer de pulmón, en circunstancias Y, para ser fumador), mientras que las leyes causales son asimétricas. Las aserciones causales generales poseen dos marcas características que no acompañan a muchas de las regularidades NS: sustentan contrafácticos y se ven confirmadas por sus casos particulares. (Si yo hubiese cuidado a un leproso -cosa que me he guardado de hacer- en circunstancias X, hubiese contraído la lepra; pero, aunque existe una buena regularidad NS que vincula el consumo inmoderado de café y el cáncer de pulmón, no es verdad que si yo hubiese sido un consumidor inmoderado de café, en circunstancias x, habría desarrollado un cáncer de pulmón.). Sin embargo, hay regularidades NS que apoyan contrafácticos y son confirmadas por sus casos particulares que no son regularidades causales. Que los aviones de un aeropuerto cualquiera hayan de recorrer un espacio mayor antes de despegar es condición NS de que las coladas tendidas en los alrededores del aeropuerto tarden más en secarse (y esta regularidad sustenta contrafácticos, y es confirmada por sus casos particulares). Pero no hay dependencia causal entre ambas propiedades, sino que ambas dependen causalmente de una misma propiedad (que la humedad ambiente en los alrededores del aeropuerto sea mayor a la habitual)¹⁶.

Ante estos hechos, existen dos posibilidades (dejando al margen la primera alternativa, a saber, olvidarse de las leyes y analizar directamente la relación causal entre acaecimientos mediante contrafácticos, dando por supuesta la relación de similaridad entre mundos posibles; esta alternativa soluciona ciertamente las dificultades mencionadas -está pensada para ello). Una es la alternativa tradicionalmente considerada «empirista», la alternativa de Hempel, Suppes o Van Fraassen -la alternativa «humeana», aunque Hume no fue ni mucho menos tan sutil: Decir que una ley causal es una regularidad que tiene, a nuestros ojos, ciertas características ideales. Por ejemplo, una regularidad que puede pasar a formar parte «del sistema de regularidades que constituiría nuestra teoría ideal del mundo»: una teoría simple,

¹⁵ La idea es de Mackie: véase *The Cement of the Universe*, Clarendon Press, Oxford, pap. rep., 1980, cap. 3. Mackie exige que C sea parte necesaria pero por sí sola insuficiente de una condición suficiente pero no necesaria de E (así, un factor causal es una condición INUS: insufficient but non-redundant part of and unnecessary but sufficient condition. Esto se ajusta bien al caso de fumar y el cáncer de pulmón, pero no en general a todos los casos. Estoy de acuerdo con J. Bennet en exigir algo menos estricto, una condición NS. Cf. *Events and their Names*, Oxford University Press, Oxford, 1988, §18. Por otro lado, para acoger la posibilidad de la causalidad en un mundo indeterminista, deberíamos decir más bien que el darse un acaecimiento tipo C, junto con los otros posibles factores X, es suficiente no para que se dé un acaecimiento de tipo E, sino para que exista una mayor o menor tendencia o propensión a que se dé un acaecimiento de tipo E. Pero es más razonable no complicar las cosas, y tampoco es preciso en este contexto.

¹⁶ El ejemplo es de Salmon; está tomado de su excelente libro *Scientific Explanation and the Causal Structure of the World*, Princeton, New Jersey, 1984.

elegante, empíricamente corroborada, etc...¹⁷ (Hume sostuvo que una ley causal es una regularidad -no NS, sino una en la que la causa es condición necesaria y suficiente del efecto- que estamos habituados a tomar como algo más fuerte que una mera regularidad).

La segunda -la «aristotélica»- consiste en suponer que en el mundo hay propiedades objetivas y relaciones de dependencia causal objetivas entre ellas, tomar como primitiva la relación de dependencia causal entre propiedades, y decir que una aserción causal general (pongamos « $C \rightarrow E$ ») no tiene las condiciones de verdad de un enunciado de primer orden (pongamos por caso, « $\forall x C(x) \wedge X(x) \rightarrow E(x)$ ») no es una mera regularidad, sino que es verdadera si y solo si, de hecho, se da la relación de dependencia causal entre las propiedades designadas por C y E. (Llamaré 'leyes causales' a estas relaciones de dependencia causal entre propiedades, y no a los enunciados que las expresan)¹⁸ La existencia de la ley causal, a su vez, implica la existencia de una regularidad NS entre acaecimientos de tipo C y acaecimientos de tipo E, simplemente en virtud de lo que entendemos por 'propiedad'.

Superstición por superstición, yo me quedo con esta última posibilidad. Si he de elegir entre tomar como primitiva la relación de similitud entre mundos posibles o la relación de dependencia causal entre propiedades, me quedo con la segunda. Si he de elegir entre tomar como primitiva la existencia de relaciones de dependencia causal entre propiedades o tomar como primitiva la existencia de relaciones de simplicidad, buena integración, alto contenido explicativo, alto grado de clausura bajo la relación de implicación, etc., entre diferentes sistemas de leyes naturales, yo me quedo con la primera¹⁹.

Los análisis de nuestros sistemas conceptuales son explicaciones que no se diferencian en ningún aspecto esencial de las explicaciones de otros fenómenos naturales: nuestro sistema conceptual es una parte más del mundo. Un buen análisis debe exponer las relaciones de dependencia entre los conceptos, como una buena teoría debe exponer las relaciones de dependencia entre propiedades. Un análisis debe ser también coherente con el resto de las cosas que sabemos sobre el mundo²⁰. Pues bien, supuesto que hay las pro-

¹⁷ Véase Bas Van Fraassen, *Laws and Symmetries*, Oxford University Press, Oxford, 1989, para la última edición de este tipo de propuesta. (Una edición confeccionada en una de las dos o tres más refinadas imprentas filosóficas del momento, dicho sea de paso).

¹⁸ Cf. Fred Dretske: 'Laws of Nature', *Philosophy of Science*, 44, 1977, págs. 248-268, y D. A. Armstrong, *What is a Law of Nature?*. Cambridge U.P., Cambridge, 1983.

¹⁹ Un realista es alguien que quiere comprender la mente, la razón, en el mismo marco en que comprende la caída de los rayos o la rotura de las presas. Para ello debe dar en la superstición de creer que el mundo está, en sí mismo, robusta y profusamente articulado. Un idealista es alguien que no quiere caer en esta superstición, y a cambio acepta la mística de una mente con facultades que nadie puede explicar de dónde proceden: hábitos o prejuicios para separar determinadas regularidades de otras -hábitos que, milagrosamente, funcionan bien-, maravillosas proyecciones trascendentales de elementos de su sistema conceptual, preferencias no menos taumaturgicas por los sistemas legaliformes fácilmente axiomatizables, etc.. (Usando la palabra 'superstición' pretendiáremediar irónicamente a Wittgenstein. Cf. los epígrafes 6.371 y 6.372 del *Tractatus Logico-Philosophicus*).

Wesley Salmon propone otra alternativa en la obra antes citada, recurriendo a la noción de proceso causal capaz de transmitir una marca. La propuesta no es aceptable, en mi opinión, por las mismas razones por las que no es aceptable la propuesta de Lewis. La noción de transmisión de marcas o información debe analizarse en términos de la noción de causa, no al revés.

²⁰ Este es, en esencia, el punto de vista tantas veces enfatizado por Quine, expuesto ya en la cita de Neurath que sirve de lema a *Word and Object*: el saber es un barco que se ha de reparar en alta mar, mientras navega; los interesados en remiendos conceptuales y los interesados en remiendos empíricos están, en igualdad de condiciones,

iedades que creemos que hay, y supuesto que hay las relaciones de dependencia causal que creemos que hay, resulta explicable el advenimiento de seres con representaciones de esas cosas. Y supuestos seres con representaciones de relaciones de dependencia causal, resulta explicable que esos seres elaborasen la noción de *mundo posible* y la relación de similaridad entre mundos posibles; resulta explicable también que elaborasen la idea de un sistema ideal de leyes causales, con características como las antes descritas: simplicidad, elegancia, elevado poder explicativo y descriptivo, máxima generalidad, etc.. Los procesos inversos no son igualmente explicables. Vamos a ver dentro de poco cómo, supuesto un análisis en la línea de la tercera opción, podemos explicar tanto los contrafácticos característicos de la segunda como el ideal de un sistema simple, máximamente general, etc., de leyes naturales, característico de la primera.

Naturalmente, no analizar la relación de dependencia causal entre propiedades no es lo mismo que decir que todo vale a la hora de postular que se da tal relación entre determinadas propiedades. Mencionaremos ejemplos paradigmáticos, e indicaremos los problemas epistemológicos involucrados en la determinación de cuándo algo ejemplifica el concepto y cuándo no. Al hacerlo, precisaremos mucho más qué entendemos por *ley causal*; explicaremos qué propiedades intervienen en ellas, y elucidaremos algo mejor la relación entre ellas. Pero no podemos hacer más. Simplemente, es un hecho bruto que en el mundo se dan relaciones de dependencia causal entre propiedades.

Tal como ha sido ofrecido hasta aquí, el análisis deja algunos importantes cabos sueltos. En particular, no está claro por qué algunas aserciones causales generales son imprecisas -contienen cláusulas *ceteris paribus* tácitas- y otras no. Para cumplir la promesa efectuada en el párrafo anterior, y formular con mayor precisión el análisis, será una buena estrategia examinar algunos ejemplos.

Formulaciones estrictas y *ceteris paribus* de las leyes causales

Consideremos primero un caso típico de la vida cotidiana. Hace tiempo que vengo notando que ingerir pasteles de la pastelería de Cardedeu me produce ardor de estómago. De hecho, ahora tengo ardor de estómago, y hace un rato he comido un pastel de esa pastelería. La ingestión del pastel causa mi actual ardor de estómago. Es decir, parece que la ley causal implicada haría depender la propiedad de tener ardor de estómago de la propiedad de haber ingerido pasteles de la pastelería de Cardedeu. Se trataría de una ley *ceteris paribus*, ciertamente. No es verdad que quien consumiera pasteles de esa pastelería tendría ardor de estómago; eso solo sería verdad si se dieran además ciertos factores orgánicos, presentes en mi propio caso, pero que soy incapaz de describir, y los pasteles continuaran estando fabricados empleando ciertas sustancias, que se emplean ahora y también soy incapaz de especificar (manteca de cerdo, quizás); esto es, sólo sería verdad que, en condiciones enteramente parejas a estas, consumir pasteles de esa pastelería produce ardores de estómago.

a bordo del mismo barco. Véase, por ejemplo, 'Epistemology Naturalized', en *Ontological Relativity and Other Essays*, Columbia University Press, New York, 1969.

Pero existe otro modo muy diferente de describir la misma situación. Nosotros pensamos que, estrictamente hablando, la *misma serie de acaecimientos causalmente relacionados* podría ser descrita de un modo más apropiado, y en ese caso las leyes causales involucradas precisarían en un grado mucho menor de cláusulas de salvaguardia *ceteris paribus*. El tipo de proceso causal que lleva de un caso de consumir pasteles de la pastelería de Cardedeu al ardor de estómago, en mi propio caso, podría ser descrito en términos puramente químicos. (Entiendo por 'proceso causal' una serie de acaecimientos particulares causalmente interrelacionados, y por 'tipo de proceso causal' lo análogo a 'tipo de acaecimiento'). Descrito en estos términos, el proceso no sólo involucraría propiedades diferentes, sino también particulares diferentes (particulares más pequeños, como células, moléculas, etc., susceptibles de ejemplificar las propiedades químicas, de los que suponemos «compuestos» a los particulares «más grandes» -pasteles, personas- que ejemplificar las primitivas propiedades). Ahora bien, el que la ley causal que justifica nuestra aserción causal singular sea *ceteris paribus* parece deberse, en buena medida, al modo en que hemos descrito el proceso. Si hubiésemos consumido algo con las características químicas de los pasteles de esa pastelería, pero no fabricado en esa pastelería, igualmente hubiésemos tenido ardor. Y si hubiésemos consumido pasteles de esa pastelería sin esas características químicas, no hubiésemos tenido ardor.

Decimos que el acaecimiento consistente en que Juan fumase durante cuarenta años una media de treinta cigarrillos diarios causó el cáncer de pulmón desarrollado por Juan. Y pensamos que el enunciado 'fumar inmoderadamente durante un largo período causa cáncer de pulmón' expresa una ley causal, la ley causal que justifica la anterior aserción causal singular. Pero pensamos también que el tipo de proceso causal que lleva del fumar inmoderadamente al cáncer de pulmón podría ser descrito en otros términos, de modo tal que las leyes causales involucradas fuesen mucho más precisas. Vagamente, sospechamos que ese proceso podría ser descrito en términos de una serie de propiedades ejemplificables por entidades «más pequeñas» que Juan, de las que Juan está compuesto. Si nos parece que el cáncer de pulmón podría ser causado por otros caminos (sin necesidad de fumar), es porque pensamos que el mismo proceso microscópico en cuestión podría ser descrito, con similar imprecisión, en otros términos macroscópicos (por ejemplo, como el estar en contacto con amianto en ciertas cantidades durante cierto tiempo).

¿Qué hace a unas leyes más estrictas que otras? La respuesta parece clara. Sea «C. $\rightarrow$ E» el enunciado de una ley causal. Una ley nos parece más estricta cuanto mayor es nuestra confianza en la verdad de los enunciados «si se ejemplificara C, se ejemplificaría después E y «no se habría ejemplificado E, si no se hubiera ejemplificado antes C». Para las leyes causales en que descansamos en la vida cotidiana (y en la medicina, la sociología, la economía, la psicología, la meteorología, y hasta la física *aplicada*) ninguna de estas afirmaciones es estrictamente verdadera. Sin embargo, sí son verdaderas, supuestas fijas ciertas condiciones que no somos capaces de establecer. No es verdad que, si dejara caer un vaso de cristal desde diez metros de altura en caída libre sobre un suelo duro, el vaso se haría añicos; pero sí lo es, fijadas ciertas condiciones que están fijas en los casos en que aplicamos esa ley (es decir, *en condiciones parejas a éstas*). Descender en el tamaño temporal y espacial de nuestra descripción de un proceso causal determinado tiene la virtud de permitirnos descansar en leyes más estrictas, sin depender de supuestos no especificados.

Las cláusulas *ceteris paribus*, en el análisis de la causalidad aquí defendido, y la referencia a la relación comparativa de similaridad entre mundos posibles con respecto al mundo real, en el análisis de Lewis, pretenden recoger esta idea de diferentes modos.

Esto implica que nuestro ideal para las leyes formuladas en términos de las propiedades ejemplificadas por los particulares más pequeños, espacial y temporalmente, que sea necesario postular, (las leyes de la física, por supuesto), es el ideal determinista laplaciano. Pero no implica conceptualmente el determinismo laplaciano, porque las leyes causales que nos son mejor conocidas no satisfacen el ideal, de modo que podemos comprender perfectamente el indeterminismo incluso al nivel microfísico²¹.

Idealmente, pues, los mismos procesos que describimos de modos tales que nos obligan a descansar en leyes que incorporan cláusulas *ceteris paribus* implícitas podrían describirse con mayor precisión, descansando en leyes libres de tales cláusulas de salvaguarda. Los términos en que lo haríamos en tal caso serían «microfísicos». (Es esta intuición, por cierto, la que hace plausible el análisis de las leyes causales como regularidades que pueden incorporarse en un sistema idealmente simple, unificado, etc... Pero, como veremos en la próxima sección, nosotros podemos explicarla a partir de nuestro análisis alternativo -en lugar de tomarla como primitiva). Esto plantea dos dudas importantes. La primera: ¿de dónde proviene esta creencia nuestra?; ¿qué la justifica?; ¿hasta qué punto es necesario darla por establecida?. La segunda: ¿qué necesidad tenemos, entonces, de las leyes «macroscópicas», mucho más imprecisas?. Dado el análisis ontológicamente cargado que hemos hecho de la causalidad, ¿no sería más satisfactorio suponer que las únicas propiedades que hay, estrictamente hablando -y, por tanto, las únicas leyes causales que hay- son las propiedades que se postulan en la física?. La respuesta a estas preguntas, como el lector puede quizás colegir, es esencial para nuestro presente problema. En las dos próximas secciones trataré de responder a ellas.

²¹ Para las leyes más estrictas que podemos pensar, yo suscribiría así una idea señalada por diversos autores (creo que la prioridad se debe conceder a Anscombe): que, en contra de lo que los filósofos (desde Hume) han dicho, las causas no tienen por qué ser suficientes para sus efectos; lo que sí han de ser es necesarias. Para que sea una ley causal estricta que E depende de C, debe ser verdad que no se ejemplificaría E a menos que se hubiese ejemplificado C; pero no debe ser verdad que si se ejemplificase C, se ejemplificaría E. Si disparo a Juan con una pistola indeterminista (una pistola que sólo dispara cuando se acciona el gatillo, pero que casi nunca dispara cuando se acciona el gatillo, no habiendo «variables ocultas» que expliquen la diferencia), y la pistola funciona, mi accionar el gatillo es causa de la muerte de Juan. Y, sin embargo, hay circunstancias posibles todo lo cercanas que se pueda pedir al mundo real en que acciono el gatillo y la bala no sale, así que Juan no muere. Por otro lado, si acciono un interruptor que hace encender una bombilla indeterminista, una bombilla que en ocasiones se encendería «porque sí», sin que el interruptor hubiera sido accionado, y la bombilla se enciende de hecho, no diríamos con seguridad que mi accionar el interruptor causó el que la bombilla se encendiera, *pese a que sea verdad que siempre que el interruptor se accionase la bombilla se encendería*.

Con esto le damos mayor cuerpo al análisis de la causalidad en términos de leyes de dependencia causal entre propiedades. Las leyes de dependencia causal establecen que la ejemplificación de ciertas propiedades (los efectos) dependen necesariamente de la ejemplificación de ciertas otras (las causas); que no se ejemplifican aquellas si no ejemplifican éstas. No establecen que sea suficiente la ejemplificación de ciertas propiedades (las causas) para la ejemplificación de ciertas otras (los efectos); en este respecto, a lo sumo puede decirse que establecen que la ejemplificación de ciertas propiedades es suficiente para que exista una tendencia o propensión en determinado número a que se ejemplifiquen ciertas otras.

¿Por qué creemos en la necesidad de leyes estrictas?

Un ejemplo más cercano a la investigación científica real servirá para fijar mejor nuestras reflexiones. Hace algún tiempo hubo en España una intoxicación colectiva, que afectó a algunos miles de individuos; a la enfermedad, identificada a través de un conjunto de síntomas más o menos característicos, se le denominó 'síndrome tóxico'. Se trataba de una serie más o menos vaga, pero en conjunto característica, de afecciones pulmonares (similares a la neumonía), musculares (calambres, etc.), en algunos casos psíquicas, etc. Se llevaron a cabo estudios típicos de etiología epidémica, y se concluyó que la causa del síndrome había sido la ingestión de aceite de colza, «desnaturalizado» primero para hacerlo no apto para el consumo humano, y adulterado después mediante un proceso destinado a eliminar las consecuencias aparentes (olor, color y sabor, hasta cierto punto) del primer proceso y poder así venderlo como aceite de oliva. No es que todos los que consumieron de ese aceite contrajesen el síndrome; pero sí existía una buena regularidad NS: consumir el aceite, en circunstancias (probablemente orgánicas) X, basta para haber contraído el síndrome. (Se dieron casos de sujetos con los síntomas característicos del síndrome que no creían haber consumido el aceite; pero, típicamente, esos casos se desestimaron -en presencia de la importante evidencia disponible- como casos de sujetos que, después de todo, sí habían consumido el aceite sin saberlo, o como casos de sujetos con otra enfermedad que presentaba síntomas parecidos).

La importancia del caso (las consecuencias penales para los causantes de la adulteración, así como la necesidad de hallar algún remedio médico para los afectados) hizo que se considerasen muchas otras hipótesis. Una hipótesis alternativa fue defendida, sin éxito, hasta el final. De acuerdo con esta hipótesis, la causa del síndrome había sido el consumo de hortalizas tratadas con ciertos productos químicos para prevenir las plagas de insectos. También en este caso existía una buena regularidad NS.

Una parte de las razones para apoyar la hipótesis del aceite era evidencia epidemiológica: se había mostrado que la red de distribución del aceite coincidía bastante mejor que la de las hortalizas con el ámbito de incidencia de la enfermedad. Pero estas pruebas nunca constituyen evidencia suficiente para sustentar una hipótesis causal. La razón es inmediata. Saber que los acaecimientos observados del tipo del efecto, y sólo los acaecimientos del tipo del efecto, han sido precedidos por acaecimientos del tipo de la presunta causa no excluye que la verdadera causa sea otra, inadvertida por nosotros, que, naturalmente, sea también una propiedad ejemplificada típicamente por los sujetos en cuestión.

Lo que sí lo excluiría sería lo siguiente: observar dos casos, que coincidiesen en *todas las propiedades* salvo en que en uno se ejemplifica la propiedad que constituye la presunta causa y en otro no se ejemplifica esa propiedad, y que son tales que en el primero se ejemplifica la propiedad que constituye el efecto pero en el segundo no. La suposición de que se da la circunstancia que acabo de describir es analíticamente inconsistente (supuesto mi análisis de las relaciones causales) con la afirmación de que no existe una ley de dependencia causal del tipo de acaecimiento del efecto respecto del tipo de acaecimiento de la (presunta hasta aquí) causa.

En las circunstancias cotidianas nunca se da una situación así; nunca podemos decir que estamos ante dos individuos que coinciden en todas las propiedades salvo que uno ejemplifica la causa y otro no la ejemplifica. Lo más que podemos hacer para acercarnos

a ese ideal (un ideal consiguiente a nuestro concepto de *causalidad*, si mi análisis es correcto) es el diseño de experimentos aleatorios. Por ejemplo, si pudiésemos dividir aleatoriamente a cien personas en dos grupos, administrar a unos la presunta causa y a otros no, y si el número de personas afectadas por algo suficientemente similar al síndrome en el primer grupo es significativamente mayor que en el segundo, podemos estar razonablemente seguros (aunque nunca completamente seguros) de que hemos dado con una ley de dependencia causal. Aún así, bien podría haber ocurrido que, casualmente, los individuos a los que se ha administrado la presunta causa compartiesen en mayor porcentaje que los del otro grupo una cierta propiedad que fuese la verdadera causa. Sea como fuere, una mera comprobación epidemiológica se aparta mucho más del ideal²².

Supongo que esto acaba de mostrar que el carácter primitivo de la relación de dependencia causal entre propiedades no permite suponer que sea una cuestión arbitraria determinar en qué casos se da y en qué casos no se da; antes bien, el análisis *explica* nuestra inseguridad práctica sobre la verdad de una aserción causal. Y lo hace sin misterios extraños, sin los misterios en que realmente consisten muchos argumentos presuntamente destinados a mostrar la falibilidad (la osadía) de la inducción. Nuestro problema epistémico no está en que una ley natural asevere lo que se da en el presente, en el pasado y en el futuro, mientras que toda la evidencia a nuestro alcance concierne a regiones espaciotemporales mucho más limitadas. Leyes naturales (no estrictas) restringidas en su ámbito de aplicación a regiones espaciotemporales a nuestro alcance epistémico plantearían las mismas dificultades (piénsese en la ley: la causa del síndrome que padecieron aquellos sujetos fue su consumo de aceite de colza desnaturalizado). Mucho menos aún radica el problema en que las leyes naturales conciernan no sólo a lo que se da en el presente, en el pasado y en el futuro, sino también a lo que se da en otros mundos posibles, unido a que no podemos «viajar» a otros mundos posibles para comprobar si, efectivamente, tienen allí vigor -como osadamente creemos. Este modo de presentar la cuestión es completamente absurdo.

El verdadero problema de la inducción está en que las leyes naturales aseveran relaciones entre propiedades, y las características causalmente eficaces no están manifiestas, no son patentes (no porque estén escondidas, sino porque hay demasiados candidatos). La posibilidad de aseverar contrafácticos es derivativa; es suficiente (y necesario) establecer relaciones de dependencia entre propiedades, para que estas relaciones nos permitan construir mundos posibles que sustentan la verdad de contrafácticos. A los mundos posibles no se viaja, porque son subproductos de las relaciones conceptuales y nómicas entre los elementos de nuestro sistema de representación.

El modo nominalista de presentar la cuestión oscurece el verdadero problema de la inducción. Se supone que los enunciados de las leyes naturales tienen las condiciones de verdad de los enunciados de la lógica de primer orden a los que los traduciríamos; esto es, que tienen las condiciones de verdad de un enunciado universal accidental, como ‘todos los

²² Ronald Giere proporciona una excelente introducción a la compleja epistemología de las relaciones causales en su *Understanding Scientific Reasoning*, Holt, Rinehart and Winston, 2nd. ed., New York, 1984. El lector comprobará que no comparto su análisis, ni, por tanto, su explicación de las virtudes epistemológicas de unos y otros tipos de evidencia causal.

asistentes al estadio llevan corbata a rayas'. Los predicados funcionan aquí como meras etiquetas: a todos aquellos a los que se aplica P se les aplica también Q. (Naturalmente, se supone también que hay rasgos *ulteriores* que separan las meras regularidades de las leyes, como la posibilidad de incluir a estas últimas en un sistema simple y elegante, etc.)

Vistas así las cosas, resulta un misterio insoluble cómo alguien puede dar en creer la verdad del enunciado, supuesto que ha comprobado que se cumple en cien, mil o diez mil casos. Sólo el examen de todos los casos sería epistémicamente aceptable. Lo que ocurre es que este no es el problema de la inducción; este es un pseudoproblema, creado por una propuesta revisionista mal justificada sobre la naturaleza de nuestro sistema conceptual²³.

Decía que en el caso del síndrome tóxico una parte de las razones que sustentaban la hipótesis de la colza eran de tipo epidemiológico. Pero había otras razones. En casos así, por razones obvias no se puede intentar un experimento aleatorio. Lo que suele hacerse es intentarlo con animales. Las ventajas epistemológicas de este tipo de experimentos para sustentar hipótesis causales se ven notablemente reducidas con ello; puede ser que sea imposible reproducir el síndrome en ausencia de ciertas condiciones orgánicas presentes únicamente en los seres humanos, pese a que la causa sea, de hecho, el aceite. O puede suceder algo inverso; puede suceder que el aceite en cuestión sea «perjudicial» para los animales pero no afecte de hecho a los seres humanos. Sin embargo, en este caso se reprodujeron algunos síntomas relativamente similares a los de los afectados en animales a los que se administró aceite de colza desnaturalizado; y no se consiguió nada similar con las hortalizas. Esto bastó para sustentar la hipótesis del aceite frente a la otra, y en general para mantener la primera hipótesis como explicación causal de los acaecimientos involucrados.

¿Por qué bastó?. Parece obvio (aunque no es en absoluto obvio por qué es obvio) que si conociésemos un mecanismo, un proceso causal, que llevase del aceite a los síntomas, toda posible duda desaparecería. Y parece obvio también que el mecanismo debe ser «microscópico» que debe describirse en el lenguaje de la biología molecular; o, mejor, en el de la química, o, mejor aún, en el de la física elemental. Los experimentos mencionados bastaron para sugerir que, de haber un mecanismo, lo había en el caso de la colza, no en el de las hortalizas. (Pero persistieron en algunos las dudas, precisamente porque no era posible asegurar que se conocían las características químicas del aceite presunto culpable; el proceso químico había sido tan «sucio», que no se podía reproducir con todas las garantías).

²³ Tanto los inductivistas, como ciertos antiinductivistas -Popper- comparten esta errónea caracterización del problema de la inducción. Popper tiene razón al recordar que si el problema de la inducción fuese el descrito, la tarea de la ciencia sería la búsqueda de nuevos casos del mismo tipo que los anteriores para acercarse a la «confirmación» ideal de la ley; mientras que los científicos en absoluto proceden así, sino que buscan casos «sorprendentes», confirmaciones «inesperadas». Pero se equivoca al creer que el problema de la inducción es el descrito. Descrito correctamente, es patente por qué comprobar en cien, mil o diez mil experimentos que, efectivamente, las tizas dejadas al vacío llegan al suelo cuando la teoría newtoniana predice, no añade un ápice a lo «confirmado» que la teoría esté para nosotros; mientras que comprobar que, efectivamente, y en contra de lo que nunca hubiésemos pensado, el cometa Halley se comporta de acuerdo con la teoría, sí le añade muchos enteros a nuestra confianza. Pero, descrito correctamente el problema de la inducción, resulta también patente por qué es absurdo pretender que la intención primera de la ciencia es proponer teorías altamente refutables.

Explicar por qué nos parecen mejores las explicaciones microscópicas que las macroscópicas no parece sencillo. Lo principal debe ser que las explicaciones microscópicas son verdaderas; algo debe contar también en favor su mayor generalidad, el que valgan sin restricciones de lugar y de tiempo; pero también parece ser importante el que nuestros casos paradigmáticos de relaciones causales (aquellos en los que las leyes causales macroscópicas más se acercan al ideal) sean los empujones, aceleraciones, golpes, detenciones, etc., de bolas de billar y objetos similares²⁴. El análisis que hemos propuesto añade un factor epistemológico fundamental a estos elementos: la posibilidad de diseñar situaciones experimentales cercanas al ideal que permitan establecer, con relativa certidumbre, relaciones de dependencia causal, es mucho mayor en el caso microscópico que en el macroscópico: las propiedades posiblemente relevantes se han reducido al mínimo. Un segundo factor -que estudiaremos después- es que separar los casos de premisión de los casos de verdadera causación exige, cuando existen dudas, descender a una escala espacial y temporalmente menor.

Es de este modo explicable que algunas de las más características explicaciones científicas no sean explicaciones de acaecimientos particulares (diré 'explicaciones causales', explicaciones que consisten en indicar las leyes causales ejemplificadas en un determinado proceso), sino propuestas para la reducción de explicaciones formuladas en ciertos términos más o menos macroscópicos a otras formuladas en términos más «básicos» (llamaré a estas explicaciones 'reducciones', explicaciones consistentes en formular leyes que identifican tipos de procesos macroscópicos con tipos de procesos relativamente más microscópicos). Propuestas, esto es, para describir procesos causales en términos que involucren leyes causales más precisas a aquellas que era preciso invocar en las viejas descripciones. El prototipo de ello es la reducción de la termodinámica fenomenológica a la mecánica estadística; pero también la reducción de la química tradicional a la química de los elementos es un ejemplo de lo mismo. Así, el calor se identifica con la energía cinética media de las partículas; y el agua, con el H_2O ²⁵. De hecho, una buena parte del trabajo del científico (el trabajo «teórico») consiste en la definición de conceptos adecuados para llevar a cabo este tipo de reducciones. Una teoría científica puede verse adecuadamente como una definición de sus términos teóricos, términos que se refieren a los particulares «más pequeños» y a las propiedades de estos particulares con ayuda de los cuales se formularán explicaciones más precisas de viejos fenómenos²⁶.

²⁴ Para la última explicación, cf. p. Strawson, 'Causation and Explanation', en Vermazen, B., y Hintikka, M., *Essays on Davidson: Actions and Events*. Clarendon Press, Oxford, 1985. Mackie, en el capítulo octavo de *The Cement of the Universe*, y Salmon, en el último capítulo de *Scientific Explanation and the Causal Structure of the World*, conceden cabal importancia a la búsqueda de «mecanismos» en el análisis de la causalidad.

²⁵ En *The Nature of Psychological Explanation*, (MIT Press, Cambridge, Mass., 1983) Robert Cummins distingue entre lo que él denomina 'transition theories' y 'property theories'. Las primeras persiguen establecer relaciones de dependencia causal; las segundas, reducir relaciones establecidas en ciertos términos a otras. Es la misma distinción que yo estoy tratando de trazar aquí, y las consecuencias para la filosofía de la mente son también similares. Recientemente he encontrado una distinción similar en el artículo de Frank Jackson y Philip Pettit 'Functionalism and Broad Content', *Mind*, 1988, págs. 381-400, esta vez en relación a la noción de explicación. 'Causal Explanation' y 'programme explanation' son sus respectivas etiquetas. El estado de la discusión filosófica en el ámbito de la filosofía de la mente explica, sin duda, estas coincidencias.

²⁶ Es así como entiendo yo la llamada «concepción semántica de las teorías», propuesta independientemente por Van Fraassen, Giere y los «estructuralistas» Sneed, Moulines y Balzer.

Es esencial a mi explicación advertir que la reducción no es sólo de propiedades, sino que también exige la reducción de los particulares que ejemplificaban las propiedades reducidas a particulares «más pequeños» que ejemplifican las nuevas y de los que los particulares «más grandes» están compuestos en modos que es preciso especificar. La reducción no es de propiedades a propiedades, sino de tipos de acaecimientos a tipos de acaecimientos. Esto no se advierte bien en el caso del agua, por cuanto 'agua' es un término de masa. Supongamos que c_1 causó e_1 , y que c_2 causó e_2 (todos ellos acaecimientos singulares), y que una misma ley causal macroscópica justifica la verdad de ambas aserciones. Bien puede ocurrir, sin embargo, que los procesos microfísicos en cada caso sean muy diferentes, que no podamos identificar «nada común» a c_1 y c_2 en la descripción microfísica. (Imagínese que la ley macroscópica es *una granizada sobre una plantación de vides causa la pérdida de la cosecha de uva*; es obvio que el detalle microfísico de cada granizada particular, y de cada pérdida de la cosecha en particular, será muy diferente. O imagínese que la ley macroscópica pertenece a la termodinámica fenomenológica, y el detalle microfísico es el que daríamos en términos de la mecánica estadística). Las leyes reductoras no identifican propiedades con propiedades, sino tipos de acaecimientos con otros tipos de acaecimientos (quizás con toda una enorme diversidad de ellos). O, mejor dicho: las leyes reductoras ofrecen un procedimiento general para, en cada caso particular, construir una descripción microfísica del proceso causal macroscópicamente descrito.

Habríamos resuelto nuestro problema, pues, si pudiésemos haber identificado el tipo de acaecimiento consistente en *consumir aceite de colza desnaturalizado* con un tipo de acaecimientos caracterizados mediante conceptos químicos (o cuánticos, aunque esto último no es preciso aquí en particular, por cuanto suponemos la posibilidad de reducir la química de modo general); si pudiésemos identificar los síntomas característicos del síndrome con fenómenos químicos (o cuánticos); y si dispusiésemos de una serie de leyes de dependencia causal que conectasen causalmente en un mismo proceso causal un acaecimiento del primer tipo con uno del segundo. Precisamente porque sospechamos que no hay tal mecanismo, creemos que haber nacido bajo el signo de Aries es una característica causalmente ineficaz en cuanto al carácter de una persona. Y es por esta razón por la que puede parecer plausible el análisis de las leyes causales como regularidades integrables en un sistema máximamente simple, etc., de leyes. Una presunta ley causal (una buena regularidad NS) no integrable en el sistema de nuestros conocimientos (a través de leyes reductoras) no es verdadera ley causal. Pero nosotros, a diferencia de lo que se hace en el análisis mal adjetivado como «empirista», explicamos por qué lo creemos así, en lugar de tomarlo como un hecho primitivo²⁷.

Naturalmente, es una cuestión contingente que esta creencia nuestra sea correcta. Hay experimentos mentales, y algunos no mentales, que muestran cómo en ocasiones nos sentimos perplejos ante la posibilidad de que se revele, después de todo, una creencia falsa.

²⁷ Mal adjetivado porque la actitud más coherente con el espíritu de lo que se llama tradicionalmente 'empirismo' exige naturalizar la mente. Para ello, es necesario no descansar en primitivos «mentales», como nuestros gustos por la simplicidad, etc., de nuestras teorías.

La mecánica cuántica parece ser una fuente inagotable de perplejidad a este respecto. Y ya antes de ello, recordamos que las dudas de los cartesianos sobre la teoría de Newton no provenían sólo del chovinismo, sino de la misma perplejidad que provocó el 'hypotheses non fingo' del propio Newton. La teoría describía bien los fenómenos, pero ¿los explicaba?. No, mientras no se comprendiese la naturaleza del mecanismo meramente descrito como la acción de la fuerza de la gravedad. Pero con el paso del tiempo y la acumulación de éxitos predictivos de la teoría, esas dudas parecen haberse aliviado. Aunque quizás podría sostenerse que no es que las dudas se aliviaran, sino que la eficacia predictiva de la teoría permitía confiar en que, de hecho, existía ese mecanismo, y en que alguna vez sería descrito. Por otro lado, sólo hemos aceptado la validez de las explicaciones teleológicas de hechos biológicos cuando, primero Darwin, y después la genética contemporánea, nos han demostrado que es razonable creer que esas explicaciones no se apoyan en procesos descriptibles de un modo puramente físico.

Lo único que sostengo, pues, es que, siendo las cosas tal y como son ahora, si hay razones para pensar que no existe un modo más preciso de reducir un proceso explicado en términos de una supuesta ley *ceteris paribus* a otro explicado en términos de leyes estrictas, esa aserción causal no está justificada.

En resumen, el análisis del concepto de *causalidad* que surge de la anterior discusión es éste: *c* causa *e*, donde *c* y *e* son acaecimientos, si y solamente si hay una ley que establece la dependencia causal de propiedades -*C*- ejemplificadas en *e* respecto de propiedades -*E*- ejemplificadas en *c*; es así que acaecimientos -tipo *E* dependen de los acaecimientos- tipo *C*, en el bien entendido de que *c* es un acaecimiento de tipo *C* y *e* es un acaecimiento de tipo *E*. La ley en cuestión puede ser una ley *ceteris paribus*; pero, en tal caso, debe haber leyes (reducciones) que suministren procedimientos generales para identificar procesos causales como el descrito con procesos causales cuyas transiciones dependan sólo de leyes estrictas.

Este no puede ser, empero, nuestro análisis definitivo, porque hay casos en que se cumple la condición aunque *c* no es causa de *e* (la condición propuesta no es suficiente) y hay otros casos en que *c* es causa de *e* pero no se cumple la condición (la condición no necesaria).

Una bola *a*, con tal aceleración y tal masa y moviéndose en tal dirección, golpea a otra, *b*, en *t*, haciendo que la segunda comience a moverse con una trayectoria tal que no podría golpear a la bola *c*, en reposo en *t*, dada la ubicación en que está *c* en *t*. En *t'*, sin embargo, se interpone en el camino de *b* la bola *d*, sin que venga a cuento de dónde proviene (podemos suponer que se ha «materializado de la nada»). Al hacerlo, *b* se desvía del camino que seguía y en *t''* golpea a *c*. Claramente, el golpe de *a* a *b* es causa del golpe de *b* a *c*; si no se hubiera producido el golpe de *a* a *b* en *t*, no se habría producido tampoco el golpe de *b* a *c* en *t''*. Pero no hay una ley causal que lleve de las características ejemplificadas en *t* a las ejemplificadas en *t''*. La condición del presunto análisis no es, pues, necesaria.

Los casos de premisión o anticipación sirven para ejemplificar la otra cara del fracaso de ese presunto análisis. Un vaso de cristal estalla por la acción de un agudo grito; pero si el grito no lo hubiese hecho estallar, de todos modos hubiese estallado, porque cuando estalló por efecto del grito estaba a punto de colisionar contra el suelo, lo que, dada su aceleración en ese momento, hubiese producido igualmente rotura. La rotura se produjo un

instante antes de entrar el vaso en contacto con el suelo, a causa del grito. Sin embargo, la posición, aceleración, etc., del vaso un instante antes de estallar constituyan una parte necesaria de una condición suficiente, completada por algunas de las restantes circunstancias de la situación, para producir que el vaso estuviese hecho añicos pasado un corto intervalo temporal; esta regularidad NS parece estar sustentada en una ley causal genuina, y, efectivamente, se dieron acaecimientos de ambos tipos. Pero el primero no fue causa del segundo, porque la posible eficacia causal del primer acaecimiento sobre el segundo fue permitida por la acción del grito. La caída del rayo causó el incendio de la paja, pero aunque el rayo no hubiese caído, la paja se hubiese incendiado porque, cuando el rayo cayó, una tea encendida estaba a punto de entrar en contacto con la paja. La eficacia causal de la caída de la tea ha sido permitida por el rayo; pero el caer una tea sobre la paja es una condición NS de que se incendie la paja, y el resto de las condiciones necesarias junto con la caída de la tea encendida se daban en el caso, e incendios como el de la paja dependen de ese conjunto de condiciones en virtud de una genuina ley causal. El supuesto análisis, pues, no proporciona una condición suficiente.

Ambos tipos de contraejemplos no muestran que la idea central del análisis, la de que las relaciones causales singulares entrañan relaciones de dependencia causal entre propiedades, sea una idea errada. Lo que muestran es que el análisis estaba incompleto hasta aquí; que es parte fundamental de nuestra noción de causalidad la idea de proceso causal. Las cadenas de acaecimientos causalmente conectados son *procesos*: la deriva continental, por ejemplo, es un proceso más que un acaecimiento. En los casos del primer tipo hay un proceso causal que lleva de *c* a *e*, aunque no hay una ley causal que los vincule; en los casos del segundo tipo (en los casos de premisión) hay una ley causal que vincula propiedades ejemplificadas en *c* con propiedades ejemplificadas en *e*, pero no hay un proceso continuo que lleve de *c* a *e*. Ahora bien, lo que vincula a los eslabones inmediatos de un proceso causal genuino es la existencia de una ley causal. Es esta idea intuitiva la que pretendemos recoger con el análisis final que proponemos.

Digamos que el acaecimiento *c* es *causa directa* del acaecimiento *e* si y sólo si existe una ley causal tal que la propiedad *E* depende causalmente de la propiedad *C*, *c* consiste en la ejemplificación de *C* y *e* consiste en la ejemplificación de *E*. (Por tanto, hay también una regularidad NS entre acaecimientos del tipo *C*, y acaecimientos del tipo *E*). Sean *c* y *e* dos acaecimientos particulares, sea *u*₁ la ubicación espaciotemporal en la que acaece *c* y *u*₂ la ubicación espaciotemporal en la que acaece *e*. Establecemos que *u*₁ y *u*₂ son regiones temporales distintas, aunque posiblemente continuas; eso basta para que causa y efecto sean «exigencias distintas». Pues bien, (*C*) *c* es causa de *e* si, y solo si,

- (i) caso de que *u*₁ y *u*₂ sean continuas, *c* es causa directa de *e*;
- (ii) caso de que *u*₁ y *u*₂ no sean continuas, hay un acaecimiento que sucede en la ubicación *u*₃, situada espaciotemporalmente entre *u*₁ y *u*₂, que causa *e* y es causado por *c*²⁸.

²⁸ La definición está inspirada en el análisis de los casos de premisión que ofrece Lewis en 'Causation' y, especialmente, en el 'Postscript' E al mismo artículo en los *Philosophical Papers*, vol. 2. También está muy influenciada por el papel que Wesley Salmon concede a los procesos causales en su análisis de la causalidad; cf., especialmente, el capítulo quinto de *Scientific Explanation and the Causal Structure of the World*. El lector encontrará alguna afinidad con el análisis de J. Bennett en *Events and their Names*; la afinidad es debida a la común inspiración en Lewis.

De acuerdo con el análisis del concepto de *causalidad* que propongo, pues, todo proceso causal correctamente caracterizado en términos químicos, meteorológicos, biológicos o geológicos es también un proceso caracterizable en términos microfísicos. Esta es la versión del fisicalismo que considero aceptable. Sin embargo, de ello no se deduce ninguna tesis reduccionista, en el sentido tradicional de la expresión (yo prefiero decir «no se sigue ninguna tesis eliminacionista», y preservar ‘reduccion’ para la actividad nada sospechosa que he descrito antes). En primer lugar, no es preciso conocer la formulación microfísica del proceso causal para confiar en la corrección de una descripción «macroscópica» del mismo. Basta con no tener razones para pensar que no existe esa formulación precisa. Es perfectamente razonable decir que fumar inmoderadamente causó el cáncer de pulmón de Juan, mientras supongamos, como suponemos, que es en principio posible describir ese proceso de otro modo. En segundo lugar, no hay que confundir las leyes causales con las explicaciones causales; las tesis reduccionistas conciernen, normalmente, a la explicación. Diré algo a continuación sobre la cuestión de la explicación, antes de volver al tema de este artículo.

Leyes causales y explicaciones causales

Si alguien nos pregunta, ‘¿cómo se rompió el jarrón?’, y respondemos ‘la petición de María a Evaristo para que le cantara una copla causó la rotura del jarrón’, mi respuesta, aunque verdadera, inducirá -cuando menos- a la confusión. La petición de María a Evaristo para que le cantara una copla causó que el jarrón se rompiera, ciertamente; pero la razón de que ello fuera así es que la petición de María a Evaristo fue el mismo acaecimiento que la emisión por María de un sonido con una altura y una intensidad elevadísimas. Lo que se nos pedía era *una explicación* de cierto acaecimiento. *Explicar*, como referir, o mejor, referirse a, no es algo que hagan las palabras, si no algo que hacen los sujetos con las palabras. Sólo en un sentido derivativo decimos que ‘Madrid’ se refiere a Madrid, o que la teoría de Newton (concebida en parte lingüísticamente) explica la aparición recurrente del cometa Halley. (Lo que no significa que sea incorrecto decir estas cosas: no lo es, mientras no se olvida su carácter derivativo).

Explicar un acaecimiento es ofrecer información sobre su etiología, sobre su historia causal²⁹. En virtud del análisis que hemos ofrecido, para ofrecer una buena información

Sé que la definición está falta de la clarificación de la noción de continuidad espaciotemporal; pero ese es justamente el problema de que adolece nuestro concepto intuitivo de causalidad. Eso es lo que explica nuestra perplejidad ante las paradojas de Zenón, especialmente ante la paradoja del estadio y la de Aquiles y la tortuga. Más que paradojas del movimiento, pueden verse como paradojas de la causalidad (lo que resulta claro si consideramos la persistencia de un particular a lo largo del tiempo como un caso particular de un proceso causal).

²⁹ De acuerdo con el modelo de la ‘covering-law’, propuesto en el clásico artículo de Hempel y Oppenheim ‘Studies in the Logic of Explanation’ -reproducido en el libro de Hempel *Aspects of Scientific Explanation*, del que hay traducción castellana con el título de *La explicación científica*, Paidós, 1979- explicar un acaecimiento es construir un argumento lógicamente válido cuya conclusión sea un enunciado que describe el acaecimiento y al menos una de cuyas premisas es una ley natural que establezca que, supuestas determinadas condiciones iniciales, debe darse (o es probable que se dé) un acaecimiento de esas características. Existen conocidas objeciones a este modelo explicativo, que en mi opinión confluyen en la omisión que en él se hace del carácter causal -en el sentido explicado en este trabajo- de las leyes supuestas. Cf. Salmon, op. cit., y David Lewis, ‘Causal Explanation’, en *Philosophical Papers*, vol. 2.

etiológica sobre un acaecimiento no basta con indicar algún otro acaecimiento que forme de hecho parte de la historia causal del primero; en general, es preciso además suministrar información sobre la ley causal en virtud de la que el primer acaecimiento forma parte de la historia causal del segundo. Si esto es lo que se supone que he de hacer, respondiendo la pregunta como lo hice alimento la confusión. Dado este supuesto, implícito en las demandas de explicación, al responder como lo hago estoy dando a entender a mi interlocutor que el hecho de que el acaecimiento en cuestión fuese del tipo de los acontecimientos en que alguien le pide a alguien otro que recite una copla fue causalmente significativo; es decir, que la propiedad de romperse un jarrón (o un objeto similar) depende causalmente de la propiedad de pedir alguien a alguien otro que le cante una copla. Y es normal que mi interlocutor no sea capaz de «ver la relación».

El problema fundamental está en que los acaecimientos, como particulares que son, pueden ser denotados de modos completamente «extrínsecos» a sus características causalmente significativas con respecto a la producción de un efecto dado. Al enunciar la existencia de una relación causal entre dos acaecimientos, podemos referirnos a ellos mediante propiedades accidentales de los acaecimientos, sin mencionar en absoluto las propiedades causalmente revelantes en la situación. a y b saben que Juan murió ayer a causa de la ingestión de setas tóxicas. a pregunta «¿qué causó la muerte, la comida o la cena?», y b responde: «el primero de esos acaecimientos causó la muerte de Juan». Naturalmente, el que ese acaecimiento fuese el primero en ser mencionado en este contexto no constituye una característica causalmente eficaz de la muerte de Juan. Es falso que si la comida *no hubiese sido mencionada en primer lugar por a* en tal contexto, Juan *no hubiese muerto*. El viaje en martes de Juan causó su muerte, y aquí no hay mecanismo que lleve de *viajar en martes a morir*. Pero lo que ocurre es que el viaje en martes de Juan fue también su viaje en un avión que transportaba una potente bomba; Y sí hay una ley causal que conecta los viajes en aviones que transportan una potente bomba con la muerte de los pasajeros.

Las explicaciones causales de acaecimientos singulares, pues, conllevan ciertos aspectos pragmáticos; como por lo demás cualquier intercambio comunicativo. Muchas aserciones posibles, pese a ser estrictamente verdaderas, de llevarse a cabo resultarían, cuando menos, inadecuadas: inadecuadas a los fines genéricos a que se destinan esos acaecimientos que consisten en la transmisión de información. Difícilmente se escuchará a un español de ahora mismo decirle a otro «Felipe González es presidente de Gobierno», salvo en ciertas circunstancias especiales. No porque no sea verdad, o porque haya algo semánticamente inadecuado en ello; simplemente, porque decirlo es «hablar por hablar»³⁰. Pero es fundamental no confundir los aspectos pragmáticos de las aserciones, con los aspectos semánticos. Si los hechos son como he indicado, la aserción efectuada por mí mediante la proferencia de «la petición de María a Evaristo para que le cantara una copla causó la rotura del jarrón» es literalmente verdadera; mi acción de llevar a cabo esa aserción con la intención de explicarle lo sucedido a quien me inquires no es, sin embargo, especialmente feliz.

³⁰ H.P. Grice ha estudiado los fines genéricos que presiden los intercambios verbales, y ha mostrado cómo nos las arreglamos para decir cosas que nuestras palabras, estrictamente hablando (semánticamente hablando) no dicen, apoyándonos para ello en el mutuo conocimiento implícito de esas máximas de la conversación. Cf. «The Logic of Conversation», en Paul Grice, *Studies in the Ways of Worlds*, Harvard U.P., Cambridge, Mass., 1989.

Una cuestión mucho más trascendente la ejemplifica la siguiente respuesta a esta demanda de explicación: ¿por qué se perdió la mayor parte de la cosecha de uva en el Penedés el año pasado?. «A causa de la granizada caída el siete de septiembre». De acuerdo con lo que hemos sostenido hasta aquí, la verdad de esta afirmación exige que ese mismo proceso pueda ser descrito en términos microfísicos, y explicado el acaecimiento efecto (una vez reducido a términos microfísicos) en términos de las leyes estrictas de la física (en lugar de invocar, como aquí se ha hecho implícitamente, la ley «las granizadas sobre vides a punto para la cosecha producen la pérdida de una parte importante de la misma»). Pues bien, ¿por qué no prescindir de la explicación macroscópica, que es mucho más imprecisa, y proponer sólo la mucho más precisa explicación microscópica?. La respuesta es inmediata, dado lo que hemos dicho antes. De las virtudes epistemológicas de las explicaciones microfísicas no se sigue en absoluto que una explicación microfísica de un proceso sea mejor que una macroscópica. En general, resulta mucho más inteligible para nosotros comprender cómo una granizada produce la pérdida de una cosecha (pese a que hay granizadas posibles sobre campos de vid madura que no producirían la pérdida de la cosecha), que comprender cómo el complejo acaecimiento en que consiste la granizada, microfísicamente descrita, da lugar al complejo acaecimiento en que consiste la pérdida de la cosecha, microfísicamente descrita.

Yo no creo que nadie tenga razones serias para decir que no existen las granizadas; el hecho de que consistan en muchos objetos más pequeños no es una prueba de su no existencia, sino todo lo contrario. Lo mismo vale para las nubes, los ríos, las piedras, los tigres y los seres humanos. Y, si existen esos objetos, tanto como sus partes más minúsculas, igualmente es razonable decir que son parte objetiva del mundo sus propiedades, tanto como puedan serlo las propiedades de los objetos más pequeños que los componen. Y si existen las propiedades de los ríos, las granizadas y los tigres, tanto como puedan existir las propiedades de los quarks, igualmente existen las leyes causales que involucran aquellas propiedades; leyes causales y propiedades navegan en el mismo barco (existen aquellas propiedades que son términos de leyes causales). Es cierto que, dado nuestro concepto de causalidad, las leyes causales macroscópicas son mucho menos epistémicamente fiables que las otras; que tienen un alcance mucho menos general, o, en fin, que contienen implícitas cláusulas *ceteris paribus* que no sabemos eliminar si no es descendiendo de nivel ontológico (es decir, formulando otras leyes). Pero no es menos cierto que los objetos macroscópicos y sus propiedades nos son cognitivamente más familiares que los más pequeños. (Ambas cuestiones son caras de una misma moneda, naturalmente. Las leyes causales que nos son cognitivamente más próximas, que nos suministran explicaciones fáciles de comprender y también las predicciones que necesitamos en la vida cotidiana, son aquellas que la Madre Naturaleza nos ha dado recursos innatos para formular; aquellas que, hasta cierto punto, también conocen los chimpancés. Es comprensible que esas leyes no tengan un fundamento epistémico muy seguro).

El punto de vista de Davidson

Tenemos ya los elementos suficientes para exponer el punto de vista de Davidson. Contrastemos, para ello, dos tipos de afirmaciones de causalidad singular verdaderas que hemos encontrado ya: (i) 'el viaje en martes de Juan causó su muerte' o 'la petición de María

a Evaristo para que le cantara una copla causó la rotura del jarrón'; (ii) 'la granizada caída el siete de septiembre pasado causó la pérdida de la mayor parte de la cosecha de uva en el Penedés'. Las aserciones son verdaderas en ambos casos; pero, mientras que la ley causal *ceteris paribus* que se sugiere como explicación en ejemplos como el de (ii) es una ley genuinamente explicativa, los predicados de casos como los de (i) no aciertan a indicar la ley causal que justifica la verdad de la afirmación causal singular. Si esas afirmaciones son verdaderas es porque los acaecimientos causa ejemplifican propiedades que no tienen nada que ver con aquellas, también ejemplificadas por ellos, mediante las que los denotamos, y esas otras propiedades son causa (*ceteris paribus*) de propiedades ejemplificadas en los acaecimientos efecto. El supersticioso que trata de sugerir que hay una ley causal que vincula viajar en martes con los accidentes o las desgracias en general está, lisa y llanamente, equivocado. Y lo que establece la diferencia es que es posible reducir *en general*, mediante leyes reductivas, acaecimientos-tipo consistentes en granizadas y acaecimientos-tipo microfísicamente descritos -acaecimientos-tipo ejemplificados también, respectivamente, por la granizada en cuestión y la pérdida en cuestión- causalmente interconectados a través de leyes estrictas; pero no es posible reducir *en general* acaecimientos-tipo consistentes en viajar en martes y acaecimientos-tipo consistentes en desgracias aéreas a acaecimientos-tipo microfísicamente descritos causalmente interconectados mediante leyes estrictas³¹.

Para mostrar la peculiaridad explicativa de lo mental, pues, no es necesario intentar mostrar que mi intención particular de hacer salir una piedra de mi zapato no puede ser causa del movimiento frenético de mi mano, en el sentido en que la colisión del vaso contra el suelo es causa de su rotura (como Habermas, Anscombe o Von Wright pretende hacer). Basta con mostrar que el predicado 'intención de hacer salir una piedra de mi zapato' es un predicado del mismo tipo que 'viaje en Martes', en el ejemplo anterior: aunque nos permite referirnos a un acaecimiento particular, que de hecho es causa del movimiento de la mano, no describe ese acaecimiento en virtud de sus características causalmente eficaces. Basta con mostrar que no puede haber leyes reductivas que ofrezcan procedimientos generales para identificar procesos descritos en términos psicológicos con procesos descritos en términos físicos. Basta con mostrar, pues, que no puede haber leyes *psicofísicas*. Y esto es, justamente, lo que Davidson pretende haber hecho.

Por otro lado, este es el supuesto de cuya plausibilidad depende la existencia o no de una diferencia cualitativa entre explicación y comprensión, de que hablé al comienzo del artículo; el precio que Davidson nos ha enseñado que hemos de pagar si queremos ver la psicología como una más de las ciencias naturales. El precio no es otro que aceptar el materialismo, en el sentido tradicional de la palabra (aunque no en el «eliminacionista», de acuerdo con lo que se ha explicado en la sección anterior: la existencia de las leyes psicofísicas no conllevaría, en modo alguno, la necesidad o la posibilidad siquiera de

³¹ Recuérdese que la posibilidad de reducción mediante ley en el caso de las granizadas y las pérdidas de cosechas no significa que, en cada caso particular en que sea una granizada lo que cause una pérdida de cosecha el proceso microfísico haya de ser idéntico. Lo que importa es que exista un procedimiento general, como el que la mecánica estadística ofrece para la termodinámica fenomenológica.

prescindir de los conceptos «mentalistas» en la explicación). Si de algún buen argumento se siguiera la inexistencia de leyes psicofísicas, entonces sería cuestión de empezar a ver nuestras explicaciones «mentalistas» del mismo modo que vemos ahora las explicaciones animistas, las de los brujos de toda laya y las de la mayoría de los partidarios de la medicina «naturalista». Eso es, justamente, lo que pretende mostrar Davidson. Según Davidson, puede decirse que mi intención de elevar el brazo causó la elevación de mi brazo, pero sólo puede decirse en el sentido en que puede decirse que el viaje en martes de Juan causó su muerte. Puede decirse porque, de hecho, esta intención particular mía es idéntica a un estado neurológico mío que sí causó, a través de leyes estrictas, la elevación de mi brazo. Pero no puede decirse porque mi intención, ese *tipo* de estado, esa *propiedad*, sea identificable con un tipo de estado o proceso microfísicamente descrito. Y, si fuera así, no habría leyes causales aceptables que involucrasen conceptos tales como el expresado por ‘intención de elevar el brazo’.

Acaecimientos mentales

Una parte de nuestro lenguaje nos sirve para describir nuestra vida mental, y la vida mental de las cosas a las que se la suponemos: nuestros semejantes, ciertos animales, algunos seres no espaciotemporales -como ángeles, demonios y dioses-, árboles y fenómenos naturales -si tenemos creencias animistas-, etc... Expresiones tales como ‘mi deseo de que la reunión comience a la hora’, ‘mi intención de ir al cine esta noche’, ‘mi recuerdo de la visita del Papa’, ‘mi opinión de que el segundo cinturón no estará listo en el 92’, etc., son típicas expresiones para la descripción de la vida mental. Recuerdos, saberes, opiniones, intenciones, pensamientos, propósitos, afectos, sinsabores, creencias y deseos son paradigmas de lo que llamaré ‘estados mentales’ o ‘procesos mentales’.

Los estados mentales son también «cosas que pasan», sucesos o acaecimientos, por supuestos nada espectaculares o insólitos sino sumamente humildes. Son «cosas que pasan», como lo son la proyección de una película, una comida familiar, el sitio de una ciudadela o la espera de una visita. No son, en cambio, «objetos»: no son entidades del tipo de una mesa, un perro o un número. Algunos, como la espera de una visita, duran un cierto tiempo «sin que apenas cambie nada»: así, por ejemplo, una creencia, un deseo o una intención. Estos son más «estados» que «procesos»³². Otros, como un recuerdo repentino, el pensamiento de la solución a un acertijo o un propósito, son más parecidos a la proyección de una película o a una comida familiar: son «procesos» más que «estados». En cualquier caso, los estados mentales ocurren en el tiempo: llegan a ser, duran con más o menos modificaciones, y desaparecen. Más difícil es decir si ocurren en el espacio: tenemos la sensación de que se dan en algún lugar en el interior de nuestro cuerpo (¿se imaginarían aquellos de nuestros antepasados que creían que el órgano del pensamiento es el corazón,

³² Hasta aquí no he dicho que, en contra del uso común, un acaecimiento no debe conllevar el que se produzca un cambio. Los acaecimientos, en nuestra ontología, son los términos de la relación causal singular; y los estados (esos «acaecimientos» en los que nada cambia) son también causas y efectos (la causa del pinchazo fue la presencia de un clavo en la carretera). Por tanto, modificamos un tanto el uso común y los consideramos acaecimientos legítimos.

que sus procesos y estados mentales ocurrieran en algún lugar del pecho del mismo modo que nosotros los imaginamos ocurriendo en algún lugar de la cabeza?). Sin embargo, parece claro que «suenan bastante raro» preguntarse por la ubicación espacial de un recuerdo, una creencia o una intención; mientras que no suenan nada raro preguntarse dónde se proyectó una película o tuvo lugar una comida familiar.

Puesto que los estados mentales transcurren a lo largo del tiempo, una parte importante de la individualización necesaria para referirnos a ellos consiste en indicar el intervalo temporal en que acaecieron. En las expresiones que antes utilicé como ejemplos de típicas descripciones de estados mentales la referencia temporal estaba implícita, como suele estarlo en el uso cotidiano del lenguaje. Pero no basta indicar el intervalo temporal en que ocurrieron para singularizarlos: en el mismo intervalo temporal, diferentes sujetos forman diferentes opiniones, diferentes deseos, etc... Por eso, a falta de una ubicación espacial definida, nos referimos cuando los especificamos al sujeto del estado mental: ‘mi recuerdo de la visita del Papa’ (se entiende mi recuerdo *actual*). Y como el mismo sujeto, en el mismo intervalo temporal, puede tener diferentes estados mentales, acabamos de especificarlos mediante una expresión para el *tipo* de estado mental (las expresiones subrayadas son típicas expresiones para indicar el *tipo* del estado mental: un *recuerdo*, un *saber*, una *opinión*, una *intención*, un *pensamiento*, un *propósito*, un *afecto*, una *creencia*, un *deseo*, etc.) y una descripción de su *contenido* (las expresiones subrayadas son típicas expresiones para indicar el *contenido* del estado mental: ‘mi deseo *de que la reunión comience a la hora*’, ‘mi intención *de ir al cine esta noche*’, ‘mi recuerdo *de la visita del Papa*’, ‘mi opinión *de que el segundo cinturón no estará listo para el 92*’).

Dos estados mentales diferentes pueden ser estados idénticos, en el mismo sentido en que dos corbatas diferentes pueden ser idénticas o dos coches diferentes idénticos: pueden ser estados diferentes del mismo tipo y con el mismo contenido. Yo puedo tener, en dos intervalos temporales diferentes, el mismo recuerdo: el recuerdo de la visita del Papa. Y dos sujetos diferentes pueden tener, quizás en el mismo intervalo temporal, el mismo recuerdo: el recuerdo de la visita del Papa. Quizás el lenguaje común sea tan poco específico, para la función de la descripción del contenido de los estados mentales, que alguno pueda sentir una cierta insatisfacción al oír que dos recuerdos de la visita del Papa puedan ser idénticos, meramente por ser recuerdos de la visita del Papa. Pero ciertamente pueden serlo: dos sujetos que vieron en directo la derrota del Madrid ante el Milán por cinco a cero desde una posición similar en el campo pueden, ciertamente, tener recuerdos *con el mismo contenido* del partido (aunque describirlos mediante el lenguaje resultaría un tanto engorroso; una película resultaría más adecuada).

La estructura común de nuestras referencias a lo mental es, pues, S-t- σ -[t], donde S indica al sujeto, τ el tipo de estado, σ la descripción del contenido y t el intervalo temporal. Los corchetes indican que la especificación del tiempo se deja habitualmente a factores contextuales. A cualquier conjunto del tipo τ - σ (‘recuerdo de la visita del Papa’ ‘creencia de que dos más dos son cuatro’, ‘deseo de que Margaret Thatcher pierda las próximas elecciones’), omitiendo específicamente el sujeto del estado y el tiempo en que ocurre, le llamaré ‘descripción mental’, o ‘descripción mentalista’. Las descripciones mentales son descripciones, no de estados mentales concretos y específicos, sino de propiedades de, o

características comunes a, estados mentales concretos: propiedades comunes a diferentes estados mentales concretos. Y los estados o procesos mentales son estados o procesos descritos mediante descripciones mentales.

Supongamos que estoy percibiendo una esfera uniformemente azul sobre un fondo uniformemente rojo (una fotografía serviría en este caso mejor que el lenguaje común para describir el contenido de mi estado mental; porque percibir es, ciertamente, un tipo de estado mental). En alguna parte de mi cuerpo pasan cosas que suponemos correlacionadas de algún modo con mi estado mental; en particular, suponemos que hay un proceso físico que comienza en mi retina y acaba en mi cerebro, de algún modo correlacionado con mi estado mental. Este proceso puede, en principio, ser descrito en términos puramente físicos (en términos neurológicos), aunque hoy no podamos especificar sus detalles: determinadas células retinianas que se excitan, otras que se inhiben, procesos bioquímicos que transmiten la excitación o la inhibición de unas neuronas a otras, paso por los cuerpos geniculados, excitación e inhibición final de determinadas células del cortex. La correlación entre el estado mental y el proceso físico es, en principio, temporal: a la vez que ocurre la percepción, ocurre el proceso físico. Como mínimo, ambos procesos son *paralelos*.

Porque también el proceso físico ocurre en el tiempo, por supuesto. Y, en este caso, no hay dudas sobre su ubicación espacial. Por otro lado, este proceso físico concreto mío, que tiene lugar durante este momento en que estoy percibiendo la esfera azul sobre el telón uniformemente rojo, también se describe de un modo general: si omito la indicación temporal, y la ubicación espacial, puedo observar que el mismo proceso físico puede ocurrir en este mismo momento en diferentes sujetos, y que el mismo proceso (un proceso de las mismas características físicas, quiero decir) puede haber tenido o tener lugar en mí en otras ocasiones. Esto es, en todo lo que respecta a las partes de sus cuerpos implicados en el proceso en cuestión, dos sujetos diferentes pueden ser idénticos (de nuevo, en el sentido en que dos automóviles diferentes pueden ser idénticos). A una descripción general de un proceso como la que he intentado ejemplificar, omitiendo los detalles espaciales y temporales, la denominaré ‘descripción física’.

Repárese en la parte de una descripción mental que nos da el contenido del estado mental descrito. En esa parte aparecen expresiones (‘el Papa’, ‘el II Cinturón’, ‘Margaret Thatcher’, ‘visitar’, ‘finalizarse [una obra]’, etc.) que usualmente se usan para describir el mundo externo. A primera vista, utilizamos, en nuestras descripciones de lo mental, objetos externos, propiedades que ejemplifican objetos externos, y relaciones en las que pueden estar objetos externos, para individualizar parcialmente nuestra vida mental. La «intencionalidad», su «estar dirigidos hacia» o, mejor, su estar caracterizados (en parte) por referencia a objetos externos y propiedades ejemplificadas por objetos externos, es la marca característica de lo mental.

La intencionalidad de los estados mentales está relacionada con el papel que juegan las descripciones mentales en nuestro esquema conceptual. Las descripciones mentales ocupan un lugar destacado en la explicación de la acción. Explicamos parte de nuestra conducta, y de la conducta de los demás, atribuyéndonos estados mentales tales como percepciones, saberes, recuerdos, conjeturas, creencias, y también intenciones, deseos, esperanzas, querer, necesidades, etc. Explicamos así aquellas de nuestras conductas que son acciones,

que no son simplemente cosas que nos pasan (como lo son el movimiento reflejo de nuestra rodilla cuando se nos golpea de cierto modo, o las cosas que «hacen» un individuo sometido a hipnosis o un sonámbulo) sino cosas que nosotros hacemos. En el caso prototípico, explicamos una acción atribuyendo al sujeto de la misma un estado del tipo de los enumerados en primer lugar (un estado *doxástico*) y un estado del tipo de los enumerados a continuación (un estado *volitivo*)

¿Por qué el acusado disparó a la víctima?. Porque la víctima estaba a punto de estrangular a la niña. Naturalmente no porque la víctima estaba a punto de estrangular a la niña, sino porque el acusado lo *creía* así (una línea prometedora para el fiscal es mostrar que, pese a ser cierto que la víctima estaba a punto de estrangular a la niña, el acusado no podía creerlo así, dado que desde su posición no podía ver completamente la escena). Y, naturalmente también, no sólo por eso: el acusado podía haber creído lo mismo y no haber disparado a la víctima, si hubiera deseado que la niña fuese estrangulada, o le hubiese resultado indiferente, o hubiese preferido cualquier cosa a usar un arma, etc. Así, disparó a la víctima porque creía que la víctima estaba a punto de estrangular a la niña, no quería que la niña sufriese tal suerte, y juzgó que un buen medio para evitarlo era disparar a quien así habría de convertirse en la víctima. Y este tipo de explicaciones valen para cualquier acción racional, por humilde que sea; no sólo para aquellas acciones que nos obligan a deliberar, dada su especial trascendencia. ¿Por qué te desviaste del camino?. Porque había un charco. Naturalmente, no porque había un charco, sino porque yo *observé* que había un charco; y no sólo por eso: porque no quería mojarme y pensaba que un medio para no mojarme apropiado era desviarme del camino.

Los estados internos que, intencionalmente descritos postulamos en las explicaciones de la acción poseen dos características fundamentales: *racionalizan* la acción (nos la hacen «inteligible» de un modo casi empático) y, en contra de Anscombe, Habermas o Von Wright, *causan* la acción. Mi desear que ésta porcelana que me regalaron se rompa, y mi pensar que la porcelana se rompería si la arrojo al suelo con cierto movimiento de mi mano no explican que realice ese movimiento, si la causa del mismo no es ese estado interno mío que acabo de describir como una combinación de cierta creencia y cierto deseo (ese estado interno que, intencionalmente descrito, racionalizaría mi conducta), sino un impulso reflejo causado por un tic nervioso, pongamos por caso, o si la causa ha sido la intención resultante de mi creencia de que romper esta porcelana en este preciso momento alejaría a los malos espíritus y mi deseo de alejar a ciertos malos espíritus³³.

Las razones de Davidson para sostener que los estados mentales que hacen inteligible (que racionalizan) un acaecimiento deben ser su causa, en el sentido usual del término, son contundentes. Pensemos en lo que nosotros llamamos ‘racionalizaciones’. El sujeto sostiene que si denunció a Pedro fue porque creía que participaba en una conspiración contra el Estado, y no estaba dispuesto a permitir que una conspiración contra el Estado continuase adelante. Nosotros pensamos que si denunció a Pedro fue porque envidiaba la posición académica de Pedro, y creía que denunciándolo la rebajaría notablemente. ¿En qué consiste nuestra conjetura?. No en que el sujeto no tuviese, de hecho, las creencias y deseos que dice

³³ Cf. D. Davidson: ‘Actions, Reasons and Causes’, y ‘Intending’. Ambos en *Essays on Actions and Events*.

tener. Podemos dar por bueno que las tiene. Lo que pensamos, sin embargo, es que esas creencias y deseos no son causalmente eficaces, en lo que a la denuncia se refiere; que, aunque no las hubiese tenido, igualmente hubiese denunciado a Pedro. ¿Existe alguna diferencia entre el sentido en que empleamos el concepto de *causalidad* en consideraciones como estas, y el sentido en que lo empleamos en consideraciones que sólo afectan a acaecimientos físicos?. Juan ingirió una plato de *amanita phalloides*; pero la ingestión no fue causalmente eficaz en su muerte, porque la muerte la produjo un repentino infarto, por completo independiente de la ingestión de las setas: aunque no hubiese comido las setas, habría muerto igual. ¿Hay alguna diferencia conceptual entre estas consideraciones y las anteriores?.

Davidson sostiene que no, y sus argumentos son considerablemente fuertes. Máxime, si, como veremos, existe otro modo de explicar esa diferencia que muchos parecen sentir en cuanto a la participación de conceptos de lo mental en explicaciones causales. Un acaecimiento físico puede no ser un «simple acaecimiento físico»; puede ser una acción. La víctima murió en un terrible accidente; una causa del mismo fue la rotura de la dirección.

Pero la rotura de la dirección no fue un «simple acaecimiento físico»; no fue un «acaecimiento fortuito». Fue la acción del marido de la víctima. Si fue la acción del marido de la víctima, existen ciertos estados mentales del marido de la víctima que racionalizan lo que hasta ahora parecía un acaecimiento fortuito, que nos permiten verlo como una acción. Pero esos estados mentales deben tener «algo que ver» con la rotura de la dirección; no basta que el sujeto desee asesinar a su esposa y crea que rompiendo la dirección lo podrá conseguir. Y, ¿qué puede ser tal «tener algo que ver», salvo *causar*?

Creo que las consideraciones anteriores constituyen un fuerte argumento en contra de Habermas, Anscombe o Von Wright. El argumento precisaría ser completado con un examen crítico de los propios argumentos de estos filósofos, un examen que muestre por qué no son aceptables; y con una explicación alternativa de las peculiaridades que ellos encuentran en las explicaciones mentalistas. No puedo llevar a cabo lo primero aquí, pero sí hemos de ocuparnos de lo segundo (naturalmente, desde el punto de vista de Davidson).

El monismo anómico

El punto fundamental es éste. Podemos aceptar (i) que los acaecimientos mentales (algunos, al menos) están causalmente relacionados con acaecimientos físicos (digamos «la interacción causal entre lo mental y lo físico», IMF)³⁴; aceptar también (ii) que las relaciones causales entre acaecimientos singulares entrañan relaciones de dependencia causal entre tipos de acaecimientos (digamos «el carácter nomológico de la causalidad», CNC), sin que ello nos obligue lógicamente a no aceptar (iii) la no existencia de leyes causales que conecten lo mental y lo físico (digamos «la anomalía de lo mental», AM).

³⁴ En el texto sólo hemos examinado una dirección de la relación causal entre acaecimientos mentales y acaecimientos físicos: la producción causal de acaecimientos físicos por acaecimientos mentales. Pero existen razones, esencialmente del mismo tipo, para argumentar que las percepciones y la memoria están causalmente conectadas con acaecimientos externos; y que las creencias deben estarlo también, si bien de un modo mucho más indirecto. Cf. Paul Grice, 'The Causal Theory of Perception', en *Studies in the Ways of Words*; David Lewis, 'Veridical Hallucination and Prosthetic Vision' en *Philosophical Papers*, vol. 2; y Fred Dretske, *Knowledge and the Flow of Information*, Basil Blackwell, Oxford, 1981.

Justamente la conjunción de las tres tesis, aparentemente contradictorias, es lo que sostiene Davidson en sus tres más conocidos artículos sobre el tema que nos ocupa³⁵. En mostrar la compatibilidad de las tres tesis reside su originalidad; y la defensa de (iii) es su versión de la tesis tradicional sobre la diferencia entre explicar y comprender.

Veamos por qué las tres tesis no son contradictorias, y qué hemos de aceptar como consecuencia de sostener las tres conjuntamente. De acuerdo con (i), hay aserciones verdaderas que establecen relaciones causales singulares entre acaecimientos mentales y acaecimientos físicos. Tengamos a la vista dos ejemplos: mi caída de rodillas en el brasero causa mi presente recuerdo de mi caída en el brasero; mi intención de levantar el brazo izquierdo causa de la elevación de mi brazo izquierdo. De acuerdo con (ii), el acaecimiento causa ejemplifica cierto acaecimiento-tipo, en ambos casos; el acaecimiento efecto ejemplifica cierto acaecimiento-tipo, también en ambos casos, y existe una ley que hace depender causalmente los acaecimientos del segundo tipo de los acaecimientos del primer tipo. Ahora bien, como ya sabemos (el caso del viaje en martes y la muerte de Juan), los conceptos en términos de los cuales formulemos la ley causal implicada no tienen por qué incluir los conceptos mediante los que caracterizamos los acaecimientos concretos a los que nos referimos; la *explicación* más informativa de que la caída cause mi recuerdo no tiene por qué mencionar recuerdos para nada. Las descripciones mediante las que nos referimos a los acaecimientos pueden ser completamente accidentales -en cuanto a su eficacia causal. En particular, resulta compatible con (iii) que la ley en cuestión haga depender ciertos procesos neurofisiológicos de la caída en el brasero; y haga depender la elevación de mi mano de ciertos procesos neurofisiológicos. Para ello, lo que necesariamente debe ocurrir es que mi intención sea *también* un cierto proceso neurofisiológico; y que mi recuerdo sea *también* un cierto proceso neurofisiológico. Es decir, la consecuencia de aceptar la conjunción de las tres tesis es el monismo: los acaecimientos mentales son también acaecimientos físicos.

Este es, en resumidas cuentas, el punto de vista de Davidson. Si clasificamos las teorías sobre la mente en *monistas* y *dualistas*, según que acepten o no la identidad entre acaecimientos mentales y acaecimientos físicos, y por otro lado en *nómicas* y *anómicas*, según que acepte o no la existencia de leyes de dependencia entre tipos de acaecimientos mentales y tipos de acaecimientos físicos, podríamos clasificar la posición de Davidson como un *monismo anómico*³⁶.

Antes de examinar con algún mayor detalle el punto de vista de Davidson, conviene reparar en que, pese a ser lógicamente coherente, es sumamente contraintuitivo. Una vez convencidos de la interacción causal entre lo mental y lo físico, resulta muy difícil aceptar la anomalía de lo mental. Veámoslo de este modo. El disparo de Juan a Pedro causó la muerte de Pedro. Ello es verdad; pero decirlo en respuesta a la pregunta '¿por qué murió Pedro?', en un contexto usual, provocaría asaz confusión; porque estaríamos sugiriendo una

³⁵ 'Mental Events', 'Psychology as Philosophy' y 'The Material Mind', los tres en la recopilación *Essays on Actions and Events* ya citada. (Hay traducción castellana del primero con el título de 'Sucesos mentales' en Cuadernos de Crítica, México, 1981; y del tercero, con el título de 'La mente material', en la recopilación *Mentes y Máquinas*, Tecnos, 1985).

³⁶ La clasificación es de Davidson. Un buen ejercicio es intentar catalogar las doctrinas tradicionales y contemporáneas sobre lo mental en este sistema clasificatorio.

«cadena causal» que no es la que funcionó en este caso. Aquí estamos ante una «cadena causal desviada»: el disparo de Pedro causó a Pedro un infarto que le produjo la muerte; la bala, por lo demás, describió milagrosamente una trayectoria por el cuerpo de Pedro que poco podría haber afectado a su salud. Una cadena causal desviada es un proceso causal en el que las leyes que conectan el acaecimiento tipo ejemplificado por el acaecimiento causa con el acaecimiento-tipo ejemplificado por el acaecimiento efecto son muy distintas a aquellas que, usualmente, conectan acaecimientos del tipo de la causa con acaecimientos del tipo del efecto.

En el análisis de fenómenos que típicamente involucran lo mental (como el concepto de acción *intencional*) debemos tomar en consideración estas cadenas causales desviadas. Aunque mi arrojar la porcelana al suelo con un movimiento de mi mano esté causado por mi desear que esta porcelana que me regalaron se rompa, y con mi pensar que la porcelana se rompería si la arrojo al suelo con cierto movimiento de mi mano, el primer acaecimiento no es una acción mía, si lo que ha ocurrido es que el formar esa creencia y ese deseo me ha producido una excitación nerviosa tal (quizás movido por cierto sentimiento de culpa) que ha causado que moviera mis brazos («accidentalmente») con el resultado indicado.

Que ciertas creencias y deseos causen cierto movimiento corporal no basta para que aquellas lo racionalicen, nos permitan verlo como una acción intencional, si la cadena causal es desviada. El mismo Davidson nos ha mostrado cómo comprender muchos fenómenos de irracionalidad en este marco, como casos de cadenas causales desviadas³⁷.

Lo importante es comprobar que la existencia de cadenas causales desviadas que involucran lo mental entraña la existencia de leyes causales que involucran lo mental si la caracterización que he hecho antes de una cadena causal desviada es correcta.

Ello no debe considerarse más que como una indicación de lo que parece obvio: que no sólo creemos que la intención cause la elevación del brazo, sino que creemos que hay leyes que hacen depender, *ceteris paribus*, las elevaciones del brazo de las intenciones de elevarlo; que no sólo creemos que escuchar un chiste verde causa sonrojo, sino que hay alguna ley causal que hace depender, *ceteris paribus*, los sonrojos de las audiciones de chistes verdes.

Pero también el supersticioso cree que el viaje en martes de Juan causó su muerte, no sólo en el sentido en que lo creemos todos; él cree también que hay una ley causal que hace depender las muertes y otras desgracias de los viajes en martes. Pero no está en lo cierto; y la razón para creerlo nos la suministra el que tenemos buenas razones para creer que no hay modo general de reducir su «ley» a leyes estrictas; que, por muchas salvedades *ceteris paribus* que le supongamos, ‘viajar en martes’ no menciona ninguna propiedad causalmente eficaz de los acaecimientos. La posición de Davidson consiste en que nuestra creencia sobre la eficacia causal de la mente es similar a la del supersticioso, porque no es razonable creer en la existencia de leyes causales *strictas* que conecten lo mental y lo físico³⁸.

³⁷ Cf. su análisis de la debilidad de la voluntad, o incontinencia, en ‘How is Weakness of the Will Possible?’, en la recopilación *Essays on Actions and Events* ya citada, y su análisis del autoengaño, el pensamiento desiderativo y otros fenómenos similares en ‘Deception and Division’, en la recopilación de LePore y MacLaughlin *Perspectives on the Philosophy of Davidson: Actions and Events* también citada.

³⁸ Este es el lugar en que advertir que los argumentos de Davidson conciernen a las disciplinas en cuestión entendidas como se las entiende habitualmente; es decir, entendidas como disciplinas que involucran leyes en que

Vamos, pues, a examinar las razones de Davidson. Observe el lector que el argumento es completamente *a priori*, puramente conceptual. Y permítame que le recuerde, una vez más, lo que el argumento *no* pretende. Tomemos una vez más un ejemplo para fijar nuestras ideas: la audición por parte de María de una historia escabrosa causa su sonrojo. Davidson intentará mostrarnos que el predicado ‘audición de una historia escabrosa’ no refiere a una propiedad causalmente eficaz en la producción del sonrojo de María; es decir, que no hay una ley causal formulable mediante cláusulas *ceteris paribus* que haga depender los sonrojos de las audiciones de historias escabrosas. Y que no la hay, porque no hay un modo más estricto de describir procesos similares haciendo depender los estados microfísicos correspondientes a sonrojos de ciertos estados neurofisiológicos correspondientes a audiciones de historias escabrosas.

Quizás el lector espere que el argumento de Davidson se apoye en la posibilidad de «múltiple realizabilidad» de los estados mentales; en que el estado neuronal correspondiente a la audición de una historia escabrosa puede ser diferente en diferentes sujetos (no digamos si incluimos entre los sujetos posibles robots y marcianos, en los que los estados físicos correspondientes no tienen por qué ser ni siquiera neurológicos); en que incluso es razonable pensar que pueda ser diferente en María en diferentes momentos de tiempo.

Davidson rechaza enfáticamente que este sea su argumento; y es natural que lo haga, porque esa posibilidad ya la habíamos tenido en cuenta. No es razonable esperar que todos los procesos consistentes en la producción de un sonrojo por la audición de una historia escabrosa sean microfísicamente idénticos; pero, como hemos visto, tampoco es razonable esperar que todos los procesos consistentes en la producción de una pérdida de cosecha por una granizada sean microfísicamente idénticos, y ello no invalida la posibilidad de la reducción. (La realidad de la reducción, en este caso).

Lo que está en cuestión es que el estado de oír una historia escabrosa, en María ahora, corresponda a un cierto estado neurológico; y que si ese estado, en María dentro de un cierto tiempo, pasa a corresponder a otro estado neurológico, ello sea explicable, quizás porque se ha producido algún cambio en el sistema conceptual de María (el aprendizaje de algún nuevo concepto, por ejemplo). Lo que está en cuestión es que exista una ley que identifica el estado de oír una historia escabrosa en sujetos que han tenido un aprendizaje de un tipo similar, en los aspectos cognitivos relevantes, al de sujetos físicamente similares a María, con un tipo de estado neurológico. Quizás el único sujeto de hecho existente con tales características sea María, y María no durante mucho tiempo. Pero la afirmación continúa teniendo carácter general, y existen posibilidades empíricas de ponerla en cuestión. Y, cuando los conceptos implicados son mucho menos dependientes de la historia personal (‘percibir una esfera gris sobre fondo rojo’, por ejemplo), las consecuencias empíricas de postular una ley psicofísica así son mucho más comprometidas.

se hace uso de términos tales como ‘creía que p’, ‘sabía que p’, ‘tenía la intención de que p’, etc... Una psicología que pretenda explicar los movimientos corporales en términos neurofisiológicos, de la que se haya desterrado la intención de explicar las ilusiones perceptivas, los errores lógicos al razonar, la incidencia desestabilizadora de nuestros deseos en nuestras creencias, etc., sí podría ser (por lo que hace a los argumentos de Davidson) una disciplina científica, pero constituiría un «cambio de tema», como dice acertadamente Davidson.

Davidson sostiene que, incluso aclarada la cuestión como lo acabamos de hacer, no es razonable, a priori, creer que existan leyes así. Su argumento es, en mi opinión, una clara elucidación de las verdaderas razones que todos tenemos para ser dualistas, para creer que el alma es «otra cosa». Las páginas de la *Crítica de la razón pura* sobre la antinomia de la libertad y la causalidad son otra elucidación de esas oscuras razones.

Este es, por fin, el argumento de Davidson, según yo lo interpreto. La atribución de estados mentales a un sujeto está sujeta a ciertos criterios de aplicación, criterios constitutivos de nuestro concepto de lo mental. No tiene sentido atribuir el concepto de rojo -o quizás el concepto de número-, pongamos por caso, a quien, sin explicación, manifiesta aplicarlo a un número. Puede que en casos particulares tenga sentido atribuir a alguien la intención de hacer $\emptyset$ en el momento de tiempo t , aunque, llegado t , y sin explicación, el sujeto no haga $\emptyset$. Pero no tiene sentido atribuir intenciones a alguien que, en general, y sin explicación, «procede» de ese modo. Tiene sentido atribuir el concepto del cuantificador universal a alguien que, en alguna ocasión, comete, sin explicación, un error lógico patente (cree que todos los árboles del bosque son pinos, y se asombra de que este árbol del bosque sea un pino). Pero no tiene sentido atribuir el concepto del cuantificador universal a quien continuamente comete este tipo de errores. Por resumir: la atribución de estados mentales a un individuo viene condicionada por criterios normativos, criterios de racionalidad teórica y práctica que el individuo debe satisfacer. Tales criterios, como todas las leyes que involucran lo mental, no pueden formularse más que de un modo general (dando por supuestas cláusulas *ceteris paribus*, claro); pero eso no les resta un sólo ápice de fuerza.

Supongamos que, en un viaje de exploración espacial, caemos en un cierto planeta; y que en ese planeta encontramos ciertos sujetos. Para seguir el experimento mental, elimine el lector toda expectativa apriorística. Imagine a los individuos en cuestión con una apariencia ajena por completo a la de los seres humanos, y no los imagine dados a actividades patentemente humanas. ¿Qué condiciones habrían de cumplirse para que fuera correcto decir de esos sujetos que creen o desean, o, más aún, para decir que llevan a cabo ciertas actividades de carácter lingüístico, mediante las cuales expresan tales y cuales pensamientos, mandatos o preguntas?. Quizás no sea necesario hacerles «creyentes (por lo general) de verdades y amantes (igualmente por lo general) de lo bueno (según nuestras propias luces, no hace falta decirlo)», como dice Davidson. Quizás sean concebibles situaciones en las que conseguimos atribuir a los sujetos creencias y deseos, así como la capacidad de expresar sus creencias y deseos lingüísticamente, aun cuando muchas de las creencias que les suponemos son falsas y muchos de los deseos que les suponemos son absurdos. Pero, en tal caso, debe ocurrir cuando menos lo siguiente: que, supuestas las experiencias que ellos han tenido, nosotros mismos hubiésemos construido un sistema de creencias tan inadecuado como el que les suponemos; y, supuestas sus creencias y sus necesidades fundamentales, nosotros hubiésemos construido un sistema de valores tan absurdo como el que les suponemos. Es decir, debemos suponerles, como mínimo, un sistema de formación de creencias correcto y un sistema de formación de valores correcto (según nuestros propios criterios, no hace falta decirlo)³⁹.

³⁹ La enmienda al modo en que Davidson formula sus principios de racionalidad la tomo de Lewis, una vez más, Cf. 'Radical Interpretation', en *Philosophical Papers*, vol. 1.

La atribución de estados mentales está sujeta, pues, a que los estados atribuidos satisfagan criterios globales de racionalidad teórica y práctica. El argumento de Davidson consiste en mostrar la implausibilidad de creer que los estados neurológicos que nuestras supuestas leyes psicofísicas hacen corresponder a los estados mentales, satisfagan *también* esos criterios globales de racionalidad. Pero si hay leyes psicofísicas, necesariamente deben satisfacerlos igualmente, aunque de modo indirecto.

Yo creo que todos los árboles de este bosque son chopos. Esta creencia mía, si hay leyes psicofísicas, es también un estado neurológico mío de un cierto tipo. Quizás no siempre que he creído tal cosa he estado en ese mismo estado físico; quizás ese estado físico, en otros seres humanos (o incluso en mí mismo, en otros momentos de tiempo) corresponde a otra creencia. Pero hay un orden nómico en todo ello. Sea ϕ el tipo de estado neurológico en cuestión. Un momento después, formo la creencia de que este objeto es un árbol de este bosque. Sea F el estado neurológico correspondiente. *Ceteris paribus*, ha de creer también que este árbol es un chopo. Sea σ el estado neurológico correspondiente. El problema que Davidson nos presenta es que, mientras los elementos normativos consiguientes a la atribución de estados mentales hacen que, con todas las salvedades genéricas que se quiera, yo *deba* tener la última creencia, supuesto que tengo las otras dos, nada, supuestas todas esas salvedades, hace que alguien que está en el estado ϕ y también en el estado Φ *haya de* estar en el estado σ .

Podríamos haber presentado la dificultad con problemas derivados de la normatividad práctica, en lugar de hacerlo con problemas derivados de la normatividad teórica de nuestros conceptos de lo mental. Mi actual intención de ir a Barcelona mañana, junto con mi opinión de que, considerados todos los factores, el tren es el mejor medio de transporte para este caso, hacen que, con salvedades que no pueden ser detalladas completamente, yo *haya de* tener la intención de ir a coger el tren mañana por la mañana. Parece sin embargo sorprendente decir que el estado neurológico que corresponda a mi actual intención, junto al que corresponda a mi opinión, *hayan de* causar el estado que corresponda a mi intención de mañana⁴⁰.

Puede verse así hasta qué punto el argumento es similar a aquellos otros argumentos que se apoyan en la perplejidad que produce decir que la producción libre, intencional, de un cierto movimiento corporal pueda verse también como la producción causal de un tipo de acaecimiento físico por otro tipo de acaecimiento físico⁴¹. Si hubiera leyes causales que hacen depender acaecimientos del tipo del acaecimiento consistente en la elevación de mi mano de acaecimientos del tipo del acaecimiento de mi intención de elevar mi mano, descritos esos acaecimientos por sus características neurológicas, ¿no podría entonces, en

⁴⁰ El artículo de Jaewon Kim 'Psychophysical Laws', en la recopilación de LePore y McLaughlin antes citada, me ha ayudado mucho a comprender el argumento de Davidson -argumento expuesto, como es habitual en él, de un modo excesivamente elusivo. (El lugar en que con una relativamente mayor claridad aparece el argumento es 'Mental Events', págs. 221-223; cf. también 'Psychology as Philosophy', pág. 223 y 239-239). De acuerdo con la práctica que he seguido a lo largo del trabajo, he tratado de exponer el argumento de Davidson sin hacerlo depender de sus propias opiniones sobre la naturaleza de la interpretación del lenguaje y de la mente. Con esa dependencia, el argumento me parece más débil de como lo he presentado.

⁴¹ El propio Davidson enfatiza la similitud en 'Mental Events'.

principio, predecirse mi acción?, y, lo que es peor, ¿no podría provocarse de modo ajeno a mi voluntad?. ¿No haría todo eso que mi acción no fuera ya libre?. Tengo que elegir entre escribir una carta esta tarde o leer un artículo. A lo largo de mi elección, voy a examinar consideraciones en favor de formar una u otra intención. Unas u otras consideraciones me inclinarán finalmente por una u otra elección. Pero, si hay una ley causal que lleva del estado neurológico consistente en mi examen de las consideraciones al estado neurológico consistente en la formación de la intención de leer el artículo, ¿por qué molestarme en examinar esas consideraciones?.

En esto radica la diferencia, según Davidson, entre la psicología (o la sociología, la lingüística, la economía y la historia, en la medida en que estas disciplinas involucran conceptos mentalistas), por un lado, y la biología, la meteorología o la geología. Aunque las leyes en que descansan las explicaciones biológicas, geológicas o meteorológicas no son estrictas, existen (o es razonable creer que existen) procedimientos generales (leyes reductoras) para describir cada proceso causal caracterizado en los términos de esas disciplinas en términos tales que involucrarían únicamente leyes estrictas -en términos microfísicos. Y esas disciplinas no son prescindibles, porque las necesidades explicativas que demandamos de ellas quedan convenientemente satisfechas con sus leyes *ceteris paribus* (mejor de lo que quedarían si prescindiésemos de las leyes de la genética de poblaciones, pongamos por caso, en favor únicamente de leyes formuladas en términos cuánticos).

Como ya he indicado, el argumento de Davidson no me parece decisivo. Sin embargo, muestra algo importante: qué precio hemos de pagar para seguir creyendo, como creemos ingenuamente, en la eficacia causal y en la virtualidad explicativa de las características mentales. Mostrar que ese precio puede pagarse exige entrar en un examen detallado de nuestros conceptos de lo mental. Habiendo rebasado con creces el espacio que se me ha concedido, debo dejar ese examen para otra ocasión.

Angel García-Carpintero
Septiembre, 1989



PANORAMA ACTUAL DE LA FILOSOFIA ANALITICA DE LA MENTE: FUNCIONALISMO Y EXPERIENCIA

Vicente Sanfélix Vidarte

I. Introducción

Según reza el programa de estas sesiones, me toca a mí bosquejarles el panorama actual que la tradición analítica presenta en ese campo al que normalmente se denomina «filosofía de la mente», «psicología filosófica» o «filosofía de la psicología». Antes de nada, quisiera hacer algunas precisiones de cara a que mi tarea quede claramente circunscrita.

La primera de ellas es que no deja de resultar problemático hablar del panorama actual de la filosofía analítica en cualquier campo de investigación, dado que es harto discutible que hoy en día exista la filosofía analítica. En realidad, no estoy muy seguro de que haya existido nunca. Lo más que me parece indiscutible es que ha habido distintos filósofos - Moore, Russell, Wittgenstein, Carnap, Ryle,... y otros muchos más a los que se suele calificar de «analíticos». Y que sigue habiendo filósofos -Putnam, Davidson, Armstrong, Shoemaker... y otros muchos- cuyo pensamiento se ha construido fundamentalmente en diálogo y discusión con aquellos. Cuando hablamos, pues, del panorama actual de la filosofía analítica, debe entenderse que estamos utilizando una expresión más o menos cómoda para referirnos a las posiciones de una serie de autores que sería más exacto calificar como los descendientes de los filósofos analíticos, y que no tienen por qué compartir, ni por lo general comparten, ni un conjunto de tesis, ni un método, ni tan siquiera una concepción de lo que la filosofía sea o deba ser.

La segunda precisión es relativa al tipo de problemas que voy a tratar. Generalmente los libros cuyos propios autores califican como de filosofía de la mente (o psicología filosófica, o filosofía de la psicología) presentan una gran variedad de temas: la naturaleza de lo mental y su relación con el cuerpo, la identidad personal y el autoconocimiento, el problema de las otras mentes, de la percepción, el carácter de la acción intencional, de la comunicativa, el problema del libre albedrío... Presentar el estado de la discusión de todos estos temas es algo que no sólo sería incapaz de hacer por falta de tiempo sino también, me temo, por falta de competencia. De manera que me limitaré a hablar de la situación de uno sólo de estos temas, el más general y a la vez central de la filosofía de la mente, aquel que ningún libro que verse sobre la misma puede obviar: el de la naturaleza de lo mental, también conocido en la literatura analítica como el problema mente-cuerpo. Obviamente se trata de un tema metafísico, pero supongo que a nadie se le escapará que tiene connotaciones epistemológicas y metodológicas claras, y que la manera como se afronte está en función, entre otras cosas, de los presupuestos lógicos de los que partamos.

Hechas estas precisiones, voy a intentar presentar la manera como los filósofos descendientes de los filósofos analíticos han afrontado el problema mente-cuerpo.

II. Origen y características generales del funcionalismo

Sin embargo, así circunscrito, el problema parece muy asequible. Se trataría simplemente de discutir la viabilidad del funcionalismo, pues el funcionalismo, si no una posición universalmente compartida entre los actuales filósofos de la mente, es, sin duda, la tesis más difundida entre los mismos (Dennett, 1978d. Fodor, 1985. Putnam, 1988), de manera que la actual discusión sobre el problema mente-cuerpo parece concretarse en una discusión sobre las virtudes o/y limitaciones del funcionalismo.

El problema estriba en que, como ya pasara con otras doctrinas de lo mental que en un momento u otro llegaron a convertirse en oficiales, tales como el dualismo, el conductismo, o el materialismo, bajo un rótulo general se esconden en realidad una enorme disparidad de puntos de vista (Block, 1980b), de manera que nada impide a un autor ser funcionalista en un sentido -por ejemplo, en el de reconocer que el funcionalismo es el único enfoque que suministra a la psicología una estrategia de investigación coherente- y no serlo en otro -negando, por ejemplo, que un análisis funcional de nuestros predicados psicológicos pueda hacer justicia al significado de los mismos (a mi entender, esta es probablemente la posición de Stich, 1983)-. Debiéramos empezar, pues, por aclarar lo que se entiende por «funcionalismo», y para ello, como siempre que se trata de analizar una posición filosófica históricamente concretada, nada mejor que remontarnos un poco a su génesis.

En realidad, el funcionalismo puede verse como la respuesta lógica a las insuficiencias que presentaban las doctrinas de lo mental más aceptadas antes de él: el conductismo lógico y el materialismo reductivo, o tesis de la identidad de los (tipos de) estados mentales con (tipos de) estados físicos. Estas insuficiencias son múltiples y pueden clasificarse de distintas maneras, pero para nuestros actuales propósitos bastará que digamos que eran básicamente de dos tipos: intuitivas y lógicas. Veamos.

Aunque el rótulo «conductismo lógico» se suele aplicar a las posiciones muy diferentes de autores muy distintos (Carnap, 1932; Hempel, 1935; Ryle, 1949; Wittgenstein, 1953; Strawson, 1959), muchos de los cuales, por ende, habrían rechazado tal calificativo caso de que se les hubiera propuesto, la tesis básica de esta posición puede sintetizarse diciendo que según la misma el significado de los predicados psicológicos debe ser analizado haciendo referencia no a procesos o entidades internas y privadas, sino a un conjunto de conductas públicamente observables efectivas o potenciales. Desde este punto de vista, decir de alguien que piensa que está lloviendo, o que tiene dolor de muelas, no es atribuirle ningún tipo de proceso interno o de experiencia privada, sino decir que se comporta o está dispuesto a comportarse de cierta manera, de forma análoga a como -por utilizar un ejemplo tópico de la literatura conductista- decir que el azúcar es soluble en el agua no es sino decir que si ponemos un terrón de azúcar en un vaso de agua, se disolverá.

Las objeciones intuitivas y lógicas del conductismo lógico, desarrolladas con profusión por los primeros funcionalistas (Putnam, 1965; Chichara y Fodor, 1965; Fodor, 1968; Armstrong, 1968), son obvias. Para empezar parece claro que cuando hablamos de procesos y de eventos mentales lo hacemos muy en serio. Cuando decimos que alguien está pensando, lo que decimos es exactamente eso: que está pensando; o sea que está realizando un proceso mental, y no simplemente que está dispuesto a contestar nuestras preguntas de cierta manera, o a comportarse de cualquier otro modo. Y de la misma manera, cuando decimos

de alguien que tiene dolor lo que decimos es algo muy diferente a que se queja o que está dispuesto a quejarse; lo que estamos haciendo es atribuirle algo más sustantivo: una sensación dolorosa. Que lo esencial en nuestra atribución de predicados psicológicos son, precisamente, los procesos y eventos internos que el conductista quiere dejar fuera de su significado es algo que nos parece muy claro desde el momento en que su inexistencia haría falsas nuestras atribuciones de los mismos a cualquier sujeto. Por más que un individuo se comporte como si estuviera pensando, o como si tuviera dolor, si no está realmente pensando o teniendo dolor será falso decir de él que esté pensando o teniendo dolor. Pero si la existencia de los procesos y los eventos internos es tan decisiva para las condiciones de verdad de las proposiciones en las que intervienen predicados psicológicos, difícilmente será creíble que no tenga nada que ver con el significado de los mismos. Puede resultarnos más o menos creíble la existencia de un proceso o evento mental que no tenga manifestación conductual alguna (Putnam, 1965), pero lo que es intuitivamente inadmisibles es que la corrección en la atribución de ciertos predicados psicológicos no dependa para nada de la existencia de ciertos procesos y eventos mentales.

Emparentada con esta intuición está una objeción lógica que los funcionalistas han solido hacer contra el conductismo (Armstrong, 1968; 1980), a saber: haber entendido de una manera errónea el concepto de disposición. En efecto, al pretender negar la significatividad de los procesos y eventos internos los conductistas se han visto obligados a concebir operacionalmente su naturaleza disposicional. Una disposición, según este punto de vista, no sería sino el conjunto de comportamientos que la actualizan. Pero, obviamente, es ésta una concepción muy reductiva de la naturaleza de las disposiciones, una concepción que sólo podría defenderse si adoptáramos una teoría verificacionista del significado difícilmente justificable. Y si no lo hacemos todo parece indicar como mucho más sensato optar por una concepción realista de las disposiciones, según la cual éstas son diferentes de y explican causalmente a, sus manifestaciones. Del mismo modo en que debemos entender la propiedad de la solubilidad del azúcar como una propiedad intrínseca del azúcar -su peculiar estructura molecular- responsable de su disolverse en el agua, debemos entender el dolor como un peculiar estado del sujeto causalmente responsable de su actuar de cierta manera.

Ahora bien, los estados y procesos mentales no son sólo causalmente eficaces respecto a la conducta. A veces también tienen como consecuencia la aparición o desaparición de otros estados o procesos mentales, como cuando mi llegar a creer que en Brasil es muy fácil pasárselo bien me hace desear ir allí, pero mi pensar en lo mal que lo pasaré durante el viaje en avión, me hace desistir de la idea. Y con esta interconexión causal de los eventos y procesos mentales está íntimamente relacionada la otra objeción clásica que los funcionalistas han desarrollado contra el conductismo (Fodor, 1968; 1975). Como hemos señalado, los conductistas lógicos opinaban que el significado de los predicados psicológicos podía ser adecuadamente analizado en términos de un conjunto de proposiciones que describieran conductas efectivas o potenciales del sujeto al que se adscriben. Sin embargo, no deja de ser notorio el que ningún conductista lógico haya suministrado nunca un ejemplo, ni tan siquiera aproximado, de un análisis tal para ningún predicado psíquico, por elemental que éste sea. Podría pensarse que quizás la razón de ello estribe en que, como señalaron los conductistas, el tipo de disposición que denotan los términos psicológicos es altamente

complejo. Pero ya Chisholm (1957) apuntó a una razón mucho más convincente de por qué tales reducciones de los predicados psicológicos a descripciones conductuales sería imposible, a saber: que para que una conducta cuente como expresión de determinado predicado psicológico debemos asumir que el sujeto que la lleva a cabo está en otros estados psicológicos diferentes del que pretendemos definir conductualmente (Putnam, 1962; Dennett, 1971; 1978b; Stich, 1983; Churchland, 1984; etc.). Por poner un ejemplo de nuevo típico: si yo pretendo establecer el conjunto de comportamientos que corresponden a la creencia de que llueve el problema es no sólo que la lista sería posiblemente infinita (cerrar las ventanas de casa, recoger la ropa, advertir a mi madre de que no olvide el paraguas, coger el coche para ir a la Facultad, etc.) sino que además todas y cada una de estas conductas sólo cuenta como una expresión adecuada de aquella creencia si se supone que son ciertos de mí otros predicados mentales, aquéllos por ejemplo que me atribuyen ciertos deseos (hago todas esas cosas porque, además de creer que llueve, no deseo que entre agua en mi casa, ni que se moje o ensucie la ropa tendida, ni que mi buena madre coja una pulmonía, etc.).

Hemos visto ahora los defectos que los funcionalistas han encontrado más usualmente en las tesis del conductismo lógico; este trabajo de desescombros nos permite, no obstante, avistar el terreno donde se asentarán los fundamentos del funcionalismo. A diferencia del conductismo, no negará sino que subrayará la realidad de los procesos y los eventos mentales, dotándolos de eficacia causal no sólo respecto a la conducta sino respecto a sí mismos. De esta manera el funcionalismo volverá a una forma de mentalismo, pero debemos tener bien claro que, como advierte Fodor, 1968; no por ello debe volver necesariamente a postulados dualistas.

Podemos pasar ahora a analizar la otra filosofía de la mente en contraposición a la cual nace el funcionalismo: el materialismo reductivo o la tesis de la identidad. A diferencia del conductismo lógico, el materialismo reductivo tuvo un periodo de gestación muy breve y su vigencia fue también efímera. Sin embargo, haríamos mal si menospreciáramos su importancia, pues aparte de que su intuición básica de que los procesos y eventos mentales son procesos y eventos materiales está hoy más vigente que nunca, su formulación dió origen a una intensa discusión que sirvió para plantear claramente los problemas que en la actualidad dominan el campo de la filosofía de la mente. Los trabajos pioneros de Place, 1957; Feigl, 1958; Smart, 1959; 1962; así como los hitos más importantes de esa discusión aludida: Shaffer, 1962; 1963; Cornman, 1962; Feyerabend, 1963; Nagel, 1965; Rorty, 1965; Malcolm, 1964; etc., han sido profusamente recopilados, siendo quizás las mejores selecciones la de O'Connor, 1969; y la de Borst, 1970.

Por contraposición a muchos de los conductistas lógicos, cuya militancia filosófica se situaba en lo que se conoce ahora como la escuela del análisis del lenguaje ordinario, los materialistas reductivos se movían prácticamente todos ellos en las coordenadas del realismo científico (Armstrong, 1983); lo que quiere decir que entendían que la ciencia natural suministra un esquema conceptual privilegiado para comprender la realidad, a la vez que constituye un polo de referencia obligado en el trabajo del filósofo. Y precisamente de la historia de la ciencia natural sacaron la idea básica que articula su posición.

En efecto, la historia nos suministra ejemplos de reducción teórica de entidades y propiedades categorizadas en el esquema del sentido común mediante la postulación de su

identidad con entidades y propiedades cuya existencia se admite en el cuerpo de una teoría científica. Todos tenemos en mente los ejemplos del agua, reducida o identificada con el H_2O , o del rayo, identificado con una descarga eléctrica, o de la temperatura de un cuerpo, que se reduce al grado de energía cinética de sus moléculas. Si estas reducciones se habían conseguido con el desarrollo de la física y la química, ¿por qué no habría que esperar que el desarrollo de la neurofisiología nos permitiera lograr reducciones semejantes de los procesos y eventos psíquicos reconocidos en nuestra psicología cotidiana?

Cuando los primeros materialistas reductivos formularon sus teorías, la neurofisiología aún distaba de ofrecer tales tipos de identificaciones reductivas -dicho sea de paso, sigue sin hacerlo y hay razones empíricas y conceptuales como para poner en duda que pueda llegar a hacerlo nunca- pero optimistas como eran respecto a las potencialidades de la ciencia, pensaron que todo era cuestión de esperar. Ya había, en su opinión, evidencias suficientes de que los procesos y los eventos mentales eran procesos y eventos neurofisiológicos o, para ser más exactos, procesos y eventos del sistema nervioso central. Siguiendo su ejemplo favorito: a la pregunta «¿qué es el dolor?» lo más sensato era responder algo como «la excitación de las -hipotéticas- fibras C».

Como ocurrió con el conductismo, los funcionalistas en su gran mayoría encuentran el materialismo reductivo intuitiva y lógicamente defectivo. Veamos por qué.

Empezaremos una vez más por las objeciones intuitivas. La más usual de todas ellas ya fue formulada claramente por Putnam, 1967. Supongamos que, en efecto, la neurofisiología llega a identificar el dolor con la excitación de cierta región neuronal de nuestro cerebro; la pregunta clave es: ¿por qué habríamos de suponer que en todos los seres que pueden experimentar dolor éste se concretará en la excitación de un mismo tipo de neuronas? En realidad, sería muy sorprendente si esto ocurriera dada la gran variedad de sistemas nerviosos estructuralmente diferentes que presentan las distintas especies de seres vivos de los que de una manera u otra podemos decir que sienten dolor. Dicho de otra forma: dado que la neurofisiología identificará básicamente los estados mentales con estados de nuestro sistema nervioso, cualquier filosofía de la mente que pretenda basarse en sus resultados para desentrañar la naturaleza de lo mental estará muy expuesta a incurrir en el pecado de chauvinismo, el pecado de tener que negar que organismos a los que por su conducta estaríamos dispuestos a atribuirles procesos y estados mentales, realmente los tengan.

La validez de ésta y otras intuiciones anti-materialistas resulta muy discutible (se pueden encontrar discusiones de este tipo de objeciones en Lewis, 1969; Kim, 1972; Wilkes, 1978; 1981; Wilson, 1979; Churchland, 1979, etc.) pero como la discusión que siguió a las tesis de los materialistas acerca de si sus propuestas identidades podrían o no cumplir la ley de Leibnitz, nos puede inducir a abrigar una sospecha que puede resultar casi ineludible a la luz de ciertas argumentaciones lógicas y epistemológicas: la de que los predicados mentales no resulten proyectables en predicados físicos (sobre la noción de proyectabilidad, Cf. N. Goodman, 1965).

Esta objeción contra el materialismo reductivo ha sido muy bien elaborada por Fodor, 1975; y también por Davidson, 1970; 1973; 1974; si bien en éste último los argumentos juegan no sólo contra el materialismo reductivo sino contra cualquier tipo de reduccionismo acerca de lo mental y, en este sentido, es muy posible que estén dirigidos también contra el mismo funcionalismo, al menos en algunas de sus versiones centrales (Cf. Lycan, 1981).

A estas alturas no nos podemos permitir el lujo de entrar en detalle en estos argumentos, pero resumiéndolos podríamos decir lo siguiente. Es obvio que cada predicado que podamos utilizar es un instrumento que nos permite clasificar las entidades del mundo. El predicado «es una moneda», por ejemplo, nos permite agrupar a ciertas entidades y distinguirlas de otras. Si quisiéramos encontrar un predicado físico que se redujera a éste predicado económico, un requisito imprescindible es que fuera co-extensivo con él. Ahora bien, ¿qué tipo de predicado físico podría cumplir esta tarea? Si consideramos que una moneda puede estar hecha de cobre, o de bronce, o de plata, o de oro, o de aluminio, o incluso de papel; que puede ser circular, octogonal, cuadrada, rectangular, etc., está claro que ese predicado físico será tremendamente complejo... y en la misma medida completamente inútil para formular con él cualquier tipo de generalización nomológica. ¿Quién se atrevería a decir, vamos a suponer, que porque todas las monedas que se hubieran encontrado hasta hoy fueran o de plata, o de bronce, o de oro, que todas las monedas con que en el futuro nos encontraremos deberán ser de planta, de bronce o de oro?. Una consideración como ésta nos sirve para ver como muy verosímil la conclusión de que resultará imposible reducir los predicados económicos a predicados físicos, y por ende la economía a física. Aún si todas las monedas son entidades físicas, de ahí no se sigue que la propiedad «ser una moneda» sea equivalente a una determinada propiedad física. ¿Por qué no habría de ocurrir lo mismo con las propiedades descritas por los predicados psicológicos?. ¿Por qué no habría de ocurrir que, aunque todo estado y proceso psicológico fuera de hecho un estado y un proceso físico, la propiedad «ser un estado o un proceso psíquico» fuera diferente a cualquier específica propiedad física?. Lo que tenemos ahora es un respaldo lógico y epistemológico para nuestra intuición de que, dado que podemos predicar los mismos tipos de estados y procesos psíquicos de organismos neurofisiológicamente muy diferentes, aquellos estados y procesos no podrán identificarse con ningún estado o proceso neurofisiológico concreto.

Como ocurriría con la revisión de los argumentos funcionalistas contra las tesis conductistas, podemos sacar ahora de la crítica al materialismo reductivo una conclusión acerca del talante del funcionalismo. Este pretenderá constituir una alternativa a la vez antichauvinista y antireductiva en filosofía de la mente. Esto es, procurará dar una caracterización de lo que son los estados y procesos mentales que no nos obliga a predicar los mismos sólo de los seres humanos (Cf. Block, 1980b; Putnam, 1967b), y a la vez defenderá para la explicación psicológica una autonomía y una irreductibilidad a cualquier otro tipo de explicación científica (Cf. Haugeland, 1978; Cummis, 1975, 1983).

Sin embargo, es preciso advertir ya mismo que estas conclusiones no serían válidas para todos los autores que suelen encuadrarse bajo el rótulo común del funcionalismo. Algunos de los considerados a veces -aunque, significativamente, no por sí mismos-, como tales -el caso de Armstrong y de Lewis (Ved. Block, 1978; Armstrong, 1983)- no estarían muy de acuerdo con la conclusividad de argumentos como los esbozados arriba, de manera que en el análisis funcional de los predicados psicológicos no verían sino un aliado metodológico y epistemológico del materialismo, que permitiría a éste soslayar alguno de los escollos con que en sus formulaciones primigenias chocó. Por supuesto, estos autores tampoco compartirían la tesis de la irreductibilidad de la psicología, y seguirían centrando sus esperanzas en el desarrollo futuro de la neurociencia. Estas consideraciones creo que pueden servir para

explicar por qué no es del todo inusual el encontrarse con autores definidamente materialistas y a los que, en consecuencia, no se les sitúa en el terreno del funcionalismo, escribiendo, no obstante, alegatos en favor de éste (Cf. los Churchland, 1981; Wilkes, 1981).

III. Variedades del funcionalismo: La tesis del funcionalismo computacional

Lo que hasta aquí sabemos es más bien poco y algo bastante general acerca del origen y las características del funcionalismo. Vamos a entrar en un poco más de detalle. Como dejamos dicho, el funcionalismo como alternativa filosófica diferenciada fue gestándose poco a poco como un intento de corregir las deficiencias tanto del conductismo lógico como del materialismo. Una objeción muy usual que los dualistas han esbozado tanto contra uno como contra el otro es lo que se denomina a veces como la objeción del esencialismo, que viene a decir: sencillamente, no podemos identificar un estado mental con un tipo de conductas o de eventos físicos por la sencilla razón de que es concebible que se dé aquél sin que se den éstos, y al revés. Este argumento, que tiene su formulación primigenia en la obra de Descartes, ha sido reformulado de una manera u otra por los dualistas contemporáneos tales como Kripke, 1971; 1972; y Nage, 1974. Tomemos uno de los ejemplos favoritos de los dualistas: el dolor. Obviamente, es esencial al dolor el que se experimente o se perciba como doloroso; pero no es de ningún modo esencial a ninguna conducta, o a ningún evento de los que ocurren en nuestro cerebro, el que sea percibido o experimentado de este modo. El *esse* de lo mental es ser percibido de cierta manera, no así el *esse* de la conducta o de las estimulaciones nerviosas, etc. (para ver algunos argumentos contra esta intuición pueden consultarse Graham, G. & Stephens, L., 1985; Nelkin, 1987). De donde se deduce que el estado mental de tener dolor, y los eventos físicos de comportarse de cierta manera, o de sufrir ciertas estimulaciones en ciertas áreas de nuestro sistema nervioso, no tienen la misma esencia. Y difícilmente podría entenderse la identidad entre entidades esencialmente diferentes (Para una discusión de los argumentos de Kripke, Cf. Feldman, 1974; Lycan, 1974).

Por otra parte, esta diferencia ontológica entre los ámbitos de lo mental y lo físico parece explicarnos ciertas diferencias epistemológicas y lógicas que se dan en nuestro lenguaje ordinario entre el discurso psicológico y el discurso sobre lo físico. Comparemos los enunciados «Tengo miedo», «Me tiemblan las manos y la voz» y «Tengo estimuladas las fibras P». Lo que nos induce a aseverar el primero de ellos es muy diferente de lo que induce a otros a referirse a nosotros con un enunciado equivalente. Pero no así en los dos casos restantes. Decimos «tengo miedo» sólo porque sentimos miedo, lo que, obviamente, no es el caso cuando decimos «tiene miedo». Pero cuando decimos «me tiemblan las manos y la voz» o «tengo estimuladas las fibras P», las razones que tenemos para hacerlo son exactamente las mismas que pueden tener otros para decir que nos tiemblan las manos y la voz, etc. A saber, el ver u oír nuestras manos y nuestras voces temblorosas, una inspección de, o una inferencia sobre, el estado de nuestras fibras P., etc. Esta asimetría que se da entre los enunciados psicológicos en primera y tercera persona, pero no entre los enunciados que describen conductas o eventos neurofisiológicos, está conectada con otra peculiaridad epistémica de los primeros frente a los segundos. La autoatribución de estados psíquicos

parece ser incorregible, pero no as   la autoatribuci  n de conductas o de estados neurofisiol  gicos. Y a su vez, esta peculiaridad epist  mica implica ciertas peculiaridades l  gicas del discurso psicol  gico; en concreto, condena al sinsentido preguntas acerca de la base evidencial en que nos apoyamos para autoatribuirnos estados mentales, mientras que, por contra, las preguntas equivalentes acerca de los enunciados en que nos autoatribuimos conductas o eventos neurol  gicos conservan su legitimidad.

La asimetr  a de los enunciados psicol  gicos en primero y tercera persona, la incorregibilidad de unos frente a los otros, y la exclusi  n gramatical de ciertas cuestiones (o en t  rminos positivos, el car  cter gramatical de ciertas afirmaciones -«S  lo yo puedo sentir mi dolor»-) constituyen las coordenadas de lo que se conoce como el rompecabezas de la privacidad. El reto es encontrar una soluci  n del mismo que no nos comprometa con posiciones dualistas. Un reto que, como es bien sabido, Wittgenstein afront   con el famoso argumento del lenguaje privado, sin que yo tenga muy claro -debo confesarlo- si lo que intent   -no digamos ya lo que consigui  - fue encontrar esa soluci  n o, m  s bien, mostrar que la alternativa dualista tampoco constitu  a ninguna alternativa v  lida para encarar el mismo.

Siete a  os m  s tarde de la publicaci  n de la obra fundamental del segundo Wittgenstein, H. Putnam afront   en un breve pero muy importante art  culo (Putnam, 1960) el mismo reto. Su estrategia era, cuando menos, mucho m  s f  cilmente detectable que la de Wittgenstein. Se trataba de mostrar que una m  quina de Turing (Cf. Kleene, 1974; cap. XIII) que pudiera autoatribuirse estados reproducir  a la situaci  n en que nos encontramos nosotros cuando nos auto-atribuimos predicados psicol  gicos. Los enunciados en que la m  quina se autoatribuir  a estados ser  an tan asim  tricos con respecto a los enunciados que podr  a producir para atribuir estados a otra m  quina, tan incorregibles y tan poco susceptibles de dar pie a ciertas cuestiones, como lo son aquellos de nuestros enunciados en que nos atribuimos estados mentales. Si lo que explica las caracter  sticas epistemol  gicas y l  gicas de los enunciados psicol  gicos es, seg  n el dualista, la peculiaridad ontol  gica del sujeto que los realiza, su ser algo diferente de una realidad material, la consecuencia que debiera sacar es que las m  quinas de Turing no son entidades materiales. Pero si no se est   dispuesto a sacar esta conclusi  n, si se piensa que las m  quinas de Turing son simples m  quinas, artefactos mec  nicos o sistemas f  sicos, entonces la conclusi  n es que la privacidad no resulta incompatible con el materialismo.

En 1960 Putnam s  lo intentaba mostrar que una m  quina de Turing modelaba la subjetividad de la conciencia humana -una tesis que defendi   de las objeciones que le fueron planteadas en 1964- con la finalidad de defender la legitimidad del materialismo en su versi  n de teor  a de la identidad de tipos. Su estrategia consist  a en mostrar que las peculiaridades epistemol  gicas y l  gicas del discurso sobre lo mental no ten  an por qu   explicarse, en contra de lo que supone el dualista que apela a argumentos esencialistas, por las tesis del acceso directo o procesos semejantes. Quiz  s por ello, y tambi  n porque el modelo de m  quina de Turing considerado era excesivamente simplificado -por poner s  lo un ejemplo: se trataba de aut  matas deterministas-, aunque Putnam expres   su admiraci  n por la analog  a existente entre los estados l  gicos de la m  quina y los estados mentales de los seres humanos, no dej   de considerarla una mera analog  a, absteni  ndose de dar el paso

adicional de identificar la mente con una máquina de Turing. Sin embargo, en posteriores artículos (1964; 1967a; sobre todo) Putnam consideró la posibilidad de compliar las máquinas de Turing -concebir las como autómatas probabilísticos, antes que deterministas; como sistemas dinámicos, antes que estáticos; dotarlas de una función de preferencia racional, así como de transductores u órganos sensoriales y de un sistema motor, de manera que no sólo pudieran interaccionar con el medio sino hacerlo en calidad de agentes racionales- para hacerlas susceptibles de modelar ya no sólo comportamientos psicológicamente elementales, como el de la atribución de sensaciones, sino otros mucho más complejos como los de expresar preferencias y creencias. El paso ulterior de identificar los estados mentales de cualquier organismo con los estados lógicos de un cierto tipo de máquina de Turing estaba servido. Y Putnam lo dió en 1967b formulando claramente, por vez primera, las tesis fundamentales de lo que, basándonos en Block 1978, 1980b, podemos denominar como funcionalismo mecánico o, quizás mejor, computacional. No vamos a entrar por el momento en las peculiaridades de este tipo de funcionalismo. Nos limitaremos a señalar cómo se alcanza en él la tesis básica común a todo tipo de funcionalismo.

Como hemos visto, lo propio del funcionalismo computacional consiste en identificar los estados mentales de un organismo con los estados lógicos de una máquina de Turing de cierto tipo. En consecuencia, la naturaleza de los estados mentales, su esencia, será coincidente con la naturaleza o la esencia de los estados lógicos de las máquinas de Turing. ¿Cuál es ésta?

Dado que un estado queda individualizado cuando se especifican sus relaciones con la serie de posibles inputs, outputs y con el resto de estados, bien podemos responder esta pregunta diciendo que la esencia de los estados lógicos es, por consiguiente, puramente relacional. El criterio de individuación de los estados lógicos de una máquina de Turing, o lo que es lo mismo, el criterio de individuación de los estados mentales de un organismo, estriba en el rol funcional que ese estado juega en el sistema (lógico-psicológico) completo. Dos sistemas que sean funcionalmente isomórficos serán también psicológicamente equivalentes; y dos estados que sean funcionalmente equivalentes, serán estados psicológicos tipológicamente isomórficos.

Que esto constituye una respuesta nueva al enigma de la naturaleza de lo mental es obvio. La esencia de lo psíquico no estriba en ninguna de las características que señalaron los dualistas -ni en su inextensión, ni en su carácter subjetivo o privado, ni en su accesibilidad epistemológica privilegiada- ni los materialistas -no es ninguna peculiaridad física la que hace de lo mental un fenómeno diferente-, ni tan siquiera los conductistas -pues el carácter relacional de los estados psicológicos no se reduce a sus conexiones con inputs y outputs simplemente-.

Por otra parte, a esta concepción de los estados mentales llegaron otros autores por derroteros muy diferentes de los que siguió Putnam; no basándose en la analogía de las máquinas de Turing, sino en un análisis del significado de los predicados psicológicos. A este tipo de funcionalismo -defendido sobre todo por Armstrong (1968, 1980), Lewis (1966, 1970, 1972), Smart (1971) y Shoemaker (1975b, 1981a, b, c)- lo denominaremos, siguiendo a este último autor, «Funcionalismo analítico». Veamos cómo se gestó para después analizar las peculiaridades que lo distinguen del funcionalismo computacional.

IV. Variedades del funcionalismo: las tesis del funcionalismo analítico

En el apartado anterior hicimos referencia a los argumentos esencialistas que filósofos como Nagel y Kripke han utilizado contra las posiciones materialistas. El valor de la crítica de este último es, sin embargo, independiente de la -a mi entender dudosa- validez de tales argumentos.

En efecto, lo que hace importantes los trabajos de Kripke (1971, 1972) es que la conclusión negativa acerca de los enfoques no dualistas en filosofía de la mente se encuentra como colofón de una magnífica aportación a los campos de la teoría del significado y de la lógica modal. No podemos entrar en el detalle de estas aportaciones. Para nuestros propósitos, basta tener presente que las conclusiones de Kripke fueron que, por una parte, los nombres no deben ser concebidos como abreviaturas de un grupo de descripciones definidas sino como designadores rígidos que refieren a la misma entidad en todos los mundos posibles en los que la misma exista y, por la otra y como consecuencia, que el status modal de las identidades teóricas era el de verdades metafísicamente necesarias.

Esta última conclusión afectaba directamente a los teóricos materialistas, dado que éstos siempre habían admitido que las propuestas identificaciones de estados, eventos o procesos psíquicos con estados, eventos o procesos físicos eran identificaciones contingentes. Luego o Kripke estaba equivocado, y cabe la posibilidad de identificaciones teóricas no necesarias metafísicamente, o si Kripke estaba en lo correcto, entonces las tesis de los materialistas eran falsas. Claro que esto último podía ocurrir por dos razones muy diferentes: bien porque las identificaciones propuestas no eran posibles -y esta era la opinión de Kripke-; bien porque esas identificaciones, siendo correctas, son no obstante necesarias, y éste es el punto de vista de los funcionalistas analíticos.

Lo que se requería para defender esta última posición era una nueva teoría de la manera como las identificaciones teóricas en general, y las identificaciones psicofísicas en particular, son llevadas a cabo. El mérito y la responsabilidad de esta nueva teoría fue básicamente de D. Lewis (1966, 1970, 1972. Aunque también son importantes a este respecto los trabajos de Chichara y Fodor, 1965; y Harman, 1973) y el *quid* de la misma estribaba en considerar de una manera nueva la relación entre términos teóricos y términos no teóricos. Veamos.

La epistemología de filiación positivista, anclada por el viejo dogma de la existencia de una distinción tajante entre el ámbito de lo analítico y de lo sintético, del significado y de lo empírico, había legado un concepto de teoreticidad por el que ésta se contraponía tajantemente a lo observacional. Según esta concepción las teorías científicas serían una combinación de términos categorialmente heterogéneos: teóricos los unos, observacionales los otros. Términos cuya diferencia estribaría en que mientras los primeros denotarían entidades, propiedades o procesos no observables y quasi-ficticios (recuérdese que, siguiendo básicamente a Russell, los positivistas lógicos pretendían, en un principio al menos, tratar a todos los términos no empíricos como *ficciones* lógicas), los segundos denotarían entidades, eventos o procesos empíricamente constatables.

Frente a esta concepción, la de Lewis es que la línea divisoria entre los términos teóricos y los no teóricos no estriba en su observabilidad o inobservabilidad, sino en el rol que juegan en nuestras teorías; en concreto la hipótesis de Lewis es que los términos teóricos son tales

porque en el seno de las teorías en los que se postulan los mismos son definidos funcionalmente, esto es: por el rol causal que se atribuye a sus *denotata* en relación a los eventos, propiedades o procesos denotados por los términos no teóricos.

En su trabajo de 1972, Lewis pone un ejemplo bastante pedestre en clave detectivesca. Para ilustrar su posición, particularmente prefiero acudir a un ejemplo extraído de la historia real de la ciencia. Después del descubrimiento de Urano, realizado por casualidad por W. Herschel en 1781, las observaciones del planeta recién descubierto empezaron a contradecir las expectativas que las tablas construidas para explicar su órbita hacían concebir. La corrección de las viejas tablas por parte de A. Bouvard, con el fin de ponerlas en concordancia con las nuevas observaciones, de nada sirvieron, pues al cabo de tan sólo 20 años la trayectoria efectivamente observada de Urano volvía a diverger significativamente de lo previsto por las tablas enmendadas. La opción que se presentaba era o bien concluir que la ley de Newton no era exacta, o presuponer la existencia de otra masa planetaria transuránica que, con su influjo, permitiera, preservando la ley, dar cuenta de las perturbaciones de la órbita de Urano. Esta última fue la opción por la que se decidió Le Verrier. Este predijo la existencia de un octavo planeta que sería *causalmente* responsable de las anomalías orbitales de Urano, y calculó su posición de manera que permitió a J. Galle descubrir, el 23 de Septiembre de 1846, la existencia de Neptuno.

De acuerdo con el punto de vista de Lewis, en la teoría de Le Verrier los únicos términos teóricos serían aquellos que utilizara para referirse al hipotético planeta responsable causalmente de los hechos a explicar; hipotético planeta que en la teoría de Le Verrier no se identificaría sino por criterios funcionales, o sea: sólo por aquellas propiedades que se considerasen necesarias como para hacerlo responsable causalmente de aquellos hechos. El resto de términos, esto es: todos aquellos que utilizamos en la caracterización de estos hechos, serían términos no teóricos (Lewis los denomina O-terms, pero advierte que la O no viene de «Observational» sino de «old», por lo que podríamos denominarlos, por nuestra parte, términos familiares). «Neptuno», o cualquier otro nombre con el que Le Verrier hubiera bautizado a su presunto planeta, o cualesquiera descripciones definidas que hubiera utilizado con la intención de referirse a él tales como «el planeta trans-uránico», o «el planeta más excéntrico del sistema solar», etc serían sus términos teóricos. Si Le Verrier hubiera vivido unos 100 años después podría haberse servido del procedimiento recomendado por Ramsey para formular su hipótesis, procedimiento que le habría permitido pasarse sin términos teóricos ya que estos quedan en este proceder sustituidos por variables ligadas a cuantificadores existenciales en las que recae la responsabilidad de referirse a las hipotéticas entidades cuya existencia resulta explicativamente fructífera de acuerdo con la teoría. En lugar, pues, de haber dicho algo como «Neptuno -o el planeta transuránico, etc-, cuya órbita y masa son tales y tales, según la ley de Newton, lo que hace que la órbita de Neptuno tenga tales y tales características observadas» podría haber dicho: «Existe un planeta x tal que su órbita y su masa son tales que, etc». Esta artimaña técnica resultará de la mayor importancia cuando pasemos a aplicar todas estas consideraciones al ámbito de la psicología. Por el momento sólo pretendemos hacer ver, merced a este ejemplo, la verosimilitud de la tesis de Lewis. Está ahora claro que los términos teóricos no son aquellos que denotan entidades, propiedades, eventos o procesos inobservables. De hecho Neptuno es observable

aunque para Le Verrier su nombre fuera un término teórico. No menos erróneo es considerar que los términos no teóricos o familiares denoten todos ellos entidades observables. Parte de los términos no teóricos para Le Verrier serían numerales, aquellos que caracterizarían la órbita de Urano. Y por muy platónico que se sea nadie está dispuesto a afirmar que los números son el tipo de entidades que podemos observar. La auténtica línea demarcatoria entre los términos teóricos utilizados por Le Verrier y los términos no-teóricos de su teoría pasa porque aquellos denotan entidades cuya existencia y propiedades son presupuestas para cumplir cierto rol en una explicación causal, de manera que podemos decir que son funcional e implícitamente definidos, mientras que este último no es el caso de los términos familiares o no teóricos. Por otra parte, la identidad de Neptuno con el pretendido referente de los términos teóricos de Le Verrier no fue una identidad postulada sino descubierta, fue el resultado de descubrir que Neptuno efectivamente realizaba la teoría de Le Verrier, lo que, llegado el caso, daría pie a que la considerásemos como una identidad metafísicamente necesaria.

¿Qué tiene que ver todo esto con la filosofía de la mente y, más concretamente, con el funcionalismo? La respuesta a esta pregunta empieza a perfilarse desde el momento mismo en que nos percatamos de que para los funcionalistas analíticos, como para algunos funcionalistas computacionales, nuestra utilización del lenguaje psicológico ordinario viene posibilitada por nuestra asunción de una teoría psicológica de sentido común (contra esta tesis argumentan Putnam 1964 y Wilkes, 1981). Según ellos lo que nos permitiría cuando observamos la conducta de un organismo atribuirle ciertos predicados psicológicos sería cierta teoría psicológica en la que aquellos predicados contarían como términos teóricos y, por consiguiente, tendrían su significado definido implícita y funcionalmente por esta teoría. Por poner un ejemplo, lo que según los funcionalistas analíticos nos permitiría atribuir a alguien que corre delante de un león estados psicológicos tales como los de ser conscientes de la presencia del león, sentir miedo y tener el deseo de salvarse, es una teoría psicológica de sentido común según la cual el ser conscientes de la presencia de un león, tener miedo y el deseo de sobrevivir son causas suficientemente explicativas de la conducta de correr delante de un león. Podríamos hacer explícita esa teoría formulando una serie de obviedades como la anteriormente citada, obviedades que, no obstante, nos permiten captar el significado de los términos psicológicos por su relación con otros estados psicológicos y con el conjunto de estímulos que los provocan y respuestas a que dan lugar. Una vez hecha explícita esa teoría podríamos transcribirla según la manera recomendada por Ramsey, de manera que desaparecieran de ella todos los términos teóricos, en este caso todos los términos psicológicos.

No obstante la importancia de esta conclusión -la eliminabilidad del lenguaje intencional- con la que se comprometen los funcionalistas analíticos, lo que nos interesa subrayar ahora es que, siguiendo un camino muy diferente al de los funcionalistas computacionales, han llegado a una conclusión muy parecida a la suya acerca de la naturaleza de los estados, procesos y eventos psicológicos. Los criterios para identificar éstos son, como los criterios utilizados para dotar de significado a los términos teóricos en general, puramente funcionales. *Prima facie* aquellos estados, procesos y eventos tienen una naturaleza puramente relacional, pues en tanto que estados, procesos y eventos psicológicos no son sino los estados, procesos y eventos causados por ciertos estímulos y que causan ciertas respuestas.

Empero, no conviene perder de vista las diferencias, serias y profundas, existentes entre las dos versiones del funcionalismo que hemos considerado. Veamos.

V. Diferencias entre las dos variantes del funcionalismo

Para empezar, es muy diferente el status epistemológico que los funcionalistas analíticos y los computacionales conceden a sus conclusiones. Para estos últimos, por lo general más cercanos a la concepción naturalizada de la epistemología en particular, y de la filosofía en general, sus conclusiones no tienen un carácter sino hipotético. Ellos no afirman que no quepa otro análisis sino el funcional de los estados mentales. En su versión más fuerte (la del primer Putnam) todo lo que dicen es que si los organismos dotados de una psicología pueden ser considerados como realizaciones de una máquina de Turing, entonces sus estados mentales funcionalmente definidos pueden identificarse tipológicamente con los estados lógicos de la máquina que realizan. Mientras que en su versión más débil este tipo de funcionalistas admite que los estados mentales no son estados lógicos de una máquina de Turing (Block & Fodor, 1972), pero insiste en que la consideración de los mismos como tales es teóricamente útil y está justificada en la medida en que existe alguna analogía interesante entre el funcionamiento de las máquinas de Turing y la mente (Cf. el mismo Putnam, en 1975; y Fodor, 1983). Así, de manera semejante a como los materialistas eliminacionistas fiaban al desarrollo de la neurociencia la justificación última de sus tesis, los funcionalistas computacionales fiaban la justificación de las suyas al desarrollo de una psicología que caracterice funcionalmente los estados, eventos y procesos psicológicos a los que apele en sus explicaciones.

Por contraposición, los funcionalistas analíticos no consideran que sus conclusiones tengan un carácter hipotético, ni deban esperar al desarrollo de ninguna ciencia para verse confirmadas. Más bien entienden que las mismas son el resultado de analizar el significado de los predicados psicológicos de nuestro lenguaje ordinario. Lo que defienden, por consiguiente, es que dado como utilizamos nuestro lenguaje, los eventos, procesos y estados mentales tienen una naturaleza funcional.

Esta divergencia es algo más que una mera divergencia metafilosófica, puesto que según la opción que adoptemos quedará limitado de una manera u otra el repertorio de estrategias argumentativas que podremos oponer al anti-funcionalista. Pongamos por caso que alguien objeta que el significado de nuestros predicados que refieren a procesos o eventos acompañados de contenido fenomenológico, no puede quedar agotado por un análisis funcional de los mismos (una objeción clásica del dualista). El funcionalismo analítico no tiene más remedio que tomarse la misma con toda seriedad e intentar demostrar por todos los medios que no es cierta. Por contra, el funcionalista computacional bien puede conceder la razón al objetor dualista (lo que no es conceder la razón al dualismo), pero añadir que su observación resulta totalmente irrelevante para la verosimilitud o inverosimilitud de su hipótesis. Y esta posible divergencia en las estrategias defensivas del funcionalismo analítico y del funcionalismo computacional frente al viejo argumento de los *qualia*, se repite de una forma u otra frente a todos los argumentos más típicos que contra el funcionalismo suelen hacerse.

La siguiente diferencia notable entre el funcionalismo computacional y el analítico es de índole ontológica, y se refiere a la diferente posición de uno y otro frente a la verosimilitud del materialismo, esto es: la verosimilitud de identificar los procesos psicológicos con las realizaciones físicas de los mismos. Como ya advertimos cuando hablamos de las características generales de las posiciones funcionalistas, los partidarios del enfoque computacional han apelado a una serie de intuiciones y de argumentos en contra del materialismo; argumentos e intuiciones a las que se han opuesto los partidarios del funcionalismo analítico. Lo que queremos añadir ahora es que esta diferencia no es de ninguna manera accidental al modelo de funcionalismo por el que se opte.

En efecto, si lo que conduce a las conclusiones funcionalistas es la comparación de la mente con una máquina de Turing, entonces el escepticismo respecto al materialismo parece inevitable, por la sencilla razón de que las susodichas máquinas no son sino autómatas abstractos, esto es: autómatas caracterizados no por ninguna propiedad material (por su *hardware*) sino única y exclusivamente por sus características lógicas (su *software*). Dicho de otra forma: lo único que se exige a un sistema para contar como una máquina de Turing, es que sea capaz de ejecutar el programa especificado por la tabla de la máquina, siendo irrelevante que ese sistema tenga una naturaleza material u otra o, incluso, el que tenga o no una naturaleza material (Putnam, 1975).

Si las mentes son (análogas a) máquinas de Turing, y éstas, como acabamos de decir, son autómatas abstractos, es obvio que la realización física que una mente vaya a tener, si es que, como la mayoría de funcionalistas computacionales piensan, va a tenerla, será completamente accidental. Añadamos a esto la premisa Kripkeana de que las identificaciones teóricas deben ser, caso de verdaderas, necesarias, y el funcionalismo computacional estará comprometido con la negación de la posibilidad de identificar los estados psicológicos por sus características físicas.

Justo lo contrario es lo que defiende el funcionalista analítico, y ello por la esencia misma del argumento que le conduce hasta sus conclusiones; pues recordemos que aunque en su opinión los términos psicológicos tal y como aparecen en nuestro lenguaje ordinario tienen un significado puramente funcional, esos mismos términos, en la medida en que no tengan sino una definición implícita en el contexto de la psicología del sentido común, no denotan sino entidades teóricas, entidades cuya verdadera naturaleza, o si se prefiere, su esencia, aún está por descubrir. Pero, añade el funcionalista analítico, ¿qué otra hipótesis sensata cabe hacer sino la de que esa naturaleza será física? Y si finalmente llegamos a descubrir que es así, como en su opinión ya lo estamos descubriendo, no cabría sino concluir que los estados mentales son estados físicos de una manera tan necesaria como el agua es H₂O.

Esta disidencia ontológica está, por otra parte, estrechamente relacionada con una divergencia de índole metodológica y, consecuentemente, con las expectativas que uno y otro tipo de funcionalismo tiene respecto a la psicología. Esto se hace muy claro desde el momento mismo en que nos percatamos de que lo que constituye el punto final de las explicaciones de los funcionalistas computacionales no sería para los funcionalistas analíticos sino el punto de partida.

Obviamente, si se adopta un enfoque computacional de los estados, eventos o procesos psicológicos, la tarea por excelencia del psicólogo no debiera ser otra que la de suministrar lo que Cummins (1975) denominó explicaciones analítico-funcionales. Se trataría de explicar procesos psicológicos complejos tales como la percepción, o el habla, o la resolución de problemas, etc., analizándolos o descomponiéndolos en procesos psicológicos más simples, de manera análoga a como el programador de un computador descompone una tarea compleja, tal como la de sumar, en una serie de tareas más simples tales como las de borrar, imprimir, etc. (de ahí que a veces se le denomine también con el rótulo de «funcionalismo homuncular»)... El funcionalismo computacional respalda así los enfoques cognitivos de la psicología que, como el de Neisser 1967, defienden la especificidad e irreductibilidad de ésta a disciplinas como la neurofisiología.

Por su parte, los funcionalistas analíticos debieran mostrar más entusiasmo por enfoques, como el de Lindsay y Norman, 1977, en los que las explicaciones analítico-funcionales de los procesos psicológicos son constantemente complementadas con explicaciones de la base física de esos procesos. O dicho de otra forma: los funcionalistas analíticos debieran mirar con mayor simpatía aquellas investigaciones psicológicas que no utilizaran una estrategia estrictamente top-down como la recomendada por sus colegas computacionales y que estuviera abierta a los enfoques bottom-up. Ello es así porque, recordémoslo una vez más, para este tipo de funcionalistas los análisis funcionales de los procesos psicológicos no constituirían sino el primer paso de la investigación, un primer paso que queda incompleto en tanto no se complementen con un análisis de la estructura física de los mismos. De esta manera, los funcionalistas analíticos defenderían, cuando menos, la necesidad de complementación de la psicología por la neurofisiología.

VI. Los problemas del funcionalismo

Expuestas las tesis funcionalistas en sus principales variantes, nos queda por analizar ahora los escrúpulos que dificultan la marcha del teórico funcionalista. Hay que advertir, para empezar, que muchos de esos escrúpulos han sido esbozados por los propios funcionalistas, y que las consecuencias de los mismos han sido muy diferentes en unos casos y otros. A algunos -el caso de Putnan 1988- las dificultades planteadas les han llevado a desdecirse de sus posiciones funcionalistas, mientras que a otros -Fodor es el ejemplo más notable- esas mismas dificultades les han llevado, simplemente, a rebajar las pretensiones del enfoque funcionalista, aunque siguen defendiendo que el mismo es el más adecuado.

En líneas generales podríamos decir que el reto que se presenta al funcionalismo es el de hasta qué punto su comprensión de los eventos, procesos o estados psicológicos se ajusta a nuestra concepción preteórica u ordinaria de los mismos. Un desajuste grave entre una y otra invalidaría *ipso-facto* las pretensiones del funcionalismo analítico dado que, recordémoslo una vez más, éste pretende ser el resultado de un análisis del significado de nuestro actual lenguaje psicológico en el que aquella concepción viene expresada. Pero aún si no tan drásticas, las consecuencias que tendría para el funcionalismo computacional, autoconcebido como una propuesta empírica, serían también graves. Pues si éste concluye proponiendo un enfoque de los fenómenos psicológicos que se aparte radicalmente del

sentido común, inmediatamente se le plantearán una serie de difíciles interrogantes como, por ejemplo: ¿cómo entender la relación entre el enfoque funcional y el propio del esquema conceptual ordinario? ¿Habría acaso que conceder la razón al materialismo eliminacionista (Quine, 1951; Feyerabend, 1963; Rorty, 1965; Churchland, 1984) y concluir que los conceptos con los que funciona tal esquema son radicalmente erróneos, de manera que debemos prescindir por completo de ellos? Y si es así ¿desde qué nueva óptica se estaría concibiendo la acción humana? ¿Y cuáles serían los intereses teóricos y prácticos de la misma? Por poner un ejemplo especialmente debatido en la literatura reciente (Armstrong, 1973; Loar, 1981; Lycan, 1981; Dennett, 1981, 1982; Stich, 1983; Baker, 1985; Fodor, 1987, etc.) consideraremos el caso de la creencia.

Muchas de las explicaciones psicológicas que damos de nuestro comportamiento o del de nuestros semejantes apelan a la atribución a nosotros mismos o a ellos de unas u otras creencias. Los intereses teóricos y prácticos de tal proceder explicativo son obvios: no sólo nos permite presentar nuestra conducta, o ver la de los demás como coherente, sino que también abre la posibilidad de intervención racional sobre la misma. Nuestra creencia de que el líquido que contiene el vaso que vemos es un refresco, explica, dada nuestra sed, la pretensión de bebernoslo. Pero también permite a nuestros amigos hacernos desistir de tal empeño informándonos de que, en realidad, lo que el vaso contiene es cianuro.

¿Qué pasaría ahora si, como muchos creen, nuestro concepto de «creencia» se muestra refractario a un tratamiento estrictamente funcional? ¿Acaso debiéramos decir que una explicación psicológica que aspire a ser científica tendrá que prescindir de este concepto? ¿Debiéramos, en consecuencia, negar toda validez a nuestras explicaciones psicológicas cotidianas y negar que es falso que nadie tenga creencias?

Y aunque no lleguemos a conclusiones tan radicales, aunque nos limitemos a decir que la explicación psicológica científica no apelará al concepto de creencia sin por ello negar validez al uso de tal concepto en las explicaciones psicológicas ordinarias, ¿bajo qué descripción de nuestra conducta explicará entonces la teoría psicológica construida en términos estrictamente funcionales la misma? ¿Qué tipo de comprensión y qué posibilidad de intervención sobre ella nos permitirá?

No es mi intención en este punto afirmar dogmáticamente que el funcionalista carecerá de respuestas satisfactorias a estos interrogantes, sino sólo justificar por qué la concordancia del análisis funcional con nuestra concepción pre-teórica del ámbito psicológico supone un desafío para cualquier variante del funcionalismo, pues caso de no darse la misma quedaría invalidada automáticamente la versión analítica del mismo y, como mínimo, claramente cuestionada la plausibilidad de la versión computacional.

Pero si éste es el problema general que el funcionalismo debe afrontar, es obvio que la gran multiplicidad de predicados psicológicos que el lenguaje ordinario pone a nuestra disposición lo hace prácticamente intratable así formulado. No podemos pronunciarnos sobre la viabilidad de un análisis funcional de todos y cada uno de los particulares conceptos psicológicos que utilizamos normalmente dado el elevado número de los mismos. Lo que sí podemos es, si conseguimos establecer una clasificación general de los estados, eventos o procesos denotados por esos conceptos, considerar la posibilidad de éxito del enfoque funcionalista en el tratamiento de los tipos que nuestros criterios taxonómicos establezcan;

una estrategia que bien puede llevarse a cabo analizando el caso de algunos conceptos que consideremos más paradigmáticos de cada uno de estos tipos.

A este respecto, inspirándonos en Wittgenstein (1967; 1980) podríamos distinguir entre aquellos predicados psicológicos a cuyos *denotata* parece consustancial un contenido fenomenológico, y aquellos a los que no. De esta manera los predicados perceptivos o sensoriales, tales como «ver», «oir», «sentir dolor», etc. se incluíran dentro del primer tipo, y muchas de las denominadas actitudes proposicionales, tales como «creer», «desear», «pensar», «juzgar», etc., en el segundo.

La problemática que rodea al funcionalismo puede redefinirse ahora como la de determinar la medida en que el mismo puede ofrecer un análisis de los estados, eventos o procesos psicológicos con contenido fenomenológico y de las actitudes proposicionales carentes del mismo.

Aún así reconducida, queda fuera del alcance de este trabajo un análisis pormenorizado de la misma. Lo que resta lo dedicaremos a considerar la viabilidad de un análisis funcional del primer tipo de conceptos psicológicos.

VII. Funcionalismo y *qualia*

Nuestros criterios para atribuir procesos sensoriales y perceptivos son múltiples. La pertinencia de las causas que los provocan, el tipo de conducta que causan, las certezas o creencias, así como otras actitudes proposicionales, que los acompañan, son relevantes a la hora de hacer tal atribución. Pero aparte de todo esto cuando atribuimos a un sujeto alguno de los predicados que se refieren a estos procesos parece que le atribuimos algo más: la conciencia de una determinada experiencia cualitativa (la conciencia de un *sense-datum*, de un *quale*, de un *raw-feel*, etc.).

Consideremos el caso del dolor. El mismo tiene unas causas típicas (una herida, la disfunción de algún órgano corporal...) y una conducta que le es propia (quejarse); el sujeto que lo sufre tiene certeza de padecerlo y, generalmente, de su localización, así como el deseo más o menos manifiesto de dejar de tenerlo. Pero además tiene una experiencia cualitativa dolorosa.

Mutatis mutandis lo mismo podríamos decir de cualquier proceso perceptivo: de la percepción visual, por ejemplo. Para decir de alguien que ha visto X tenemos en cuenta la posibilidad de que X haya sido la causa de su proceso perceptivo, el conjunto de creencias que ha adquirido mediante tal proceso sobre X, la conducta discriminativa que a consecuencia del mismo puede exhibir, etc. Pero también suponemos que ha tenido una experiencia cualitativa; una experiencia visual.

Es esta experiencia o conciencia cualitativa, que distingue a los procesos perceptivos y sensoriales de la mayoría de esos otros eventos psíquicos que son las actitudes proposicionales, la que mayores quebraderos de cabeza da al teórico funcionalista. Para empezar, ¿cómo podríamos reconstruir esta diferencia en términos estrictamente funcionales? ¿Puede el análisis funcionalista justificar una taxonomía de los estados mentales que, como ésta, nos parece tan obvia?

Muchos han respondido negativamente a esta pregunta, y como justificación de su pesimismo han añadido la observación de que muchas veces somos capaces de discriminar

estados por su contenido fenomenológico aún antes de que seamos conscientes de ninguna diferencia funcional entre los mismos, lo que parece reforzar la conclusión de que lo que identifica al menos a algunos estados psicológicos no es su rol funcional sino algo completamente diferente del mismo, a saber: su contenido cualitativo.

La conclusión de la inapresabilidad de este carácter fenomenológico de algunos de nuestros estados mentales por el análisis funcional de los mismos se ve reforzada además por la consideración de ciertos experimentos mentales, tales como el famosísimo del espectro invertido (cif. Shoemaker, 1981c). ¿Qué ocurriría si mañana alguien se levantara y comprobase que todos los objetos que hasta hoy le producían la sensación del color rojo sistemáticamente le producían ahora la sensación de verde, y viceversa?. Lo razonable parece que sería esperar que tras un período de adaptación el sujeto afectado por esta inversión de sus sensaciones cromáticas consiguiera adaptar su conducta a la nueva situación, lo que demuestra que la sustancialidad cualitativa de ciertos estados mentales es diferente e independiente del rol funcional que los mismos puedan desempeñar. Pero si esta independencia existe ¿por qué admitir sólo la posibilidad de que dos estados funcionalmente isomórficos puedan ser cualitativamente diferentes?; ¿por qué, siendo consecuentes, no admitir también la posibilidad de que dos estados puedan ser funcionalmente equivalentes, y que, sin embargo, mientras a uno de ellos le corresponde un estado cualitativo el otro no tenga asociado ninguno? He aquí el argumento denominado del quale ausente.

De acuerdo con los Churchland (1981) podríamos desgarnar el desafío que los qualia representan para el funcionalismo en cuatro problemas: el de justificar funcionalmente, o al menos explicar cómo es posible la taxonomía común de los estados mentales que distingue aquellos dotados de contenido fenomenológico de aquellos que carecen de él; el de explicar cómo es posible la discriminación de estados al margen de sus características funcionales; el problema de la inversión del espectro cualitativo y, por último, el problema de la total ausencia cualitativa. Hay que decir que los autores funcionalistas no han prestado la misma atención a todos estos problemas y que, como era de esperar dada la disparidad de posiciones que se esconde bajo el rótulo común de funcionalismo, han afrontado los mismos con estrategias harto diferentes. Veamos.

Si fuéramos funcionalistas podríamos, inspirándonos en Fodor y Block (Fodor y Block, 1972; Block, 1978; Fodor, 1981a; etc.) tomar una alternativa que cortara por lo sano. Podríamos reconocer que no cabe, apelando a criterios funcionales, construir una taxonomía de los estados mentales que se asemeje para nada a lo que hemos denominado la taxonomía común. Ni que tampoco podemos dar cuenta, en términos funcionales, de nuestra capacidad para discriminar cualitativamente algunos estados psíquicos de otros. Ni, en definitiva, dar cuenta en absoluto de la presencia o la ausencia de los *qualia*. Pero una vez admitido todo esto podríamos añadir que ello no resulta de la menor importancia, dado que las únicas teorías psicológicas fructíferas serán aquellas que caractericen a los estados psicológicos en términos estrictamente funcionales. Dicho más sencillamente, la réplica fodoriana al problema de los qualia consistiría en, aún reconociendo la dimensión fenomenológica de algunos de nuestros estados mentales, señalar la intrascendencia de la misma para la construcción de una teoría psicológica adecuada. Que la psicología no explique alguna de las propiedades de los fenómenos que estudia no es, por otra parte, algo sorprendente; al fin

y al cabo ninguna ciencia explica todas las propiedades de los fenómenos de su dominio, y hay propiedades de las cosas que no parecen ser de interés para ninguna ciencia. Bien pudieran ser los *qualia* uno de estos fenómenos científicamente irrelevantes.

En principio, tal estrategia defensiva del funcionalismo suena plausible, pero debemos ser conscientes del precio que nos exige pagar. Y éste no es otro que el de reconvertir el funcionalismo de una filosofía de la mente en una filosofía de la psicología. El funcionalista no nos estaría diciendo qué es la mente, sino sólo cómo la concebirá la psicología. Claro que el funcionalista partidario de esta opción podría añadir que tal precio es bastante módico, dado que eso de la filosofía de la mente sería una quimera tan inalcanzable como cualquier otra empresa metafísica. Quienes pretenden decir qué es la mente no estarían diciendo sino cómo se entienden los fenómenos psicológicos desde el punto de vista del sentido común. También ellos estarían haciendo filosofía de la psicología, sólo que no de la psicología científica sino de la cotidiana.

Sin embargo, y aún aceptando estas consideraciones metafísicas, la estrategia Fodoriana resulta discutible por razones puramente intrínsecas. Para empezar, no creo que pueda decirse sin más que el problema de los *qualia* es irrelevante para la psicología. Desde luego no lo ha sido (véase Hamlyn, 1957) y creo evidente que entre una teoría psicológica que pudiera explicarnos las peculiaridades fenomenológicas de algunos de los procesos mentales, y otra que no nos suministrara la menor idea al respecto, nos quedaríamos indudablemente con la primera. Además me parece que el tema de las propiedades fenomenológicas resulta difícilmente soslayable cuando se estudian ciertos procesos psíquicos tales como el de la percepción, aunque sea con una óptica cognitivista, dado que una explicación homuncular del mismo parece exigir una definición respecto a las posibles o efectivas interacciones de los sistemas cognitivos periféricos (transductores, módulos sensoriales o como quiera denominárselos) con los sistemas cognitivos centrales (memoria, resolución de problemas, etc.), definición que difícilmente podrá darse sin, a su vez, posicionarse sobre la naturaleza de los out-puts de aquéllos. A este respecto resulta instructivo el caso del propio Fodor (1983), quien enfrentado al problema de determinar cuál es la naturaleza de los productos de los sistemas de entrada se ve obligado a recurrir al criterio de «accesibilidad fenomenológica» para zanjar la cuestión.

Pero si la estrategia Fodoriana no nos convence -y a mí, desde luego, no me convence- aún podemos recurrir a otras estrategias de características similarmente evasivas, utilizadas por autores particularmente funcionalistas o filomaterialistas (Dennett, 1978a; los Churchland, 1981; Putnan, 1981; Davidson, 1988) para soslayar el problema de los *qualia*. En líneas generales la misma consiste en insistir en que, lejos de lo que presupone el dualista, los *qualia* distan de ser entidades perfectamente definidas y, por consiguiente, no puede apelarse a ellos para explicar el significado de los predicados psicológicos -hasta aquí el acuerdo de todos estos autores- quedando éste determinado por criterios funcionales -y ésta es la tesis específica del subgrupo de los funcionalistas-.

En efecto, el dualista parece comprometido con la tesis de que lo que determina la aplicación de un mismo predicado a dos estados o eventos psíquicos es la semejanza que pueda existir en su carácter cualitativo. Pero tal condición no se cumple ni tan siquiera en el caso de aquellos estados o eventos a los que nadie osaría negar un componente

fenomenológico. Consideremos el caso del dolor. ¿No existen acaso una gran cantidad de diferencias cualitativas entre las experiencias a las que catalogamos como dolorosas?. Y justo lo inverso podemos decir de otros fenómenos psíquicos como las emociones, por ejemplo. Muchas de ellas van acompañadas de sensaciones particulares, pero una y la misma de estas sensaciones puede acompañar a emociones muy distintas. Por otra parte, es un lugar común desde hace mucho tiempo en la filosofía de la percepción (Quinton, 1955) que, bajo condiciones normales, no prestamos la menor atención al específico carácter cualitativo de nuestras experiencias sensoriales cuando atribuimos cualidades objetivas a las cosas. Como se encarga de recordarnos Putnan, 1981, ¿podríamos asegurar que nuestras experiencias fueron semejantes cuando en dos ocasiones diferentes vimos el mismo mantel rojo?.

Las conclusiones lógicas y epistemológicas que se siguen de estas observaciones parecen claras. El contenido intencional de nuestros procesos perceptivos o sensoriales no parece poder explicarse por el contenido cualitativo de los mismos. Los criterios de significatividad de nuestros predicados psicológicos no parecen coincidir con las semejanzas y desemejanzas cualitativas de las experiencias que puedan acompañar a los estados que denotan, y lo que nos legitima a utilizarlos no es, tampoco, la consideración de la naturaleza cualitativa de tales experiencias. Proponer el rol funcional que tales estados juegan como candidato para desempeñar este rol lógico-epistémico es un paso adicional posible, aunque no obligatorio.

A mi entender, hay algo de correcto en estas observaciones, pero dudo que sea lo suficiente para librar al funcionalismo de la objeción de los *qualia*. Pues, al fin y al cabo, ¿qué es lo que todas estas observaciones demuestran?.

Volvamos al caso del dolor. Estrictamente, lo que debíamos concluir de cuanto se nos ha dicho es que cuando nos autoatribuimos sensaciones dolorosas, por lo general no discriminamos muy sofisticadamente sus características cualitativas. Pero decir ésto implica conceder que tales características existen y que, aún si de hecho no nos dedicamos a discriminarlas muy detenidamente, tales discriminaciones son posibles -de lo contrario, ¿cómo sabríamos que lo que llamamos doloroso es muy diferente en unas ocasiones y en otras?-.

Por otra parte, aún cuando sea cierto que las finas discriminaciones del carácter cualitativo de nuestras experiencias estén por lo general de más, no es menos cierto que una discriminación, por muy grosera que sea, de la naturaleza cualitativa de las mismas es absolutamente imprescindible en la aplicación de algunos de nuestros predicados psicológicos. ¿Acaso si no experimentásemos sensación alguna nos atreveríamos a decir que estamos emocionados?. Y si un estado no lo experimentásemos con ciertas propiedades ¿nos atreveríamos a calificarlo como doloroso?. ¿Y qué decir de nuestros juicios perceptivos?. Si mis sensaciones cromáticas del mantel no hubieran sido estrechamente semejantes, ¿habría podido reconocerlo en las dos ocasiones en que lo vi como el mismo mantel rojo?.

Estas últimas cuestiones me parecen imponer una matización sobre las observaciones escépticas que los funcionalistas y ciertos materialistas hacen acerca de la determinación del carácter cualitativo de nuestras experiencias psíquicas, y del papel que el mismo juega en la aplicación de nuestros predicados psicológicos. No creo que el grado de semejanza o

desemejanza cualitativa de nuestras experiencias sea igualmente irrelevante en todos los casos. Posiblemente, el mismo tenga muy poca importancia en el caso de la auto-atribución de emociones, pero es mucho más relevante a la hora de auto-atribuirnos estados dolorosos, y fundamentalmente cuando hacemos juicios perceptivos visuales.

El funcionalista puede replicar en este punto que en cualquier caso, y aún concediendo la existencia de un contenido cualitativo dotado de propiedades intrínsecas, el mismo es lógica y epistemológicamente parasitario del contenido intencional de nuestros procesos perceptivos y sensoriales. Dicho de otro modo, que sólo podemos discriminar y describir las propiedades fenomenológicas de nuestra experiencia si ya somos capaces de discriminar y describir las propiedades empíricas y objetivas de los eventos o las cosas del mundo, pues cuando decimos, por ejemplo, que experimentamos un *qualia* visual rojo lo que estamos expresando abreviadamente es, en realidad, algo como que nuestra experiencia es del tipo de las que normalmente nos causa la percepción de las cosas rojas.

Estoy por completo de acuerdo con esta observación, pero a mi entender habría que señalar también que aquí el orden ontológico es el inverso del lógico y del epistémico, y que la existencia y naturaleza de nuestra experiencia cualitativa determina los posibles contenidos intencionales de nuestros procesos perceptivos y sensoriales. Consideremos el caso del daltónico. A quien padece una enfermedad semejante le están vedados los procesos perceptivos visuales cuyo contenido intencional consista en discriminar entre el color rojo y el verde. La única hipótesis verosímil al respecto es que tal deficiencia no lo es de aprendizaje -no depende, por consiguiente, del funcionamiento de los procesos cognitivos centrales- sino que depende del funcionamiento de los transductores ópticos... de la naturaleza cualitativa del *ouï-pui* que éstos suministran a aquéllos. La conclusión es obvia: el contenido cualitativo de nuestros procesos perceptivos y sensoriales es un condicionante causal de su posible contenido intencional. Es por ello por lo que, a pesar de su dependencia lógica y epistémica respecto al contenido intencional objetivo, un análisis del significado de nuestros predicados psicológicos encargados de referir a aquellos procesos no puede eludir la consideración de la naturaleza cualitativa de nuestra experiencia.

Pero retomemos el hilo de nuestra discusión. Como conclusión valorativa de la estrategia defensiva del funcionalismo que hemos estado considerando creo que podríamos decir que la misma fracasa por varias razones. Aún suponiendo que fuera correcta en todos sus términos, lo más que mostraría es que los *qualia* no determinan el significado ni la aplicación de nuestros predicados psicológicos ordinarios, no que los *qualia* no existan y que no caractericen a algunos de nuestros estados mentales. Podríamos formular, pues, un reproche análogo al que dirigimos contra la posición de Fodor. Admitir que el análisis funcional no puede dar cuenta del carácter cualitativo de nuestras experiencias es admitir una limitación del funcionalismo en tanto que filosofía de la mente. Y por otra parte -y esto es señalar que la tesis fundamental en torno a la cual se articulan las posiciones que estamos considerando es errónea- no parece que el análisis de algunos de los predicados psicológicos que se refieren a procesos sensoriales y perceptivos pueda pasarse sin apelar a la naturaleza cualitativa de nuestra experiencia, pues la misma parece ser una condición causalmente relevante para la existencia de aquellos procesos y, por consiguiente, un factor relevante para determinar la corrección de la atribución de aquellos predicados que los refieren. Si dos

experiencias no fueran desagradables, el individuo que las padece no estaría legitimado a calificarlas como dolorosas, y si sus sensaciones cromáticas no fueran pertinentemente semejantes mentiría cuando afirmara que está viendo que dos objetos (o uno en diferentes ocasiones) tienen el mismo color. Todo lo cual no hace sino reforzar el segundo de nuestros reproches a Fodor, a saber: que es más que dudoso que una teoría psicológica de ciertos procesos mentales pueda pasarse sin la consideración de sus características fenomenológicas.

La situación parece ser, pues, ésta. La existencia de los *qualia* resulta difícilmente disputable, y es sensato pensar que los mismos juegan un papel importante en la determinación del significado de algunos de nuestros predicados psicológicos. Por consiguiente, todas y cada una de las objeciones que se plantearon contra el funcionalismo siguen en pie, y no queda sino intentar lidiar con ellas.

Así lo han comprendido tanto algunos de quienes militan en el campo del funcionalismo computacional (el mismo Dennett, en 1978e; Rey, 1989), cuanto algunos de los que se sitúan en los dominios del funcionalismo analítico (Lewis, 1978; y, sobre todo, Shoemaker, 1975a; 1975b; 1981b; 1981c). No podemos aquí hacer justicia a los detalles de sus posiciones, y las reflexiones críticas que a continuación vamos a exponer tendrían muy diferente alcance según contra quien las dirigiéramos.

Dada la concepción puramente relacional de los eventos, estados y procesos mentales que los funcionalistas profesan, la única estrategia coherente que les cabe es la de intentar mostrar que el contenido cualitativo puede reducirse a, y explicarse en términos de, el rol funcional que el mismo desempeña. Dicho con otras palabras: la naturaleza cualitativa de nuestra experiencia está en función de los otros estados, eventos o procesos mentales con los que su procesamiento característico la relaciona, así como de los estímulos que la provocan y la conducta a la que, de manera mediatizada, da lugar.

Antes de entrar en un análisis más pormenorizado quisiera hacer dos consideraciones que pueden predisponernos escépticamente respecto a la viabilidad de esta estrategia. Recordemos que una de las críticas que los funcionalistas dirigían a los materialistas era la de que su análisis conducía indefectiblemente a una posición *chauvinista*: si el dolor debe identificarse con la estimulación de las fibras C, de cualquier organismo que careciera de las mismas habría que negar que pudiera padecer dolor. Mi pregunta ahora es si el funcionalismo no incurre en un *chauvinismo* incluso más grave.

Es obvio que la economía funcional de la mente de los animales es muy diferente de la nuestra. Los out-puts de sus órganos sensoriales, por atenernos a la diferencia más palmaria, no pueden interactuar con ningún centro de control lingüístico. Esto significa que incluso a animales muy cercanos a nosotros en la escala evolutiva, no digamos ya a los que se hallan situados muy lejos de ella, tendríamos que adscribirles contenidos cualitativos completamente diferentes de los nuestros. Y posiblemente no sólo contenidos cualitativos, lo que aún podría resultar verosímil, sino también contenidos intencionales. Pues no menos que el cualitativo, el contenido intencional de los procesos perceptivos también quedaría determinado por la economía funcional global de la que forman parte. De manera que si adoptáramos una posición funcionalista estricta no sería legítimo decir, por ejemplo, que el gato salió corriendo porque vio al perro abalanzarse contra él. Pues ver perros sería potestad de aquellos que compartieran una estructura psíquica análoga a la nuestra.

Realmente la situación para el funcionalista es incluso peor que ésta, pues una conclusión lógica de su posición es que la divergencia en los contenidos cualitativos e intencionales se produce en el seno mismo del género humano. No sólo los animales son incapaces de expresarse lingüísticamente -o al menos con un lenguaje parecido al nuestro-. Los niños no nacen hablando. Por consiguiente, también habría que decir que el contenido cualitativo de sus actos perceptivos y sensoriales es diferente del nuestro. Y lo mismo habría que decir de aquéllos que conceptualicen la experiencia de una manera alternativa a la nuestra. Dado que ante unos mismos estímulos visuales el esquimal adquiere creencias y tiene conductas discriminatorias diferentes a las nuestras, habría que concluir que el carácter cualitativo de su experiencia visual es diferente al nuestro. Y puesto que -no sé si esto es cierto, pero basta con que sea verosímil- tampoco ellos hacen distinciones entre el verde y el rojo, quizás haya que decir que más bien es semejante a la del daltónico.

En resumen pues, que si estoy en lo correcto el criterio funcional de adscripción de contenido cualitativo e intencional para los procesos perceptivos y sensoriales nos llevará a la adopción de posiciones mucho más *chauvinistas* que aquéllas con las que nos comprometía el materialismo.

La segunda observación que quiero hacer es la siguiente. Según los principios funcionalistas, el contenido cualitativo de los procesos perceptivos y sensoriales queda determinado por las relaciones (computacionales) que mantiene con otros eventos, estados y procesos mentales, así como por la conducta que (mediatizadamente) puedan causar. Hay aquí, a mi entender, una especie de reduccionismo análogo a aquél del que pecaban las teorías más rudimentarias que pretendían explicar la percepción en términos de adquisición de creencias (Armstrong, 1961; Pitcher, 1971). Sin entrar en la cuestión de la existencia de un percibir no epistémico (Cf. Dretske, 1969), es duro llegar a aceptar que sólo vemos aquello de lo que nos formamos una creencia merced al uso de nuestros sentidos, y esta misma intuición subsiste aún cuando compliquemos cuanto queramos el número y la naturaleza del resto de procesos psíquicos con los que interrelacionar nuestra experiencia. Si la riqueza de ésta no puede reducirse a las creencias que en ella se originan, tampoco pueden reducirse a esas creencias los recuerdos a corto o largo plazo, los *out-puts* hacia la unidad de control lingüístico, etc. Sencillamente no procesamos todo lo que vemos.

Vamos a pasar por alto estas observaciones y a suponer que el teórico funcionalista tiene réplicas convincentes a las mismas. ¿Cómo se resolverían entonces los problemas relacionados con los *qualia*?

El criterio funcional de determinación de contenido sugiere, para empezar, una respuesta obvia al problema de la total ausencia de carácter cualitativo de la experiencia (Cf. Shoemaker, 1975b; Churchland, 1981). Recordemos que éste se planteaba por la posibilidad de que un sistema que fuera funcionalmente isomórfico con nosotros no fuera, sin embargo, consciente de *qualia* alguno. Pero si presuponemos la validez de aquel criterio tal posibilidad se desvanece de manera inmediata. Pues para que un sistema fuera funcionalmente isomórfico con nosotros sería necesario no sólo que se comportara análogamente a como nosotros lo hacemos, sino que también el resto de sus operaciones psíquicas resultaran equivalentes a las nuestras; entre ellas la introspección. Para que se diera el isomorfismo funcional estricto no bastaría, pues, que ante estímulos semejantes el otro sistema fuera

capaz de hacer discriminaciones objetivas análogas a las nuestras, sino que también debiera ser capaz de dar informes introspectivos sobre la naturaleza cualitativa de su experiencia como nosotros somos capaces de hacerlo, pues esta operación forma parte esencial del conjunto de eventos, procesos y estados psíquicos que dotan de contenido cualitativo a nuestras experiencias. Y si fuera así, ¿qué más necesitaríamos para conceder la conciencia de *qualia* a ese sistema?.

Es muy importante darse cuenta de que, como Shoemaker subrayó (1975b), este argumento no es de corte verificacionista, y que más bien lo único que asume es una teoría causal del conocimiento y de la referencia. El argumento no afirma que no tiene sentido dudar de la existencia de un *quale* cuando éste da lugar a determinadas creencias sobre su carácter cualitativo o/y a determinada conducta, porque la evidencia introspectiva y la conductual son los únicos métodos de verificación de los que disponemos para aseverar la existencia de los *qualia*. Más bien el argumento es de este tipo: si la conciencia introspectiva sobre el carácter cualitativo de nuestros estados mentales merece algún crédito epistemológico, o si los términos con que expresamos las creencias a las que tal conciencia da lugar se refieren a algo, entonces el carácter cualitativo de nuestros procesos perceptivos y sensoriales debe ser causalmente responsable de aquella conciencia. Si un sistema es funcionalmente isomórfico con nosotros, entonces entre sus estados, procesos y eventos psicológicos deberá darse una relación causal semejante a la que se da entre los nuestros; en concreto, sus creencias introspectivas también deberán, como en nuestro propio caso, estar causadas por el carácter cualitativo que la experiencia tenga. Quiero que se retenga este punto porque -ya volveré sobre ello- la asunción de esta teoría causal del conocimiento y de la referencia y, en general, de la mente, creo que nos permite construir un argumento que muestra la inviabilidad de mantener una posición estrictamente funcionalista en torno al problema de los *qualia*.

Está claro que esta solución al problema de los *qualia* ausentes sugiere respuestas obvias a los problemas de taxonomía y discriminación. Recordemos que el primero consistía en dar una caracterización funcional de la distinción entre los estados mentales con contenido fenomenológico y los estados mentales carentes del mismo, mientras que el segundo nos pedía una explicación de cómo es posible la discriminación de ciertos estados mentales aún antes de que seamos conscientes de ninguna de sus peculiaridades funcionales.

A esto último podríamos responder que una cosa tal es radicalmente imposible. Si somos capaces de discriminar entre dos estados mentales, como mínimo debemos ser capaces de tener creencias introspectivas diferentes acerca de la naturaleza cualitativa de los mismos, pero dado que las creencias introspectivas a la que los estados mentales puedan dar lugar ya son características funcionales de los mismos, de ahí se sigue que no hay discriminación posible de estados mentales que no sea ya una discriminación en virtud de sus características funcionales. Podemos discriminar entre estados mentales siendo ignorantes de algunas de sus características funcionales (por ejemplo, siendo ignorantes de sus relaciones con *inputs* estímulares o *outputs* conductuales), pero no siendo ignorantes de todas sus características.

Por lo que hace al problema taxonómico estará ahora claro que la clasificación de los estados mentales por su posesión o carencia de características cualitativas puede ser construida funcionalmente, como una clasificación de los estados por su capacidad para dar lugar a estados introspectivos de cierto tipo.

Es más, esta solución funcionalista nos permite explicar en un cierto estilo Wittgensteiniano algunas de las peculiaridades epistemológicas de nuestra experiencia: la aparente autoridad de la primera persona sobre la misma, por ejemplo. Si nuestro testimonio acerca de las propiedades de nuestras experiencias perceptivas o sensoriales es fiable ello obedece al simple hecho de que la explicitación de su carácter cualitativo forma parte esencial de los procesos que determinan su naturaleza. Como el ¡Ay! con respecto al dolor, habría que entender la relación entre los informes en primera persona y las características de la experiencia perceptiva o sensorial más como causal y expresiva que como descriptiva.

En cualquier caso, es obvio que el tema es bastante diferente con respecto al argumento del espectro invertido. Supongamos que yo tengo una creencia introspectiva acerca del color del que tengo experiencia ahora, mientras miro el mar, y que considero que la manera correcta de manifestar mi creencia es afirmando que soy consciente de algo azul. Mi vecino puede en este momento tener una creencia introspectiva que expresa con los mismos términos. Admitamos que si ambos tenemos conciencia introspectiva del carácter cualitativo de un estado mental, entonces este estado debe tener un tal carácter, pero ¿qué garantiza que un mismo informe introspectivo haya de ser causado en todas las ocasiones por estados cualitativamente idénticos?. Los defensores del argumento del espectro invertido insisten en que es una posibilidad lógica que si yo fuera a tener conciencia del estado de mi vecino, entonces estimara que la descripción correcta del carácter cualitativo del mismo fuera otra que la que estimamos lo es cuando somos conscientes del nuestro. Y si esto es una posibilidad lógica entonces lo que se demuestra es una limitación del funcionalismo. Si bien podríamos dar una caracterización funcional de la existencia de los *qualia*, paradójicamente no podríamos definir funcionalmente sus relaciones de semejanza y desemejanza.

Podríamos, como estrategia defensiva, insistir en que la posibilidad del espectro invertido es una mera posibilidad lógica empíricamente muy difícil de concretarse (Cf. Harrison, 1973; 1984), pero no creo que la misma nos llevara muy lejos dado que sabemos de la existencia de ciertos casos que vienen a apuntar a la misma conclusión a la que el argumento de la inversión espectral apunta, tales como el de algunos de los pacientes afectados de dicromatismo.

En efecto, la merma en su capacidad discriminatoria de los colores, en estas personas en las que algún tipo de conos retinianos ha sufrido un proceso de deterioro, puede corregirse mediante la aplicación a las mismas de lentes de contacto coloreadas que actúan como filtros de luz, de manera que los enfermos dicromáticos pueden volver a distinguir las cosas rojas de las verdes... no obstante lo cual en sus informes introspectivos aseguran que su experiencia de estos colores es cualitativamente diferente de lo que lo era antes. Este fenómeno bien contrastado (Cf. Huvich, 1981) viene a indicar lo mismo que el argumento del hipotético espectro invertido: cierta autonomía del contenido cualitativo respecto al contenido intencional de nuestros procesos perceptivos y sensoriales. Pero si ello es así ya no podemos pensar que los criterios de identidad del contenido cualitativo de nuestros procesos perceptivos y sensoriales puedan reducirse a criterios puramente funcionales.

Para que esta reducción fuera factible debiéramos suponer que la identidad de roles funcionales de dos experiencias determinaría la identidad cualitativa de las mismas. Y precisamente esto es lo que pone en tela de juicio el caso de los enfermos dicromáticos.

¿Acaso no se sirven de experiencias cualitativamente diferentes de las de los perceptores normales (incluidas las de ellos mismos cuando lo eran) para hacer los mismos juicios que éstos sobre los colores de las cosas?.

Sin embargo, el funcionalista estricto aún puede protestar a estas alturas, y argumentar que en el caso de los dicromatos o, para el caso es lo mismo, en el de los hipotéticos afectados por la inversión del espectro, la verosimilitud de atribuirles una diferencia en el carácter cualitativo de sus experiencias es explicable en términos funcionales. Al fin y al cabo los dicromatos hacen juicios introspectivos («mi experiencia del color es ahora diferente de la experiencia que recuerdo tenía de los mismos») que los perceptores normales no hacemos, y es de suponer que las diferencias no sólo en los informes introspectivos sino también en el plano conductual serían mucho más severas en el caso de los afectados por la inversión espectral. Luego las diferencias en el contenido cualitativo de la experiencia se corresponden con diferencias en el rol funcional que desempeñan.

Esta réplica no deja de ser verdadera, pero podemos complicar nuestros ejemplos hasta hacer perder al funcionalista toda esperanza de poder mantener dignamente su posición. Supongamos que a uno de los pacientes afectados de dicromatismo, para instalarle las lentes coloreadas que le van a permitir recobrar sus capacidades discriminativas, lo anestesiarnos totalmente, momento que aprovechamos para inyectarle una sustancia que borre su memoria cualitativa (si esto suena muy inverosímil conviene tener en cuenta que los afectados de acromatopsia cerebral adquirida no sólo pierden su capacidad para percibir el color de las cosas, sino también para imaginárselo, o para soñarlo. Cf. Wasserman & Sacks, 1988). Cuando nuestro pobre dicromato despierte su experiencia cualitativa será diferente a la que tenía antes de que sus conos se deteriorasen... pero lo ignorará. No sólo su conducta discriminatoria sino sus propios informes introspectivos serán de todo punto indistinguibles de los nuestros.

Si el funcionalista se empeña en mantener su posición, lo único que le queda decir es que dado que la experiencia juega el mismo rol funcional antes del advenimiento de la enfermedad y después del sometimiento a la operación, en uno y otro momento tiene el mismo carácter cualitativo. Mientras la experiencia del dicromato operado con nuestra técnica no sufriría alteración alguna respecto a la experiencia que tenía cuando estaba sano, no podría decirse lo mismo respecto a la experiencia del dicromato al que se le ha instalado las lentes de contacto por el método habitual. Pero ¿cómo una mera diferencia en la memoria puede explicar una diferencia en la naturaleza de la experiencia? Si el dicromato tiene buena memoria, el contenido cualitativo de sus procesos perceptivos y sensoriales será diferente; si la tiene mala será idéntica a antes de caer enfermo.

Parece que el funcionalista estricto está aquí poniendo la carreta delante de los bueyes. Pues es conceptualmente lícito defender que la naturaleza de los eventos determina, si éste es fiel, el recuerdo que tengamos de ellos, pero no veo cómo se pueda defender que el falso recuerdo o, aún peor, la falta de recuerdo, determina la naturaleza de los eventos. ¿Se atrevería el funcionalista estricto a defender algo semejante para el mundo físico? ¿Se atrevería a defender, por ejemplo, que dado que no recordamos que el coche del vecino era de color rojo hace cinco días, éste no puede haberlo mandado pintar con el color azul que luce ahora? Desde luego que si el vecino fuera funcionalista y malmemoriado se podría

ahorrar la factura del planchista.... Y si no se atreve a defender esto -¡esperemos que no se atreva!- ¿será acaso que mientras es sensatamente realista respecto al mundo físico es fenomenista respecto al mental? ¿Defenderá acaso que en el mundo mental el *esse est percipi*, de manera que, como defendió Hume, las cosas son como nos parecen, y nos parecen tal como son?.

Mucho me temo que la posición dramática que el funcionalista se ve obligado a tomar no sólo es intuitivamente indefendible -el sentido común nos dice que los dicromatos, independientemente de cómo anden de memoria, tienen una experiencia cualitativamente diferente de la de los perceptores normales- sino también conceptualmente absurda.

Volvamos al lecho rocoso de las intuiciones. El sentido común nos empuja a decir que los dicromatos con lentes coloreadas tienen una experiencia cualitativamente diferente a la de los perceptores normales. ¿Por qué?. Bueno, para empezar tenemos su testimonio introspectivo, una diferencia funcionalmente caracterizable. Pero esto no explica el caso del dicromato cualitativamente amnésico. ¿Por qué decimos de éste que su experiencia es semejante a la de sus compañeros de enfermedad y diferente de la de los perceptores sanos? La respuesta obvia es: porque el estado de sus órganos sensoriales es semejante al de sus compañeros de patología y diferente al de los perceptores normales.

Supongo que adonde quiero ir a parar ya lo verán claro. Lo que quiero indicar es que, sin negar que los criterios funcionales y conductistas (que no son sino un subgrupo de los funcionales) forman parte de los criterios que utilizamos para determinar las relaciones de semejanza o diferencia de los contenidos cualitativos de los procesos perceptivos y sensoriales, los criterios físicos no son menos ineludibles. Es el presupuesto de una naturaleza biológica común, y no sólo de una conducta o una economía psico-funcional común, el que nos justifica a adscribir contenidos cualitativos semejantes a los procesos perceptivos y sensoriales de todos nuestros congéneres.

Podríamos extendernos ahora sobre cómo este reconocimiento de la dimensión materialista que tienen los criterios ordinarios de adscripción de predicados psicológicos dotados de contenido cualitativo o fenomenológico permite soslayar, o por lo menos debilitar, el peligro de *chauvinismo* que denunciábamos acechaba al funcionalismo estricto; cómo, también, permite dar una respuesta que suene menos tautológica que la dada por el funcionalista al problema taxonómico, discriminatorio y del *quale* ausente; o cómo, en contra de las apariencias, no tiene por qué suponer una grave amenaza contra la autoridad de la primera persona; pero preferimos utilizar el poco tiempo que nos queda en construir otro argumento que conduce a la misma conclusión: la necesidad de completar los criterios funcionales con criterios materialistas.

Volvamos al argumento funcionalista contra la objeción de la ausencia de carácter cualitativo de nuestras experiencias. Recordemos que el mismo pendía de la asunción de una teoría causal del conocimiento y de la referencia. Si se tienen creencias acerca del carácter cualitativo de los estados psíquicos, y aquellas creencias tienen alguna fiabilidad epistémica, entonces las mismas deben ser causadas por estados dotados de ciertas características cualitativas. Ahora bien, un problema común tanto a las teorías causales del conocimiento, como a las teorías causales de la referencia, de la acción, etc.; o, en general, un desafío usual a cualquier teoría causal de la mente, lo constituye la posibilidad de la

existencia de cadenas causalmente desviadas. Esta posibilidad es la que obliga a matizar que una creencia sólo es conocimiento, o una acción sólo es intencional, o una expresión sólo refiere, etc., si está causalmente conectada de la manera habitual con la evidencia que la sustenta, o con los deseos y creencias que la provocan, o con la entidad que designa, etc. (véase a este respecto, Goldman, 1967; Martin & Deutscher, 1966 y Grice, 1961 y, en general, los trabajos recopilados en Davis, 1983). Pues bien, este mismo problema afecta a la solución que el funcionalista da al argumento del *quale* ausente.

¿Qué ocurriría en el caso de que alguien tuviera acceso al out-put de alguno de los sistemas modulares de entrada de otra persona?. Supongamos, por ejemplo, que la ceguera al color que alguien padece es total y no se debe a una deficiencia de sus conos retinianos sino a alguna lesión de su cortex visual. Puestos a suponer, sigamos suponiendo que los neurofisiólogos han descubierto alguna manera de transmitir la información que el ojo de este paciente suministra al nervio óptico de un perceptor normal donde la misma es procesada de la manera habitual por el módulo sensorial visual, y que el output de éste vuelve a ser transmitido al cerebro del perceptor lesionado, en el que puede interactuar con los sistemas cognitivos centrales. Este paciente podría no sólo hacer discriminaciones acerca de las propiedades empíricas objetivas que ve, sino también informes introspectivos acerca del carácter cualitativo de una experiencia que es causa de aquellas discriminaciones y de estos informes... ¡pero de una experiencia que él no tiene!.

Lo que este experimento mental muestra, a mi entender, es que no basta para asegurar la posesión de estados mentales cualitativamente caracterizados el que tengamos creencias acerca de sus características cualitativas causadas por los mismos. Además debemos suponer que tal causación se produce en la manera habitual; lo que no significa sino que los procesos, eventos y estados pertinentes tienen una instanciación física normal. La explicación funcional sólo funciona, una vez más, si asumimos la normalidad física o biológica del sujeto al que atribuimos predicados psicológicos.

Si estoy en lo correcto, pues, no sólo para resolver el problema de la semejanza cualitativa de los procesos perceptivos y sensoriales, sino también para resolver el problema mismo de los *qualia* ausentes debemos completar las definiciones funcionales con intuiciones materialistas.

Por otra parte, creo que este tipo de problemas que se plantean al funcionalismo está relacionado con una característica muy básica del mismo que bien puede ser completamente errónea: su enfoque atomista del sujeto psicológico. El funcionalismo pretende explicar a este sujeto como el producto del ensamblaje de distintas estructuras funcionales, dando prioridad lógica a las partes sobre la unidad que conforman. Es este presupuesto, a la vez metodológico y ontológico, el que lo hace vulnerable a las paradojas de las realizaciones extravagantes (las habitaciones chinas de Searle, 1980, 1984; las naciones pensantes de Block, 1978; el robot con cabezas de hidra de Putnan, 1981, etc.), paradojas que pueden resolverse si, más de acuerdo con nuestros criterios preteóricos de adscripción de predicados psicológicos, consideramos que el sujeto propio de atribución de tales predicados son los organismos, los seres vivos a los que resulta consustancial su unidad, y no los sistemas cuya unidad es el resultado de un ensamblaje accidental de estructuras funcionales.

Podemos ahora concluir nuestra exposición. Las posibilidades de que el funcionalismo pueda suministrar un tratamiento adecuado de los procesos psicológicos dotados de características cualitativas son más bien escasas. No creo que una estrategia estrictamente funcionalista pueda dar cuenta de los problemas relacionados con los *qualia*. Para enfrentarse a los mismos será necesario atender también a los criterios materialistas en los que se apoya nuestra adscripción de los predicados que denotan aquellos procesos. Por otra parte, estoy convencido de que las cosas no irían mucho mejor por lo que respecta a las actitudes proposicionales. Sólo que aquí para superar las limitaciones del funcionalismo nos veríamos obligados a la consideración del sujeto psicológico no sólo como un organismo sino como un organismo que interacciona con el mundo natural y, en el caso de las personas, con el social. Nuestra conclusión se acerca, pues, a la alcanzada por Putnam, 1973. Si las consideraciones biológicas son relevantes para entender la naturaleza de ciertos procesos psíquicos, las sociológicas pueden serlo no menos esenciales para la del resto.

Vicente Sanfélix Vidarte
Septiembre, 1989



NOTA: Discut   un primer borrador de este trabajo con mi buen amigo J.L. Prades Celma, y la forma definitiva del mismo se debe en buena parte a algunas de sus observaciones cr  ticas. No obstante, no debe colegirse de esto que   l comparta los puntos de vista que aqu   he expresado.

Bibliograf  a

- ARMSTRONG, D. (1961), *Perception and the Physical World*. Routledge and Kegan Paul. London.
- ARMSTRONG, D. (1968), *A Materialist Theory of The Mind*. Routledge & Kegan Paul. London.
- ARMSTRONG, D. (1973), *Belief, Truth and Knowledge*. University Press. Cambridge.
- ARMSTRONG, D. (1980), «The Nature of Mind», en Block (1980).
- ARMSTRONG, D. (1983), «Recent work on the relation of mind and brain», en Floistad, G. (1983).
- BAKER, L.R. (1985), «A Farewell to Functionalism», *Philosophical Studies*, vol. 48.
- BIRO, J.I. & SHAHAN, W. (Edts.) (1982), *Mind, Brain and Function*, The Harvester Press. Sussex.
- BLOCK, N. & FODOR, J.A. (1972), «What Psychological States are not» en Block, 1980.
- BLOCK, N. (1978), «Troubles with Functionalism» en Block (1980).
- Block, N. (Edt.) (1980a), *Readings in Philosophy of Psychology*. Methuen. London.
- BLOCK, N. (1980b), «What is Functionalism?», en Block, N. (Edt.) (1980).
- CARNAP, R. (1932), «Die Physikalische Sprache als Universalsprache der Wissenschaft», *Erkenntnis*, II.
- CORNMAN, J. (1962), «The Identity of Mind and Body» en Borst (1970).
- CHICHARA, CH. & FODOR, J.A. (1965), «Operationalism and Ordinary Language», en Fodor, (1981).
- CHISHOLM, R. (1957), *Perceiving*. Cornell University Press. New York.
- CHURCHLAND, P.M. (1979), *Scientific Realism and the Plasticity of Mind*, University Press. Cambridge.
- CHURCHLAND, P.M. & CHURCHLAND, P.S. (1981), «Functionalism, Qualia and Intentionality», en Biro & Shahan (1982).
- CHURCHLAND, P.M. (1984), *Matter and Consciousness*. MIT. Cambridge (Mass).
- CUMMINS, R. (1975), «Functional Analysis» en Block (1980).
- CUMMINS, R. (1983), *The nature of Psychological Explanation*. MIT, Bradford.
- DAVIDSON, D. (1970), «Mental Events» en Davidson, 1980.
- DAVIDSON, D. & HARMAN, G. (Edts.) (1972), *Semantics of Natural Language*. Reidel, Dordrecht.
- DAVIDSON, D. (1973), «The Material Mind» en Davidson 1980.
- DAVIDSON, D. (1974), «Psychology as Philosophy» en Davidson 1980.
- DAVIDSON, D. (1980), *Essays on Actions and Events*. Clarendon Press. Oxford.
- DAVIDSON, D. (1988), «The Myth of the subjective». No publicado.
- DAVIS, S. (Edt.) (1983), *Causal Theories of Mind*. De Gruyter. Berlin.
- DENNETT, D. (1971), «Intentional Systems» en Dennett, D. (1978).

- DENNETT, D. (1978a), «Why yo can't make a computer that feels pain» en Dennett (1978c).
- DENNETT, D. (1978b), «Skinner Skinned» en Dennett, D. (1978).
- DENNETT, D. (1978c), *Brainstorms*, MIT. Press. Cambridge (Mass).
- DENNETT, D. (1978d), «Current Issues in the Philosophy of Mind», *American Philosophical Quarterly*, 15: 4.
- DENNETT, D. (1978e), «Towards a cognitive theory of consciousness» en Dennett (1978c).
- DENNETT, D. (1981), «True Believers» en Dennett (1987).
- DENNETT, D. (1982), «Beyond Belief» en Dennett (1987).
- DENNETT, D. (1987), *The Intentional Stance*. MIT. Cambridge.
- DRETSKE, F.I. (1969), *Seeing and Knowing*. Routledge and Kegan Paul. London.
- FEIGL, H. (1958), «Mind-Body, Not a Pseudo-Problem», en Borst (1970).
- FELDMAN, F. (1974), «Identity, Necessity and Events», en Block (1980a).
- FEYERABEND, P. (1963), «Materialism and the Mind-Body Problem», en Borst (1970).
- FLOISTAD, G. (Edt.) (1983), *Contemporary Philosophy*, Vol. IV. Martinus Nijhof, The Hague.
- FODOR, J.A. (1968), *Psychological Explanation*, Random House, New York.
- FODOR, J.A. (1975), *The Language of Thought*. Crowell, New York.
- FODOR, J.A. (1981a), «Something of the state of the art», en Fodor (1981b).
- FODOR, J.A. (1981b), *Representations*. MIT. Press. Cambridge. (Mass).
- FODOR, J.A. (1983), *The modularity of Mind*. MIT. Cambridge.
- FODOR, J.A. (1985), «Fodor's Guide to Mental Representations: The intelligent Auntie's Vade-Mecum», *Mind*, vol. XCIV.
- FODOR, J.A. (1987), *Psychosemantics*, MIT. Cambridge, (Mass).
- GOLDMAN, A.I. (1973), «A Causal Theory of Knowing» en Davis, S. (1983).
- GOODMAN, N. (1965), *Fact, Fiction and Forecast*. Bobbs Merrill. Indianápolis. Indiana.
- GRAHAM, G. & STEPHENS, L. (1985), «Are Qualia a pain in the neck for functionalist?», *American Philosophical Quarterly*. Vol. 22.
- GRICE, H.P. (1961), «The Causal theory of Perception» en Davis, S. (1983).
- HAMLYN, D.W. (1957), *The Psychology of Perception*. Routledge and Kegan Paul. London.
- HARMAN, G. (1973), *Thought*. University Press. Princeton.
- HARRISON, B. (1973), *Form and Content*. Blackwell. Oxford.
- HARRISON, B. (1982), «Identity, Predication and Color», *American Philosophical Quarterly*. Vol. 23.
- HAUGELAND, J. (1978), «The Nature and Plausibility of Cognitivism» en Haugeland (1981).
- HAUGELAND, J. (Edt.), (1981), *Mind Design*. MIT. Cambridge.
- HEMPEL, H. (1935), «The Logical Analysis of Psychology» en Block (1980).
- HURVICH, L.M. (1981), *Color Vision*. Sinauer. Sunderland (Mass).
- KIM, J. (1972), «Physicalism and the Multiple Realizability of Mental States» en Block (1980).

- KLEENE, S.C. (1974), *Introducción a la Metamatemática*. Tecnos. Madrid.
- KRIPKE, S. (1971), «Identity and Necessity» en Munitz 1971.
- KRIPKE, S. (1972), «Naming and Necessity» en Davidson & Harman (1972).
- LEWIS, D. (1966), «An argument for the Identity Theory», *Journal of Philosophy*, 63.
- LEWIS, D. (1969), «Review of Putnan» en Block (1980).
- LEWIS, D. (1970), «How to define Theoretical Terms», *Journal of Philosophy*, 67.
- LEWIS, D. (1972), «Psychophysical and Theoretical Identifications», en Block, 1980.
- LEWIS, D. (1978), «Mad Pain and Martian Pain» en Block (1980).
- LINDSAY, P.H. & NORMAN, D.A. (1977), *Human Information Processing*. Academic Press. New York.
- LOAR, B. (1981), *Mind and Meaning*. University Press. Cambridge.
- LYCAN, W.B. (1974), «Kripke and the Materialists», *Journal of Philosophy*, n. 79.
- LYCAN, W.G. (1981), «Psychological Laws» en Biro & Shahan (1982).
- MALCOLM, N. (1964), «Scientific Materialism and the Identity Theory» en O'Connor (1969).
- MARTIN, C.B. & DEUTSCHER, M. (1966), «Remembering» en Davis, S. (1983).
- MUNITZ, M.K. (Edt.) (1971), *Identity and Individuation*. University Press. New York.
- NAGEL, T. (1965), «Physicalism» en O'Connor (1969).
- NAGEL, T. (1974), «What is it like to be a bat?» en Nagel (1979).
- NAGEL, T. (1979), *Mortal Questions*. University Press. Cambridge.
- NEISSER, U. (1967), *Cognitive Psychology*. Meredith Publishing. New York.
- NELKIN, N. (1987), «How sensations get their names», *Philosophical Studies*, nº 51.
- O'CONNOR, J. (Edt.) (1969), *Modern Materialism: Readings on Mind-Body Identity*. Harcourt. New York.
- PITCHER, G. (1971), *A Theory of Perception*. University Press. Princeton.
- PLACE, U.T. (1957), «Is consciousness a brain processes», en Borst, C.V. (Edt.) (1970).
- PUTNAN, H. (1960), «Minds and Machines» en Putnan (1975).
- PUTNAN, H. (1962), «Dreaming and 'Depth Grammar'» en Putnan, H. (1975).
- PUTNAN, H. (1964), «Robots: machines or artificially created life?», en Putnan (1975).
- PUTNAN, H. (1965), «Brains and Behavior» en Block (1980).
- PUTNAN, H. (1967a), «The Mental Life of some Machines» en Putnan (1975).
- PUTNAN, H. (1967b), «The Nature of mental states» en Putnan (1975).
- PUTNAN, H. (1973), «Reductionism and Psychology», en Haugeland (1981).
- PUTNAN, H. (1975), *Mind, Language and Reality*, University Press. Cambridge.
- PUTNAN, H. (1981), *Reason, Truth and History*. University Press. Cambridge.
- PUTNAN, H. (1988), *Representation and Reality*, MIT, Press. Cambridge (Mass).
- QUINE, W.V.O. (1951), «On mental entities» en O'Connor (1969).
- QUINTON, A. (1955), «The Problem of Perception», *Mind*, vol. 64.
- RORTY, R. (1965), «The Problem of Perception», *Mind*, vol. 64.
- RORTY, R. (1965), «Mind-Body Identity, Privacy and Categories» en O'Connor (1969).
- RYLE, G. (1949), *The Concept of Mind*. Hutchinson. London.
- REY, G. (1989), «Sensational Sentences», No publicado.

- SEARLE, J.R. (1980), «Minds, Brains and Programs» en Haugeland (1981).
- SEARLE, J.R. (1984), *Minds, Brains and Science*. BBC. London.
- SHAFFER, J. (1962), «Could mental states be brain processes?», en Borst (1970).
- SHAFFER, J. (1963), «Mental events and the Brain», en Borst (1970).
- SHOEMAKER, S. (1975a), «Phenomenal Similarity» en Shoemaker (1984).
- SHOEMAKER, S. (1975b), «Functionalism and Qualia» en Shoemaker (1984).
- SHOEMAKER, S. (1981a), «Some Varieties of Functionalism» en Shoemaker (1984).
- SHOEMAKER, S. (1981b), «Absent Qualia are Impossible» en Shoemaker (1984).
- SHOEMAKER, S. (1981c), «The Inverted Spectrum» en Shoemaker (1984).
- SHOEMAKER, S. (1984), *Identity, cause and Mind*. University Press. Cambridge.
- SMART, J.J.C. (1959), «Sensations and Brain Processes», en O'Connor, (1969).
- SMART, J.J.C. (1962), «Brain Processes and Incorrigibility» en Borst (1970).
- SMART, J.J.C. (1971), «Reports of Immediate Experience», *Synthese* nº 33.
- STICH, S. (1983), *From Folk Psychology to Cognitive Science*. MIT. Press. Cambridge (Mass).
- STRAWSON, P.F. (1959), *Individuals*, Methuen, London.
- WASSERMAN, R. & SACKS, O. (1988), «El caso del pintor ciego a los colores», *El Paseante*, nº 8.
- WILKES, K.V. (1978), *Physicalism*. Routledge & Kegan Paul. London.
- WILKES, K.W. (1981), «Functionalism, Psychology, and the Philosophy of Mind», en Biro & Shahan (1982).
- WILSON, W. (1979), *The Mental as Physical*. Routledge & Kegan Paul. London.
- WITTGENSTEIN, L. (1953), *Philosophical Investigations*. Blackwell. London.
- WITTGENSTEIN, L. (1967), *Zettel* Blackwell. Oxford.
- WITTGENSTEIN, L. (1980), *Bemerkungen über die Philosophie der Psychologie*. Blackwell. Oxford.



ESCRITURA Y «PHONE»: EL DESPLAZAMIENTO DERRIDIANO DE LA TRADICION ANALITICA

Jorge Pérez de Tudela Velasco

«... *Les choses ne sont pas si simples*»¹

Comenzaré por el comienzo, es decir por el título. Por el propio título que encabeza o registra esta ponencia, pero, a la vez, por el fundamento de legitimación de las pretensiones del mismo. Ahora bien, lo primero que de él se desprende es lo desmesurado de sus ambiciones. «Escritura y phonè»: he aquí un contenido que sólo podría contener, esto es, apresar o sujetar, un discurso de al menos tanta dilatación como la propia aventura de escritura (incansable, interminable, paciente) que públicamente aparece firmada por «Jacques Derrida», «J.D.». Y en cuanto a lo arriba mencionado como «tradición analítica», no es sólo que el proyecto de operar con tal concepto (si es que lo es) se enfrentaría a las mismas o parecidas dificultades de fondo; es que la propia polémica de la que aquí trataré abre precisa -y no injustificadamente- su fuego con la advertencia hecha por uno de los contendientes de que nunca se entienda lo que Derrida tenga que decir al respecto como un auténtico choque entre tradiciones prominentes -como un choque entre legados correctamente documentados, y transmitidos ante notario². Mi intervención no tiene, pues, propósito de instalarse en el «y» al que alude el título «Escritura y phonè»: pues ese «entre» dibujaría, propondría el diagrama del «lugar» atópico en que ya siempre se pactó lo que, a falta de mejor locución, tantos llaman «Occidente». Tampoco me propondré levantar acta de un no testimoniado debate entre los textos derridianos y otros textos de los que una supuesta «tradición analítica» pudiera o quisiera responsabilizarse. Más modestamente, pretendo sólo repetir, citar de nuevo un enfrentamiento que sí textualmente (nunca mejor dicho) llegó a tener lugar, dando lugar a que algunos textos derridianos hubieran de trabarse en lucha con textos en los cuales se expone (manifiesta y corre peligro) la doctrina de los actos de habla. Que este juego intertextual logra poner en juego, sin embargo, motivos de pensamiento cuya resonancia acaso haga menos escandaloso un título como el propuesto, ésto no es un juicio precipitado y autocomplaciente de valor, sino una descripción llana del contexto en que se inscribieron a sí mismas las posiciones en liza.

¹ DERRIDA, JACQUES: «Psyché. Invention de l'autre», en *Psyché. Inventions de l'autre*. Galilée, Paris, 1987, pp.11-61, p. 33.

² La advertencia es de Searle, al comienzo mismo de la *Reply* (vid. *supra*, en el texto). Derrida mismo, por lo demás, lo acepta, si bien por razones diferentes (de hecho, comenta, es lo único que acepta de la *Reply: Limited Inc a b c...*, pp. 9-10, *ed.cit.*, *supra*, en el texto).

Parecería, pues, que el título no puede ostentar sino muy dudosos títulos para sus pretensiones, y que sólo la ulterior precisión de sus perfiles logrará hacerlo operativo. Hay, empero, un término en él cuya inclusión sí tiene, a mi juicio, bases para su defensa: es el término «desplazamiento», cita aquí de esa palabra, *déplacement*, que tantas veces acontece (y que al acontecer quiere desplazar tantas cosas) en las encrucijadas sensibles del trabajo de Derrida. Desplazamiento, y no «de(s)construcción», pero sin que la preferencia que se marca en esta elección marque a su vez ningún género de reconducción de la indomeñable cadena de los términos subrayados, anotados, acuñados por Derrida (par exemple «écriture», «trace», «différance», «supplément», «hymen», «pharmakon», «marge», «entame», «parergon», etc.³, cada uno con su respectivo entorno grafemático de ocurrencia) a ningún supuesto término-matriz de los que aquellos puedan «seguirse» como las bellotas del roble⁴. No hay contraseña liminar que dé entrada al edificio derridiano -ni a su plano, ni a sus ruinas-; y sin embargo, indecible como es el funcionamiento de la práctica derridiana, y siendo la indecidibilidad la propiedad «común» a cada uno de sus emblemas, acaso «desplazamiento» sea entre todos, lo subrayo de nuevo, el nombre a subrayar desde esta múltiple operación desestabilizadora que pregunta, lo leeremos, si no es toda inscripción, por inscripción, indefinidamente desplazable -y que inscribe en el tímpano mismo de un «libro» la sentencia que dice, cito: «Ecrire autrement»⁵.

El desplazamiento derridiano de la tradición analítica, «caso» del desplazamiento, de la sollicitación, del desestabilizamiento infinitamente fiel con que Derrida desplaza el texto abierto de la metafísica mediante su ¿mera? repetición (y que repitiéndolo lo divide, lo altera o hace otro), ese desplazamiento será pues nuestro tema. Asunto de textos, y que pone en cuestión el estatuto mismo de la textualidad, la polémica tiene un estricto *iter* escrito, cuyos mojonos más notables (los mismos que aquí enmarcan la discusión) son los siguientes:

a) «Signature Événement Contexte», protocolo recogido en *Marges de la philosophie* de una comunicación presentada por Derrida al Congreso Internacional de las Sociedades de Filosofía de lengua francesa celebrado (Montreal, 1971) en torno a un tema, «La Comunicación», cuyo concepto mismo es (como no podía ser menos en Derrida, y como viene también ocurriendo en este mismo trabajo) el primero de los problemas a que se enfrenta dicha comunicación⁶.

b) «Reiterating the Differences: A Reply to Derrida», once páginas protegidas por un Copyright (subrayar ese hecho, explotar su riqueza con la más seria de las ironías, no debería serme atribuido a mí, sino a Derrida) a nombre de John R. Searle y publicadas en el número 1 de *Glyph*, 1977, pp. 198 y ss.

³ DERRIDA, J.: «Lettre à un ami japonais», en *Psyché*, cit., p. 392.

⁴ Según una imagen hegeliana a la que Derrida ha tenido ocasión de aludir: «Tener oído para la filosofía», entrevista con J.D. de Lucette Finas. Texto aparecido en *La Quinzaine Littéraire* (16-30 Nov. 1972), recogido luego en AA.VV., *Ecartis* (Fayard, París, 1973), traducc. esp. de C. de Peretti en *Jacques Derrida. ¿Cómo no hablar?* y otros textos. (Suplementos *Anthropos* nº 13), Barcelona, 1989, pp. 89-94.

⁵ DERRIDA, J.: *Marges de la philosophie*. Les Editions de Minuit, París, 1972, p. XX.

⁶ Traducc. esp.: *Márgenes de la filosofía*. Traducc. e Introducc. Carmen González Marín. Cátedra, Madrid, 1989, pp. 347-372.

c) *Limited Inc. a b c...*, amplia, morosa, para sus compañeros de ruta aplastante contestación derridiana al texto anterior de Searle, extendida a lo largo de más de setenta páginas fechadas en 1977, y como suplemento de *Glyph*, nº 1 publicado por la Johns Hopkins University Press de Baltimore and London.

d) «The Word Turned Upside Down», una reseña del «On Deconstruction» de Culler (vid. nota 8) firmada por Searle y aparecida el 27 de Octubre de 1983 en la *New York Review of Books*⁷.

Estos son, insisto, señas de identificación incluídas, los textos capitales entre los que discurre la cuestión. Otros autores han intervenido con posterioridad en la polémica por uno, otro o ningún bando (casos del citado Jonathan Culler, de Jürgen Habermas, entre nosotros del profesor Acero)⁸; pero las limitaciones usuales de tiempo, tinta y espacio aconsejan no ceñirse en este escrito de ahora, salvo las alusiones que parezcan oportunas, sino a los cuatro documentos básicos que acabo de citar. Así pues, comencemos, y comencemos, de nuevo, por el principio: esto es, por el texto inaugural.

(Suele decirse, sin embargo, que son las exigencias de la cosa misma las que deben guiar el procedimiento de exposición; tratando de Derrida, nuestra exposición no será simple, ni lineal: intentaremos, por el contrario, con un gesto modestamente doble, con una ciencia modestamente doble, hacer resonar ocasionalmente las notas de la polémica al compás mismo en que se desgrana la posición de nuestro autor. Ello puede complicar la explicación -pero también simplificarla. Verdaderamente, las cosas nunca son tan simples).

«Signature Evenement Contexte» tiene una primera -y fácil- caracterización general: según expresión literal de su firmante «analiza las premisas metafísicas de la teoría anglosajona- y profundamente moralizante- del realizativo, de los *speech acts*, de los actos o acontecimientos discursivos. Dichas premisas sustentan en Francia -creo yo- la hermenéutica de Ricoeur y la arqueología de Foucault»⁹ (así como, podría haber añadido, del trabajo de Searle en particular)¹⁰. Es quizá también, dicho más periodísticamente, la huella de una decepción anunciada: la decepción experimentada por Derrida ante el grado de valor y de fuerza deconstructiva del edificio metafísico que pueda contenerse, en última instancia, en la teoría austiniana de los actos de habla. Decepción, por cuanto el propio Derrida ha admitido sentirse «à beaucoup d'égards très proche d'Austin»¹¹ y por cuanto que debido («al menos», dice) a cuatro razones (se diría que Austin no considera los actos de discurso sino en cuanto actos de comunicación; y comunicar, en el caso de los realizativos, no sería sino comunicar una fuerza por el impulso de una marca; el análisis del realizativo

⁷ Volumen XXX, nº 16, pp. 74-79. Debo al profesor Acero, de la Universidad de Granada, la primera noticia acerca de la existencia de este texto.

⁸ CULLER, JONATHAN: *On Deconstruction: Theory and Criticism after Structuralism*. Cornell University Press, 1982 (Trad. esp.: *Sobre la Deconstrucción*. Cátedra, Madrid, 1984). Jürgen Habermas: *Der philosophische Diskurs der Moderne*. Suhrkamp, Frankfurt A.M., 1985 (Traducc. esp.: *El discurso filosófico de la Modernidad*. Taurus, Madrid, 1989). Juan José Acero: «Derrida vs. Austin-Searle: ¿dos tradiciones en pugna?», en *J. Derrida. ¿Cómo no hablar?... cit.*, pp. 123-126.

⁹ DERRIDA, J.: «Tener oído...», *cit.*, pp. 92-93.

¹⁰ DERRIDA, J.: *Limited Inc...* *cit.*, pp. 9-10.

¹¹ *Op.cit.*, *loc. cit.*

se sustraería a la autoridad del «valor de verdad», a la oposición V/F, sustituyéndola a veces por el valor de la fuerza, o de diferencia de fuerza, ilocucionaria o perlocucionaria; el enunciado realizativo, a diferencia del constataivo, no tiene un referente externo: no describe nada colocado ante él y preexistente a él, más bien produce una situación o la transforma, en suma, opera), por estas razones, digo, podría pensarse que con Austin salta por los aires el concepto clásico de comunicación como comunicación puramente semiótica, lingüística o simbólica, según una cierta pre-determinación típicamente «metafísica» de la idea de «comunicación» que la entendería como «le véhicule, le transport ou le lieu de passage d'un sens et d'un sens un»¹². Y sí hay, aparentemente, algo de nietzscheano en la aventura de Austin, algo casi humeano (y por ende, según algún intérprete, deconstructivo o afín a la deconstrucción)¹³ en esa búsqueda austiniana de una clarificación firme y distinta de lo que acaso se resista por principio (y así lo habría enseñado esa misma búsqueda) a la categorización definida, al dominio conceptual: es decir, el reino de la *proferencia*, heredero del de la *conciencia*, heredero del de la *esencia*, formas todas de la *presencia*¹⁴. Pero al ser interrogada de nuevo, la empresa austiniana revela que contiene mucho más que esa mera puesta en cuestión del lingüisticismo y de la autoridad del código, hecha desde la propia analítica del lenguaje, que inicialmente ha llamado la atención de su comentador¹⁵. Y eso de más que contiene le hará (inevitablemente) reinscribir un análisis «patiente, ouverte, aporétique, en constante transformation»¹⁶, y que pudo haber sido radical, intramuros otra vez del modo de saber cuya forma teleológico-ideal viene determinada por «la» metafísica. ¿En qué sentido ocurre tal cosa? Véamoslo brevemente.

Forma ya parte del caudal común de la cultura «filosófica» de hoy que la llamada «deconstrucción» denuncia, disecciona, subvierte y a la postre trata de desplazar la denominación de un gesto filosófico por excelencia, a cuyo imperio apenas cabe encontrar, de Platón a Heidegger, un sólo contraejemplo: el gesto que establece a) una oposición irreductible y neta, limpia, entre pares de conceptos de máxima generalidad y sentido encontrado: conceptos como interior/exterior, material/espiritual, simple/complejo, presencia/ausencia, etc. Más en particular, la oposición signifiante/significado¹⁷ que ha animado tantos proyectos filosóficos de corte platónico-transcendental. A renglón seguido, b) una primacía jerárquica y una correspondiente caracterización positiva de uno de los dos polos, el polo del significado, de lo eidético y lo espiritual frente al polo del signifiante, lo sensorial y material. Primacía jerárquica que atribuye al primero de dichos polos las notas de lo «originario» y «firme», lo auto-transparente y autárquico; en tanto que del lado del segundo se colocará lo «accidental», «impuro», «secundario» y «derivado». En tal sentido, c) el acontecimiento deconstructivo solicita (*sollicitat*: provoca, agita, perturba) una

¹² DERRIDA, J.: «Signature...», *cit.*, p. 367 (trad. esp., p. 349).

¹³ HIRSCH, E.D. Jr: «Derrida's Axioms», en *London Review of Books*, 21 July - 3 August 1983, pp. 17-18, p. 17: «Hume, the deconstructionist *par excellence*...»

¹⁴ DERRIDA, J.: *Limited Inc...* *cit.*, p. 24.

¹⁵ DERRIDA, J.: «Signature...», *cit.*, p. 390 (trad. esp., p. 370).

¹⁶ *Op.cit.*, p. 383 (trad. esp. p. 363).

¹⁷ O, todavía más generalmente, la oposición realidad/signo. Vid.: Cristina de Peretti: *Jacques Derrida. Texto y Deconstrucción*. Anthropos, Barcelona, 1989, p. 30.

ordenación «platónica» que aquí estaría sellada por la mūltisecular y en absoluto inocente subordinación de la escritura a la palabra hablada, de la *grammè* a la *phonè*, de la comunicación escrita a la comunicación oral.

Ahora bien, supuesta esta forma de comprensión deconstructiva de la metafísica, la apresurada voz pública no ha querido leer, bajo el complicado dispositivo derridiano, sino el intento de invertir, sencillamente, el tal estado de cosas; inversión que, respetando la estructura misma de la separación entre pares de conceptos, daría pues paso a la posibilidad de un dominio, ahora, de la escritura sobre el habla, de la materia sobre la forma -o del significante sobre los significados. Nada más alejado de la textualidad. Rebelde como es, por definición, a sujetarse a ningún conjunto finito de caracterizaciones, y menos un conjunto de caracterizaciones clásicas («doctrina», «método», «teoría», «operación»...), la deconstrucción insiste ante todo en el cuidado con que evita reproducir o re-iterar, mediante su puesta en funcionamiento, el sistema mismo jerárquico, binario y teleológico que pretende, *en el límite y desde los bordes*, deconstruir. Pero propugnar algo así como una nueva (y vengativa) preponderancia del significante ¿quebraría realmente el gesto filosófico, la instauración de la dicotomía -también axiológica- «fundamento»/«fundado»? Y así empieza, bajo esta sospecha sospechada cierta, la «interminable interpretación»¹⁸ que si conduce, como lo hace, a subrayar la palabra «escritura», no lo hace refiriéndose a la escritura en el sentido habitual del término, sino a una cierta «archi-escritura» (de la que las escritura y habla usuales son sólo «casos» o «efectos»); lo cual no es sino el «nombre» dado a la estructura caracterizante de todo signo en general como ya siempre desplazado respecto de sí mismo y de los otros signos, necesariamente dividido y auto-dividido, presencia ya siempre escindida, trabajada por la ausencia, posibilidad infinita de reiteración de una mismidad siempre distinta.

(Signos convencionalmente «escritos» y signos convencionalmente «orales», por tanto, no podrían diferenciarse nunca por la nota de la «iterabilidad»; pues en ésta tienen, por el contrario, una comunidad de estructura. Aserto, este último, que se diría muy atendible, por cuanto forma la sustancia misma del «argumento» que Searle, en su «Réplica», pretende extrañamente oponer a Derrida -o, mejor dicho, oponer a una pretensión gratuitamente atribuida a Derrida de «diferenciar» signos orales y signos escritos a base de una supuesta iterabilidad exclusiva de los segundos. Dejemos aparte la anécdota, por cierto bastante divertida -léanse los párrafos derridianos sobre la perplejidad que le produce tener que defenderse de su «propio» argumento...-. Desde Derrida o desde Searle, la opinión es sustancialmente la misma: cualquier elemento lingüístico, cualquier «elemento» de un sistema de representación regulado por reglas, o es iterable, repetible, o esas mismas reglas que lo gobiernan no tendrían ocasión de ser aplicadas. Las consecuencias de este «simple» hecho, consecuencias extraídas por Derrida en decenas de páginas, distan mucho de ser tan simples¹⁹.

¹⁸ Vid. la entrevista con Christian Descamps publicada en «Jacques Derrida. Una teoría de la escritura, la estrategia de la deconstrucción», nº 93 de la Revista *Anthropos*, Barcelona, 1989, pp. 26-29.

¹⁹ SEARLE, J.R.: «Reiterating...», *cit.*, p. 199. J. Derrida: *Limited Inc...* *cit.*, p. 19.

En este equívoco campo se juega pues el leer, reescribiéndolo, el texto austiniano. Pero este campo es «general», y de ahí que la lectura derridiana de Austin se inscribe en el contexto de una re-exposición de la posición general derridiana en torno a un problema que es también (volveremos sobre ello) el problema de la relación entre lo «serio» y lo «ficticio», lo «verdadero» y lo «metafórico». No es casual, en tal sentido, no *podía* quizá ser casual, que la explicación con Austin se entreteja con la cuestión de la posibilidad, o no, de delimitar un núcleo unitario de sentido para el término «comunicación»; y ambas, a su vez, con la cuestión de la posibilidad, o no, de comprender eso que la filosofía llama «la experiencia» mediante generalización de los rasgos que posibilitan una estructura grafemática o «escriturística» en general. De esas dos cuestiones, la primera es (aparentemente) abandonada de inmediato (quizá recorra, de hecho, la totalidad de este escrito, y de otros muchos de nuestro autor): ¿a qué preguntarse, en efecto, si de los dos campos de sentido de la palabra «comunicación», el «semio-lingüístico» y el «físico», uno es el «originario» y «propio», el otro el «derivado» y «metafórico», cuando el concepto mismo de «metáfora» o transporte metafórico que aquí habría que emplear está constituido a su vez por un valor de desplazamiento o transporte, esto es, de «comunicación»?²⁰. Es toda la extraña «lógica» de lo que ya siempre se antecede a sí mismo, de la interna metaforicidad de todo discurso acerca de la metáfora, la que aquí irrumpe, naturalmente, sobre la página. Ahora bien, qué sea la comunicación y cuál su estatuto, a los efectos de nuestro autor, no es indiferente, porque la doctrina clásica acerca de la escritura ha considerado ésta, ante todo, como un medio de comunicación. El medio, justamente, de que se valen los hombres cuando quieren asegurar y extender la transmisión de un contenido uno y unívoco de significado, presente como tal a una conciencia, que primera y primariamente se transmite por la voz o por el gesto -y que debe transmitirse de forma que se altere lo menos posible la pureza de su intención²¹.

El sistema filosófico interprete pues a la escritura en el contexto del paradigma de la (primacía de la) *presencia*, en cuyo marco se elaborará la teoría del signo como representación de las «ideas» (que, a su vez, representan a los objetos de la intuición). Desde este horizonte (que es el horizonte mismo del interpretar clásico como búsqueda del sentido en el horizonte de la inteligibilidad), la escritura aprecherà como una especie del género «comunicación», en donde esta última es concebida, insisto, como la transmisión de un contenido ideal, y donde la primera, representación de una representación, extendería de forma continua, homogénea y económica los poderes iniciales de una «primera» forma de representar. Siendo el predicado de «ausencia» la nota peculiar que distinguiría dicho medio «secundario» de representación, mimesis de mimesis, de las otras formas posibles de vehicular un sentido inmutable²².

Ahora bien, si llegara a mostrarse, frente a esta interpretación predominante, que la «ausencia» supuestamente peculiar al signo «escrito» caracteriza, por el contrario, no sólo a *todo signo*, a *todo lenguaje*, sino incluso a la estructura de *toda «experiencia»*, Derrida

²⁰ DERRIDA, J.: «Signature...», *cit.*, p. 368 (trad. esp., p. 350).

²¹ *Op.cit.*, pp. 369-372 (trad. esp., pp. 351-353).

²² *Op.cit.*, *loc.cit.*

tendría allanado el camino para un «desplazamiento general» que trasladaría el peso de la determinación de los conceptos justamente a una «escritura» ya no dominada ni dominable por el horizonte transcendental-hermenéutico de la verdad del sentido, sino «dominadora» ella, ahora, de las categorías («signo», «concepto», la propia «comunicación») a las que una cierta tradición («la» tradición) habría querido someterla. Tal es el *télos* de la argumentación²³, y a desarrollarlo dedica en efecto Derrida algunas páginas. Se trata, en esencia, de hacer ver:

1) que es un rasgo estructural, constituyente, de la marca escrita, el poder ejercer efectos, funcionar, *in absentia*: en ausencia y más allá de la «muerte» del emisor, del receptor, el referente y aún (como pudo reconocer un Husserl) en ausencia del «significado» de la expresión. «Toute écriture doit donc, pour être ce qu'elle est, pouvoir fonctionner en l'absence radicale de tout destinataire empiriquement déterminé en général»²⁴. Dicho con otras palabras: es esencial que una marca escrita pueda seguir siendo *legible* con independencia, con total separabilidad de la intención-de-significación que la «animaba» en el momento de ser emitida. Para que una marca funcione como tal es pues estructuralmente *necesaria* la *posibilidad* de funcionar al margen de y sin la intención o la atención *absolutamente* actual y presente, *plena*, del «primer» emisor²⁵. Toda intención, digámoslo brevemente, es a priori *différente*²⁶, *ninguna* puede ser consciente del todo²⁷. (En vista de ello, cuando Searle «opone» al supuesto argumento de Derrida que la separabilidad de significante y significado, de significante e intencionalidad, no es un rasgo específicamente «grafemático», sino algo propio de *todo* sistema de representación, repite el mismo gesto, antes aludido, de poner en juego «contra» Derrida la argumentación misma que éste había esgrimido ya, y con -inconscientemente- el mismo resultado)²⁸. La posibilidad de la escritura, por tanto, se encuentra directamente vinculada a un absoluto de la ausencia que ya no podrá ser regido por el paradigma de la presencia y la re-presentación que presuntamente «suple» a la primera, junto con sus derivados de costumbre (normal/anormal, esencial/accesorio, etc.)

2) Que, por todo ello, pertenece a la estructura constituyente de la «escritura» una esencial *repetibilidad* o *iterabilidad*. «Escribir», dijimos antes, es tanto como hacer separable lo escrito del «querer decir»²⁹ inicialmente responsable absoluto del mismo y

²³ Supuesto que lo sea, lo cual no todos los comentaristas de Derrida aceptarían: «I endeavour, to some extent against his express intent, to treat the work of Jacques Derrida as containing philosophical opinions and arguments to support them» (Staffan CARLSHAMRE: *Language and Time. An Attempt to Arrest the Thought of Jacques Derrida*. Acta Universitatis Gothoburgensis, 1986, preface). Derrida mismo ha hablado de «demostración». A mi juicio, la naturaleza exacta de los hilos que anudan las propuestas derridianas está por precisar.

²⁴ DERRIDA, J.: «Signature...», *cit.*, p. 375.

²⁵ *Op.cit.*, pp. 376-377 (trad. esp., pp. 357-358).

²⁶ DERRIDA, J.: *Limited... cit.*, p. 28.

²⁷ *Op.cit.*, p. 45.

²⁸ SEARLE, J.R.: «Reiterating...», *cit.*, p. 201.

²⁹ «Querir-decir», como se sabe, es la expresión elegida por Derrida para «definir, si no traducir» (J. Derrida: *La voz y el fenómeno*. Traducc. P. Peñalver, Pre-textos, Valencia, 1985, p. 58) el *bedeuten* husserliano. Vid.: «La forme et le vouloir-dire», en *Marges...*, *cit.*, pp. 185-209.

entregarlo a una deriva, a una errancia estructural³⁰. Emancipándose así del «padre», es decir de la conciencia, del *Lógos*, el grafema deviene algo susceptible de ser identificado y re-identificado como siempre el «mismo» en una infinidad de contextos diferentes³¹. Desde la terminología derridiana, también pertenece, pues, al signo escrito la *citabilidad* (o citacionalidad), esto es: la omnipresente posibilidad de poder ser puesto entre comillas, de ser citado. Ahora bien, siendo esencialmente «citable», deberá igualmente sostenerse.

4) Que el signo escrito tiene como condición ineludible de posibilidad de ser «separado» del contexto inicial en que fue emitido e «injertado» en una indefinida pluralidad de contextos distintos. Circunstancia que irá de la mano con a) la deconstrucción del principio de determinabilidad o saturabilidad total de los contextos; b) la deconstrucción de la autoridad de los códigos como sistemas finitos de reglas; c) paralelamente, la irrupción de una fuerza de ruptura con el contexto que toda marca grafemática lleva consigo, fuerza de ruptura que sustrae al signo escrito a la dominación de ningún contexto o código que pretenda «cerrarse sobre él» y controlar la diseminación -que no polisemia- de sus efectos de sentido, producidos al inscribirse en el seno de nuevas cadenas significantes³².

De nuevo, es una extraña «lógica» (o, mejor, «gráfica»)³³ la que aquí se dibuja: con un dramatismo que quizá desagrade al habituado a la precisión formal (salvo que acaso éste no sea justo con la materia), Derrida reconoce (lo ha hecho muchas veces) que delineando así las condiciones de posibilidad de la escritura también delinea, en rigor, las condiciones de su propia imposibilidad, de la imposibilidad de su pureza³⁴. Pero ésta es, si se me permite la expresión, la más afirmativa de las dialécticas negativas³⁵. Y su ambición, repito, es marcar el contorno no de la marca escrita, sino de todo signo en general: la posibilidad estructural (esto es, necesaria) de una *cierta* no-presencia del emisor, del receptor, del contexto y del significado, conviene a *toda marca*, así sea oral o gestual; y conviniéndole la convierte en grafema, esto es, «la *restance* non-présente d'une marque différentielle coupée de sa prétendue 'production' ou origine»³⁶. Estructura de toda marca, la iterabilidad (alteridad y repetición) es además, y ya en suma, la marca de *la propia experiencia*³⁷. En la medida, en efecto, en que puede aceptarse «qu'il n'y a pas d'expérience de *pure présence* mais seulement des chaînes de marques différentielles»³⁸, los caracteres del signo *son* los

³⁰ DERRIDA, J.: «Signature...», *cit.*, p. 381 (trad. esp., p. 361); vid. nota siguiente.

³¹ *Op.cit.*, p. 376, (trad. esp., p. 357).

³² *Op.cit.*, p. 377 (trad. esp., p. 358).

³³ DERRIDA, J.: *Limited...cit.*, p. 20.

³⁴ Vid. *Limited... cit.*, pp. 53 y ss., y *passim*.

³⁵ Expresión que, voluntariamente paradójica, tampoco pretende zanjar este espinoso asunto. Como se sabe, Habermas ha encontrado paralelismos entre el trabajo deconstructor y la dialéctica adorniana. Otros han hablado de su parentesco con la teología negativa (vid. a este respecto «Comment ne pas parler», en *Psyché*, *cit.*, pp. 535-597). La deconstrucción, que ante todo es, señala Derrida, un *acontecimiento*, escapa a todas estas caracterizaciones, sin embargo (aunque no les sea del todo ajena).

³⁶ DERRIDA, J.: «Signature...», *cit.*, p. 378 (trad. esp., p. 359). En orden a evitar errores como los cometidos por Searle, *Limited Inc. a b c...* dedica buen espacio a resaltar la no-identidad del concepto de «restance» y el concepto de «permanencia» (vid. p. 23).

³⁷ *Op.cit.*, *loc. cit.*

³⁸ *Ibidem*.

caracteres de una «experiencia» que viene a ser así «determinada», como la escritura, por el *espaciamento* (*espacement*): espaciamiento que escinde, divide, aleja a cada elemento del texto-mundo («Omnis mundi creatura quasi liber est et pictura...») de todos los demás elementos que con él se encadenan (contexto interno), de todo referente presente, objetivo o subjetivo (referente externo); pero, a la postre, también de sí mismo. El texto que es el mundo estaría así constituido por unidades de iterabilidad, unidades que, constituidas en su identidad por la iterabilidad, esto es, por la posibilidad de re-injertarse, re-inscribirse o re-insertarse en nuevos e inevitables contextos, arrastran al mismo tiempo la no-identidad consigo mismos: elementos ya siempre diferidos, desplazados y aplazados, deportados, escisiones de sí mismos y de los demás, presencias ya siempre habitadas por la ausencia; rasgos y rasgaduras cuyo origen, diciéndolo con Derrida, se perdió desde siempre en un camino³⁹. Todo lo cual significa tanto como decir que lo que aquí se da a pensar es la paradójica estructura del «acontecimiento» en general, de lo que pretende ser único, singular, pleno y puro, y no puede, sin embargo, no ser re-iterable, él mismo ya siempre una cita, y citable⁴⁰.

El gesto derridiano ¿le ha colocado en el «exterior» de toda la tradición llamada «filosófica»? Obviamente, no. Le coloca (si es que tienen aún sentido las metáforas, topológicas u otras; si es que es imposible no hablar *en* metáforas) en una «región» más ambigua: la frontera, o el límite mismo. No el «borde» (inexistente, como inexistente es la escritura) de «lo» (si es que todavía cabe apelar a la distinción singular/plural; si es que es imposible no recurrir a ella) metafísico, sino la *travesía* que perfora, diagonalmente, aviesamente, la región de lo metafísico. Comprendemos ahora, pues, cómo esa «posición»⁴¹ pudo, en un momento, en un signo, no parecerle tan alejada de la posición de Austin. ¿Cómo no sentirse próximo (en estas cuestiones, que también son cuestiones de proximidad y de lejanía) al Austin que denuncia la «falacia descriptiva», a la que tan proclive ha sido la filosofía -esto es: la arbitraria reducción de los usos cognitivamente valiosos del lenguaje a la transmisión de información y el registro de estados de cosas, de hechos, en detrimento de los otros muchos restantes? ¿O al Austin que obstinándose en encontrar una demarcación fiable entre los constatativos y los realizativos deja entrever cada vez más, en esa misma operación, que esa demarcación buscada nunca puede ni podrá estar clara, y que en una nota a pie de página de *How to do things...* registra al fin la sospecha de que acaso sean todas las proposiciones realizativas (y el «describir» un caso particular del «hacer»)? ¿O al Austin que discute la moralidad de una teoría que permita distinguir las emisiones formales de promesa de las supuestas intenciones reales -y contrarias a lo prometido que interiormente

³⁹ Para ser exactos, «dont l'origine ne saurait être perdue en chemin» (*Op.cit.*, p. 381).

⁴⁰ DERRIDA, J.: *Op.cit.*, p. 391 (trad. esp., p. 370) (a propósito, vid. *infra*, de la posibilidad de la firma). Es cierto, sin embargo, en general, que «... la question la plus insistante, dans Sec (je propose cette abréviation pur «Signature événement contexte») c'est de savoir ce que peut être un événement...» (J. Derrida: *Limited Inc... cit.*, p.9).

⁴¹ Si acaso es cierto que Derrida no mantiene «doctrina» o «teoría» ni tiene «método» alguno, sí que defiende, en cambio, «posiciones». Vid.: *Positions*. Les Editions de Minuit, Paris, 1972. (Trad. esp., en Pre-textos, Valencia, 1977).

diga luego tener el aparentemente (realmente) comprometido «por su palabra»?⁴². Por todo ello, el programa austiniano apunta, en efecto, hacia una «superación» analítica del propio análisis del lenguaje, gesto doble, ciencia doble que la deconstrucción mira con simpatía. ¿Por qué, sin embargo, este proyecto inicialmente prometedor no tarda prácticamente nada en retornar al hogar metafísico? Para Derrida, por una razón sobre todo: Austin habría ignorado el sistema de predicados «grafemáticos» que opera en la propia estructura de la locución⁴³; dicho en otros términos, habría ignorado (o, lo veremos, más bien *querido* ignorar) la función de la estructura parasitaria que ya desde siempre actúa en toda marca como su propia condición de posibilidad-imposibilidad. ¿Qué quiere decir esto?

Quiere decir, en primer término, que la empresa austiniana estaría ya de antemano condenada al fracaso (en cuanto a una supuesta capacidad de transgresión de lo metafísico) desde el momento en que su tarea analítica se escribe en el marco de una doble *exclusión*:

a) la exclusión de las *infelicités* que pueden presentarse acompañando a la emisión tanto de realizativos como, en general, a la ejecución de cualquier otro tipo de acciones. Austin reconoce abiertamente que en toda acción se da una posibilidad omnipresente de fracaso; pero a la vez aparta de su consideración ese fenómeno como una suerte de «accidente» ajeno a la delimitación ideal de una esencia⁴⁴.

b) Si en el punto anterior se ha observado la exclusión de negatividades que, afectando a toda acción, afectarán también (constantemente, de hecho) a la emisión de los realizativos, ahora se excluirán aquellas negatividades estructuralmente inevitables que, afectando a toda expresión, afectan también a los realizativos en tanto que éstos son expresiones: a saber, la posibilidad que tiene toda expresión de ser formulada de forma «hueca o vacía», de un modo peculiar, por ejemplo por un actor en un escenario, incluida en un poema o en un soliloquio. Según Austin, en estos casos el lenguaje no es usado realmente en serio, sino de formas que dependen del uso normal, formas incluidas en una doctrina acerca de las «decoloraciones» de que el lenguaje es capaz⁴⁵.

Este doble gesto excluyente, no hace falta decirlo, resulta solidario, coherente. Apunta al retorno austiniano a la seguridad del lenguaje «ordinario», a las circunstancias «ordinarias» de emisión de las expresiones, a la «normal» *presencia* del «auténtico» querer-decir del emisor, a una comunicación perfectamente transparente de su intención⁴⁶. No es pues

⁴² AUSTIN, J.L.: *How to do things with Words*. Oxford University Press, N. York, 1962 (trad. esp., *Cómo hacer cosas con palabras*. Paidós, Barcelona, Buenos Aires, 1982), pp. 3, 103, 9-11 (trad. esp., pp. 43, 147 y 50-53).

⁴³ DERRIDA, J.: «Signature...», *cit.*, p. 383 (trad. esp., p. 363).

⁴⁴ Para ser exactos: A) la *infelicity* es la posible enfermedad que la emisión de un realizativo comparte con todos los actos convencionales, con todos los actos que tienen un carácter genéricamente ritual o convencional; b) con las otras acciones, comparte no propiamente la posibilidad de *infelicity*, sino ciertas «whole dimensions of unsatisfactoriness» que son distintas, o distinguibles, de las *infelicités*. Sin embargo, c) Austin añade que acaso una doctrina general podría abarcar *infelicités* y «these other 'unhappy features' of the doing of actions» y que, en todo caso, «we must just remember, thought, that features of this sort can and do constantly obtrude into any case we are discussing» (Austin: *op.cit.*, pp. 18-21, trad. esp., pp. 60-63).

⁴⁵ *Op.cit.*, pp. 21-22 (trad. esp., p. 63).

⁴⁶ DERRIDA, J.: «Signature...», *cit.*, p. 384 (trad. esp., p. 364).

esta, como Searle arguye, una mera exclusión metodológica, estratégica, quizá provisional, de Austin⁴⁷. Sobre que la oposición «decisión estratégica»/«presupuesto metafísico» no contaría con el apoyo de Derrida, es que, para éste, la citada doble exclusión, exclusión de la omnipresente posibilidad de fracaso, exclusión de la omnipresente posibilidad de ser citadas que tienen las preferencias, se transforma en el signo de cómo reina la metafísica más tradicional sobre la herencia austiniana⁴⁸, con todas sus implicaciones. Austin, pues, habría subordinado, y pretendidamente «sacrificado», por una decisión en apariencia metodológica, estratégica, pero tan «metafísica» como todas las de esta calaña, lo «anormal» a lo «normal», lo «parasitario» a lo «propio», lo «no serio» a lo «serio», lo «vacío» a lo «pleno», lo «no literal» a lo «literal»... Signos, una vez más, de todo ello:

1) El intento austiniano por justificar una inicial preferencia por la primera persona del singular del presente de indicativo en la voz activa dentro del marco de una búsqueda de algún criterio gramatical preciso que sirva para distinguir las expresiones realizativas. Esta preferencia estaría fundada en el hecho de que, tratándose de acciones, las acciones son llevadas a cabo por personas; así pues, el «yo» *debe* aparecer, lo cual se consigue, según Austin, en las expresiones orales por el hecho de que quien habla es la persona que emite la expresión, en las escritas por el hecho de que el autor del texto suele estampar su firma en él⁴⁹. Parece inútil insistir sobre la transcendencia de esta reinstauración del «yo» como centro presente y activo de sus propios actos.

2) El hecho de que, en expresión austiniana, «the total speech act in the total speech situation is the *only actual* phenomenon which, in the last resort, we are engaged in elucidating»⁵⁰. Expresión que remitiría a un tercer problema y consiguiente tercera piedra de discordia entre ambos autores, a saber, el de la posibilidad, o no, de una saturación o determinación exhaustiva de los contextos de producción de las preferencias; siendo así que para Austin, leemos en Derrida, los análisis exigen «en permanence une valeur de *contexte*, et même de contexte exhaustivement déterminable, en droit et téléologiquement»⁵¹. Supuesto este, de la determinabilidad del contexto, cuyo rechazo constituye, insisto, una pieza a resaltar en la cadena de las posiciones derridianas.

En un sólo ademán múltiple, pues, Austin a) excluye el riesgo esencial de fracaso de toda preferencia; b) excluye la omnipresente posibilidad de ser puesta entre comillas de toda marca; c) reintroduce la idea de comunicación de un sentido intencional emanado de una conciencia plenamente presente a sí misma y última responsable de todos y de la totalidad de sus actos; caso particular, de todos y la totalidad de sus actos de habla. Pero ¿cómo justificar, salvo por imperativo de un viejísimo ideal ético-teleológico de «pureza», exclusiones que se perfilan sobre un fondo muy diferente de este otro: el fondo de un reconocimiento explícito de la intrínseca posibilidad que tiene todo acto de fracasar, y toda

⁴⁷ SEARLE, J.: «Reiterating...», *cit.*, p. 205.

⁴⁸ DERRIDA, J.: *Limited Inc...* *cit.*, p. 65.

⁴⁹ AUSTIN, J.L.: *Op.cit.*, pp. 55-67 (trad. esp., pp. 99-111).

⁵⁰ *Op.cit.*, p. 147 (trad. esp., p. 196).

⁵¹ DERRIDA, J.: «Signature...», *cit.*, p. 383.

expresión de ser usada «en vano»? Lo que Austin habría tratado, poniéndolo a la vez de manifiesto, de ocultar o acallar, sería eso mismo que ya hemos denominado aquí más de una vez, con Derrida, estructura grafemática de todo signo y, por generalización, de toda «experiencia». Estructura grafemática que también puede nombrarse como estructura parasitaria, y que tiende a hacer valer la iterabilidad diferencial como rasgo insoslayable de la «archi-escritura», o escritura en general⁵².

Porque *todo* grafema *tiene* que *poder* funcionar con referencias a un absoluto de la ausencia, a un absoluto también de posibilidades de re-iteración, la estructura grafemática de toda comunicación es sencillamente inexplicable con exclusión o al margen de una teoría general del parasitismo estructural de las marcas. Y en virtud de esa misma grafematicidad estructural, ningún realizativo podría tampoco tener la más mínima probabilidad de éxito si no se tratara, ya, él mismo, de la repetición, de la «cita» de una fórmula, de un rito; un «rito» que, en cuanto tal, es intrínsecamente convencional, intrínsecamente vulnerable, sometido a una permanente amenaza no-accidental de fracaso⁵³. Y por último, también en virtud de la estructura grafemática de la comunicación podrá sostenerse (contra el presupuesto metodológico austiniano) que la condición de posibilidad de que se produzca un efecto normal de firma es justamente su condición de imposibilidad, de la imposibilidad de que se produzca en un sentido «puro» o «pleno». Pues no hay firma que funcione sin ser, ya, esencialmente imitable, falsificable, repetible. «C'est là l'originalité énigmatique de tous les paraphes. Pour que le rattachement à la source se produise, il faut donc que soit retenue la singularité absolue d'un événement de signature et d'une forme de signature: la reproductibilité pure d'un événement pur»⁵⁴. Y así, poniendo de relieve las dificultades enormes de la idea de «acontecimiento», ese acontecimiento que sólo es el «mismo» en tanto que diferente y auto-diferente, que sólo repite su repetibilidad, es como ésto que todavía puede (y debe) conservar el nombre de «escritura» hace estallar la comprensión clásica de la experiencia en el horizonte del sentido, de la verdad, la conciencia, el presente, la determinación teleológica y (sombras de Leibniz) la plenitud.

...

Considerado, como suele decirse, «en conjunto» (si es que tal cosa puede aún decirse aquí), «Signature...» aparece como una gran interrogación. Lo interrogado puede ser el estatuto y la función del copyright, el valor de acontecimiento (y el valor de acto) o bien «en quoi le discours philosophique traditionnel relève-t-il de la fiction»⁵⁵. Derrida mismo ha aludido a cada una de estas posibilidades, y también nosotros hemos tenido ocasión de subrayar, al paso, el relieve adquirido en estos textos por esos eslabones de la cadena «conceptual» derridiana. Hay otros pasajes, sin embargo, en que irrumpe y se inscribe, dentro siempre del texto derridiano, una cuestión de quizá mayor calado aún que las propuestas, y que a mi juicio señala más precisamente el marco de referencia del trabajo

⁵² *Op.cit.*, pp. 388-390 (trad. esp., 368-369).

⁵³ *Op.cit.*, pp. 387-388 (trad. esp., pp. 366-368).

⁵⁴ *Op.cit.*, p. 391.

⁵⁵ DERRIDA, J.: *Limited Inc*, cit., p. 49.

deconstructor. Según esta indicación a la que me refiero, Derrida no se pregunta, aquí, sino esto: qué es el idealismo, cuál es su fuerza y su necesidad, cuál, también, su límite interno⁵⁶. No podría menospreciarse fácilmente el alcance y el peso de tal interrogación. Pero no insistiré en ello. Sí insistiré, en cambio, en el hecho de que a la pregunta por no exactamente el «fundamento» del *proyecto* de idealización o de la abstracción idealizante, pero sí por aquello que lo permite⁵⁷, Derrida contesta especialmente, en *Limited Inc...*, con ese mismo término al que ya hemos tenido ocasión de hacer referencia, y que en mi opinión proporciona no poco apoyo para entender su punto de vista: «iterabilidad», «estructura de iterabilidad» o, incluso, «la ley de la iterabilidad»⁵⁸. Bajo ningún concepto podrán entenderse, sin embargo, estos «términos», al modo de nuevos nombres para un fundamento o un principio lógico o transcendental, para una «raíz» de nada. Incapaz de ser apresada, a su vez, por el *proyecto* idealizador, la «raíz» que aquí se introduce, a modo de cuña, en el texto filosófico, sería en todo caso una raíz «doble», una no-simplicidad insusceptible de identificarse o de reapropiarse que, en cambio, resulta ser aquello mismo que permite la idealización⁵⁹. Una estructura o una ley de iterabilidad, en efecto, es lo único que permitirá «pensar» un «texto general» («experiencia») constituido por «marcas» en las que no como límite, sino como condición positiva de funcionamiento, han de haber *necesariamente* dejado su marcas (re-marcándolas) las múltiples ausencias de cuya *posibilidad* vive aquélla (puesto que no nace sino para, eventualmente, suplirlas)⁶⁰. La iterabilidad, en otros términos, es lo único que permitirá «pensar» la extraña identidad auto-escindida de «elementos» que, esencialmente reiterables, se fracturan en el mismo golpe⁶¹ en que se constituyen, y no tienen otra mismidad que la que la propia iterabilidad permite: esto es, la identidad repetible, e identificable en, a través, e incluso con vistas a la alteración⁶². La estructura de iterabilidad, en este sentido, es un raro «fundamento», que no funda más identidad que la que diferencia, ni más idealidad que la que arruina⁶³, ni más pureza que la que contamina, ni más claridad que la que ensombrece. El idealismo permitido por ella, pues, se asienta sobre el deslizante terreno minado que le proporciona un «concepto» como el de «iterabilidad diferencial», bajo cuyo dominio ese idealismo ve ante todo amenazada su característica básica: la pureza de las determinaciones, la simplicidad de los límites, lo tajante de las distinciones; la identidad plena y la presencia total. Por eso es por lo que en este «espacio», que no sólo es un espacio «teórico», abierto por la iterabilidad, acontecen, enlazadas pero (siempre en contra de las sospechas de Searle) no confundidas, cosas como estas:

⁵⁶ *Op.cit.*, p. 66.

⁵⁷ *Op.cit.*, p. 64, p. 42.

⁵⁸ *Op.cit.*, p. 64.

⁵⁹ *Ibidem*.

⁶⁰ *Op.cit.*, p. 22.

⁶¹ Término cuyas implicaciones ha explorado a fondo Derrida. Vid. por ejemplo *La dissémination*, Seuil, Paris, 1972, p. 151, nota 54.

⁶² DERRIDA, J.: *Limited... cit.*, pp. 24-25.

⁶³ *Op.cit.*, p. 48.

a) se cuestiona, y radicalmente, no tanto la intención o la intencionalidad como su *télos*, es decir, la posibilidad de su cumplimiento. Como ya hemos tenido oportunidad de hacer ver, la intención o la atención no podrá alcanzar jamás eso mismo a lo que sin embargo tiende, a saber, la auto-presencia total a sí misma y a su objeto. La misma iterabilidad que la identifica, por el contrario, la aparta ya siempre de sí misma, la divide y deporta; expropia ya siempre la presencia «ideal» a sí mismo del querer-decir. La intención dirigida hacia un iterable y determinada por él a la iterabilidad, pues, no puede no ser una intención abierta a la contaminación, a la «parasitación». De ahí que pueda decirse que siempre se «quiere-decir» otra cosa que lo que se quiere decir, siempre se dice otra cosa que lo que se dice, siempre se comprende otra cosa que lo que se comprende... etc. La estructura de las marcas, en suma, es tal que a esas marcas les pertenece la posibilidad general de, dicho en inglés, el «mis-» («misunderstanding»... etc.)⁶⁴.

b) la iterabilidad de la marca emborriona, arruina, amenaza todas y cada una de las oposiciones filosóficas. De modo que el límite que (identifica -y-) diferencia lo «serio» de lo «no serio», lo «literal» de lo «metafórico», lo «puro» de lo «corrupto», etc., es ya un límite estructuralmente corrompible -y, por ende, algo que nunca podrá ser rigurosamente determinado. Caso particular de este estado de cosas, el «límite» que se diría tendría que correr entre la «literatura» y la «filosofía», entre la «ficción» y algo así como una descripción de las condiciones de posibilidad del conocimiento empírico, etc., no es un límite precisable⁶⁵.

c) En estas condiciones, tampoco resulta hacedero determinar de modo incontestable dónde se encuentra el límite entre ese empleo del lenguaje considerado «normal» por la teoría de los actos de habla, a saber, el empleo o modo en el cual el lenguaje es «usado», y ese otro modo, «anormal» o «parasitario», en el cual el lenguaje es «mencionado» o se hace, «anormalmente», objeto para sí mismo (constituyéndose así, según la lógica que ya conocemos, en ejemplo de una anormalidad a la que quien la padece está ya siempre y por principio abierto)⁶⁶.

Una vez más, hay que señalar la magnitud del alcance de lo que se está discutiendo. Pero esa relevancia no nos fuerza sino a precisar todavía más, aún a riesgo de ser reiterativos, la posición exacta de Derrida. Análisis como los anteriores, en efecto, han propiciado una lectura de dicha posición que, en la polémica posterior (y ruidosamente) levantada a este respecto, ha tendido a fijar las opiniones derridianas en un sentido parecido a este: Derrida trataría, simplemente, de «invertir» el orden tradicional. De este modo, en lugar de sostener, como siempre se ha sostenido, que las diferencias innegables entre los elementos (lingüísticos) dependen de las presencias que «son» esos elementos, habría sostenido justamente lo

⁶⁴ *Op.cit.*, p. 12, p. 9.

⁶⁵ *Op.cit.*, p. 42.

⁶⁶ «Comparing these sentences two facts are obvious: first, the same word begins both sentences, and second, the role that the word plays in an utterance of the sentence is quite different in the two cases, since in 1 it has its normal use to refer to a particular man and in 2 it does not have its normal use, but rather is talked about -as is indicated by the present of the quotation marks» (J.R. Searle: *Speech Acts. An Essay in the Philosophy of Language*. Cambridge University Press, 1969, p. 74 (traducc. esp., *Actos de Habla*. Cátedra, Madrid, 1980, p. 82). Los subrayados son míos.

contrario: las identidades, las presencias, como efectos de un supuesto sistema «más fundamental» de diferencias. Igualmente, en vez de sostener, como siempre se ha sostenido (y como Searle ha sostenido)⁶⁷, que la posibilidad de emitir una (falsa) promesa, por ejemplo sobre las tablas de un teatro, depende de su emisibilidad «en serio» (en circunstancias adecuadas, con los testigos pertinentes, etc.), habría sostenido que lo contrario es lo cierto: si puedo emitir un juramento válido, es *porque* puedo emitirlo en vano. O, tercero, en vez de sostener que si puedo hacer ficción es porque puedo describir el mundo, especular «veritativamente» acerca de él (hipótesis también de Searle), habría sostenido que, por el contrario, es gracias a los poderes narrativos del lenguaje como puede efectuarse la narración que llamamos «filosofía», o «ciencia». O, cuarto y último (aquí): no sería la «mención» un caso del «uso», sino el «uso» un caso de la «mención»...⁶⁸.

Muchas veces, por cierto, ha sido entendida así la postura de Derrida; el propio Searle, por ejemplo, en el cuarto de los textos que al principio se citaron; alguno de sus seguidores oficiales lo han expuesto así también⁶⁹. Se comprende entonces, en este contexto, reacciones como la de Habermas, que en su reciente intervención en el debate ha advertido de los riesgos que una filosofización de la crítica literaria y una literaturización de la filosofía comportarían para el rendimiento de ambas disciplinas⁷⁰. Ocurre, sin embargo, que a mi juicio lo anterior no sería una exposición inatacable de los asertos de Derrida. Lo que éste más bien ha tratado de decir es: ¿cómo debemos entender el hecho de que las formas «fingidas», por ejemplo de una promesa, no pueden fingirse sino en la medida en que el llamado acto «normal», «serio», «literal», «standard», de promesa que se repite, reproduce, mima o finge, en ese acto, es reproducido, mimado, fingido, en ese momento, justamente *como standard*, esto es, justamente, *como reproducible*, como ya desde siempre parasitable? Se trata, una vez más, de buscar no la condición de posibilidad⁷¹, sino la raíz doble que

⁶⁷ SEARLE, J.: «Reiterating...», *cit.*, p. 205.

⁶⁸ SEARLE, J.: «The Word...», *cit.*, p. 74: «The deconstructionist wants to undermine these oppositions, and so undermine logocentrism, by first reversing the hierarchy, by trying to show that the right-hand term is really the prior term and that the left-hand term is just a special case of the righthand term; the right-hand term is the condition of possibility of the left-hand term. This move gives some very curious results. It turns out that speech is really a form of writing, understanding a form of misunderstanding, and that what we think of as meaningful language is just a free play of signifiers or an endless process of grafting texts onto texts». Searle, sin embargo, añade correctamente a continuación: «But this move is only a part of a two-step procedure... the aim is to resituate, undo, or displace the entire system of values expressed by the classical opposition».

⁶⁹ Como Searle hace ver (*Op.cit.*, p. 77), es el caso del propio Culler (sin perjuicio de que la reconstrucción de este último resulte obviamente mucho más próxima a Derrida que la que Searle propone).

⁷⁰ HABERMAS, J.: *Op.cit.*, p. 246: «Das philosophische Denken wird, wenn es gemäß Derridas Empfehlungen von der Pflicht, Probleme zu lösen, entbunden und literaturkritisch umfunktioniert wird, nicht nur seines Ernstes, sondern seiner Produktivität und Leistungsfähigkeit beraubt. Umgekehrt büsst auch die literaturkritische Urteilskraft ihre Potenz ein, wenn sie, wie es Derridas Gefolgschaft in den literary departments vorschwebt, von der Aneignung ästhetischer Erfahrungsgehalte auf Metaphysikkritik umgestellt wird. Die falsche Assimilation der einen Unternehmung an die andere raubt beiden ihre Substanz» (trad. esp., *cit.*, p. 252).

⁷¹ DERRIDA, J.: *Limited... cit.*, p. 72: «L'itérabilité n'est pas davantage une condition transcendente de possibilité...».

permite la idealidad. Pero la idealidad se desenvuelve precisamente al modo aporético que se acaba de recordar. Luego si un acto «standard» depende de la posibilidad de ser repetido, y por tanto eventualmente mimado, fingido, citado, simulado, parasitado... también es cierto que la recíproca es igualmente atendible, y que esta segunda posibilidad «depende» también de la primera: lo serio depende de lo lúdico como lo lúdico de lo serio porque ambos, a su vez, «dependen» de la estructura de iterabilidad que, como hemos visto, arruina la simplicidad de la línea que separaría las oposiciones de la metafísica, por ejemplo la oposición entre lo ideal y lo no-ideal, por ejemplo la oposición entre el uso y la mención⁷². Pero seguramente el borrarse y emborronarse del límite no lo anula; más bien lo constituye. A partir de este momento, lo que ocurre es que está excluida no la exclusión, sino la simplicidad de una exclusión como, efectivamente, la que excluiría lo literario del dominio filosófico, o viceversa. Lo que ocurre, en otras palabras, es que sólo cabe dibujar cómo «rien de tout cela n'est ni simple ni assuré»⁷³.

Nada es simple, ni seguro, pero sí lo es la rotundidad con que Derrida, en respuesta a una de las críticas de Searle, y reiterando su propia posición originaria, rechaza toda la posibilidad de a) confundir ninguno de los diversos fenómenos que aquí se traen a colación (iterabilidad, citabilidad, parasitabilidad...)74; b) concebir la iterabilidad como una suerte de «género» del cual la «cita», por ejemplo, fuera una «especie». No puede serlo, entre otras cosas, porque lo que la estructura de iterabilidad conlleva es ante todo, además de lo dicho, la puesta radical en cuestión de un imperativo típicamente occidental, típicamente filosófico: *la ley del género*, la ley que prohíbe mezclar los géneros, traspasar las fronteras que marcan su mutua exclusión⁷⁵. El imperativo se presenta como «lógico»; pero obviamente (sombras de Pierce) atesora mucho de ético y no poco de estético. Impone que desde el primer momento en que un género se anuncia, se respete una norma, se excluya al monstruo, la coyunda antinatural⁷⁶. Y nuevamente aquí, en el «campo sin límites de una textualidad general»⁷⁷, la sospecha derridiana es si, pese a todo, no será sencillamente imposible mezclar los géneros. Si ningún texto, pese a no haber texto sin género, puede, en rigor, *pertenecer* a género alguno, aún cuando todo texto *participe*, insiste, de uno o de varios géneros. Porque la ley del género acaso tenga en su interior, como «condición de posibilidad», una contra-ley, un axioma de imposibilidad o una ley de impureza, de contaminación, en cuya virtud los géneros pueden siempre pasar del uno al otro, del uno a sus otros. Y otra vez, el «resorte» común a esta «perversión» o «descomposición» íntima de toda composición, mantenida en cuanto tal, tiene un nombre: repetición, la misma repetición ya antes saludada que hace indecidible la determinación, en cada caso, de los

⁷² *Op.cit.*, p. 64. Por eso es por lo que «... le partage ne passe pas entre une lecture tropique et une lecture appropriée ou littérale, juste et vrai, mais entre des capacités tropiques» (J. Derrida: «Le retrait de la métaphore», en *Psyché*, *cit.*, pp. 63-93. Traducc. esp. en Jacques Derrida: *La desconstrucción en las fronteras de la filosofía*. Introd. y traducc. Patricio Peñalver, Paidós, I.C.E.-U.A.B., Barcelona, 1989, pp. 74 y 50, respectivamente).

⁷³ *Op.cit.*, p. 4.

⁷⁴ SEARLE, J.R.: «Reiterating...», *cit.*, pp. 206-7, J. Derrida: *Limited Inc... cit.*, p. 72.

⁷⁵ DERRIDA, J.: «La loi du genre», en *Parages*. Galilée, Paris, 1986, pp. 2489-287.

⁷⁶ *Op.cit.*, p. 253.

⁷⁷ *Op.cit.*, p. 262.

acaeceres, como «acontecimientos» o «relatos», relatos de acontecimientos o acontecimientos de relato⁷⁸. ¿Por dónde pasa, inquiere Derrida, el límite entre el uso y la mención? ¿Entre constatar y realizar? ¿Es tan difícil como supone «la» tradición (y dentro de ella, la teoría de los actos de habla) encontrar ejemplos de «mezcla» de géneros a este respecto? Tomemos el caso de *invenio*: «invenio» significa «inventar» y «descubrir», y entonces es constatativo; pero significa también «producir», y entonces es realizativo...⁷⁹. ¿Cómo decidir? Más en general: ¿Qué ocurre cuando la primera línea de un poema -Derrida analiza «Fable», de Ponge- comienza así: «Par le mot *par* commence donc ce texte»? ¿No ocurre, acaso, que nos encontraríamos ante un ejemplo de realizativo que simultáneamente describe y efectúa su propio engendramiento? ¿Que *hace* lo que *dice*? Ciertamente que no todos los realizativos tienen este carácter reflexivo, esta cualidad de no «describir» otra cosa que ellos mismos, y que se describen a sí mismos en su calidad, precisamente, de realizativos. Pero ¿puede quedar explicado el lenguaje sin atender a la omnipresente posibilidad, aquí acontecida, de que en un sólo verso un enunciado confunda el uso y la mención, la auto-referencia y la hetero-referencia? La frase pongiana, según Derrida, no cita otra cosa que la propia frase citándose -y por ello los valores de uso, en ella, no son sino subconjuntos del valor de mención. La frase «cita» su propia «re-citación»; se inventa inventando el relato de su invención; une el hacer y el decir, la descripción y la acción, en un único acontecimiento inventivo que por lo mismo es siempre ya doble, escindido; o en el cual, mejor dicho, se produciría una oscilación infinitamente rápida entre lo realizativo y lo constatativo, el lenguaje y el metalenguaje, la ficción y la no ficción. Esencialmente inestable, la frase es pues inmediatamente meta-lingüística, pero con el agravante de que aquí no hay propiamente un lenguaje «anterior», un lenguaje-objeto al que ese metalenguaje pudiera referirse desde un supuesto nivel «superior» de consideración. *Todo* en la mencionada primera línea es así a la vez lenguaje y metalenguaje; y *nada* lo es, el metalenguaje resulta tan inevitable como imposible⁸⁰. «Sans la ruiner totalement, puisque aussi bien elle a besoin d'elle pour provoquer ce singulier événement, l'économie fabuleuse d'une petite phrase très simple... *déconstruit* spontanément la logique oppositionnelle qui s'en tient à la distinction intouchable du performatif et du constatif et à tant d'autres distinctions connexes»⁸¹.

El análisis derridiano de esto que efectivamente es «sólo» una simple frase puede parecer excesivamente micrológico, trivial incluso. Podría acusársele también (léase a Searle)⁸² de una cierta mala fe, la mala fe de quien voluntariamente oscurece sus proposiciones, como un terrorista del intelecto, para facilitarse luego a sí mismo el poder llamar idiotas a los demás. Pero en cuanto a esto segundo: ¿es posible, aún con la mejor fe del mundo, exponer simplemente la no-simplicidad? Y en cuanto a lo primero: «apocalípticas»

⁷⁸ *Op.cit.*, pp. 254-255, 264, 274.

⁷⁹ DERRIDA, J.: *Psyché*, cit., p. 23.

⁸⁰ *Op.cit.*, p. 22-25.

⁸¹ *Op.cit.*, p. 25.

⁸² SEARLE, J.R.: «The Word...», cit. p. 77. Searle apela, en apoyo de su queja, a Michel Foucault.

o no⁸³, falsamente hinchadas o no, las propuestas derridianas sí apuntan, a mi juicio, más allá de esta aparente micrologicidad de su empeño, a problemas de envergadura; problemas de excesivo calado como para poder ser aquí tratados con todo el detalle que merecen; pero excesivamente acuciantes, al tiempo, como para poder dejarlos pasar en silencio. Pues lo que está en juego, me parece, en esta polémica y en los textos que la roedan, no es sólo (aunque también) la posibilidad de delinear más «correctamente» la aventura deconstructiva como algo que no temerá enfrentarse a su propia imposibilidad, puesto que su interés no es otro que el de «una cierta experiencia de lo imposible»; ni como algo que nunca, en su rigor, se ha presentado como extraño a lo «literario»⁸⁴. O como algo, por último, para quien la capacidad de ser, o no, «inventiva», es una cuestión de auténtico ser o no ser (aunque no haya más invención posible, añade Derrida, que la invención de lo imposible); la deconstrucción camina, inventa, afirma, sí, pero ante todo deja venir lo otro, la aventura, el acontecimiento de lo totalmente otro; y, en tal sentido, no es performativa, realizativa, más que en la medida en que no es sólo ni únicamente realizativa, sino aquello mismo que perturba las propias condiciones de la performatividad⁸⁵.

Además de todo ello, y en modo alguno al margen de todo ello, se pone aquí de relieve la extraña «ley» que parece «centrar» descentradamente el campo de la experiencia, el campo general de la textualidad o, en el sentido derridiano del término, del «lenguaje». Pues lo que demuestra, y demuestra en acto, un acontecimiento como la frase de Ponge, según Derrida, no es otra cosa que la propiedad que tiene el lenguaje de poder siempre, sin poderlo; hablar de sí mismo⁸⁶. Ahora bien, este *dictum* nada tiene de trivial. Pone de manifiesto, por el contrario, un poder y un aspecto que tiene, para Derrida, no sólo el lenguaje, sino el «lenguaje» en que, por generalización de los rasgos del primero, se ha transformado ahora el todo de la experiencia. Y hay también razones para apoyar lo que se dice, una demostración sobreentendida. Late en efecto en todo esto, a mi juicio, una argumentación subterránea que, por lo que sé (pero puedo equivocarme), Derrida mismo nunca ha hecho explícita, pero cuyas líneas maestras podrían quizá reconstruirse. Pues ciertamente en la sentencia anterior el lenguaje ha dicho algo acerca de sí mismo, a saber, su posibilidad y su imposibilidad de hablar de sí mismo. Pero *el lenguaje no puede predicar del todo de la realidad nada que no pueda predicar de sí mismo*. En otros términos, lo que el lenguaje predica de sí mismo forma *necesariamente* parte del acervo de atributos que aquél *puede* ulteriormente atribuir a lo demás. Para un lenguaje que así se pronuncie acerca de una totalidad que inevitablemente le incluya, *él mismo* es pues siempre la constante y primera piedra en que contrastar sus opiniones, la primera y constante contraprueba que ante él mismo puede presentarse (su límite). Por ello, un lenguaje omni-referente, esto es, a la vez auto- y hetero- referente, *no puede no predicar del todo de la realidad, al menos y como mínimo*, lo que predica de sí. *Al menos y como mínimo*, un lenguaje auto-concebido como compuesto por «marcas» ya siempre «remarcadas» por otras marcas, huellas de huellas que

⁸³ En expresión usada, aunque de modo gramaticalmente distinto, por Searle: *Op.cit.*, p. 75.

⁸⁴ DERRIDA, J.: «Psyché...», *cit.*, pp. 26-27.

⁸⁵ *Op.cit.*, pp. 35, 59-61.

⁸⁶ *Op.cit.*, p. 24.

son ya siempre huellas de sí mismas, *no puede no predicar* estos mismos caracteres (u otros que no les sean contrarios) del todo de lo que hay.

Al menos y como mínimo, permítaseme insistir, pero también *sólo y nada más que eso* en el supuesto de un lenguaje en el cual, como puede que sea el caso del lenguaje derridiano, se defiende la generalización a todo el campo de la «experiencia» de los caracteres de los signos «escritos» y la comprensión de dicha «experiencia» en términos de textualidad. (Lo cual, entiéndase bien, no significaría en esta argumentación postular la supuesta «verdad» de lo que entonces se diga, sino, sólo, la *coherencia* interna del tal lenguaje, del sistema o la cadena de sus opiniones). Ahora bien, si esto es así, cuando el lenguaje derridiano predica de sí mismo la indecidibilidad del límite que corre entre lo posible y lo imposible, ese mismo lenguaje ya no podrá, si quiere predicar algo del todo de la experiencia, predicar nada que contradiga lo que así ha sido dicho y ya limita sus posibilidades; y no podrá, simplemente, dejar de incluir esa nota, si es que supone ilimitadamente generalizable el modelo del lenguaje. Por una u otra vía, pues, desde el lenguaje derridiano lo indecible del límite entre lo posible y lo imposible no puede no ser considerado «esencial». Es, una vez más, el dominio sin hegemonía de una extraña situación. Situación que quizá no pueda ser comprendida con ayuda de ningún modelo propio del pensar «lógico», «metafísico» o «representacional». Pero que acaso por eso mismo ayude a explicar otro «modelo» no sé si «anterior» o más «potente» que cualquiera de los «puramente» filosóficos, pero cuya innegable riqueza atraviesa el tiempo: *el mito de Edipo*, fundacional por tantas razones en Occidente⁸⁷.

En ninguna otra parte, en efecto, se ha plasmado con tal fuerza la ambigüedad de los fundamentos, la duplicidad de los caminos, lo indecible del resultado de los métodos. El mito narra, en efecto, cómo el salvador de la ciudad atrae sobre ella la peste, cómo el hombre que podría codearse con los dioses resulta ser un monstruo abyecto, cómo el que vio se quita, por haber visto, su propia vista -y todas las demás dimensiones de la tragedia a que desde antiguo se viene aludiendo. Pero el mito, a mi juicio, hace más: muestra la razón profunda de estos acontecimientos dobles, hermanando a la vez esta razón con la «razón» que señorea la vida. Y es que Edipo salva a la ciudad resolviendo un enigma que es a la vez el enigma del hombre y el enigma del tiempo. Y para resolverlo acude a mostrar cómo lo mismo puede ser diferente, como el ser dotado de una sola voz puede, sin embargo, no caminar siempre del mismo modo; cómo el presente, el pasado y el futuro, diferentes entre sí, se comunican constantemente entre sí. Edipo establece el orden y el ritmo, discierne las épocas en el interior de una unidad temporal superior y compleja, distingue las edades por las que, sin embargo, es un mismo hombre el que va pasando. Su «método» salva así a la ciudad, descifrando los signos dobles que le proponen un enigma mortal. Y sin embargo, la respuesta al enigma también es: él mismo, Edipo o el hombre, y en él la lógica de los géneros que continuamente se transforman, diferentes como son, el uno en el otro, no puede sino

⁸⁷ En un sentido diferente al mío, la pertinencia de acudir a la tragedia de Sófocles a fin de aclarar, en lo posible, estas cuestiones, ha sido señalada por Henry Staten: *Wittgenstein and Derrida*. Basil Blackwell, Oxford, 1984, pp. 127-130.

adquirir un tinte terrible. Edipo, que separó los géneros al mostrar su unidad, los mezcla ahora en sí mismo, en su propia vida, y puede así convertirse, encarnación viva del enigma «resuelto», en padre de sus propios hermanos, marido de su madre, hermano de su propio padre (en cuanto usufructuador del mismo espacio que aquél usufructuó). Estas posibilidades, por cierto, estaban ya siempre inscritas en la estructura de un ser uno-y-plural que, por ello, puede siempre, como el lenguaje, violar la ley del género que le constituye. ¿No es Derrida quien ha señalado la solidaridad de fondo que reúne el problema de la idealización con el problema de la temporalización?⁸⁸ ¿Y ambos con el problema general de la estructura de marca, de huella? ¿Y no es en cierto sentido esto mismo lo que puede leerse en el contenido de un relato ancestral en el que se muestra cómo quien resuelve el enigma del tiempo, del signo y de sí mismo no puede hacerlo si no es acudiendo a una «lógica» cuyo desvelamiento libera a la vez los poderes de destrucción y de auto-destrucción de esa misma «lógica» que, sin embargo, cura y permite vivir? Queden aquí por ahora abiertos estos interrogantes. Tiempo habrá más adelante de ocuparse por extenso de estos temas, que por el momento sólo puedo dejar apuntados.

El firmante de este trabajo tuvo el privilegio, y el placer, de presentar algunas de estas cuestiones ante un auditorio de amigos y compañeros reunidos, Abril de 1989, en la Universidad de Santiago de Compostela. La publicación que el lector tiene entre las manos quiere ser justamente el eco de aquellos dos días intensos. Y si la memoria no me falla, a la exposición por mi parte de las ideas de Derrida se le objetó -bien en el coloquio inmediatamente posterior, bien en la mesa redonda que clausuró las jornadas- el que, habiendo mostrado yo mi distancia crítica respecto a Derrida, no explicitara a mi vez cuál era la tradición a la que me adscribía, o el credo filosófico que acataba; o cuál, en definitiva, el lugar teórico desde el cual hablaba. Imperativos de índole material, especialmente lo intempestivo, ya, de la hora, no me permitieron contestar en aquel momento al profesor Blanco, que formuló la pregunta. Quiero aprovechar ahora lo que me resta aquí de espacio (puesto que aquel tiempo ya se agotó) para responderle, siquiera sea abocetadamente, formulando yo también una pregunta. Mi pregunta/respuesta es: ¿dónde debe situarse un lugar o camino, un cruce, mejor, de reflexión, que en el interior del camino analítico busca ante todo (comprender) la posibilidad de sobrepasar, analíticamente, el dogma analítico; en el interior del camino hermenéutico, la posibilidad de sobrepasar, hermenéuticamente, la hermenéutica; en la deconstrucción, la deconstrucción de la deconstrucción; en el pragmatismo, la semiótica o la filosofía trascendental, la fuerza de auto-superación que anida, por y mediante la simple fidelidad a sí mismas, en cada una de estas rutas? Esté donde esté, o donde no esté, yo reivindico ese espacio. Reivindico, incluso, la necesidad que hoy tenemos de hacerlo operar en cada una de nuestras re-exposiciones de la tradición, en cada uno de nuestros interrogantes, en cada una de las respuestas que vamos arriesgando a esos interrogantes. Para mí, ese espacio no tiene, hoy por hoy, una etiqueta singular y reconocible que lo identifique ante todos. Desconociendo su nombre, sí conozco, en cambio, sus efectos

⁸⁸ DERRIDA, J.: *Limited Inc... cit.*, p. 33.

(aunque no todas las tradiciones acepten tal distinción): y el ubicarme en ese espacio, que acaso impida considerar *nada* como ajeno, impide desde luego considerar nada del conocimiento como ajeno, nada «filosófico» como ajeno. O dicho de otra forma: nada en absoluto como algo a lo que exclusivamente yo pertenezca o me pertenezca, salvo: el poder mismo de la reflexión. Reflexión auto-difundida en innumerables caminos que, seguidos hasta el fin, sólo llevan, de nuevo, al camino.

¿Que todo esto está por precisar? Naturalmente que sí. Sólo añadiré que, en todo caso, lo que sí me asiste siempre es la esperanza davidsoniana, la esperanza de que uno no se quedará tan fácilmente sin trabajo... (y ésto, desde mi «posición», también es una broma - pero una broma que debería ser tomada completamente en serio).

Jorge Pérez de Tudela Velasco
Septiembre, 1989





PUTNAM, O LAS PERPLEJIDADES DE LA FILOSOFÍA*

Mercedes Torrevejano

Una reflexión que se propusiera tomar el pulso al «encuentro de tradiciones» que se constata en el momento presente, había de encontrar en Hilary Putnam un significativo exponente de tal situación, en el seno de la tradición analítica.

Hilary Putnam es, desde los avanzados años 50, en que aparecieron sus primeros trabajos, uno de los filósofos más importante del mundo anglosajón. Su inicial formación en matemáticas y física le llevó a la filosofía de las matemáticas y de la ciencia, hasta venir a parar a una revisión de las raíces de la tradición analítica, para venir a centrarse, cada vez más, -a través de los básicos temas de la filosofía del lenguaje y de la mente- en el problema mismo de la filosofía, en la identificación de sus ineludibles tareas, en la corrección, o mejor, contención de sus parcialismos, o -si se quiere ver así- en la comprensión integrada de la diversidad de sus instancias. Todo ello perfilado, a la postre, en un asunto de asuntos: la racionalidad humana.

Una leve ojeada sobre su producción intelectual nos orienta al respecto¹.

Es muy poco frecuente, cuando se habla de los filósofos pertenecientes a la tradición analítica, el intento de comprenderlos globalmente, el intento de reconducir en alguna medida sus temas y tesis a función de un proyecto filosófico, o de un entendimiento de la tarea filosófica que ofrezca cierta perspectiva sistemática. La tradición analítica es muy consciente -y ha hecho profesión de ello en fuerza de su convicción de que es lo único con

(*) Este trabajo abre un conjunto de tres en torno a Hilary Putnam. Los tres muestran un cierto balance del seminario que sobre el autor celebramos semanalmente un grupo de profesores y postgraduados de la universidad de Santiago en 1989, como se ha indicado en la presentación. Quede aquí constancia de la gran aportación que supuso el diálogo asiduo con todos ellos: Angel Alvarez, José Luís Falguera, Concepción Martínez, María Uxía Rivas, Alejandro Sobrino, Juan Vázquez. Los profesores José Miguel Sagüillo, Luís Villegas y yo misma establecimos un acuerdo para proyectar en páginas escritas algunas de las ideas discutidas entre nosotros. A todos, y muy particularmente a los dos últimos mi agradecimiento por su valiosa ayuda.

¹ Dicha producción está constituida, básicamente, por siete volúmenes, tres de ellos (PHILOSOPHICAL PAPERS I. *Mathematics, Matter and Method* (1975) II. *Mind, Language, Reality* (1975) III *Realism and Reason* (1983), todos en Cambridge U.Press) son recopilación de trabajos, las más de las veces publicados anteriormente. El resto de los libros es: *Meaning and the Moral Sciences*. London, Henley an Boston, Routledge and Kegan Paul, 1978. *Reason, Truth and History*. Cambridge U.Press, 1981. *The many Faces of Realism*. La Salle. Illinois, Open Court, 1987, y *Representation and Reality*. The MIT Press. 1988.

Hay buen número de trabajos de Putnam publicados en revistas, homenajes y colectivos que quedan fuera de este conjunto. No debe confundirse el mero título general de PHILOSOPHICAL PAPERS III (*Realism and Reason*) con un trabajo del mismo título que constituye la parte IV de *Meaning and Moral Sciences*. Aparte de frecuentes solapamientos temáticos que se producen en todo este conjunto de páginas, una efectiva repetición (que no llega a ser completa) se registra entre: cap. V de *Reason, Truth and History* y Ph.P. III, 11: «Filósofos y entendimiento humano».

sentido-, de que el suyo es un estilo de filosofar que no plantea posiciones sistemáticas, o integradas, sino que elabora, propone argumentos sobre problemas filosóficos concretos - «consists of piece-meal problems solving»-, dice exactamente el mismo Putnam².

Sin embargo Putnam, que se siente absolutamente dentro de esa tradición, se ha ido mostrando cada vez más, percibiendo que la filosofía analítica ha vivido su propia autocomprensión demasiado poco críticamente (*uncritically*, dice Putnam); y reacciona, diríamos, inconformistamente ante esa falta de crítica. De hecho, buena parte de sus trabajos han sido justificados por él mismo como resultados de revisiones, como expresiones de *cambios* que quiere introducir en los estados de cuestión vigentes y/o en sus propios enfoques de los mismos³.

Pues bien, sabido esto, cabe decir que en Putnam -como en el caso de Wittgenstein, si se me permite la comparación- deben distinguirse dos épocas: *una* hasta 1975 y la *segunda* a partir de 1976. Esta fecha puede considerarse que está representada por «Models and Reality», trabajo de ese año, que abre la colección PhP.III (1983). Por tanto, *Meaning and moral Sciences*, de 1978, y *Reason, Truth and History* de 1981 están plenamente en la segunda época⁴.

El «hilo conductor» de tal evolución es sin duda la cuestión del *realismo*, una cuestión que imbrica otras muchas, pero que no es una más entre los innumerables temas estudiados por Putnam, en la medida en que una clarificación de este asunto se impone como condición de posibilidad y de sentido de cualquier otro. En función de ello, Putnam centra el trabajo de la filosofía en la elaboración de una *semántica referencial* que se abre camino sorteando alternativas insatisfactorias. Si hubiéramos de describir «desde fuera» las vicisitudes de ese hilo conductor, cabría decir que la cuestión del realismo afloró en Putnam como la reivindicación «de una específica perspectiva *realista* en filosofía de la ciencia». Esto

² Cfr. Ph.P.III, 10: «Convention, a theme in Philosophy», p.180. Putnam presenta esa manera de proceder como un hecho; y está claro que lo describe desde su personal distanciamiento crítico. DUMMETT en un trabajo titulado «Can analytical Philosophy be systematic, and ought it to be?» reflexiona sobre el mismo hecho, desde la comprensión de la filosofía como «doctrina» sobre el significado y no doctrina cognoscitiva *more scientia*. Cfr. en M.DUMMETT: *Truth and other enigms*. Oxford, Duckworth, 1978, p.437.

³ En 1975 presenta Ph.P.I (trabajos que se remontan a 1960) como exponente de un período de quince años de reflexión durante el cual «my views underwent a number of changes...» (Introduction vii). Pero algo más adelante, en 1983, vuelve a ofrecer otra colección de trabajos (Ph.P.III) que abarcan hacia atrás desde 1976, y vuelve a decimos que reflejan un período «of rethinking and reconsidering much of my philosophical position» (Introduction vii). Y -por citar el último libro que conocemos- de nuevo en 1988 Putnam nos advierte que la obra debe ser entendida como respuesta crítica (criticism) a tesis filosóficas que están de moda, «incluyendo por supuesto ideas filosóficas mías anteriores» (*Representation and Reality*. Preface, ix).

⁴ Hay que hacer esta división, naturalmente, desde el punto de vista de las publicaciones. Es muy difícil determinar lo que en el transcurso de un pensamiento filosófico es *nuevo* o estaba prepensado en consideraciones o tesis anteriores. Lo normal es que *subjetivamente* el pensador que vive ese transcurso tienda a «tender hilos» de continuidad entre las expresiones diversas de su pensamiento, reinterpretando sus tesis anteriores en el marco general donde se protegen las últimas... Putnam refleja esta tendencia y se suele remitir a escritos de su primera época para retro-traer a ellos las razones, en último término válidas, de sus posiciones maduras. Por otra parte, Putnam se ve a sí mismo como habiendo cambiado constantemente sus posiciones. Y sale al paso de las críticas por ello: «The fact that I change my mind in Philosophy has been viewed as a character defect». Replica que ello ha sido común en los grandes filósofos, como Russell, Carnap... a quien ve como ejemplo «of a human being who puts the search for truth higher than personal vanity» (*Representation and Reality*. Introduction, xii).

quiere decir que en un principio la cuestión del realismo le interesó a Putnam en cuanto ciertas cuestiones de la filosofía de la ciencia implicaban -a su juicio- la necesidad de comprometerle con un *realismo* más allá de lo que los filósofos de la ciencia realistas de los años sesenta estaban dispuestos a conceder. Este realismo se anclaba en tres tesis o convicciones: la de que los enunciados científicos son o verdaderos o falsos (aunque no sepamos cual es el caso); la de que su verdad o falsedad no consiste en ser métodos -altamente derivados- de describir regularidades en la experiencia humana; y la de que 'realidad' no es parte de la mente, sino al revés. El *realismo* de Putnam en estos años es un realismo relativo no sólo a los objetos materiales sino también a las entidades matemáticas y asimismo a entidades tales como magnitudes físicas y campos.

Tal postura se perfila como un distanciamiento básico del verificacionismo (tan vinculado históricamente al positivismo), al que Putnam ve abocado a dejar en el vacío la cuestión del significado. Pues bien, tras esa primera posición, la cuestión del realismo se explicita cada vez más, se afronta directamente y viene a convertirse en la crítica machacona de un inconsistente y malentendido *realismo*, el «realismo metafísico»; crítica y negación en la que *se decantará* -en una suerte de depuración- el espíritu o perspectiva siempre realista de Putnam, ahora bien identificado como *realismo interno*⁵.

Dicho ésto y dejando de lado por ahora este tema central quiero hacer notar, algo más ampliamente, otra cara de esta evolución, que me parece relevante y muy significativa en Putnam: lo que he insinuado ya como su conciencia crítica, su tener en cuenta las raíces históricas de la tradición analítica, la caracterización situacional que hace de los problemas; todo lo cual le abocará, como he adelantado, a reflexionar sobre las instancias temáticas que estructuran la filosofía, y a cifrar en último término su tarea en la pregunta por la racionalidad humana.

Ese carácter crítico-histórico-situacional se muestra en los trabajos de Putnam de un modo oblicuo aunque progresivamente menos oblicuo-, en forma de observaciones históricas y de breves *ex cursus* que parecen incidentales y poco relevantes (supuesto que lo relevante es el desarrollo de los argumentos que hacen al caso en el tópico que se está tratando). Sin embargo son consideraciones muy dignas de atención y más que meramente ilustrativas.

Haciendo un esfuerzo por organizarlas creo que se podrían distinguir tres series de ellas: *unas* relativas a las raíces históricas del Análisis, o a su condicionamiento histórico-cultural; *una segunda serie* sería la sistematización (o sistematizaciones) a las que recurre sobre las perplejidades internas de la tradición analítica, sobre los callejones sin salida de sus doctrinas más o menos escolastizadas; *en tercer lugar* aparecen con decisión diagnósticos sobre las serias aporías culturales y filosóficas del momento presente. Esta tercera línea de crítica implica la atención explícita de Putnam a los tópicos y demandas de las sociedades

⁵ Tiene cierta gracia la expresiva explicación de Putnam cuando en 1975 nos entrega sus Ph.D.: «mi interés en los últimos quince años no ha sido partirme el pecho por defender la corrección del realismo...» (Lo directamente en cuestión eran problemas de filosofía de la ciencia, insistamos). Pero la verdad es que la producción de Putnam desde 1978 podría ser descrita gráficamente y sin ninguna ironía como una lucha a brazo partido contra el realismo metafísico y en favor del realismo interno.

de hoy -una clara actitud de «intelectual», de velado perfil- y una atención también explícita a las alternativas filosóficas «continentales» de los últimos años.

1. Putnam interpreta el fenómeno llamado «Filosofía Analítica» como parte del fenómeno llamado *Modernismo*. El espíritu modernista de rechazo de la tradición se refleja en esa nota tan distintiva de la comunidad de los filósofos analíticos («even now», dice Putnam en 1983), que consiste en considerar a la «filosofía tradicional» como algo *contra lo que* hay que tomar posición y a lo que hay que combatir. Pero no se entienda tal semejanza a la ligera: el pensamiento filosófico -en este caso el Análisis- pertenece al modernismo y emerge del mismo *humus* que sus propuestas; de ahí que comparte con él una misma cierta forma. Putnam se detiene en establecer una analogía entre los conceptos, o los sistemas de creencias que configuraron, por ejemplo, los proyectos y experiencias arquitectónicos del modernismo -el sueño de casas y «ciudades radiantes», reflejos de armonía y de justicia, de simplicidad y belleza racionalmente afirmada, de adaptación al, y dominio del medio- con los ideales de tersura y nitidez científicas que propugna Carnap, por citar un egregio ejemplo. (Putnam utiliza la metáfora de «arquitectura futurista» para la obra de Carnap y su ideal cientifista)⁶. Esto parece obvio, pero al lado de esta optimista y tal vez facilona analogía, Putnam pone el acento en «la otra ala», el *otro lado* antropológico-pesimista que tal movimiento proyectó: el que inspiraba aversión a la tradición y desesperanza hacia la idea de un hombre nuevo... Y aquí es Musil el invocado como buen denunciador de ambas facetas -también la pesimista-⁷. Esta dualidad que la época ha traspasado al Análisis es como una esquizofrenia. Y los dos caminos «están cansados» a estas alturas... «El fuerte rechazo de todo lo *tradicional* y la gran esperanza de que sacudiéndonos la tradición le posibilitaríamos al hombre o una utopía futura (modernismo optimista), o el reconocimiento, por fin, de la grandeza y terror de la vida, devolviéndole la inmediatez última que nos robaron el cientifismo y el mito del progreso (modernismo quietista)...» esa esperanza se desvanece. Los dos caminos «are both beginning to look very tired»⁸.

Unas cuantas páginas, alusiones sueltas dedicadas a este tipo de consideraciones no son mero recurso retórico. Putnam toma nota -como estamos viendo- de estas circunstancias en la medida en que adopta la tesis del «condicionamiento histórico» -digamos cautamente- de las ideas y de su verdad, dispuesto a extraer de ello consecuencias, las que fuere menester; dispuesto a afrontar el debate sobre lo ideográfico en su oposición a lo nomotético, «lo natural» en su oposición a «lo calculado» o «planeado», etc... De ahí que, en fuerza de la trascendencia que coherentemente debe tener la entrada en este tipo de consideraciones situacionales, habrá de ponderar hasta dónde está dispuesto a comprometerse con planteamientos historicistas⁹. Volveremos sobre esto.

⁶ Cfr. Ph.P.III,10, p.179 y ss.

⁷ Putnam nos invita sutilmente a asociar al pesimista Wittgenstein con el personaje Thomas de *Die Schwärmer* de Musil: nunca ha sabido decir nada sobre qué hacer con la vida, se le reprocha... «Así se va deambulando por la vida» -contestará Thomas-. «Toda esa gente que sigue adelante haciendo su preciso camino te es hostil, mientras tu vas por ese mismo mundo recorriendo el mísero e indefinido viaje del espíritu... Y sin embargo, a esa gente perteneces de algún modo» (*Ibidem*, p.182)

⁸ Cfr. Ph.P.III,16, "Beyond historicism", p. 303.

⁹ Cfr. *Ibidem*, p.287.

2. «El alma dividida del positivista lógico» -The divided logical-positivist soul- :ésta no es tampoco una mera frase retórica de Putnam¹⁰. Por muy concreto que parezca el tema a propósito del cual está pronunciada literalmente -en esta ocasión se está hablando de filosofía de la mente- la idea constituye una constante interpretativa por parte de nuestro autor. Otra vez aquí, aunque los temas tratados por Putnam son muchos, hay otro hilo conductor: la interna tensión agazapada en el espíritu «cientista» de esta filosofía, tensión que Putnam pone de manifiesto retrotrayendo los problemas y su revisión al *verificacionismo*, es decir, al principio del significado tal como se debate en el seno del primer Análisis.

Por lo pronto Putnam piensa que la temprana objeción hecha al principio de verificación (de no ser él mismo verificable) fue tomada poco en serio por el movimiento analítico¹¹. Y denunciará la ligereza del positivismo al creer que tal objeción se ha ventilado cuando se interpreta que, en realidad, la teoría empirista del significado no es sino una *pro- puesta* que como tal no es ni verdadera ni falsa. «Propuestas -dirá Putnam en 1981, refiriéndose a *esta* evasiva- presuponen fines y valores. Y es doctrina esencial del positivismo que bondad y maldad de fines últimos y de valores es algo enteramente subjetivo»¹².

En su primera época -ha confesado más tarde- atribuyó los errores de la filosofía analítica al verificacionismo más «ingenuo» en unos casos, o más «sofisticado» o avisado de las dificultades del principio, en otros (Reichenbach y Carnap son reconocidos como «sofisticados» cuando expone, compartiéndolas, las objeciones al verificacionismo ingenuo)¹³. Putnam ve al verificacionismo, en líneas generales, remitiendo la explicación de la *referencia* al *significado*. A este respecto, Putnam entiende que tal explicación implica conceder que «los significados están en la cabeza». Pero que el significado sea un asunto de «estar en un cierto estado psicológico» no le parece en absoluto pertinente. Pero frente a esa «intuición fundamental» -en el sentido de acercamiento inicial al asunto- está, dice Putnam, la contraria, la del realista que *enfatisa la referencia* en la determinación del significado. El realista diríase que hace el camino en la dirección opuesta... Pues bien, en la sintonía de ese otro camino se sitúa Putnam ya en su primera época.

En su rastreo del problema interno del verificacionismo Putnam ve a Reichenbach «echando un cable» al primitivo o ingenuo verificacionismo (enredado en la tesis del *convencionalismo lingüístico* explicativo de la relación entre los enunciados científicos y los hechos que los verifican; en la aguda dicotomía entre *hechos* y *reglas del lenguaje*, etc...) explicando aquella relación como una *inferencia* probabilística en el seno de una teoría. Ve a Peirce anticipando las consecuencias que de ello se derivan, en el enunciado de que «every change in a person's information is a change in the meaning of his words». Ve a Quine

¹⁰ Ph.P.II,22: «Logical Positivism and the Philosophy of Mind», p.441.

¹¹ Esta crítica «had very little impact on the logical positivism and did little to impede the growth of their movement. I want to suggest that the neglect of this particular philosophical gambit was a great mistake; that the gambit is not only correct, but contains a deep lesson and not just a lesson about logical-positivism» Ph.P.III, 11,p.184-185.

¹² *Ibidem*,16,p.288. El trabajo donde están estas afirmaciones es de 1981. Recogeremos más adelante las consecuencias de esta crítica de cara a discutir la concepción criterial de la racionalidad que ha arrastrado el positivismo.

¹³ Cfr.Ph.P.II. Introduction,viii-ix.

en concordancia fundamental con todo este camino renunciando paladinamente, con la mejor de las lógicas, a la noción de significado. Pues esa es la cuestión: «as Quine has long urged, we should abandon the notion of meaning altogether». Y debe hacerse lógicamente -insiste Putnam-¹⁴ si no se está dispuesto a mantener una distinción entre el significado de las palabras de una persona y lo que cree sobre los hechos; sin embargo esa distinción sobre desacuerdo en el significado de las palabras y desacuerdo sobre los hechos es central para cualquier concepto de significado lingüístico.

La incomodidad de Putnam ante este proceso es evidente: representar lo que son realmente las inferencias probabilísticas en el seno de las teorías como equivalencias lógicas, es una seria distorsión; y todavía más lo es representar estas inferencias como una equivalencia significativa puramente convencional¹⁵.

Putnam percibe el verificacionismo (o fenomenalismo, pues todo verificacionismo es un fenomenalismo «at heart»)¹⁶ cazado en un dilema sin salida: entre el deseo de seguir hablando del significado un tanto a la manera tradicional y su adhesión a la teoría reticular del significado, la cual, tomada seriamente implica la inhabilitación de la noción del significado lingüístico.

Repitamos que su instinto o «intuición» realista coloca a Putnam ya desde siempre en la dirección de considerar al significado ampliamente determinado por la referencia y a la referencia ampliamente determinada por conexiones causales, al menos, y principalmente respecto a los nombres propios y a los términos de géneros naturales. Esas ideas son las que en la primera época quedan ampliamente expuestas en «Explanation and reference», pero sobre todo en «The meaning of 'meaning'»¹⁷.

Este enfoque realista soporta el hecho de que nuestras teorías cambien, a cambio de reconocer un «momento», o un elemento del significado de las palabras no cambiante: justamente lo que constituye la referencia. La identificación de la referencia podemos hacerla al margen de cualquier cambio que sufra la teoría -por violento que éste sea- distinguiendo tal momento (la referencia) como aquello que es causalmente responsable de tales o tales efectos. La teoría intenta preservar la referencia a través de una -digamos- cadena histórica de usos y hace a la fijación de la referencia función del trabajo social-lingüístico; con lo cual hace posible el uso de términos, exitosamente, sin que ninguna descripción definida esté asociada al término por todos los hablantes. Los individuos son singularizados por una «ceremonia bautismal» (*naming ceremonie*), pudiendo inferirse con éxito la existencia de entidades teóricas definidas (a las que se puede otorgar nombres). («Indexicabilidad» de la referencia)¹⁸.

¹⁴ *Ibidem*.

¹⁵ Estas afirmaciones tienen su justificación argumentada en el seno de concretos trabajos de Ph.P.II. Cfr. 2: «The analytic and the synthetic» y en general los que venimos citando: 11, 13, 22...

¹⁶ Ph.P.II, 13, "Language and Reality", p.273.

¹⁷ Ambos en Ph.P.II, 11 y 12 respectivamente. El segundo de estos trabajos fue traducido en TEOREMA, vol. XIV, 3/4, 1984, p.345-406: "El significado de 'significado'".

¹⁸ En este contexto Putnam se distancia explícitamente de los «realistas tradicionales», a los que describe como quienes piensan que «la referencia estaba determinada por entidades mentales o platónicas, por intensiones». Pero esta doctrina acerca de significados «fijados» (bien en la cabeza o siendo un reino de entidades abstractas, conectadas de algún modo con la cabeza), significados fijados que determinan la referencia de una vez por todas «es susceptible de las mismas objeciones que pueden hacerse al verificacionismo». Ph.P.II, Introduction, ix-x.

El conocimiento científico es *convergente*, ha de pensar Putnam. La sustitución de una teoría (anterior) por otra teoría (sustitutoria) debe conceder, si la sustitutoria pretende ser verdadera, que la primera está contenida en la sustitutoria, como un límite de ella. Y si los términos de la primera tienen referencia se deberá poder asignar referentes a esos términos, desde la teoría sustitutoria. (Tal es el «principio de caridad» que hay que conceder para que no nos quedemos sin verdad)¹⁹. El teorizador anterior aceptaría los términos del teorizador sustitutorio para sus referentes...

A Putnam lo que le preocupa es preservar en este enfoque una teoría de la verdad extra-teórica, una teoría donde el significado no sea intrateórico, o teórico-dependiente. «Verdad y falsedad son los términos fundamentales de la crítica racional y una filosofía que se precie debe dar cuenta de ellos, o si no lo hace, mostrar que puede pasarse sin ellos»²⁰.

...

La revisión del verificacionismo que hace Putnam no queda en esto. Los problemas del significado que atañen a la construcción de nuestras teorías acerca del mundo y que implican teorías acerca de la realidad, de los hechos o entidades u objetos que constituyen o construyen o «amueblan» el mundo arrastran -como una sombra- otra cuestión: la de la naturaleza de lo mental o, en términos más amplios y generales, el problema mente-cuerpo. Efectivamente, la perspectiva de la filosofía de la mente es el segundo gran caballo de batalla de Putnam. Otro frente que se abre en su crítica de la tradición analítica y del principio verificacionista que la conforma.

En este punto Putnam se enfrenta a la tesis reduccionista que acaba identificando cualquier estado mental con un tipo de estado físico. Materialismo es tal reducción, aunque de nuevo aquí a lo largo de los años Putnam ve a este reduccionismo sofisticándose por mor de la tensión interna que provoca la dicotomía observacional y teórica del problema (un primer acercamiento a la cuestión, intuitivo, podríamos decir, no encarrila el asunto hacia reducción de ningún tipo. Si hablamos de sentimientos o pensamientos, etc... lo hacemos bien en serio). Dicotomía que de nuevo divide el alma del positivista lógico...

La postura *realista* de Putnam en este punto se va a concretar en *funcionalismo*; funcionalismo «no como doctrina sobre el significado de las palabras psicológicas, sino funcionalismo como una hipótesis sintética sobre la naturaleza de los estados mentales». La realidad de los procesos mentales, su eficacia causal, etc. se explica -propone Putnam- no en términos de física o química, sino de *programa* (de una máquina) por más que tal

¹⁹ Cfr. Ph.P. II, 13, p.274. En realidad Putnam lo llama *the principle of benefit of doubt*, pero concede que podría llamársele principio de caridad, nombre usado por N.L. Wilson en «Substances without substrata», en REVIEWS OF METAPHYSIC, 12 (June) 1959, p. 521-539. Averte que el sentido del principio en Wilson es incompatible con su punto de vista.

²⁰ «When I say that truth and falsity are the fundamental terms of rational criticism, I don't mean that we always are able to judge that a theory or doctrine is true or false; often we are lucky to be able to say that something is 'probably true' or 'approximately true'. But the semantics of probable truth and approximate truth presupposes *the semantics of truth and falsity*; these notions make non-sense if truth and falsity make no sense». *Ibidem*, nota a pie de página. Subrayado nuestro.

programa sea realizado por una particular física o química. La psicología -piensa Putnam- tiene una función explanatoria autónoma²¹.

Según todo ésto, Putnam entiende, digamos de pasada, que su postura en este asunto es sorprendentemente la misma de Aristóteles²².

EL SEGUNDO PUTNAM

3. A partir de 1976 -hemos dicho- la obra de Putnam adopta un nuevo acento. No ha sido nuestro objetivo en estas páginas perseguir el detalle de sus temas, o la estrategia particular de ninguno de sus argumentos cuando Putnam oficia de filósofo de la lógica, o del lenguaje, o de la ciencia, o de la mente. Desde el comienzo hemos optado más bien por presentar el sentido general de su trabajo filosófico; de ahí que tengamos ahora que caracterizar ese *nuevo acento*.

Haciendo un esfuerzo de síntesis, creo que tal nuevo acento se podría calibrar en el marco de un par de coordenadas, que ya fueron avistadas al comienzo: *en primer lugar* habría que decir que Putnam muestra una cierta radicalización, o tal vez más exactamente, nos ofrece la impresión de quien «ha tocado fondo» cuando vivencia la situación presente como de crisis filosófica, insinuándonos además que sólo en la *pa-tencia* de esa crisis nuevas hipótesis y propuestas filosóficas tienen -deben y pueden tener- su oportunidad²³. *En segundo lugar*, Putnam proclama y confiesa, por una parte la necesidad, por otra la evidencia fáctica de haberse situado en una suerte de «reconformación» de la comprensión de la filosofía; lo cual, con relación a la tradición analítica de la que proviene, significa el desenmascaramiento, ¡al fin posible!, por parte de la filosofía, de su genuína vocación sistemática o «integrada», una vez superado un pasado ideologizado y tendencial²⁴. En este punto la filosofía analítica ha venido -está viniendo- a confluir con el modo de comprender la filosofía que ha sido característico de la tradición continental²⁵.

²¹ Cfr. Ph.P.II, Introduction, xiii. A lo largo de su primera época el concreto modelo computacional (estados de una máquina de Turing) propuesto en principio por Putnam es sometido a crítica por el mismo Putnam. Vid. *ibidem*, 14: «Philosophy and our mental life». Pero los términos generales del funcionalismo que Putnam defiende se mantienen, acompañados al *realismo* de esta primera época.

²² «(Having come this far I was pleasantly surprised to find that my view was substantially the same as Aristotle's, although stated a bit more precisely with the aid of the vocabulary of contemporary scientific methodology and cybernetic)» *Ibidem*. Sería muy interesante analizar el aristotelismo de Putnam, tanto en filosofía de la mente como en ética, como más adelante mencionaré.

²³ «...it is only by seeing that the currently fashionable views do not work that we can begin to see what the tasks of philosophy might really be». *Representation and Reality* Preface.

«There is a certain duality in our ideology that we are not going to get rid of. If we are not going to get rid of, let us accept it and proceed from it, I suggest, instead of continuing the attempt to breathe life into failed intellectual positions» Ph.P.II, 16, p.290.

²⁴ «The desire for integration is so central to Philosophy, I think, that no philosophical tendency will long endure without it» *Ibidem*, p.303.

²⁵ Putnam explica, en la introducción de Ph.P.III que su mayor sensibilidad para ciertos aspectos de los problemas ha tenido sus consecuencias en los trabajos que presenta, «consequences which I did not expect. It led me to think about questions which are thought to be more the province of 'continental philosophy' than of 'analytical philosophy'». Cfr. p.vii.

«Analytic philosophy claimed to be piece-meal philosophy. It gave up, or said it gave up, the dream of an integrated view, which is so characteristic of what people refer to as continental Philosophy». *Ibidem*., 16, p.303.

Cierto que muchos analíticos han sido *de facto* grandes sistemáticos; de hecho los positivistas lógicos lo fueron («...had very much an integrated view, and Quine has very much an integrated view»); pero Putnam entiende que el positivismo lógico y el atomismo lógico inicial, en la medida en que fueron el *motor* de la filosofía analítica -sabido que no todo filósofo analítico haya sido positivista- han «sellado» con un tono fuertemente ideológico y como *tendencia*, esa tradición.

Putnam muestra una rebeldía (que casi trasluce enojo) ante la tendencia de esa tradición a sugerir que las presuntas *profundidades* de la filosofía son paranoias estúpidas, puesto que no hay tales problemas. «Pero hay preguntas que *son* profundas y lo que resulta estúpido son las respuestas fáciles... los filósofos *están* comenzando a hablar acerca de grandes temas otra vez, y a sentir que algo puede *decirse* acerca de ellos, aunque no haya soluciones definitivamente espectaculares»²⁶ (La filosofía, pues, en la prosecución de su genuína «esencia», en la *catarsis* o *emendatio* de su proyecto: Putnam nos aparece saltando hacia atrás para verse seguro en el camino de la tradición moderna, autenticándose con acentos kantianos y fenomenológicos)²⁷.

Cierto también que recuperar la vocación sistemática de la filosofía no es repetir ningún pasado. La conciencia de nuestro autor es lúcida sobre los límites de la tarea, límites sobre los que esa misma tradición moderna nos ha enseñado bastante: «On the other hand, every attempt at integration which has been too grand has collapsed... Como filósofos, nos vemos atrapados entre nuestro deseo de integración y nuestra conciencia de la dificultad del empeño...»²⁸.

El lugar referencial de la filosofía: inmanencia/transcendencia de la razón.

Filósofos y no filósofos -piensa Putnam- padecen hoy una «atmósfera sofocante» que Putnam cifra en el desgarramiento entre «subjetivismo» y «objetivismo», tanto a propósito de la verdad, como a propósito de la cuestión del valor o de los valores: dos campos diversos donde la «sofocante presión» adopta tonalidades distintas pero cuyas claves acaban llevándonos al mismo fondo.

Cuando se trata de la verdad, parece un hecho que las posiciones de los pensadores hoy se disparan en direcciones tan opuestas que la oposición se hace *casi ideológica*. Los defensores de que la verdad significa el atenimiento a los hechos como ellos son, al margen del espíritu que los «registra», no están dispuestos a conceder que cualquier otra postura pueda ser otra cosa que «capitulación» o aceptación de que todos los «cuadros del

²⁶ Cfr. Ph.P. III, 13, p.236. Ya en Ph.P.II, p.273 decía: «Philosophy of language cannot solve the traditional problems of philosophy by showing them to be pseudo-problems». Y un tanto irónicamente comenta que en esto Wittgenstein y Carnap fracasaron por cierto. No nos han convencido de que ciertos problemas no lo son. Esto se conecta con la crítica de Putnam a la concepción criterial de la racionalidad del positivismo como indicaremos más adelante.

²⁷ Considero acento «fenomenológico» el claro intento de recuperación de la tarea filosófica en medio de una conciencia de crisis donde cientismo, historicismo, relativismo reflejan de manera entrecruzada y contradictoria rasgos contradictorios de la cultura que se vive en una cierta analogía con la experiencia y consiguiente demanda husserliana. De hecho Husserl y su *Crisis*... han sido rememoradas por Putnam en el contexto de sus replanteamientos filosóficos. Cfr. *The many faces of realism*, p.29.

²⁸ *Ibidem*.

pensamiento» son desesperadamente «subjetivos». Una minoría audaz por su parte ha recogido el desafío y se ha apuntado a la alternativa. Kuhn, Feyerabend, Foucault son representantes paradigmáticos al respecto. Describiéndolos en principio *per modum unius* Putnam nos sintetiza su posición: frente a ingenuas teorías de la verdad-copia, los sistemas de pensamiento, sean ideologías, sean incluso teorías científicas -como es el caso de Kuhn y Feyerabend- sólo se explican en último término como subjetivos. Y concluye para que no quede duda de cómo los interpreta: «se han tomado en serio el desarrollar una concepción relativista y subjetiva»²⁹.

En otros contextos este relativismo-subjetivismo es tematizado en términos de *historicismo*, lo cual es muy interesante a mi juicio. Pues lo que esa oposición (objetivismo/ subjetivismo o relativismo) pone en evidencia es la necesidad de revisar la concepción misma de racionalidad que subyace a las distintas posiciones. En líneas generales «el objetivismo» de que habla Putnam, representado por la tradición positivista secuestra a la racionalidad en una concepción criterial de la misma; en una concepción que entiende que *hay normas institucionalizadas* que definen lo que es o no es racionalmente aceptable. Por relación a esto, las «rebeliones» subjetivistas se sitúan en la alternativa de que la institucionalización de las normas tiene que ver con la situación cultural e histórica que viene marcada por el desarrollo humano... Pero lo importante es ver que en este punto salta el tema de la racionalidad como el problema filosófico básico y salta la urgencia del diálogo y la contrastación con alternativas extra-analíticas³⁰.

Cuando se trata de problemas de valor, evidentemente la situación es más pregnante o «mundana» -por hablar en términos kantianos-. Resulta relativamente indiferente para la vida la disensión teórica sobre hasta dónde existe un mundo real etc...

Pero la necesidad de la vida impone elegir, y toda persona pensante debe tener una opinión sobre la distinción entre hechos y valores, opinión fáctica que a lo mejor no corresponde con su opinión nocional.

Pues bien, Putnam denuncia en este asunto un hecho: la tesis de que hay una irreductible dicotomía entre hechos y valores, *se ha institucionalizado*, se ha convertido en una cierta *forma mentis* del hombre medianamente culto de hoy; de otra manera, está asumida como un dogma la creencia de que no hay respuesta posible o, al menos, no hay respuesta evidente a «la cuestión de saber si las cosas son peores o mejores, buenas o malas...». Lo cual quiere decir que cualquier reflexión crítica que alguien se propusiera hacer sobre lo infundado de esa dicotomía (está pensando en sí mismo), o más exactamente sobre lo infundamentado de la irreductibilidad de la dicotomía, diríase que tiene perdida la partida de antemano.

²⁹ *Reason, Truth and History*. En este caso utilizamos la edición francesa: *Raison, Vérité et Histoire*. Paris, Les Éditions de Minuit, 1984. Cfr. Préface, p.7.

³⁰ En el fondo de este subjetivismo-relativismo está el *ciencismo*. Aunque la tesis de que la racionalidad se mide por la situación cultural no es adoptada descaradamente por los anarquistas del método (aquí Kuhn es siempre la posición más conservadora) Putnam defiende que el límite natural de esta tendencia es un reduccionismo cientifista. «La teoría según la cual la racionalidad se define por un programa de ordenador ideal es una teoría cientista inspirada por las ciencias exactas. La teoría según la cual es definida por las normas culturales es una teoría cientista inspirada por la antropología». *Raison, vérité, et histoire*, p.142-143.

Cualquier hombre de hoy medianamente razonador, ante cualquier debate sobre cualquier tema de interés está preparado para -*more methodologico*-, digamos- controlar su debate en el momento oportuno con la pregunta a su interlocutor: «esto que dice usted ¿es afirmación sobre hechos o es juicio de valor?»³¹.

Estas descripciones a dos bandas como estamos viendo -teórica y práctica- nos muestran a un filósofo preocupado por la situación de *impasse* que afecta a nuestras teorías y a nuestras formas de vida, en la medida en que un relativismo más o menos velado nos acosa por todas partes.

Prosigamos con la cuestión del valor. Es un profundo error -piensa Putnam- la idea de que la moral es algo subjetivo, de que es un plano renuente al análisis y/o la reflexión racional. La idea de «razonamiento moral» se considera una *contradictio in terminis*, lo cual es «terribly destructive not just for morality itself but for all of the culture»³².

Correspondientemente se proyecta sobre las ciencias humanas o sociales un afán imperioso de ciencismo, un afán de perseguir en ellas «un ideal enteramente falso de objetividad», que trae como consecuencia vacío y esterilidad para ellas mismas, olvidadas de sus conexiones, incapaces de ir más allá de la descripción de nuestra sociedad... Muchos pensadores de hoy ven a la ciencia como una filosofía; concedido que lo fuera, se trata de una filosofía que se reduce a metafísica y en nada es ética... «y metafísica sin ética es ciega» nos dice Putnam en un remedo kantianizante³³.

Radical separación de ciencia y reflexión moral; postergación de la reflexión moral en la comprensión de la filosofía; necesidad de retomar los problemas más unitariamente: *Meaning and Moral Sciences* es, en 1978 la apertura explícita de esta intención, aunque sólo en torno a 1980 los trabajos de Putnam perfilan lo que venimos llamando el *desplazamiento* sistemático de su atención. La teoría del significado en que viene consistiendo la filosofía ha de asumirse en el contexto más amplio y sofisticado de una teoría sobre la racionalidad humana.

Se trataría por tanto -creo que piensa Putnam- de propiciar una tesitura filosófica donde 1) se afronten las cuestiones de justificación racional y criterialidad; y 2) se suprima el dogmatismo que prejuzga que lo que no puede ser proyectado más allá de controversias no es cognitivo.

Esta doble actitud comporta un decidido encuentro de las cuestiones límites relativas a la teoría (o a la constitución de nuestros juicios sobre los hechos) y las cuestiones relativas al bien y a los valores. Pues creo que no cabe duda que la inevitable hipótesis de una racionalidad *no criterial* se alía con la hipótesis de reconocer razón y racionalidad en el corazón de lo que es controvertible de suyo (y por supuesto que lo es, incluso *prima facie* el orden de la moral).

En realidad son dos perspectivas sumamente trabadas. Si dijéramos que la primera la desarrolla Putnam en todo lo que es su replanteamiento del *realismo interno* (en oposición a, y contraste con, *realismo metafísico*) y que aquí prosigue su línea de clarificaciones y

³¹ Cfr. *Ibidem*, p.145 y ss.

³² *Meaning and Moral Sciences*, p.94.

³³ *Ibidem* p.92.

críticas respecto de la propia tradición analítica, diríamos algo correcto. Sin embargo el tema realismo interno es algo así como el *tema englobante*, el que da paso a sus diálogos *ad extra* porque *compendia* su nueva manera de acercarse a la cuestión de la racionalidad. De ahí que el desarrollo del tema en la segunda época abra paso sin solución de continuidad a la cuestión ética. *The many faces of realism*, más que título de un libro compendia el sentido globalizante y sistemático de lo que el realismo interno comporta: una teoría de la razón, o de la racionalidad más acá o más allá de posiciones de un teórico de la ciencia; o más exactamente, una teoría de la razón a la que ha de abrirse un teórico de la ciencia que no se empeñe dogmáticamente en tratar cualquier posición filosófica como un error que «consistiera en fallos en un test aritmético».

¿Quiere esto decir que hay un lugar neutral desde el que proveernos de una teoría de la racionalidad desde la cual decidir?. No se trata de eso, sino más bien del inevitable carácter de problema transcendental que en el límite tiene el asunto. Putnam entiende que si hay algo así como racionalidad -y es evidente que nos obligamos a aceptarlo desde punto y hora en que estamos comprometidos en actividades de lenguaje o de habla y de argumentación- es autocontradictorio (*selfrefuting*, dice Putnam con frecuencia) argüir para mostrarlo que la racionalidad se identifica con, o está contenida en lo que las normas institucionalizadas -de la cultura, si es el caso- determinan como instancias o ejemplos de esa racionalidad.

El carácter autorefutante de una teoría criterial de la racionalidad también se muestra por los hechos: véanse los resultados que nos ofrecen los positivistas, Wittgenstein y aun el mismo Quine, tal vez: nos han brindado filosofías que no dejan espacio para la actividad racional de la filosofía. Por otra parte los recursos al postulado de que la filosofía ha de ser análisis conceptual, y de que los *conceptos mismos* determinan qué argumentos filosóficos son justos; o la doctrina de que los conceptos son normas o reglas que subyacen a las prácticas lingüísticas públicas no son sino forma encubierta de abogar por una concepción criterial de la racionalidad, la cual intenta acompasar, o mejor, homologar la verdad filosófica a la verdad científica o intrarracional.

En suma, «argüir sobre la naturaleza de la racionalidad es una actividad que *presupone* una noción de justificación racional más amplia que la positivista; en cualquier caso, más amplia que la racionalidad criterial institucionalizada». O para decirlo con palabras más clásicas: la racionalidad en general es condición de posibilidad de la racionalidad científica, y no ésta el canon de aquella³⁴.

Es según esto como la perspectiva historicista destila su gotita de luz, pese a situarse en el otro extremo de la aporía. Putnam simpatiza con la oposición de racionalidad paradigmática/no paradigmática de Kuhn, en la medida en que se puede corresponder con la que él está diseñando (criterial/no criterial); pero en fila de frente el extremo relativismo que se expresa en la tesis -manifiesta ciertamente también en Feyerabend- de la inconmensurabilidad de las teorías. Esto significa llevar al extremo la manera en que las diferentes culturas y épocas históricas producen diferentes paradigmas de racionalidad o institucionalizan normas criterios. Foucault es también avistado en este contexto.

³⁴ Ph.P.III, 11, p.191.

Insistamos como resumen que este asunto de la racionalidad al que se ve abocado desde el estado de cuestiones que estamos viendo, es el que potencia el replanteamiento del problema del realismo; es el que le lleva de inmediato a introducir la ética como momento esencial del problema filosófico; finalmente es lo que le lleva a reconsiderar en filosofía de la mente el funcionalismo de su primera época³⁵.

Realismo interno, verdad, racionalidad.

Es el tema de Putnam más estudiado. Y es la cuestión que trata el profesor Sagüillo en el trabajo que sigue a éste. No vamos a detenernos en él.

Digamos simplemente que la nueva posición se trenza al hilo de dos preguntas que brotaban ante la perplejidad de la situación reduccionista-relativista que hemos considerado: ¿Deben identificarse verdad y racionalidad?, ¿inconmensurabilidad o convergencia de nuestras teorías?. Es claro que el tema del significado y de su relación con la referencia vuelve a desafiar. Putnam lo retoma de nuevo y lo perfila, metodológicamente, por vía de negación: desenmascarando el realismo metafísico.

La verdad es que cuando Putnam describe al *realista metafísico* es difícil identificar claramente ninguna posición filosófica históricamente dada, sobre todo si nos situamos en la tradición moderna. Hay que pensar que «realismo metafísico» es un modelo o situación filosófica límite que significa el *centro de gravedad* de los planteamientos filosóficos no idealistas y, por supuesto, no trascendentalistas o de tradición kantiana (pues Kant es el punto de referencia que polariza nuestra orientación en el asunto frente a cualquier otra filosofía, por muy tipificada que la podamos tener). Esos planteamientos filosóficos tienen su piedra de toque en un par de «fantasías metafísicas» (una u otra) más o menos conscientemente asumidas: 1) la fantasía metafísica de un mundo prefabricado (a ready-made Word) y 2) la fantasía metafísica de un mundo nouménico. Esta segunda, aunque no lo parezca, se aloja, a juicio de Putnam en las teorías coherentistas de la verdad y en la tesis de la relatividad ontológica. Pero ambas fantasías implican la idea de la verdad como correspondencia.

La primera fantasía está representada *grosso modo* por la tradición realista -llamada «ingenua»- clásica que habla de esencias o formas de las cosas, es decir de objetos autoidentificados; y que entiende que la relación signíca es construída dentro de la naturaleza (recuérdese la forma que actualiza al entendimiento, la doctrina de las especies, etc); pero también esta fantasía está incoherentemente alojada en esa actual metafísica materialista moderna que quiere ser una metafísica natural dentro de los límites de la ciencia. Esta metafísica está atrapada en una teoría de la verdad como correspondencia, una teoría causal del significado, siendo así que ha abandonado toda teoría de propiedades esenciales o de estructura intrínseca... ¿Qué razones esgrimir para el privilegio de esa relación de correspondencia que se llama verdad?. La diatriba contra estas posiciones es la que Putnam lleva adelante en *Por qué no hay un mundo prefabricado*.

Más sutil es y más necesario parece evidenciar la fantasía metafísica de un mundo noumenal que subyace a la tesis de la relatividad ontológica y a la epistemología

³⁵ A esta corrección o crítica está dedicado el libro *Representation and Reality*.

naturalizada de Quine. Quine conserva la idea de correspondencia, pero abandona la idea de que hay una correspondencia que es fijada, una relación intencional de referencia. Aquí se proyecta otra imagen de la mente, como posibilidad de una serie de correspondencias, pero donde ninguna puede fijarse como la correspondencia intencional entre palabra y mundo. Aquí nos quedamos sin criterios transcendentales, digamos, o metalingüísticos a los que esté sometido el significado. La palabra «mesa», por ejemplo, se refiere a un conjunto definido de cosas, en un sentido 'empírico' ordinario, pero no en un sentido transcendental. Paradójicamente entonces la noción de «objeto» se convierte en una noción «totalmente metafísica. ¿Qué hacer -dice Putnam- con un X que es una «mesa», o un «gato», o un «agujero negro», o «el número tres»...?»³⁶. La doctrina de la relatividad ontológica, piensa Putnam, se escapa del Scila del esencialismo medieval para venir a parar al Caribdis de las aporías metafísicas kantianas³⁷.

Con estas revisiones Putnam está dando un giro hacia un antirrealismo de entidades platónicas o nouménicas, aunque queriendo quedarse en un «realismo pragmático» que reconoce una diferencia entre «p» y «pienso que p»; es decir un realismo donde la cuestión de la verdad no se identifique sin más con la de la racionalidad.

La inflexión putnamiana a su realismo anterior consiste entonces fundamentalmente en la defensa ahora decidida de un concepto epistémico de verdad. A ello le conduce la consideración de la semántica no realista propuesta por Dummett.

Dummett considera que aprender un lenguaje es aprender una práctica y no una serie de correspondencias. Lo que sabe el que habla una determinada lengua son las condiciones bajo las que las sentencias en esa lengua son afirmables; e identifica a su vez afirmabilidad con conocer la justificación de la sentencia. Ser verdad una sentencia es estar justificada. Quiere esto decir que la referencia no es anterior a la verdad. Objeto y referencia emergen, *ex aequo* del discurso, no le anteceden. (Y Putnam observa que muchas posiciones filosóficas del siglo XX analíticas y continentales -cita a Gadamer- tienden a este enfoque).

Si la afirmación de una frase implica su justificación, la verdad no se reduce a la redundancia de lo dicho en la frase, o tal vez más exactamente a la desentrecomillación de la frase. (La teoría desentrecomilladora de la verdad nos dice que verdadero se entiende cuando se dice, por ejemplo: «la nieve es blanca» es verdadero, si y sólo si la nieve es blanca). Esta concepción de la verdad (Tarski) tiene que decir que la verdad es correspondencia. Sin embargo retrotraer la verdad a condiciones de justificación permite la sutil distinción -tal vez no negada, sino simplemente no explícita en el concepto tarskiano de verdad³⁸- entre condiciones necesarias y condiciones suficientes de verdad, algo que nos lleva a planteamientos kantianos. Y digo planteamientos kantianos porque Putnam recoge la sugerencia de Dummett, evitando el verificacionismo de éste, si justificación se entiende como *justificación en una evidencia presente*. Dicho de otro modo, Putnam evita identi-

³⁶ El X al que se refiere en cualquier caso la mente, al margen de cualquier determinación, sería más exactamente el *objeto transcendental* kantiano.

³⁷ Cfr. Ph.P.III, Introduction, xii-xiii. A Quine dedica Putnam *Por qué no puede ser naturalizada la razón* en Ph.P.III, 13. *Por qué no hay un mundo prefabricado* es Ph.P.III, 12. Los dos trabajos fueron traducidos en «Cuadernos Teorema», 65. Madrid, Revista TEOREMA, 1985. Los trabajos datan de 1981.

ficarse *tout court* con el justificacionismo por los malentendidos que implica. Verdad es justificación, pero *justificación idealizada*, dice Putnam. ¿Qué puntualización es ésta?

Putnam afronta con tal puntualización el hecho de que no es posible un método general o una regla universal que establezca cuáles son las mejores o peores condiciones de justificación de un arbitrario juicio empírico. Hay que reconocer un hiato entre justificación y verdad (a gap between justification and truth). Lo que Dummett no ha llegado a ver claro es que lo que aprendemos cuando aprendemos un lenguaje no es un conocimiento que pueda ser aplicado como un algoritmo. Ello es así porque la imposibilidad *práctica* de formalizar las condiciones de afirmabilidad de sentencias empíricas no es más que el reflejo o la traducción de la imposibilidad de formalizar la inteligencia general, en si misma.

Por tanto el concepto de justificación es como una función que se satura de acuerdo con las condiciones epistémicas que nuestro conocimiento empírico va marcando a medida que se desarrolla. «El meollo de mi propio realismo es que hay condiciones epistémicas peores y mejores para la mayoría de los juicios, y la pura verdad en cuanto a lo que sería el veredicto si las condiciones fueran suficientemente buenas es un veredicto hacia el que la opinión convergería si fuéramos suficientemente razonables³⁹. La pura verdad del veredicto es *more kantiano* una idea reguladora que rige: las mejores condiciones epistémicas de las situaciones mentales más razonables. Esta es la clase de realismo que le concierne a la clase de racionalidad que es la humana⁴⁰.

Y en tanto que humana Putnam la caracteriza absolutamente al margen de un más o menos consciente telón de fondo sobre el que se han perfilado siempre las «fantasías metafísicas» descritas y sistematizadas por Putnam más arriba: el telón de fondo del *punto de vista absoluto* sobre la realidad y la verdad, el supuesto del entendimiento puro o arquetipo: *the God's Eye view*, que dice repetidamente Putnam. En una palabra, la cuestión acerca de «qué entidades hay» es internamente dependiente del *límite dentro del cual la pregunta tiene sentido*.

Según todo esto Putnam tiene que matizar el cierto «idealismo» que implica su posición por una parte, y pronunciarse en el tema «relativismo» adversus «convergencia» de las teorías, apostando por un sofisticado sentido de la convergencia -que excluye, por supuesto, el concepto regulador de una teoría ideal- para que no sufra un colapso la cuestión de la referencia y de la verdad. Los contextos de tratamiento de estos temas son siempre concretos y las críticas o matizaciones operan siempre como negaciones determinadas. Kuhn y Feyerabend son interlocutores a la vista en estas dos cuestiones, en las cuales también aquel «principio de la duda razonable» o de «caridad» del primer Putnam viene a jugar su papel.

³⁸ La verdad es que la teoría tarskiana abre el espacio de esa distinción en la medida en que para expresar correspondencias del lenguaje y hechos hay que retrotraerse a un metalenguaje que habla de los enunciados y de los hechos.

³⁹ Cfr. Ph.P.III, Introduction, xviii.

⁴⁰ Considero importante hacer notar que Putnam no concede que su realismo pueda ser llamado «científico» si con ello se entiende que todo conocimiento que merece el nombre de tal es parte de la ciencia. Cfr. *Meaning and Moral Sciences*, p.19-20.

A base de todas esas negaciones la concepción putnamiana de la racionalidad se nos va perfilando en su carácter dinámico, activo, de proceso, en una palabra, como *tarea* (Aufgabe), en el sentido kantiano⁴¹.

Putnam adopta un aire evidentemente kantiano, pero desentendiéndose críticamente de cualquier compromiso con el concepto de cosa en sí; pero además apurando el concepto kantiano de racionalidad, yo diría que hasta un límite máximo, pues la hace (a la razón) autorreferente respecto de la posición de fines. Dicho de otra manera, si la razón kantiana, en su interés práctico (promotor, por ser interés primero del interés teórico) se vertía en una dimensión transcendente (objetos metafísicos propiamente dichos), la racionalidad en Putnam abre la prosecución de su propio bien, en el ejercicio de su racionalidad, en «ser racional». Todo esto requeriría un examen más minucioso, pero yo diría que de una finitud o limitación de la razón (en Kant) que lo era precisamente por esa escisión o doble referencia (empírica-metasensible) la racionalidad limitada en Putnam lo es *tout court* sólo por relación a sí, siempre e históricamente abierta al trabajo de la justificación (idealizada) y, en esa medida, siempre concernida por la realidad y por la verdad. Pero toda esta cuestión final merece un nuevo trabajo.

Mercedes Torrevejano
Septiembre 1989

⁴¹ Es un tema importante, porque clarificaría mucho, el estudio pormenorizado de la relación Putnam-Kant. Desde 1978 aparecen muchas alusiones y aun interpretaciones de textos kantianos, siempre con la intención de corroborar la unidad de planteamiento. Bien es cierto que Putnam plantea divergencias con Kant, sobre todo en ética, donde comulga más de grado con Aristóteles, con lo que llamaríamos un *inmanentismo* o *internalismo ético*. Pero el compromiso con el mundo nouménico de la segunda Crítica le parece inasumible. Sin embargo también el Kant ético será reconocido como «forefather» en un sentido muy preciso: en cuanto Kant a juicio de Putnam ha consagrado filosóficamente el concepto de *igualdad* de los hombres, que levantó la revolución francesa y que es tal vez «la única contribución que la religión judía ha aportado a la cultura occidental» (*The many faces...*, Lecture III). (Y tal vez sea interesante indicar aquí la importancia que Putnam otorga a Rawls y a la publicación en 1970 de *A Theory of Justice*, como factor convulsiónante de la tradición analítica. Hasta esta publicación el ocuparse de cuestiones éticas era filosofía de segunda clase. Rawls mostró que «giving up the dream of totalistic epistemology is not the same thing as giving up the dream of philosophy». Es evidente que Rawls ha jugado un papel mediador importante para Putnam. Es precisamente a través de la interpretación de Rawls como Putnam se autointerpreta como kantiano en ética, en alguna medida).

EL «REALISMO» EN LA TRADICION ANALITICA. EL CASO DE HILARY PUTNAM

J.M. Sagüillo

1. ¿Qué es el Realismo Metafísico? ¿Qué critica Putnam?

En (Putnam, 1978, p. 123) se ataca la posición del Realismo Metafísico caracterizado como un modelo «visual» frente a la posición del Realismo Interno que Putnam defiende y entiende como una macro-hipótesis empírica:

«Metaphysical realism, on the other hand, is less an empirical theory than a model -in the 'colliding billiard balls' sense of model. It is, or purports to be, a model of the relation of any correct theory to all or part of The World. I have come to the conclusion that this model is incoherent».

Si lo que Putnam va a criticar es un modelo, podemos avanzar dos posibilidades en las que el resultado de la modelización puede ser insatisfactorio: 1) porque exista una incongruencia más o menos sustantiva entre el modelo y lo modelado, o, 2) porque supuesto 1, el modelo además sea internamente inadecuado o incoherente. Obviamos la posibilidad de que se dé 2 sin que se dé 1, bajo el supuesto de que lo modelado en la realidad no es inadecuado o inconsistente. Aunque quepa predicar la inconsistencia de posiciones filosóficas podemos dejar de lado el caso, toda vez que, como veremos, el argumento de Putnam se centra en el modelo y deja pendiente el punto crítico de justificar que el modelo refleja adecuadamente lo que pretende representar. Si entendemos correctamente a Putnam, su crítica se centra en 2. Según nuestro autor el Realismo Metafísico queda especificado en las siguientes tres tesis, expresando la idea de una fijación ideal de la referencia en el sentido de cómo el «ojo divino» contemplaría el nexo necesario entre términos y predicados de un lenguaje y objetos y propiedades del mundo:

T1: Existe un mundo independiente de la mente.

T2: Existe una única teoría verdadera acerca del mundo.

T3: La verdad en T2 se entiende en el sentido de correspondencia.

(Cf. Putnam, 1982, p. 49).

Todo modelo en el sentido que Putnam atribuye al Realismo Metafísico (de ahora en adelante R.M.) introduce casi siempre una dosis de idealización simplificadora que implica un distanciamiento entre el modelo y lo modelado en la realidad.

En el ámbito analítico que nos ocupa, la obtención de una explicación filosófica se basa en un principio metodológico carnapiiano fundamental. En este sentido, explicar consiste en introducir un concepto nuevo y exacto (el explicatum) para sustituir a un concepto inexacto (el explicandum). En cuanto a los requisitos de adecuación de un explicatum con respecto a un explicandum, Carnap nos señala cuatro: Debe ser *similar* en cuanto a su significado, sin que ello equivalga a la sinonimia, ni tan siquiera a la equivalencia extensional; i.e., no se trata de sustituir todos los usos del explicandum por los del explicatum sino más bien que se darán ciertos usos del explicandum que no serán usos del explicatum. También debe ser *exacto* incardinándose en un marco conceptual formal, rigurosamente definido en algún lenguaje regimentado. Debe ser *fructífero* en cuanto a su utilidad explicativa en sistemas más amplios que permitan la formulación de leyes y principios científicamente interesantes. Finalmente, debe ser *simple* en función de la disponibilidad técnico-formal y la comparación de rendimientos entre distintos sistemas de explicaciones.

Es de capital importancia, y Carnap así lo indica, no confundir la tarea genuina de la explicación con la clarificación previa del explicandum. Sólo de esta clarificación previa se puede esperar una clara y precisa presentación del problema a dilucidar o la cuestión a explicar.

Atendiendo a esta actitud filosófica, podemos entender que las tesis 1, 2 y 3 arriba mencionadas constituyen el explicandum del R.M. Deseamos subrayar en este punto, que el argumento crítico de Putnam no se centra en la idea filosófica "presistemática" del R.M. sino en su correlato formal o explicatum, y en la demostración de la incoherencia interna de este último. En este sentido, *la teoría de modelos* constituye el correlato formal o explicatum que da cuenta de las relaciones entre lenguaje y mundo. De este modo, en nuestra interpretación de Putnam, lo que él denomina el «modelo» del R.M. se corresponde a modo de explicatum carnapiiano con la teoría de modelos, que constituye una explicación analítica de lo modelado o explicandum del R.M. como quedó especificado en T1, T2 y T3 anteriormente.

Por supuesto, por lenguaje ahora no vamos a entender un conjunto de símbolos con significado, sino el lenguaje asignificativo de la lógica de primer orden, y por mundo entenderemos una porción de la teoría de conjuntos (usualmente una estructura relacional también asignificativa en sí misma). La teoría de modelos nos proporciona pues un modelo «sintético» de representación en el sentido de que las expresiones bien formadas de su sintaxis adquieren «sentido» vía representación conjuntística del mundo que adquiere sentido por su parte por definición nominal estipulativa. Así, podemos entender la teoría de modelos como la rama de la lógica matemática que incluye la metológica y metamatemática y que nos proporciona el modelo o explicatum del R.M. en sus tesis anteriores, y de acuerdo con Putnam.

Uno de los teoremas fundamentales de la teoría de modelos es el teorema de Löwenheim-Skolem a partir del cual Putnam extrae una serie de consecuencias filosóficas negativas para

¹ Es interesante anticipar aquí que los resultados obtenidos por Putnam, no son propiamente hablando una aplicación del teorema de Lowenheim-Skolem, sino una aplicación de lo que las técnicas para probar el teorema

el R.M.¹. La idea básica e indeseable que establece «la paradoja de Putnam» es que la verdad de una oración o conjunto de oraciones bajo una interpretación (el concepto de verdad entendido como satisfacción modelo-teorética) puede preservarse bajo otra interpretación distinta a la primera mientras que la referencia de los términos y predicados componentes de dicha oración se modifica. De este modo, la verdad subdetermina la referencia en la teoría de modelos.

La cuestión que queremos analizar es -utilizando una analogía de Lewis- si la «bomba» de Putnam realmente conlleva al final de la filosofía realista, o, si por el contrario, lo que de su argumento se puede concluir es la ineficacia de la teoría de modelos como explicatum-modelo del R.M.

2. Algunas matizaciones a la caracterización de Putnam del R.M.

Tenemos pues un explicandum y un explicatum o correlato sistemático basado en la teoría de modelos. Veremos a continuación que ambos polos han sido sometidos a críticas.

Con respecto al explicandum (Field, 1982) ha señalado sus dudas acerca de la selección de tesis que configuran el R.M. según Putnam. Field nos indica, creemos que acertadamente, que T1, T2 y T3 son lógicamente independientes y que un realista metafísico puede comprometerse con alguna de ellas sin tener que comprometerse con las demás.

En primer lugar, Field pone en tela de juicio que para un realista metafísico sólo pueda existir una sólo descripción completa verdadera del mundo, en el sentido de que los conceptos que utilizamos para describir el mundo no son inevitables. Es perfectamente imaginable que de tener otra estructura cognitiva de la que poseemos, nuestros predicados en el lenguaje que empleáramos podrían diferir en extensión de la extensión de los actuales; o más aún, podemos concebir un lenguaje sin predicados, pero con artilugios referenciales de distinto tipo. En esta circunstancia perfectamente plausible se requeriría -para mantener la necesidad de T2 dado T1 según Putnam- un criterio nada sencillo de traducción que justificase que realmente se está hablando de una única descripción. En ausencia de tal criterio, Field entiende que T2 no es una tesis imprescindible del R.M.

(Putnam, 1985) en réplica a Field argumenta que T2 sólo tiene sentido *suponiendo* la verdad de T1, esto es, a nuestro entender Putnam en su réplica indica que la relación lógica correcta entre T1 y T2 es la expresada por lo que se denomina un condicional «genuino», en el sentido de que dicho condicional sería indefinido cuando el antecedente (lo supuesto) sea falso. Ello hace evidente que Putnam pretendía una correlación entre sus tesis más «blanda» que la relación de consecuencia lógica, en el sentido de ser veritativo funcionalmente definible (aunque con huecos de valor):

revelan acerca de la naturaleza de los sistemas formales en general. Esta matización es importante porque el teorema de Lowenheim-Skolem sólo se aplica a los lenguajes de primer orden (Cf. Lakoff, 1987, p. 235).

«¿Qué sentido puede tener hablar de una única descripción completa y verdadera del mundo?. Puedo dar sentido a esto último si supongo el realismo metafísico. En este caso existe un conjunto definido *I* de individuos que forman el mundo (pongamos puntos espacio-temporales) y existe también un conjunto definido -llamémosle *P*- que abarca todas las propiedades y relaciones (de cada uno de los tipos, para evitar paradojas, pero consideremos, para simplificar, que son todas del tipo más bajo).

Consideremos un lenguaje ideal que tenga un nombre para cada miembro de *I* y un predicado para cada miembro de *P* (...) Un lenguaje ideal de este tipo no es un lenguaje enumerable (...), pero es único (hasta el isomorfismo) y la teoría del mundo -el conjunto de las oraciones verdaderas, hasta cualquier tipo definido- es igualmente único (...).

A la inversa, si suponemos que hay una teoría ideal del mundo, la noción de una *totalidad fija* de todos los individuos y la de una *totalidad fija* de las propiedades y relaciones de estos individuos queda naturalmente clarificada al identificar la totalidad de los individuos con el rango de las variables individuales y la totalidad de las propiedades y relaciones (de cada tipo) con el rango de las variables de predicado (de este tipo). Los realismos metafísicos *uno, dos y tres* no tienen contenido tomados aisladamente, uno por uno» (Putnam, 1985a, pp. 1-2).

La extensión de esta cita se justifica por dos cuestiones: 1º, corrobora nuestra interpretación del párrafo precedente, y 2º subyacen en ella dos elementos de enorme significación. El primero corrobora nuestra intuición de que Putnam ha solapado la tesis T1, T2 y T3 del R.M. como explicandum con su explicatum modelo-teórico, en el sentido de que su caracterización del mundo se basa en conceptos de la teoría de conjuntos y su caracterización de teoría verdadera se basa en un lenguaje ideal que por su simplicidad admitimos con Putnam, de la lógica de primer orden.

Con respecto a esto último, creemos que Putnam no está hablando de lo mismo que Field. Entendemos que presistemáticamente, los comentarios de Field son atingentes a un explicandum e indican su disconformidad con el mismo. La réplica de Putnam pierde su fuerza en virtud de que no da en el blanco requerido, sino que se centra en el pretendido explicatum de lo que caracterizó presistemáticamente bajo T1, T2 y T3 como el R.M.. Field se ocupa de lo diferenciado a modelar, mientras que Putnam se centra en el modelo.

También es necesario subrayar aquí, a efectos de una discusión ulterior que retomaremos de la mano de las críticas de Merrill (1980), que Putnam parece sugerir que el mundo modelo-teórico no sólo se constituye de individuos *sino también de propiedades*, que estarían previamente especificadas en el modelo sobre el cual se interpreta el lenguaje, punto éste que en su argumento modelo-teórico (Putnam, 1978) parece no asumir, dejando que sea la función interpretativa la constituyente de estructura del dominio del modelo a

partir del lenguaje ideal que se trate (Cf. las críticas de orden ontológico, más abajo).

Field subraya un segundo punto crítico interesante, al defender la separabilidad de T1 y T3; i.e., al apoyar la independencia del mundo sin la noción de verdad por correspondencia:

«For instance, there are many philosophers who think that the redundancy theory of truth or sophisticated variants of it capture all there is to say about truth, that talk of correspondence relations is irrelevant»

(Field, 1982, pp. 554-555).

La teoría de la redundancia se basa en el principio de que es equivalente afirmar que un enunciado es verdadero y afirmar simplemente dicho enunciado. Dicho de otro modo, que la verdad no es una propiedad. Este principio, es el punto inicial por su parte de la noción desentrecomillada de la verdad de la semántica de Tarski y de versiones más sofisticadas ulteriores. Tanto el principio como la noción tarskiana son -en opinión de Putnam- filosóficamente neutrales. Pero el problema filosófico real acerca de la verdad entonces permanece mientras no se aclare la noción de «aserción», puesto que para los partidarios de la redundancia nuestra comprensión de «verdadero» implica la internalización de *condiciones de asertabilidad* y no implica como defendería un realista, sus condiciones de verdad:

«If asserting *P* is merely *uttering P*, then the theory in effect, denies that our thoughts and assertions are *thoughts* and *assertions*. If «assertion» is to be taken in a suitable «thick» sense, however, then we have to recognize that assertions are correct and incorrect. But then the problem of truth reappears when we ask what it is for an assertion to be correct and what it is for it to be incorrect».

(Putnam, 1982b, p. 576).

Putnam nos está planteando en su respuesta a Field que el valor semántico de una preferencia no puede reducirse a «ruido» y que el que defiende que no hay verdad (que no existe tal propiedad) puede «salvarla» mediante el expediente de una teoría desentrecomilladora. Sin embargo, el problema filosófico permanece abierto cuando nos planteamos bajo qué condiciones sería correcto aseverar que *P*.

En este punto, entendemos que la réplica de Putnam a Field es contundente. Si Field se compromete con T1 (como de hecho hace) y admite que T1 no expresa meros «ruidos», entonces T1 reúne para Field las condiciones requeridas de asertabilidad y que por tanto, forma parte de un corpus total de conocimiento que se acepta. Pero ésto, nos advierte Putnam, la convierte en una especie de aserción «inmanente» a la que sólo se le puede asignar algún valor de coherencia relativo a tal corpus. Resulta obvio pues, que la posición realista -bajo esta perspectiva- queda sustancialmente desdibujada. En otros términos, mantener T1 y una teoría de la redundancia resulta incomprensible.

3. Críticas al explicatum modelo-teorético de Putnam: Teoría de modelos y el Teorema de Löwenheim-Skolem

En los ámbitos científicos, resulta muy frecuente el diseño y uso de modelos que se construyen a partir del material que provee la moderna matemática. En este sentido, se entiende que la lógica matemática es una rama aplicada de la matemática cuya finalidad es la construcción de modelos a partir del lenguaje matemático para comprender y explicar mejor el fenómeno lógico implícito en un lenguaje presistemático (como ejemplifica una lengua natural por ejemplo).

Así, un lógico matemático diseña para beneficio de sus investigaciones lógicas «*sistemas formalizados*» que resultan similares a los modelos matemáticos de los sistemas solares, cuerdas vibrantes o átomos en la física matemática, y a los modelos matemáticos de computadores en la teoría de los autómatas. (Cf. Corcoran, 1974).

Es ya hoy tradicional presentar sucintamente un modelo tal de la lógica matemática como un sistema $\langle L, D, S \rangle$ compuesto por tres partes, donde L designa un lenguaje; i.e., un subsistema sintáctico o gramática cuya finalidad es capturar sólo la forma lógica de los enunciados, D designa a su vez otro subsistema sintáctico que es el cálculo deductivo y que contiene elementos que representan principios de razonamiento reales o ideales basados en L , y finalmente S designa la semántica del sistema -típicamente se trata de una estructura conjuntística- cuya finalidad es capturar el fenómeno del significado; i.e., denotación, valores de verdad, etc.

Como indicamos anteriormente, la teoría de modelos es una rama de la lógica matemática que incluye la metodológica y la metamatemática y que da cuenta de las relaciones entre lenguaje y mundo, teniendo presente que por lenguaje entendemos un conjunto de símbolos asignificativos (L) y que por mundo entenderemos una porción de la teoría de conjuntos (S), también asignificativa en sí misma. *Dicho elemento S como estructura ontosemática, sólo determina las condiciones de verdad de las oraciones de L en virtud de la denotación de las expresiones que las componen.* Quede claro, que la semántica de un lenguaje formal no ofrece la «comprehensión» del significado de las expresiones no-lógicas (ya que tradicionalmente se considera que los significados determinan la denotación), pero es suficiente para lidiar el problema de la significación de las expresiones lógicas. *Esto sugeriría que lo característico de una expresión lógica es que su significado viene dado por sus condiciones de verdad.*

Usualmente, en L tenemos constantes lógicas y símbolos no-lógicos: un conjunto enumerable de variables, constantes, predicados n -ádicos, etc. Una teoría T construida a partir de L recibe una interpretación fijando un universo del discurso D sobre el que las variables toman valores y atribuyendo a través de reglas de designación, a cada constante individual un individuo miembro de D , a cada predicado monádico un subconjunto de D , y en general una relación o secuencia ordenada de n -tuplos a cada predicado n -ádico ($n > 1$). En la terminología actual, la secuencia ordenada de D seguida de propiedades y relaciones se denomina estructura o posible realización de T . Finalmente las reglas de verdad determinan bajo qué condiciones una oración de T es verdadera en una estructura tal.

Decimos que una *interpretación satisface* una teoría T (obtenemos un modelo de T) si todas las oraciones de T resultan verdaderas en la estructura en relación a una asignación de valores semánticos a los componentes sintácticos de las fórmulas de T. (Cf. Fraenkel, Bar Hillel & Levy, 1984).

3.1. Controles de Adecuación de los Sistemas Formalizados

La metamatemática provee los criterios por los cuales los sistemas formalizados o modelos explicata $\langle L, D, S \rangle$ pueden ser evaluados internamente.

Algunos de ellos son criterios de *consistencia*. Un sistema formalizado para demostrar sentencias es *consistente con respecto de la negación* si y sólo si ninguna sentencia A y su negación resultan ambas demostrables. Un sistema es *relativamente consistente* si y sólo si resulta interpretable en otro; i.e., los términos fundamentales (primitivos) no lógicos del primer sistema son definibles dentro del segundo sistema. De este modo, si el segundo es consistente el primero también lo es. Un sistema es *absolutamente consistente* si y sólo si hay una sentencia que no es demostrable.

Otro criterio para evaluar la consistencia de un sistema formalizado es el de corrección (soundness). La corrección es una propiedad semántica relativa a una interpretación o clase de interpretaciones del lenguaje formalizado. Se trata de comprobar si el sistema establece sólo resultados «lógicamente distinguidos» con respecto a la pretendida interpretación o clase de interpretaciones. Claramente un sistema que no sea correcto/consistente es insatisfactorio e inútil.

Otro criterio evaluativo de los sistemas formalizados es el criterio de *completud*. Algunos criterios de completud son sintácticos y otros semánticos. Todos ellos permiten detectar si un cálculo deductivo establece «suficientes» resultados. Un criterio sintáctico establece que un sistema formalizado es completo si obtiene todos los resultados que sean posibles bajo la condición de la consistencia. Por ejemplo, un sistema es *completo en relación a la negación* si al introducir una sentencia no demostrable en el sistema a modo de nuevo axioma da lugar a un sistema inconsistente en relación a la negación.

De distinto modo, el criterio *semántico de completud* requiere que el sistema sea capaz de establecer todo resultado distinguido como válido del ámbito semántico que el cálculo deductivo pretende capturar.

Aunque un sistema formalizado debe ser consistente y correcto, no siempre resulta posible obtener sistemas completos. Por supuesto, siempre que sea posible la completud es deseable.

En 1930, Kurt Gödel estableció la completud semántica del cálculo de primer orden; i.e., probó que toda fórmula válida es demostrable, o en su forma equivalente, que toda fórmula es o bien satisfactible en un dominio enumerable, o bien refutable. (Cf. Gödel, 1930, p. 585). De este modo Gödel obtuvo por una vía distinta el teorema de Löwenheim-Skolem que había sido establecido años atrás.

El teorema de Löwenheim (1915)-Skolem (1920) dice que si un sistema de la lógica de primer orden tiene un modelo, entonces tiene un modelo cuyo dominio es denumerable. De modo más preciso, para toda fórmula cerrada ϕ de la lógica de primer orden, si ϕ tiene un

modelo $\langle D, f \rangle$ donde D es un conjunto no vacío y f es una función interpretativa entonces $\emptyset$ tiene un modelo enumerable $\langle D', f' \rangle$ tal que $\langle D', f' \rangle$ es un submodelo de $\langle D, f \rangle$ donde $\langle D', f' \rangle$ es un submodelo de $\langle D, f \rangle$ si y sólo si, $D' \subseteq D$ y para cada letra predicativa P de $\emptyset$, $f'(P)$ es la restricción de $f(P)$ a D' . (Cf. Fraenkel, Bar Hillel & Levy, 1984, p. 302 y Van Heijenoort, 1971, p. 253).

Un corolario natural de la prueba de Gödel consiste en mostrar que tener un modelo; i.e., hallar una interpretación que satisfaga un conjunto de oraciones de la lógica de primer orden, equivale a la propiedad de la consistencia semántica. Por tanto, de acuerdo con Gödel y aunando el teorema de Löwenheim-Skolem, la consistencia de la lógica de primer orden equivale a tener un modelo, y tener un modelo implica tener un modelo enumerable².

El teorema de Löwenheim-Skolem tiene consecuencias sorprendentes, una de ellas descubierta y señalada por el propio (Skolem, 1922) hoy se conoce como «la paradoja de Skolem», y resulta de la aplicación del teorema a la teoría de conjuntos. Veamos en qué consiste dicha paradoja.

3.2. La Paradoja de Skolem

Si adoptamos la presentación axiomática de Zermelo-Fraenkel como nuestra «versión» favorita de la teoría de conjuntos tenemos que sus axiomas se comprometen con la existencia de conjuntos infinitos no-denumerables. Así por ejemplo, el axioma 5° , o axioma de conjunto potencia, si aplicado al conjunto ω (el símbolo utilizado por Hilbert entre otros para el tipo del infinito de los números naturales, cuya denumeración ha sido completada en un orden definido; i.e., el ordinal infinito menor), obtenemos que $P\omega = 2^\omega$, cuya cardinalidad es de un infinito mayor que el de los números naturales. Puesto que estamos hablando de un axioma, « $2^\omega > \omega$ », es verdadero en todo modelo de Z.F. Pero entonces todos los modelos de Z.F. deben ser no denumerables. Pero por el teorema de Löwenheim-Skolem ninguna teoría axiomatizada en la lógica de primer orden puede tener *sólo* modelos no-denumerables, sino que también tiene modelos numerables si tiene un modelo cualquiera CONTRADICCIÓN

El propio Skolem y toda la tradición de la lógica contemporánea incluyendo a Putnam, coinciden en señalar que la paradoja es sólo aparente y no real, y ello se constata si reflexionamos brevemente sobre el concepto de modelo: sea M un modelo cuyo dominio D es numerable que satisface los axiomas de Z.F. (la existencia de este modelo está implicada por el teorema de Löwenheim-Skolem). Si el dominio D de nuestro modelo implicado por el teorema de Löwenheim-Skolem es denumerable, entonces tenemos entre manos un modelo «no-pretendido» (cuya existencia no entraba dentro de las expectativas ni intuiciones acerca de los conjuntos) toda vez que para satisfacer los axiomas correspondientes a conjuntos no-denumerables, el compromiso existencial parece conducirnos a dominios de interpretación no-contables. Pero por otro lado, sabemos que el modelo no

² De modo más preciso, el cálculo de primer orden es completo con respecto a todas las interpretaciones de este cálculo y aún con respecto a todas las estructuras denumerables; i.e., toda sentencia del cálculo de primer orden que sea verdadera en toda estructura denumerable es demostrable en el cálculo de primer orden (Cf. Fraenkel, Bar Hillel & Levy, 1984, p. 296).

pretendido de Z.F., si lo es, lo es en virtud de que hace todos sus axiomas verdaderos, de modo que dicho modelo satisface entre otros, el axioma 5º. ¿Cómo es ésto posible? ¿Cómo puede el conjunto de todos los conjuntos de ω pertenecer a un modelo enumerable?

Para entenderlo es preciso tener presente que en el formalismo axiomático de la lógica de primer orden de Z.F., así como en su interpretación en términos de modelos, un conjunto no significa una colección definida; i.e., los conjuntos no son más que objetos que están conectados entre sí a través de ciertas relaciones expresadas en los axiomas. Por lo tanto, no existe contradicción al afirmar que un conjunto B del dominio de D es no-denumerable. Para ello basta que para el subconjunto B de D no exista una biyección con respecto a cualquier subconjunto de los números naturales. Si ello es así, B resulta no denumerable en ese modelo. De este modo, nuestra ontología «capturada» por el modelo nos dice que no existe en D una correspondencia uno a uno entre B y ω (satisfaciendo el axioma 5º), pero por otro lado, la ontología «fugitiva» que apela a la intuición matemática y no a lo cautivo en el modelo nos dice que puesto que D es denumerable, existe una correspondencia tal entre B y N , y que por tanto B no es «en realidad» supernumerable. (Cf. Skolem, 1922, Putnam, 1933, Hamilton, 1980, Mendelson, 1979, Fraenkel, Bar Hillel & Levy, 1984 y Manzano, 1989).

Skolem se refiere a este fenómeno como «la relatividad de las nociones conjuntísticas»:

«Thus axiomatizing set theory leads to a relativity of set theoretic notions, and this relativity is inseparably bound up with every throughgoing axiomatization».

(Skolem, 1922, p. 296)

4. La paradoja de Putnam

La paradoja de Putnam es el resultado de una generalización a partir de las técnicas empleadas en la demostración del teorema de Löwenheim- Skolem (Putnam, 1978) y de su ulterior extensión a lógicas intensionales (Putnam, 1982a).

Según su punto de vista, un pilar básico del modelo realista metafísico es la caracterización del concepto de verdad como radicalmente no-epistémico; i.e., una teoría ideal desde todo punto de vista podría ser no obstante falsa en la realidad. En otras palabras, un realista metafísico distingue tajantemente entre «tener por verdadero» (ámbito epistémico) y «ser verdadero» (ámbito óntico). De este modo, su argumento modelo-teorético tiene el efecto de hacer incomprensible la posibilidad de que una teoría ideal desde el punto de vista epistémico en el sentido de que reúne todas las restricciones teóricas y operacionales (T es consistente, completa, predice, es simple, bella, etc.), sea ónticamente falsa. Putnam mantiene que una formalización total del corpus del conocimiento científico en lógica de primer orden nunca podría obviar la existencia de modelos no pretendidos, haciendo ininteligible la afirmación de que dicha teoría epistémicamente ideal pudiese ser falsa en la realidad:

So let T1 be an ideal theory, by our lights (...)

I assume THE WORLD has (or can be broken into) infinitely many pieces. I also assume T1 says there are infinitely many things (so in *this* respect T1 is «objectively right» about THE WORLD). Now T1 is *consistent* (by hypothesis) and has (only) infinite models. So by the completeness theorem (in its model theoretic form), T1 has a model of every infinite cardinality. Pick a model M of the same cardinality as THE WORLD. Map the individuals of M one-to-one into pieces of THE WORLD, and use the mapping to define relations of M directly in the world. The result is a satisfaction relation SAT -a «correspondence» between the terms of L and sets of pieces of THE WORLD- such that the theory T1 comes out *true* -true of THE WORLD- provided we just interpret «true» as TRUE (SAT). So what becomes of the claim that even the *ideal* theory T1, might *really* be false?.

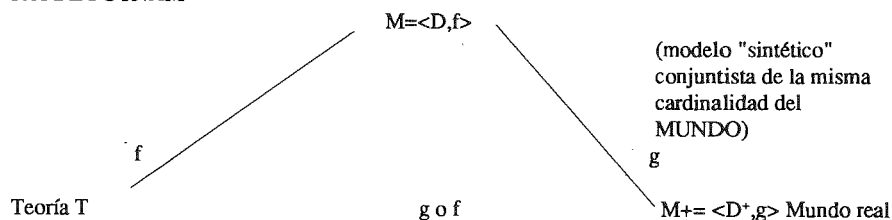
(Putnam, 1978, pp. 125-126)

El interesante argumento del párrafo citado implica la no diferenciación entre LA VERDAD del R.M. y la aceptabilidad racional, y el colapso por fusión de la primera noción en favor de la segunda, vía reducción al absurdo a partir del supuesto realista de VERDAD no-epistémica.

Así, supongamos que LA VERDAD y la aceptabilidad racional son nociones independientes. Si esto es así, entonces es posible que una teoría ideal en todos los aspectos (tanto teóricos como operacionales) pueda ser no obstante falsa en la realidad. Paso seguido, asumamos que el mundo tiene n (infinitos) individuos y que T es compatible con la idea de que existan n -individuos. Puesto que T es consistente, por el teorema de Completud o lo que es lo mismo en este contexto, por el teorema de Löwenheim-Skolem, T tendrá un modelo de cada cardinalidad. Sea M un modelo de la misma cardinalidad que EL MUNDO; i.e., M SAT T. Si M tiene la misma cardinalidad que EL MUNDO, existirá una correspondencia g uno a uno entre D y D^+ , de donde g induciría una relación de SAT entre M^+ y M. Pero a su vez, por una simple composición de funciones de interpretación g o f , tenemos que la relación SAT así obtenida nos permite afirmar la verdad de T con respecto al MUNDO. De este modo, las sentencias de T son «realmente» verdaderas acerca del mundo.

En un diagrama sencillo, la idea sería la siguiente:

PARADOJA DE PUTNAM



(cumple requisitos de adecuación teóricos y operacionales)

Así, $g(f(T))$ es una posible proyección del mundo real en T que por tanto resulta verdadera y que siempre será posible construir. Por lo tanto, la teoría T nunca podría ser falsa, contradiciéndose la suposición y solapándose por tanto la VERDAD y la aceptabilidad racional en un mismo concepto.

Con este pasaje se abren tres fisuras en el modelo realista metafísico que se materializan a nuestro entender en tres tipos de críticas: críticas de orden epistémico manteniéndose la inconsistencia del concepto de verdad modelo-teorético entendida de un modo no epistémico como verdad por correspondencia; críticas de orden ontológico manteniéndose la inconsistencia del concepto de MUNDO (único) modelo-teorético, y finalmente, críticas de orden semántico, implicando la inconsistencia de la semántica modelo-teorética como una teoría del significado. Los siguientes apartados examinan detalladamente estos tres puntos.

4.1. Críticas de Orden Epistémico: inconsistencia del concepto de verdad modelo-teorético entendida como verdad por correspondencia

El argumento de Putnam acerca de la no inteligibilidad del concepto de verdad no epistémico modelo teorético tiene un eslabón fundamental en su construcción de relevancia filosófica-formal. Se trata de la tesis de que la definición del concepto de verdad tarskiano para los lenguajes formalizados (verdad en L) como noción desentrecomilladora, que cumpliría el requisito de ser formalmente correcta y materialmente adecuada, no nos aportaría sin embargo, un concepto filosófico de dicha noción.

Este punto de vista tiene un defensor también en (Field, 1972) que considera que la labor de Tarski es de indudable valor para la semántica de la teoría de modelos reduciendo la noción semántica de verdad a otras tres:

It explains what it is for a sentence to be true in terms of certain semantic features of the primitive components of the sentence: in terms of what it is for a name to denote something, what it is for a predicate to apply to something, and what it is for a function symbol to be fulfilled by some pair of things (...).

We see then that *Tarski's semantics explains the semantic properties of complex expressions* (e.g. truth value for sentences, denotation for complex singular terms) *in terms of semantic properties of their primitive components.*

(Field, 1972, pp. 350-351)

De este modo, de acuerdo con Field, Tarski define el concepto de verdad en términos de denotación primitiva, pero la denotación referencial de dichos primitivos sigue siendo una cuestión filosóficamente abierta. Por lo tanto, lo único que tenemos con la semántica tarskiana es una lista que especifica cómo la correspondencia «aséptica» de la teoría de modelos se fija.

La utilidad obvia del trabajo de Tarski radica en que provee de un artilugio de «ascenso semántico» que permite evitar las paradojas semánticas, pero su famosa «convención T»

es sólo una relación de equivalencia entre sentencias que no nos compromete con lo que se quiera significar con dicha equivalencia o el modo de verificar o refutar la misma (Cf. Putnam, 1978, lección 2ª).

Field por su parte, propone que la denotación primitiva debe ser dilucidada filosóficamente en términos de la doctrina fiscalista; i.e., según Field, los hechos semánticos serían explicable en términos de hechos físicos. Así, por ejemplo, la teoría causal de la referencia incorporaría este punto de vista planteando una posible solución a la relación de correspondencia a partir de una relación causal. De este modo, la referencia del lenguaje para los objetos es exitosa en virtud de que existe una interacción física con ellos.

El talón de Aquiles de estas doctrinas radica -en opinión de Putnam- en la especificación del tipo de interrelación causal referida. (Cf. Putnam 1982c y 1985). Por otra parte, Putnam mantiene que no todas las propiedades semánticas o «intencionales» pueden ser reducidas a propiedades físicas según el siguiente esquema fiscalista:

« x refiere a y si y sólo si x mantiene la relación R con y » (dónde R se supone una relación física de tipo causal).

En efecto, podemos decir que «perro» refiere a los perros si y sólo si «perro» mantiene una relación causal con los perros, pero no podemos utilizar dicho esquema reductivo para explicar la relación causal del término «causa» puesto que estaríamos afirmando que «causa» refiere a las causas si y sólo si «causa» mantiene una relación causal con causas; i.e., « R » refiere a R si y sólo si « R » mantiene la relación R con R . De este modo, deberíamos admitir el esquema R (« R », R) incluyendo la relación en su propio contradominio cometiéndose un flagrante error de tipo (Cf. Putnam, 1982).

(Glymour, 1982) argumenta que de la posición de Putnam no se concluye que las relaciones causales no funcionen a la hora de fijar la referencia de «causa», sino simplemente que las sentencias que contengan la palabra «causa» fracasan al representar la fijación de la referencia de «causa»:

I do not believe that causal theorists hold that reference is ultimately fixed by *sentences* containing the word «causes»; they hold that reference is fixed by causes.

(Glymour, 1982, p. 177)

En principio este punto podría señalar una confusión entre lenguaje objeto y metalenguaje en Putnam. En efecto, una cosa es la relación física que fija la referencia y otra distinta es hablar acerca de un esquema de sentencia reductiva en un lenguaje.

Glymour sostiene sin embargo, que la circularidad de Putnam tiene su interés puesto que sugeriría una especie de diálogo ad infinitum donde cualquier afirmación acerca de lo que fija la referencia nunca es suficiente, porque tal afirmación como parte de la teoría T ideal del argumento modelo teórico sería verdadera bajo g o f y también bajo cualquier otra interpretación sobre el mundo según se deduce del teorema de Löwenheim-Skolem. Así, si mi competencia lingüística es causada por una serie de eventos $E_1, E_2, \dots$ que tienen lugar en

el mundo, entonces también es verdadero que mi competencia lingüística es causada + por $E_1^+, E_2^+, \dots$ (donde «+» denota la entidad correspondiente en un modelo no pretendido) (Cf. Putnam, 1982b).

Desde nuestra perspectiva, entendemos que si Glymour admite la no estricta circularidad «viciosa» de la tesis de Putnam, entonces le concede validez a su argumento puesto que no habría manera de determinar la verdad de la oración «la referencia se fija causalmente» de modo independiente, fuera de alguna interpretación.

Este tipo de dilema es un elemento recurrente en el quehacer filosófico, y han sido descritos desde la corriente hermenéutica como «círculos de la comprensión». Dejando de lado la cuestión de si dichos dilemas representan o no una limitación estructural genuina del entendimiento, el propio Glymour ha discutido un problema similar en el contexto de la reconstrucción estructuralista de las teorías científicas, donde la comprensión de las funciones teóricas (la obtención de valores para las mismas), presupone la comprensión global de la teoría donde aparecen, en particular la comprensión de los axiomas donde dichas funciones aparecen.

Análogamente ahora tenemos que para afirmar algo con sentido acerca de lo que fija la referencia, debemos presuponer que la referencia de nuestras expresiones está previamente fijada, produciéndose el tipo de circularidad antes aludido.

De este modo Putnam concluye que no existe ningún modo de dar cuenta de modo coherente del concepto de verdad por correspondencia, simplemente porque existen demasiadas correspondencias y la selección de una de ellas sólo podría resolverse de un modo «mágico» apelando a algún tipo de «pegamento metafísico» (Cf. Putnam, 1985):

Hay muchas maneras diferentes de distribuir por un lado los signos de un lenguaje y por otro las cosas de un conjunto S , de modo que estén en una cierta correspondencia entre sí (...) ¿Cómo podemos seleccionar *una* única correspondencia entre nuestras palabras (o pensamientos) y las supuestas cosas independientes de nuestra mente si no tenemos un acceso directo a las cosas independientes de nuestra mente?

(Putnam, 1985, p. 12).

4.2. Críticas de Orden Ontológico: inconsistencia del concepto modelo teórico de MUNDO (único)

Es evidente que el solapamiento entre tener por verdadero y ser verdadero implicado por el argumento modelo-teórico no sólo concluye en una relatividad de índole epistémica sino que tiene una contrapartida de índole ontológica. Esto es, es deducible de sus tesis que no sólo hay infinitas versiones (completas) verdaderas acerca del mundo (tesis epistémica), sino que existen infinitos mundos (tesis óntica).

Nelson Goodman, cuyos puntos de vista el propio Putnam señala como afines a los suyos, había afirmado con respecto al modo de ser del mundo que «ninguna descripción nos dice *el* modo de ser del mundo, pero cada una de ellas nos dice *un* modo de ser del mundo»

(Goodman, 1972, p. 31). Veremos a continuación que en virtud de una argumentación formal, basada en la concepción de Putnam de los modelos de interpretación de una teoría T , que la posición que nos ocupa nos permite un parafraseamiento más extremista del dictum goodmaniano. En efecto, comprobaremos en nuestra lectura de Putnam, que no se trata de que cada descripción nos diga *un* modo de ser *del* mundo, sino que cada una de ellas nos dice *el* modo de ser de *un* mundo.

Existe en la literatura modelo teórica actual una ambigüedad sistemática en el uso del término «modelo», que en algunos casos es entendido como una estructura relacional o dominio estructurado (de ahora en adelante *estructura modelo* o *modelo₁*) y por otro, como una *interpretación* entendida como una función de asignación de valores sobre un dominio de objetos (*modelo₂*). En el caso de una *estructura modelo* tenemos una secuencia $\langle D, P, R \rangle$ donde $D \neq \emptyset$ (conjunto dominio), P es un subconjunto del conjunto potencia de D y R es un conjunto de relaciones definidas sobre D . Por otro lado, una *interpretación* en el sentido mencionado es simplemente un par ordenado $\langle D, f \rangle$, donde $D \neq \emptyset$ y f es una función de asignación que toma como argumentos expresiones de un lenguaje y como valores una extensión apropiada en D .

(J. Etchemendy, 1988) ha apuntado las posibles raíces históricas de esta confusión en el sentido de que la noción de modelo como estructura modelo tendría una conexión más o menos directa con el concepto wittgensteiniano de representación pictórica, mientras que la noción de modelo como interpretación sería propiamente un concepto tarskiano.

El punto de vista de Etchemendy es interesante y de suficiente enjundia como para ser discutido en extensión, pero lo esencial de su posición es que obtener un *modelo₂* en sentido tarskiano significa convertir las sentencias de un lenguaje en funciones sentenciales (sustituyendo adecuadamente términos por variables) obteniéndose un lenguaje sin interpretación. La interpretación o realización se obtiene mediante las secuencias de Tarski para las cuales se provee de la correspondiente definición de SAT. Nótese que en este caso, es la *función interpretativa* la que otorga estructura al dominio del modelo a partir de un conjunto de axiomas, asignando contenido/extensión a las expresiones del lenguaje que se supone previamente dado. En otras palabras, el «mundo» que obtenemos en el modelo así construido no es independiente del lenguaje, obteniéndose una interpretación distinta del mismo en la medida en que nos desplazamos de un *modelo₂* a otro en virtud de que tengan lugar distintas asignaciones.

Si se contrasta nuestro esquema de la paradoja de Putnam en la p. 16, se comprueba que Putnam maneja la noción de *modelo₂* en su argumento.

La noción que Etchemendy denomina wittgensteiniana es matemática y filosóficamente distinta (*modelo₁*). En efecto, tenemos ahora una estructura relacional -una representación del mundo «amueblado» de antemano -cuya característica reseñable sería, no su dependencia lingüística, sino su capacidad de representación del mundo. Sólo en un segundo momento, la labor del lógico consiste en diseñar un conjunto de axiomas que resulten adecuados (en el sentido de corrección y completud, etc.) con respecto a dicho enfoque semántico. De este modo, cada *modelo₁* sería una representación abstracta del modo de ser que tiene o que pudiera tener el mundo (tesis de Goodman).

Reinterpreting the language is one thing; canvassing alternative ways the world might be is something else entirely.
(Etchemendy, 1988, p. 104).

Así, en el enfoque Tarski-Putnam la propuesta es cambiar la interpretación del lenguaje de acuerdo con el cambio de un modelo₂ $\langle D, f \rangle$ por otro. Si una sentencia de L es verdadera o falsa, lo es en virtud no de una construcción representativa del mundo, sino en virtud de una interpretación de los objetos sintácticos de L.

Hay otras consideraciones de orden técnico que podemos adelantar por nuestra parte para obtener una distinción aún más determinante si cabe. Kripke distingue de un modo claro entre *estructura modelo* por un lado, y *modelo* por otro. Una estructura modelo kripkeana recoge el sentido arriba asignado a dicha expresión, y un modelo resulta de introducir en la estructura modelo una función interpretativa. Tendríamos por tanto entre manos ahora, una tercera acepción de modelo o *modelo*₃ que combinaría los dos sentidos previos.

La distinción kripkeana resulta fina y perspicaz puesto que nos permite diferenciar claramente dos elementos componentes de cualquier enfoque semántico, ya que por un lado, podemos modificar las estructuras modelo sin variar el tipo de asignaciones, o bien, podemos variar el tipo de asignaciones sin modificar la estructura modelo, o finalmente, podemos modificar ambos elementos a la vez. Este concepto *modelo*₃ nos permitiría -a nuestro entender- reestablecer la distinción óntico-epistémica que una caracterización apropiada del R.M. parecería requerir. Así, una estructura modelo nos daría el segmento óntico de la realidad; el mundo con «su» amueblamiento. La dimensión epistémica que nos da la vía de acceso a ese mundo, queda capturada por un modelo que en virtud de su función de asignación nos proporciona un «anclaje» entre las expresiones de L y las «piezas» del mundo de este modo «capturadas».

En esta línea (Merrill, 1980) ha argumentado que la construcción putnamiana de los modelos (entendidos como modelo₂ = $\langle D, f \rangle$ en nuestra terminología) no hace justicia a un planteamiento legítimamente realista, resultando por tanto, en una mala caracterización de la posición filosófica que pretende criticar.

La presentación formulada del argumento modelo teórico otorga un papel determinante a las restricciones operacionales en la fijación del modelo pretendido, en la medida en que requiere al nivel de los fenómenos observables una correspondencia entre el lenguaje de la teoría y el mundo. Así, si la teoría T ideal cumple los requisitos operacionales, no tiene sentido plantear su falsedad, al menos al nivel del lenguaje observacional. De ahí, que la pretendida interpretación de los términos no-teóricos es clara. De este modo, la única vía de falsedad posible de T tiene que provenir de su correspondencia con el mundo no observable. Pero como hoy es ampliamente aceptado, la experiencia infradetermina la teoría, al menos al nivel de sus términos teóricos. Si estos se refieren a fenómenos no observables; i.e., fenómenos no sujetos a contrastación directa, entonces no se puede pretender que sean relevantes en la fijación de una supuesta correspondencia *con un modelo pretendido*. Si ello es así, entonces no tiene sentido afirmar que el modelo capturado que SAT T no es el modelo

pretendido; i.e., que SAT no estable la pretendida correspondencia entre T y el mundo³. Nuevamente comprobamos que la idea de una teoría ideal falsa en el sentido realista no tiene sentido, cuando menos apelando a que el modelo escogido sea un modelo no pretendido.

Merrill plantea por su parte una reconstrucción modelo-teorética del R.M. que pretende escapar a la aparente ininteligibilidad de que una teoría ideal pueda ser falsa en realidad, *sin apelar a la noción de modelo pretendido*. En su opinión, la caracterización del mundo de cualquier realista metafísico supone la existencia de un dominio estructurado de objetos; i.e., objetos que mantienen ciertas relaciones entre ellos. De este modo, el mundo no puede ser reconstruido en el sentido de modelo₂ sino que para hacer justicia al R.M. el concepto adecuado es el de modelo₃:

To describe *the world* is to describe the entities (or kinds of entities) in it *and their relations to one another*). Again, it is this stronger position which the realist traditionally holds, and as we shall see it is this stronger position which is immune to Putnam's attack.

(Merrill, 1978, p. 72)

De este modo, Merrill plantea que la inmunidad a la crítica putnamiana viene dada porque el mundo tiene estructura (independientemente de que sepamos cuál es ésta e independiente del lenguaje). Su representación modelo teorética correcta sería la de modelo₃ entendido como una estructura relacional parcial a modo representacional en términos extensionales de la teoría de conjuntos, y una función interpretativa I pretendida, cuyo rango está incluido en $D \cup P \cup R$ (donde $D \neq \emptyset$, $P \subseteq P(D)$ y R es un conjunto de relaciones). De este modo, una interpretación en un dominio estructurado asigna a variables y predicados del lenguaje sólo objetos, conjuntos y relaciones existentes en el dominio. En otros términos, la diferencia entre la noción de modelo₂ y modelo₃ es básica y fundamental para la posición del R.M. En el primer caso partimos de un lenguaje que estructura un dominio en virtud de una función de interpretación. En el segundo, tenemos una estructura relacional que representa el mundo, independiente del lenguaje y de la interpretación que «aparea» el lenguaje a esa estructura. De ello resulta obvio concluir, que en el sentido de Putnam, podemos tener tantos mundos como modelos puesto que es la función de asignación la que configura estructura al dominio de objetos (sea de la cardinalidad que sea) con respecto al lenguaje de T.

Manteniendo la noción de modelo₂ putnamiana, (Tuomela, 1979) mantiene, creemos que acertadamente, que habría una circularidad implícita en el argumento puesto que al existir infinitas correspondencias diferentes entre M y el mundo, el argumento presupone lo que quiere demostrar; i.e., que la verdad es fundamentalmente epistémica. Si esto es correcto, podemos indicar que esta reconstrucción modelo teorética no hace justicia a las intuiciones del realista (Cf. sección 4.1. anterior).

³ Merrill dilucida un razonamiento similar que ya Tuomela, 1979, echara en falta en Putnam, 1978, al no establecer con más meticulosidad su noción de teoría ideal en el sentido operacional (epistémico en términos de Tuomela) en conexión con la posible crítica del argumento basada en la noción de que la teoría T es verdadera en un sentido no pretendido.

Al mismo tiempo, Merrill abre una vía clara que lleva a la posición de que una teoría ideal puede no obstante ser falsa en la realidad. Ello se daría cuando no exista una interpretación de T en el mundo que haga a toda sentencia de T verdadera; i.e., cuando T no describa correctamente la estructura del mundo. La cuestión ahora es si es posible construir g (en nuestro esquema) como un isomorfismo, partiendo de que el mundo $+ = \langle D+, P+, R+ \rangle$ tiene una estructura, de modo que obtengamos la mencionada composición $g \circ f$ que haga a T verdadera del MUNDO. g como tal isomorfismo puede no existir, haciéndose por tanto inteligible la idea de la posibilidad de la falsedad de T . Tenemos ahora una prioridad del mundo que *sólo admite* ciertas proyecciones vía $\bar{g}$ (la inversa de g) en modelos sintéticos M de T .

Es importante, Merrill nos indica, subrayar que EL MUNDO tiene una *estructura conjuntística parcial e incompleta* con respecto a la combinatoria de posibilidades extensionales que se generan en un dominio con respecto a propiedades y relaciones; i.e., EL MUNDO si descrito por su dominio $D+$ más el conjunto potencia de $D+$, en unión con $(D+^2)$, unión $(D+^n)$ se convierte en un mundo máximamente estructurado que aseguraría la existencia del isomorfismo g entre M y $M+$ haciendo siempre a T verdadera acerca del mundo⁴.

4.3. Críticas de Orden Semántico: la inconsistencia de la semántica modelo teórica como teoría del significado

Finalmente vamos a analizar las consecuencias nefastas que se seguirían del argumento modelo teórico para cualquier teoría del significado que se fundamente en la teoría de modelos. Vamos a continuación, a pulsar qué respuesta puede esgrimirse de modo razonable a la cuestión acerca de la conexión que intuitivamente creemos que existe entre un término para un hablante, con el objeto denotado como su pretendida interpretación. Se trata de analizar cómo es posible usar el lenguaje referencial y significativamente. El punto de vista de Putnam es que no existe respuesta a semejante cuestión de la mano de la semántica modelo teórica.

Quede claro, que para «repasar» la construcción putnamiana tenemos que recordar que en la semántica modelo teórica, la noción de significado se entiende de un modo extensional como se especificó en la sección 3ª. De modo que el significado de un término

⁴ Con ello no se quiere indicar que el mundo sea incompleto metafísicamente, sino que resulta incompleto con respecto a la clausura del dominio bajo todas las operaciones conjuntísticas usuales. Esto sugeriría una diferencia entre propiedades «genuinas» frente a propiedades «lógicas» que posteriormente (Lewis, 1984), explotó filosóficamente propugnando una defensa «naturalista» (Cf. Lakoff, 1987, p. 241) del R.M. que consiste en distinguir propiedades genuinas frente a propiedades artificiales o pseudo propiedades generadas por la teoría de conjuntos. Una propuesta extrema con respecto a la artificialidad de las propiedades conjuntísticas es la que apunta (Glymour, 1982, p.176) en el sentido de que si no somos realistas metafísicos acerca de objetos abstractos, no tenemos porque creer en la existencia de los conjuntos, y por tanto poner en tela de juicio la existencia del isomorfismo de Putnam. Este tipo de realismo para con los objetos físicos pero no para con los objetos abstractos se propugna en Field (1981).

singular es un objeto del dominio, el de un predicado de una clase, y el de una sentencia un valor de verdad, con respecto a cualquier mundo posible a partir del cual estemos computando.

La consecuencia que Putnam extrae es que es inútil pretender explicar el significado vía relación entre símbolos de un lenguaje L por un lado, y objetos en un mundo conjuntístico por otro, vía funciones de asignación.

Cualquier teoría del significado que pretenda tener visos de adecuación debe cumplir el requisito básico de la *composicionalidad*, i.e., el significado de una oración es el resultado del significado de sus partes componentes, de modo que cualquier modificación en alguna de sus partes implica una modificación en el significado global de la oración. Así por ejemplo, en el significado de «el gato está sobre la estera», la palabra «gato» y la palabra «estera» aportan significado de modo determinante para el significado global de la oración. Desde esta perspectiva, sería impensable una teoría del significado que permitiese una modificación «radical» del significado de «gato» y «estera» sin que ocasionase una modificación en el significado global de «el gato está sobre la estera» (Cf. Lakoff, 1987, p. 230).

La posición de Putnam implica una indeterminación «rampante» de la referencia puesto que su crítica se aplica a cualquier oración de L en el sentido de que viola el principio de composicionalidad. Así, -según Putnam- comprobaremos que modificando la referencia denotativa de los términos se mantiene no obstante, el significado veritativo funcional de la oración que los contiene. Podemos ilustrar el problema con un ejemplo sencillo que el propio Putnam utiliza de modo intuitivo. Se trata de la parábola de la Torre de Babel.

La idea de la parábola es que Dios se veía sometido frecuentemente a ataques de aburrimiento que son en última instancia la causa o motivación de la creación del hombre y el mundo. En el tiempo de la Torre de Babel se volvió a aburrir, y para combatirlo, no sólo hizo que aquel puñado de hombres hablase lenguas diferentes sino que comenzó a jugar con las «flechas de asignación» ó «relaciones de satisfacción» que establecen la correspondencia entre expresiones y referentes en el mundo:

To understand what. He did, pretend that English was one of the languages in existence back then. Imagine that C_1 and C_2 are two admissible «correspondences» (satisfaction relations); i.e., that C_1 (respectively C_2) is the satisfaction relation that one gets if M_1 (respectively M_2) is the model that one uses to interpret English, where M_1 and M_2 are both models which satisfy all the operational and theoretical constraints that our practice imposes. Then what He did (to Hebrew, to Assyrian, to Coptic,...) was to specify that when a *man* used a word, the word would stand for its image or images under the correspondence C_1 , and that when a *woman* used a word, the word would stand for its image or images under the correspondence C_2 .

(Putnam, 1983b, pp. ix,x)

Un ejemplo sencillo modelo teórico de la parábola puede aclarar que Putnam nos está diciendo lo siguiente:

Sea $W = \{m_i/m_i \text{ es un mundo imposible}\}$

Sea $Dm_a = \{a^+, b^+, c^+, d^+, e^+\}$

i.e., Dm_a es el dominio de objetos del mundo actual

Sea «Pa» una sentencia de L

Sea v y m dos funciones interpretativas o correspondencias (v para varones y m para mujeres) tales que efectúan con respecto al mundo actual, las siguientes asignaciones:

$v(P)m_a = \{a^+, b^+\}$; $v(a)m_a = a^+$

Luego «Pa» resulta verdadera para esta interpretación. Sea ahora

$m(P)m_a = \{d^+, e^+, c^+\}$; $m(a)m_a = c^+$

Luego «Pa» resulta verdadera para esta interpretación.

Comprobamos pues, que para el caso de varones y mujeres se viola el principio de composicionalidad en el sentido de que v y m , modifican el «significado» modelo-teórico de las partes de una oración sin que modifiquen el significado modelo teórico de la expresión global. Tenemos pues, que la referencia modelo teórica infra-determina la verdad modelo teórica.

(Lakoff, 1987) ha puntualizado acertadamente que la crítica de Putnam no es contra la teoría de modelos «per se», sino que el problema surge de la utilización de la teoría de modelos al servicio de una teoría del significado:

In short, if Putnam is right, model theory cannot be made into a theory of meaning at all. Model theory is, of course, the natural mathematization of objectivist semantics. *What Putnam is suggesting is that there can be no such possible mathematization. That is, objectivist semantics cannot be made precise without contradiction*
(Lakoff, 1987, pp. 230-231)

Nuestro subrayado pretende poner de manifiesto algo ya aludido en la sección introductoria y es que los problemas putnamianos se erigen sobre el explicatum o modelo matemático que se creía natural a la semántica realista (u objetivista en términos de Lakoff) y no necesariamente a los planteamientos programáticos y de principios de dicha semántica. La inconsistencia es interna al modelo matemático (teoría de modelos) utilizado en la explicación y formalización de tales principios.

A pesar de la aparente claridad del párrafo citado, lo cierto es que Putnam constantemente en todos sus argumentos fracasa al dilucidar claramente entre cuestiones presistemáticas o intuitivas frente a cuestiones formales (Cf. Pearde & Rantala, 1982b). Como ya hemos insistido, en nuestra opinión, la inconsistencia modelo teórica no es

extrapolable al concepto presistemático del R.M., mientras no se demuestre que la teoría de modelos, como pretendido explicatum de dicha posición, suponga una representación fidedigna de la misma.

5. Conclusiones

1. En primer lugar, queremos destacar que todos los criterios de adecuación de los sistemas formalizados que se desarrollan en la metalógica como parte de la teoría de modelos (Cf. sección 3.1. anterior) son test de índole interno a dichos sistemas. Se trata de comprobar en última instancia la deseada mutua adecuación entre sintaxis de L y sus figuras de inferencia asociadas (D) con la semántica correspondiente (S). Sin embargo, no existe a nuestro entender, criterios de adecuación «externos» a la tripleta $\langle L, D, S \rangle$ que permitan evaluar la pertinencia de dicho modelo de la lógica matemática con respecto a lo que pretende modelizar. Estamos reclamando en este punto el desarrollo de una teoría de la modelización que hoy no está desarrollada.

Por supuesto, cualquier modelo de la lógica matemática $\langle L, D, S \rangle$ puede ser «capturado» extrasistemáticamente de un modo preciso por una estructura matemática (algebraica) de algún tipo. Así, un modelo (en el sentido de maqueta) resulta ser un modelo (en el sentido teórico conjuntista) de algún tipo de estructura. Sin embargo, esta caracterización del tipo de estructura que estamos manejando que nos proporciona el álgebra, no nos aporta más que la «la cédula de identidad formal» del constructo, y en ningún caso un criterio de evaluación con respecto a lo que pretende representar. En nuestra opinión, lo que hace patente el argumento de Putnam es el descubrimiento (quizá re-descubrimiento, véase punto 4 abajo) de que lo que se creía predestinado a ser la teoría matemática de la semántica objetivista (al menos en opinión de Putnam) no resulta tal en virtud de la incompatibilidad manifiesta entre el principio de composicionalidad del significado y la referencia extensional propia de la lógica matemática. Lleno aún más lejos, autores como (Pearce & Rantala, 1982a y 1982b) han defendido de modo contundente que la adopción de la semántica modelo teórica es compatible tanto con una visión realista como con una visión instrumentalista de las teorías de la ciencia.

2. En segundo lugar, nadie parece poner en duda que en el lenguaje natural podemos distinguir nítidamente entre interpretaciones pretendidas y no pretendidas. El problema surge cuando en teoría de modelos empleamos símbolos asignificativos (teoría de conjuntos) para aportar significado a un léxico también asignificativo. Sólo en este contexto se puede hablar de la inevitabilidad de las interpretaciones no pretendidas.

3. En tercer lugar, la adopción de un tipo de lógica determinada supone una decisión de carácter metodológico. Esta decisión puede ser más o menos fructífera, pero en el caso de la lógica de primer orden es obvio que resultados como el de Löwenheim-Skolem son conocidos desde hace más de sesenta años. Por tanto, podemos apuntar que no existe nada especialmente paradójico en el resultado de Putnam, puesto que el compromiso adoptado por la decisión metodológica de emplear lógica de primer orden implica la existencia de modelos no pretendidos.

4. En cuarto lugar, podemos citar a R. Carnap, quien defendió el carácter reconstructivo de la filosofía a partir de un principio metodológico que podemos denominar «*criterio de formalización*» y que se sustenta en un principio más general denominado «*principio de tolerancia de las formas lingüísticas*» (Cf. Carnap, 1971, pp. 51-52).

Este principio postula una tesis fuertemente convencional en lo que se refiere al lenguaje fundamental de la ciencia. A partir de esta idea, las tradicionales disputas metafísicas entre realismo y antirealismo podrían encontrar un cauce racional bajo la única condición de que se hicieran explícitas en el lenguaje adoptado las presuposiciones filosóficas que las sustentan. De este modo, las disputas metafísicas, como la que plantea Putnam, tendrían un cauce racional para el debate, convirtiéndose en esta actitud reconstructiva, en cuestiones de índole *metateórica* en la que se podrían esgrimir argumentos de tipo práctico, evaluación comparativa de rendimientos, etc.

Desde esta perspectiva, ¿qué conclusiones válidas pueden extraerse de argumentos de la teoría de modelos aplicada a la lógica de primer orden? Desde nuestro punto de vista, y sin ulteriores especificaciones acerca de los compromisos ontológicos adoptados (que Putnam no facilita) por las características de la lógica empleada, poco margen de maniobra queda para una valoración directa de las doctrinas del realismo desde la teoría de modelos. Por ejemplo, la lógica de primer orden, si limitada oportunamente, puede servir de base subyacente a la aritmética intuicionista, sobre la cual es perfectamente admisible y metodológicamente recomendable aplicar la teoría de modelos como debentadora de los criterios evaluativos de adecuación del constructo formal así obtenido.

Así, pues, con respecto a los teoremas de limitación de la lógica como el teorema de Löwenheim-Skolem, en el sentido de que expresan cierta limitación en el poder expresivo de la lógica de primer orden, rara vez puede delimitarse de modo preciso su incidencia para con ciertas tesis filosóficas, toda vez que éstas no están explicitadas de modo formal como para ser contradichas por los teoremas aludidos. Tampoco por supuesto, lo pueden ser -y este es el punto débil del argumento de Putnam- mientras no existan criterios de evaluación que midan la atingencia entre lo modelado/explicandum (tesis del R.M.) y el modelo/explicatum (lógica de primer orden más teoría de modelos).

Finalmente, queda pendiente a nuestro entender, la cuestión terminal acerca de la incidencia de los teoremas de limitación en supuestas y análogas limitaciones del entendimiento. La lógica filosófica desarrollada por Putnam pretende tener consecuencias importantes con respecto al R.M. Pero sólo si las pruebas matemáticas de resultados limitativos pueden extrapolarse a limitaciones formales del aparato cognitivo humano, el giro kantiano de la filosofía de Putnam podría ser de recibo, como consecuencia del argumento aquí analizado.

J.M. Sagüillo
Septiembre, 1989

Agradecimientos

Este artículo es en gran parte el resultado de un seminario interdisciplinar celebrado durante el curso 1989 en el que intervinieron profesores y estudiantes de tercer ciclo del área de lógica, de metafísica y de teoría del conocimiento de la Facultad de Filosofía y C.C. de la Educación de la Universidad de Santiago. A todos ellos debo mi reconocimiento por la dilucidación y discusión de los problemas expuestos y muy especialmente al profesor Luis Villegas.

Bibliografía

- Carnap, R., 1971, *The Logical Syntax of Language*. Routledge & Kegan Paul, 1ª ed. 1937.
- Corcoran, J., 1974, «Aristotle's Natural Deduction System», *Ancient Logic and its Modern Interpretations*. Corcoran ed., Reidel Pu. Co.
- Etchemendy, J., 1988, «Models, Semantics and Logical Truth», *Linguistics & Philosophy*, 11, pp. 91-106.
- Field, 1972, «Tarski Theory of Truth», *The Journal of Philosophy*, Vol. LXIX, nº 13, pp. 347-375.
- Field, 1981, *Science Without Numbers*, Princeton University Press.
- Field, 1982, «Realism and Relativism», *Journal of Philosophy*, pp. 553-567.
- Fraenkel, A., Bar Hillel, Y. & Levy, 1984, *Foundations of Set Theory*. Amsterdam.
- Glymour, C., 1982, «Conceptual Scheming or Confessions of a Metaphysical Realist», *Synthese*, 51, pp. 169-180.
- Godel, 1930, «The Completeness of the Axioms of the Functional Calculus of Logic», en *Van Heijenoort 1971*, pp. 582-591.
- Goodman, 1972, «The Way the World is», *Problems and Projects*. Bobbs Merrill Co. Indianapolis, N.Y.
- Hamilton, A., 1980, *Logic for Mathematicians*, Cambridge U. Press, 2ª ed.
- Lakoff, G., 1987, *Women, Fire and Dangerous Things*, University of Chicago Press.
- Löwenheim, 1915, «On Possibilities in the Calculus of Relatives», *Van Heijenoort*, 1971, pp. 228-251.
- Lewis, 1984, «Putnam's Paradox», *Australian Journal of Philosophy*, 62, nº 3, pp. 221-236.
- Manzano, M., 1989, *Teoría de Modelos*, Alianza Editorial, Madrid.
- Mendelson, E., 1979, *Introduction to Mathematical Logic*, Van Nostrand Co., N.Y., (2ª ed.).
- Merrill, G., 1980, «The Model Theoretic Argument Against Realism», *Philosophy of Science*, 47, pp. 69-81.
- Pearce, D., & Rantala, V., 1982a, «Realism and Formal Semantics», *Synthese*, 52, pp. 39-53.
- Pearce, D. & Rantala, V., 1982b, «Realism and Reference», *Synthese*, 52, pp. 439-448.
- Putnam, 1978, *Meaning and The Moral Sciences*, Routledge and Kegan Paul. London.
- Putnam, 1982a, *Reason, Truth and History*, Cambridge U. Press (2ª ed.).

- Putnam, 1982b, «Reply to Two Realists», *Journal of Philosophy*, pp. 575-580.
- Putnam, 1982c, «Three Kind of Scientific Realism», *The Philosophical Quarterly*, vol. 32, nº 128, pp. 195-200.
- Putnam, 1983a, «Models and Reality», *Putnam, 1983b*, pp. 1-25.
- Putnam, 1983b, *Realism and Reason*, Philosophical Papers vol. 3. Cambridge U. Press.
- Putnam, 1985a, «Defensa del Realismo Interno», ejemplar mimeográfico. Conferencia de la U. Complutense de Madrid.
- Putnam, 1985b, *Racionalidad y Metafísica*, Teorema, Madrid.
- Skolem, 1920, «Lógico Combinatorial Investigations in the Satisfiability or Provability of Mathematical Propositions: A simplified proof of a Theorem by L. Löwenheim and Generalizations of the Theorem», *Van Heijenoort 1971*, pp. 252-263.
- Skolem, 1922, «Some Remarks on Axiomatized Set Theory», *Van Heijenoort 1971*, pp. 290-301.
- Tuomela, 1979, «Putnam's Realisms», *Theoria*, nº 45, pp. 114-126.
- Van Heijenoort, 1971, *From Frege to Gödel*, Harvard U. Press Cambridge Mass.



ANTOLOGIA DE TEXTOS DE HILARY PUTNAM SOBRE EL REALISMO¹

Luis Villegas

1. «ADVERSUS» REALISMO METAFISICO

1.1. Caracterización general del realismo metafísico o perspectiva «externalista»

1.1.1. *Meaning and the moral sciences (1978)*

a) «El realismo «*metafísico*»... es menos una teoría empírica que un modelo -en el sentido de «modelo» de «bolas de billar en colisión»-... de la relación entre *cualquier* teoría correcta y la totalidad o parte del MUNDO... Este modelo es incoherente.» (pp. 123-124).

b) «... según el modelo del realismo metafísico ha de haber una determinada relación de «referencia» entre los términos de [el lenguaje]L, y piezas (o conjuntos de piezas) del MUNDO, tanto si se considera que *comprender* L consiste en «conocer» esta relación, como en caso contrario. Lo que diferencia esta perspectiva de [la del] realismo «interno» (que usa una descripción similar pero *dentro* de una teoría) es que: (1) Se supone que la descripción se aplica de una vez a todas las teorías correctas (...i.e. trasciende a la formalización completa en cualquier teoría determinada); y (2) supone que el MUNDO es independiente de cualquier representación particular que de EL se tenga-... podríamos ser absolutamente incapaces de representar el MUNDO correctamente (por ejemplo, todos podríamos ser «cerebros en una cubeta», nos dice el realista metafísico).» (p.125).

c) «La consecuencia más importante del realismo metafísico es que se considera a la *verdad* como radicalmente no-epistémica, y así, la teoría que es «ideal» desde el punto de vista de utilidad operacional, belleza y elegancias internas, «plausibilidad», simplicidad, «conservadurismo», etc., *podría ser* falsa. De acuerdo con la perspectiva metafísicamente realista, incluso, en el límite ideal, «verificada» no implica «verdadera».» (p. 125).

d) «Los «realistas sofisticados»... conceden que el relato 1 («hay puntos») y el relato 2 («no hay puntos») son «descripciones equivalentes» [del MUNDO, siendo éste una *línea recta*]... En el pasado, yo aduje que esto no es un problema para el realista-justamente es como el hecho de que la Tierra pueda ser representada mediante «proyecciones» diferentes.» (p.132).

e) «En particular, -yo creía-, puede suceder que lo que nosotros representamos como términos «incompatibles» puede ser proyectado sobre el mismo objeto real -aunque no, por

¹ Con esta antología se ha pretendido recoger algunos de los textos más relevantes de Hilary Putnam sobre el problema del Realismo. Se ha procurado seleccionar tanto aquellos donde el autor describe y critica las versiones metafísicas del Realismo, cuanto otros, donde Putnam expone una alternativa, que denomina «Realismo Interno». Los criterios de selección de texto, sistematización, distribución y denominación de los distintos apartados son de la exclusiva responsabilidad del Antologista.

supuesto, dentro de la misma teoría. Así el objeto real que es rotulado como «punto» en una teoría puede ser rotulado como «conjunto de segmentos de línea convergentes» en otra teoría. Y el *mismo* término puede ser proyectado sobre un objeto real en una teoría y sobre un objeto real diferente en otra teoría. Es una propiedad del mundo en sí mismo,...-i.e., una propiedad del MUNDO en sí mismo- el que «admita estas diferentes proyecciones».» (p. 132).

1.1.2. «Vagueness and alternative logic» (1980). En *Philosophical papers. Vol. III*²

«Dummet usa el término «realismo» para la persona que:

(1) Asume un mundo consistente en una totalidad definida de objetos y propiedades independientes del discurso.

(2) Asume la «bivalencia fuerte», i.e., que en un objeto posee o *carece* determinadamente de cualquier propiedad que puede ser predicada significativamente de ese objeto; y

(3) Asume la teoría de la verdad como correspondencia en un sentido realista fuerte de «correspondencia», i.e., un predicado corresponde a un único conjunto de objetos, y un enunciado corresponde a un único estado de cosas, implicando las propiedades y objetos mencionados en (1), y es verdadero si ese estado de cosas se obtiene, y falso si no se obtiene.» (p. 272).

1.1.3. *Reason, Truth and History* (1981) [Versión castellana: 1988]³

a) «...realismo metafísico. Según esta perspectiva, el mundo consta de alguna totalidad fija de objetos independientes de la mente. Hay exactamente una descripción verdadera y completa de «cómo es el mundo». La verdad supone una especie de relación de correspondencia entre palabras o signos mentales y cosas o conjunto de cosas externas. A esta perspectiva la llamaré *externalista*, ya que su punto de vista predilecto es el Ojo de Dios». (p. 49; p. 59)

b) «La suposición de que podría haber un mundo en el que todos los seres sintientes fueran cerebros en una cubeta presupone desde el principio la visión de la verdad del Ojo Divino -o con más precisión, la verdad del No -Ojo-, la verdad como algo totalmente independiente de los observadores». (p. 50; p. 60)

1.1.4. «Why there isn't a Ready-Made World» (1982) En *Phil. Papers. vol. III. (1983)* [Versión castellana en *Racionalidad y Metafísica. (1985)*]

a) «El realista metafísico sostiene que podemos pensar y hablar acerca de las cosas como existiendo independientemente de nuestra mente, y que podemos hacerlo en virtud de una relación de «correspondencia» entre los términos de nuestro lenguaje y alguna clase de entidades independientes de nuestra mente». (p. 205; p. 9).

² La traducción de los textos donde no hay referencia a previas versiones castellanas se debe al antologista.

³ En todos los textos donde figura una referencia a la versión castellana hay una doble paginación: la primera, corresponde a la versión inglesa, la segunda, a la traducción castellana de referencia.

b) «Lo que hace que un realista metafísico sea un realista «metafísico» es su creencia en que en alguna parte hay «una teoría verdadera» (dos teorías que fueran descripciones completas y verdaderas del mundo, serían para él meras variantes notacionales). Junto con una teoría de la verdad como correspondencia, esta creencia en una teoría verdadera requiere un mundo *prefabricado* [«ready-made»]: El mundo como tal ha de tener una estructura «dada», ya que de lo contrario, teorías con estructuras diferentes podrían «copiar» correctamente el mundo (desde diferentes perspectivas) y la verdad perdería su carácter absoluto (no perspectivista)...[y] la «correspondencia» entre nuestros símbolos y algo que no tiene una estructura determinada es una noción que difícilmente puede ser mantenida». (p. 211; pp. 19-20)

1.1.5. *Representation and Reality* (1988)

a) «...Durante largo tiempo he sentido un conflicto de atracción/repulsión respecto del «realismo metafísico». En varios lugares he descrito al realismo metafísico como un haz de ideas filosóficas sobre la verdad íntimamente asociadas: las ideas de que la verdad es un asunto de Correspondencia y que exhibe Independencia (de lo que los humanos descubran o puedan descubrir), Bivalencia, y Unicidad (no puede haber más de una descripción verdadera y completa de la realidad)». (p. 107).

b) «Lo que yo usualmente encontraba atractivo del realismo metafísico es la idea de que el modo de resolver problemas filosóficos es construir una mejor visión científica del mundo. Esta idea retiene el antiguo principio de que SER es *prior* a CONOCER... De esta idea me vi llevado naturalmente a pensar que la ciencia debería comprenderse sin «reinterpretaciones filosóficas». Desde una tal perspectiva, Independencia, Unicidad, Bivalencia y Correspondencia son ideas regulativas de las que se espera que se nutra la imagen científica final, tanto como supuestos que garantizan que debe ser posible una tal resolución científica final de todos los problemas filosóficos». (p. 108).

1.2. Argumentos contra el Realismo Metafísico en general

1.2.1. *El argumento modelo-teórico básico*

1.2.1.1. *Meaning and Moral Sciences* (1978)

a) «Sea T_1 una teoría ideal... (tal que) podemos imaginar que tiene todas las propiedades que deseamos *excepto verdad objetiva*. Así, T_1 puede ser... completa, consistente, puede predecir correctamente todas las oraciones observacionales y satisfacer cualquier «constricción operacional» que haya, puede ser «bella», «pausable», «simple», etc. El supuesto bajo consideración es que T_1 puede ser todo esto y además ser falsa». (p. 125).

b) «Asumo que el MUNDO consta de (o puede ser desmenuzado en) infinitamente muchos trozos. Asumo también que T_1 dice que hay infinitamente muchas cosas (así en este respecto, T_1 es «correcta objetivamente» acerca del MUNDO). Ahora bien T_1 es consistente (por hipótesis) y tiene (sólo) modelos infinitos. De modo que, por el Teorema de Completud (en su forma modelo-teórica) T_1 posee un modelo de cardinalidad infinita. Tómese un modelo M de la misma cardinalidad del MUNDO. Proyétense uno a uno los

individuos de M en las piezas del MUNDO. El resultado es una relación de satisfacción, SAT -una «correspondencia» entre los términos de L y conjuntos o piezas del MUNDO- tal que la Teoría, T_1 deviene *verdadera* -verdadera del MUNDO, supuesto que interpretemos «verdadera» precisamente como VERDADERA (SAT). Así, ¿qué pasa con la afirmación de que incluso la teoría ideal, T_1 , podría *realmente* ser falsa?». (p. 126).

c) «...la interpretación de la «referencia» en L como SAT cumple ciertamente las «constricciones operacionales» sobre la referencia. Pero la interpretación de «referencia» como SAT cumple ciertamente (también) todas las constricciones teóricas sobre la referencia. (Todo ello) hace que la Teoría ideal, T_1 , devenga verdadera». (p. 126).

d) «...¿qué otras constricciones sobre la referencia puede haber, de modo que distingan alguna otra interpretación como (la única) «pretendida», y determinen a SAT como una interpretación «indeseada» (en el sentido modelo-teórico de «interpretación»? El supuesto de que incluso una «teoría ideal» podría *realmente* ser falsa parece sumirse en la ininteligibilidad». (p. 126).

e) «El problema es que ...(el) relato (del realista «sofisticado») puede retener el MUNDO pero al precio de renunciar a cualquier noción inteligible de cómo sea el MUNDO. Cualquier oración que cambie el valor de verdad al pasar de una teoría correcta a otra, -p. e., una descripción equivalente- expresará sólo una propiedad del MUNDO relativa-a-la-teoría. Y cuantas más oraciones de esta clase haya, más propiedades del MUNDO se volverán relativas -a-la-teoría». (p. 132).

f) «El hecho es que, *en la medida en que muchas propiedades del mundo* -partiendo precisamente de las *categoriales*, como cardinalidad, particulares o universales- resulten ser relativas-a-la-teoría, el MUNDO acaba siendo como un mundo «nouménico» kantiano, una *mera* cosa en sí. Si no se puede decir cómo es el MUNDO, independientemente de las categorías, resulta vacuo hablar de estas teorías como «descripciones del mundo»». (p. 133).

g) «Podemos decir que el lenguaje tiene *más de un modo correcto* de proyectarse en el MUNDO (...ya que tiene una manera de ser proyectado correctamente en otro lenguaje que se proyecta, *él mismo* correctamente en el mundo). Pero en este momento, toda comprensión de la perspectiva (realista) parece desvanecerse: si lo que es un único conjunto de cosas *dentro de una teoría correcta* puede no ser un único conjunto de cosas «en realidad», entonces el corazón auténtico de la perspectiva es arrancado.» (p. 135).

1.2.1.2. «Models and Reality» (1980). *En phil. Papers. Vol. III.*(1983).

a) «A lo que Skolem apuntaba realmente era a esto: ninguna teoría interesante (en el sentido de una teoría de primer orden) puede, en y por sí mismas determinar sus propios objetos, salvo isomorfismo. El argumento de Skolem puede extenderse, como veremos, hasta mostrar que si las constricciones teóricas no determinan la referencia, entonces la adición de constricciones operacionales tampoco lo hará. En este punto es donde la referencia comienza a parecer «oculta»; comienza a parecer que no se puede ser tipo alguno de realista sin ser un creyente en poderes mentales innaturales...» (p. 23).

b) «Desde el punto de vista del no-realista, es trivial decir que un modelo en el que, como pudiera ser, el conjunto de los gatos y el conjunto de los perros son permutados (i. e.; se

asigna como extensión de «gato» el conjunto de los perros, y como extensión de «perro», el conjunto de gatos), es «indeseado» incluso si ajustes correspondientes en la extensión de todos los demás predicados hacen que al final sean «preservadas» todas las constricciones operacionales y teóricas de la ciencia total o de la total creencia. Un modelo tal sería «indeseado» a causa de que *no pretendemos* que la palabra «gato» se refiera a los perros. Pero desde el punto de vista del realista metafísico, esta respuesta no funciona; él relanza la cuestión precisamente al metalenguaje. El axioma del metalenguaje: «gato» se refiere a los gatos no puede eliminar una interpretación indeseada tal, a menos que el metalenguaje mismo haya seleccionado ya su interpretación pretendida; pero estamos en el mismo predicamento respecto del metalenguaje... Sin embargo, desde el punto de vista de la semántica «no-realista» el metalenguaje es completamente comprendido, y también el lenguaje objeto. Así, podemos decir y comprender ««gato» refiere a gatos». Aunque el modelo referido satisfaga la teoría, etc.; es «indeseado» *a partir de la descripción mediante la cual es dado*. Los modelos no son niños nouménicos buscando alguien que los nombre; son construcciones dentro de nuestra propia teoría, y tienen nombres desde su nacimiento.» (pp. 24-25).

1.2.2. Argumentos Epistemológicos

1.2.2.1. Reason, Truth and History (1981). [Trad. cast. 1988].

a) «...¿por qué hemos de concluir que debe abandonarse la teoría de la verdad como correspondencia? Aún cuando la noción de «similitud» entre nuestros conceptos y sus referentes no solucione el problema, ¿no podría haber algún tipo de isomorfismo abstracto... algún tipo de proyección de conceptos sobre cosas en el mundo (independiente de la mente)? ¿No podría definirse la verdad en términos de tal isomorfismo o «proyección»...? (Pero se da) la dificultad, no de que no existan correspondencias entre palabras o conceptos y otras entidades, sino de que existen *demasiadas correspondencias*. Para seleccionar exactamente *una* correspondencia entre las palabras o signos mentales y las cosas independientes de la mente deberíamos de antemano tener acceso referencial a las cosas independientes de la mente. No podemos seleccionar una correspondencia entre dos cosas tomando como punto de apoyo a *una* de ellas (o haciendo cualquier otra cosa sólo a una de ellas).» (p. 72; p. 81)

b) « [Para un internalista]⁴ ...no sólo puede existir una correspondencia entre objetos y (lo que nosotros consideramos) teorías incompatibles (es decir, *los mismos objetos* pueden ser lo que los lógicos llaman «modelo» para teorías incompatibles), sino que, aún cuando fijemos la teoría y los objetos, hay una cantidad infinita (si el número de objetos es infinito) de formas diferentes en las que pueden usarse los mismos objetos para construir un modelo de una teoría dada. Esto es, simplemente, enunciar en lenguaje matemático el hecho intuitivo de que para seleccionar una correspondencia entre dos dominios necesitamos tener acceso independiente a ambos dominios». (pp. 73-74; p.82)

⁴ El uso de corchetes indica una interpolación aclaratoria del antologista.

1.2.2.2. «*Why there isn't a Ready-Made World?*» (1982). En *Phil. Papers*.

Vol. III. (1983) [Trad. cast. en *Racionalidad y Metafísica*. (1985)]

a) «¿Cómo podemos seleccionar una única correspondencia entre nuestras palabras (o pensamientos) y las supuestas cosas independientes de nuestra mente si *no tenemos un acceso directo a las cosas independientes de nuestra mente*? (la Filosofía Alemana casi siempre empezó con una respuesta concreta a esta pregunta -la respuesta después de Kant es «no podemos»).» (p.207; p. 12).

b) «... un acto de desear (o una intención) no resolverá el problema. Simplemente no puedo elegir una correspondencia determinada C y desear (o estipular) que C tenga que ser la relación de correspondencia designada, porque para hacer eso necesitaríamos *de antemano* ser capaces de pensar sobre la correspondencia C -y C, siendo una relación de cosas que son externas e indepentens de mi mente, es ella misma algo fuera de la mente, algo «externo». En resumen, si la mente no tiene la capacidad de aprehender las cosas o formas externas directamente, entonces ningún acto *mental* puede darle la capacidad de seleccionar una única correspondencia (o cualquier cosa exterior).» (p. 207; pp. 12-13).

1.2.2.3. *Representation and Reality*. (1988).

a) «Lo que Gödel mostró es,..., que no podemos formalizar totalmente nuestra propia capacidad matemática a causa de que forma parte de esta misma capacidad matemática el (hecho de) que trasciende cualquier cosa que ella pueda formalizar. Similarmente, mi extensión de la técnica gödeliana a la lógica inductiva mostro que forma parte de nuestra noción de justificación en general el (hecho de) que la razón puede trascender cualquier cosa que la razón pueda formalizar.» (p. 118).

b) «La conexión entre las mencionadas propuestas epistemológicas y las cuestiones acerca de la referencia y el significado viene asegurada por la verdad del holismo del significado... La referencia no es precisamente un asunto de «conexiones causales», es una cuestión de *interpretación*. Y la interpretación es un asunto esencialmente holístico. Una «formalización» completa de la interpretación.... es un proyecto tan utópico como una «formalización» completa de la Fijación de la Creencia... (pues)... saber lo que significan las palabras de un lenguaje (y sin saber lo que significan, no se puede decir a *lo que se refieren*) es cuestión de captar el modo en que son *usadas*. Pero el uso es holístico; pues saber cómo son usadas las palabras involucra saber cómo fijar las creencias que contiene a estas palabras, y la fijación de la creencia es holística.» (pp. 118-119).

c) «Esto no significa que (como Dummett, Fodor, Derrida, y otros han argüido) si aceptamos el holismo del significado, tenemos que decir que cada vez que cambian nuestros procedimientos de fijación de la Creencia, cambia el significado de cada palabra en el lenguaje... Pues «el significado es el uso» no es una *definición* de significado. Los significados no son objetos en un museo, a los cuales quedan prendidas de alguna manera las palabras: decir que dos palabras tienen «el mismo significado» (y/o «la misma referencia») es decir precisamente que es una buena práctica interpretativa igualar sus significados (o sus referencias). Pero una práctica interpretativa sofisticada presupone una sofisticada comprensión del modo como son usadas las palabras por la comunidad cuyas palabras se están interpretando... En suma, intentar examinar el «significado» o la referencia» fracasa por la

misma razón que el intento de examinar la razón misma: la razón puede trascender cualquier cosa que ella pueda examinar.» (p. 119).

d) «El argumento epistemológico recién esbozado no significa que los programas de Traducción Automática, los «programas de procesamiento del lenguaje natural» y similares sean imposibles. Del mismo modo como puede haber sistemas formales de matemáticas más o menos potentes y sistemas formales de lógica inductiva más o menos potentes, puede haber también programas más o menos potentes para interpretar preferencias en un lenguaje natural. Lo que este argumento sugiere (si la analogía sobre la que se basa se mantiene sin problemas) es que, precisamente como ningún sistema formal de matemáticas puede *definir* qué sea una prueba matemática, y ningún sistema formal de lógica inductiva puede definir qué sea ser «confirmado», ningún programa para interpretar preferencias en un lenguaje natural puede definir qué sea para las preferencias ser sinónimas o incluso «correferenciales». Una *caracterización* computacional completa de «prueba», «confirmación», «sinonimia» y demás, será siempre una imposibilidad.» (p. 119).

e) «... desde mi punto de vista lo «epistemológico» y lo «ontológico» están íntimamente relacionados. La verdad y la referencia están íntimamente conectadas con nociones epistémicas: la textura abierta de la noción de un objeto, la textura abierta de la noción de referencia, la textura abierta de la noción de significado, y la textura abierta de la razón misma se encuentran completamente interconectadas. Es de estas interconexiones de donde tiene que surgir un trabajo filosófico serio sobre dichas nociones.»

1.3. Caracterización del Realismo Metafísico Materialista o Fisicalista.

1.3.1. *Reason Truth and History* (1981). [Trad. Cast. 1988]

a) «La sugerencia que hoy está de moda es que el mismo proceso evolutivo ha producido, de un modo u otro, una correspondencia entre nuestras palabras, las representaciones mentales y las cosas externas; se dice que no habríamos *sobrevivido* si no se hubiese dado tal correspondencia, y que ésta constituye la relación de referencia, por lo menos en un nivel rudimentario.» (p. 38; p. 49).

b) «En opinión de Hartry Field, la referencia es una «relación fisicalista», es decir, una relación causal compleja entre palabras o representaciones mentales y objetos o conjuntos de objetos... (y) corresponde a la ciencia empírica descubrir lo que sea tal relación fisicalista. ... Hay una posible definición fisicalista o naturalista de la referencia:

(1) *x se refiere a y* si sólo si *x mantiene R con y* es verdadero, donde R es una relación definible con el vocabulario de la ciencia natural, sin usar nociones semánticas... Si (1) es verdadera y empíricamente verificable, entonces (1) es una oración verdadera hasta en la teoría según la cual la referencia se fija en tanto que (y sólo en tanto que) es determinada por los constreñimientos operacionales más los teóricos. Así pues, (1) es una oración que formaría parte de nuestra teoría del mundo en «límite ideal.» (p. 45; p. 56)

c) «... el problema es el siguiente: ahí fuera hay esos objetos. Aquí la mente-cerebro, llevando a cabo su pensamiento-computación. ¿Cómo entran los símbolos del sujeto pensante (o los de su mente-cerebro) en una correspondencia única con los objetos y conjuntos de objetos de ahí fuera?... La réplica... (de)... los externalistas es que, pese a que

en realidad ningún signo corresponde *necesariamente* a ningún conjunto de cosas más que a otros, las conexiones *contextuales* entre los signos y las cosas externas (en particular, las conexiones causales) harán explicable la naturaleza de la referencia.» (p. 51, p. 61.).

d) «El externalista renuncia al requisito de que tengamos alguna conexión real (por ejemplo, una conexión causal) con *todas las cosas* a las que podamos referirnos, exigiendo sólo que los términos básicos se refieran a géneros de cosas (y relaciones) con los que tenemos alguna relación real. El externalista afirma que utilizando los términos básicos en combinaciones complejas podemos construir expresiones descriptivas que se refieran a géneros de cosas con las que no tenemos ninguna conexión real, pudiendo incluso, no existir (por ejemplo extraterrestes).» (pp. 52-53; p. 62).

e) «La formulación del problema [«cuando usamos la palabra «caballo» no sólo nos referimos a los caballos con los que tenemos conexión real, sino también a todas las demás cosas del mismo tipo»] por parte del realista metafísico... [es]... como si hubiera que empezar con todos esos objetos en sí mismos, adquiriendo entonces un tipo peculiar de lazo con algunos pocos (los caballos con los que tengo conexión real, *vía* cadena causal de tipo apropiado), topandome entonces con el problema de conseguir que mi palabra («caballo») cubra no sólo aquellos objetos que he «enlazado», sino también aquellos otros que no puedo enlazar, bien porque están demasiado lejos en el espacio-tiempo, bien por cualquier otra causa. Y la «solución» del realista metafísico es decir que la palabra cubre *automáticamente* no sólo los objetos que he enlazado, sino también los objetos que son del *mismo tipo* -del mismo tipo en *sí mismo*. Pero entonces, lo que se afirma... es que el *mundo* consta de Objetos que se Autoidentifican, pues afirmar que es el *mundo* y no los sujetos pensantes el que clasifica las cosas en géneros, significa precisamente esto.» (p. 53; p. 63).

1.3.2. «Why there isn't a Ready - Made World» (1982) *En Phil. Papers. Vol. III.* (1983) (Trad. Cast. en *Racionalidad y Metafísica*. (1985)

a) «Hoy, los objetos materiales se consideran como paradigmas de entidades independientes de nuestra mente, y la «correspondencia» como una especie de relación causal. Por ejemplo... lo que hace que pueda referirme a las sillas es que he interactuado causalmente con ellas, y no podría proferir expresiones que contuvieran la palabra «silla» ... si no hubiera establecido una conexión causal «del tipo apropiado» con las sillas. Esta compleja relación -el hecho de estar conectada con *x* por una cadena causal del tipo apropiado- entre mi palabra (o la manera de usar la palabra) y *x* constituye la *correspondencia* revelante entre mi palabra y *x*.» (p. 10)

b) «La tensión o incompatibilidad entre el realismo metafísico y la negociación de las propiedades intrínsecas no ha pasado inadvertida por los materialistas modernos. Y por esta razón encontramos ahora muchos materialistas empleando un vocabulario metafísico que guarda todavía el sabor del siglo XIV: materialistas que hablan de «poderes causales», de semejanzas y diferencias entre cosas que «se dan» en la naturaleza, incluso materialistas que hablan descaradamente de *esencias*.» (pp. 13-14).

c) «... Kant advierte de que el deseo de especulación metafísica, el deseo de construir una teoría sobre el mobiliario del mundo, está profundamente anclado en nuestra naturaleza... El materialismo y el cientifismo contemporáneos son un reflejo de este afán en dos

sentidos. Por una parte, el materialismo mantiene que la Física es una aproximación al bosquejo de la teoría verdadera, la verdadera y completa descripción del mobiliario del mundo... Por otra parte, replica al argumento epistemológico contra la Metafísica al mantener que no *necesitamos* una intuición intelectual para hacer *su* clase de metafísica: es como si su metafísica fuera abierta, infinitamente revisable y falible al igual que la ciencia misma. De hecho *es* ciencia (interpretada como pretensión de verdad absoluta, o, más bien, como *convergen*cencia hacia la verdad absoluta).» (pp. 18-19).

d) «El metafísico materialista... toma las relaciones causales como ejemplos de estructura dada (del mundo). Los hechos tienen causas; los objetos tienen «poderes causales». Y orgullosamente proclaman su realismo sobre estas cosas, su fe de que ellas están «en» el mundo mismo, en el sentido del realista metafísico.» (p. 20).

e) «Una teoría metafísica en la que «causalidad» y «explicación causal» están dadas en el mundo mismo es una teoría en la que la explicación está arrancada de lo que el profesor Frederick Will (1974) ha llamado «la institución del conocimiento» -es decir, la tradición heredada que nos define qué es una condición relevante y qué es un parámetro variable relevante- y se ha proyectado en la estructura de la realidad. Boyd replicaría probablemente que «la estructura causal» de la realidad *explica* el éxito de la institución del conocimiento: las explicaciones que tienen éxito copian simplemente la estructura causal.» (p.27).

f) «En nuestra cultura filosófica, el rechazo de las propiedades «esenciales» comenzó con ejemplos como el de una cosa cuya forma es una propiedad «esencial» en una descripción («esta estatua») pero no lo es en una descripción diferente («este pedazo de arcilla»).... Según Kripke, la estatua y el pedazo de arcilla son dos objetos, no uno. El hecho de que el pedazo de arcilla tenga una propiedad modal, a saber, «ser un a cosa que *podría haber sido* esférica», que no tiene la estatua... muestra ya que los objetos no pueden ser idénticos.» (pp. 32-33).

g) «Pero ahora empieza a parecer como si los objetos, individuados propiamente *tuviesen* esencias, *tuvieran* algunas *propiedades esenciales* en una forma especial... Un materialista cuya ontología incluya «mundos posibles» puede introducir objetos intensionales apropiados identificándolos con funciones de mundos posibles en regiones espacio-temporales. Así, por ejemplo, la estatua sería una función definida en todo el mundo posible Y en que la estatua existe y cuyo valor en Y es la región espacio-temporal ocupada por la estatua en Y. Con esto tendríamos que, de hecho, la «estatua» y el «pedazo de arcilla» son «objetos» distintos (son construcciones lógicas diferentes) incluso si ocupan la misma región espacio-temporal en el mundo real, puesto que hay otros mundos posibles en que no ocupan el mismo ámbito espacio-temporal.» (p.34).

h) «El materialista metafísico contemporáneo piensa en la referencia del siguiente modo: el cerebro es una computadora. Sus cálculos hacen intervenir *representaciones*. Algunas, quizás todas, son «proposicionales»: se asemejan a las oraciones en una *lingua mentis* interna. Algunas podrían ser oraciones en un lenguaje público como ocurre cuando desarrollamos un monólogo interior. Una persona se refiere a algo cuando... piensa «el gato está sobre la estera» y la situación global organismo-con entorno es de tal tipo que las palabras «el gato» en ese análogo de oración concreto (el de la *lingua mentis*) están en una relación física R (la relación de referencia) con un gato, y las palabras «la estera» están en la relación R con una estera determinada.» (p. 38).

i) «...hay un modo en que *nadie*, por lo que yo sé, ha intentado definir R, y es dando una lista de todas las posibles situaciones de referencia... (para un conjunto de lenguajes, incluyendo la *lingua mentis* del modo que sigue:

X se refiere a Y si sólo si X es una palabra (emitida) o un análogo de palabra e Y es un objeto o suceso y la situación global (incluyendo al organismo que produce X y el entorno que contiene a Y es S_1 o S_2 o S_3 o... (una lista infinita -posiblemente infinita no enumerable- de situaciones descriptas en términos físicos).» (p. 38).

j) «...David Lewis (1974) propone tratar la referencia como una propiedad funcional de la situación -organismo con entorno... Tener un programa concreto... es una propiedad funcional (en jerga de informáticos, una propiedad del «software») frente a una propiedad física corriente, de primer orden (una propiedad del «hardware»).... La sugerencia de Lewis es que la *referencia* es una de esas series de propiedades: no se trata de propiedades funcionales del organismo, sino de propiedades funcionales del sistema organismo-entorno. Si se pudiera demostrar esto, respondería a la cuestión de qué «tienen en común» las distintas situaciones en que algo se refiere a algo;... algo abstracto como un programa, un esquema o pauta formal de relaciones causa-efecto.» (pp.41-42).

k) «Hay dos caminos tradicionales para intentar vencer las evidentes dificultades que presenta una teoría de la verdad como correspondería. Una de ellas postula una facultad mental especial, una *intellektuell Anschauung*, que hace posible que la mente tenga acceso a las «formas»... entonces no se plantea el problema de cómo pueda ponerlas en correspondencia con sus «signos». El otro camino es postular una estructura dada en el mundo, un conjunto de esencias, y decir (lo que ciertamente es un oscuro decir) que esta misma estructura selecciona una correspondencia entre signos y sus objetos.» (p. 44).

l) «El planteamiento al que he dedicado este artículo afirma que hay una realidad «trascendental» en el sentido Kantiano, absolutamente independiente de nuestras mentes, que el ideal regulativo del conocimiento es copiarla o poner nuestros pensamientos en «correspondencia» con ella, pero (y esto es lo que la hace metafísica «natural») no necesitamos *intellektuelle Anschauung* para hacerlo: el «método científico hará el trabajo por nosotros. «Metafísico dentro de los límites de la sola ciencia» podría ser su slogan.» (p.46).

1.4. Argumentos contra el Realismo Metafísico Materialista.

1.4.1. Argumentos semánticos

1.4.1.1. Contra la Teoría «Causal» de la referencia, en general

1.4.1.1.1. Reason Truth and History. (1981). (Trad. Cast. 1988).

a) «La réplica que está hoy de moda entre los externalistas es que, pese a que en realidad ningún signo corresponde *necesariamente* a ningún conjunto de cosas más bien que a otro, las conexiones *contextuales* entre los signos y las cosas externas (en particular, las conexiones causales) harán explicable la naturaleza de la referencia. Pero así no se resuelve el problema. Por ejemplo, es probable que la causa predominante de mis creencias sobre los electrones sean diversos *manuales*. Pero mis preferencias de la palabra «electron», pese a tener en este

sentido una firme conexión con los manuales, no se *refieren* a estos. Los objetos que son la causa predominante de mis creencias contenientes de cierto signo pueden no ser los referentes de ese signo.» (p. 51; p. 61).

b) «La idea de que es necesaria una conexión causal es refutada por el hecho de que «extraterrestre» se refiere sin duda a extraterrestres, hayamos interactuado causalmente con extraterrestres o no.

... el filósofo externalista replicaría que podemos referirnos a extraterrestres pese a no haber interactuado con ninguno de ellos (que sepamos) gracias a que hemos actuado con *terrestres*... Y podemos *definir* extraterrestre como ser inteligente que no es del mismo planeta que los terrestres.... El externalista renuncia de este modo al requisito de que tengamos alguna conexión real (por ejemplo, una conexión causal) con *todas las cosas* a las que nos podamos referir, exigiendo sólo que los términos *básicos* se refieran a géneros de cosas (y relaciones) con los que tenemos alguna conexión real. El externalista afirma que utilizando los términos básicos en combinaciones complejas podemos construir expresiones descriptivas que se refieran a géneros de cosas con los que no tenemos ninguna conexión real y que pudieran incluso no existir.» (p. 52; p. 62).

c) «...[el externalista] podría haberse dado cuenta de que la extensión de palabras tan simples como «conejo» o «caballo» incluyen muchas cosas con las que no hemos interactuado causalmente (por ejemplo, conejos o caballos *futuros*, o conejos y caballos que nunca interactuaron con seres humanos). Cuando usamos la palabra «caballo» no sólo nos referimos a los caballos con los que tenemos conexión real, sino también a todas las demás cosas del *mismo tipo*.» (p. 52; p. 62).

d) «... dificultad de la teoría de... «la cadena causal del tipo apropiado»,... Sí afirmo meramente: «La palabra ‘caballo’ se refiere a los objetivos que tiene la propiedad cuya ocurrencia provoca en mí, en ciertas ocasiones, la preferencia ‘Hay un caballo dentro de mí’», entonces la dificultad radica en que hay demasiadas propiedades de esta índole. Por ejemplo, sea A - C («Aparición de caballo») la propiedad de todas las situaciones perceptivas que provocan la respuesta «Hay un caballo delante de mí» (aún cuando esté experimentando una ilusión); pero ‘caballo’ no se refiere a situaciones con esa propiedad, sino más bien a ciertos animales. La presencia de un animal con la propiedad de pertenecer a un género natural particular y la presencia de una situación perceptiva con la propiedad A - C, están *ambas* conectadas con mi preferencia «Hay un caballo delante de mí» mediante cadenas causales. De hecho, la presencia de caballos en la Edad de Piedra está conectada con mi preferencia «Hay un caballo delante de mí» mediante una cadena causal. Y... hay *demasiadas* cadenas causales como para que la referencia sea una cuestión de cadenas causales». (pp. 65-66; pp. 74-75).

1.4.1.2. Contra la circularidad implícita en ciertas justificaciones naturalizadas de la referencia

1.4.1.2.1. El argumento del «ascenso semántico»

1.4.1.2.1.1. «Models and Reality» (1980) En *Realism and Reason. Phil. Papers. Vol. III.* (1983).

a) «El problema es que añadir a nuestro hipotético lenguaje formalizado de la ciencia un cuerpo de teoría denominada «Teoría Causal de la Referencia» es precisamente añadir más *teoría*. Pero el argumento de Skolem, y nuestras extensiones de él, no quedan afectados

al implementar la teoría. Sin duda, usted puede aún considerar que la teoría consiste en *todas las oraciones verdaderas*, y habrá muchos modelos...; que satisfagan a la teoría entera. Si «refiere» puede ser definida en términos de algún o algunos predicados causales en el metalenguaje de nuestra teoría, entonces, puesto que cada modelo del Lenguaje-Objeto se extiende de una manera obvia al modelo correspondiente en el metalenguaje, sucederá que, *en cada modelo M*, referencia_M se define en términos de causa_M; pero, a menos de que la palabra 'causa' (o cualesquiera que puedan ser el predicado o los predicados causales) se encuentre ya pegada con pegamento metafísico a una relación determinada, no queda en absoluto fijada una determinada extensión para 'refiere'.» (p.18).

1.4.1.2.1.2. *Reason, Truth and History* (1981) (Trad. Cast. 1988).

a) «Lo importante aquí es que el hecho de que tengamos la posibilidad de *construir una máquina que examine cosas y diga si son gatos*, sólo diferencia a los gatos de los gatos* si podemos estar seguros de que 'examine' y 'diga' se refieren a examinar y decir, y no es más fácil indicar cómo se fija la referencia de estas palabras que indicar cómo se fija la de 'gato'.» (p. 37; p. 48).

b) «Si la referencia sólo está determinada por los constreñimientos operacionales y teóricos, entonces la referencia de '*x mantiene R* (fiscalistamente) *con y*' está ella misma indeterminada, y, por ende, saber que (1) [*x refiere a y* si y sólo si *x mantiene R con y*] será de poca ayuda. Cada modelo admisible de nuestro lenguaje-objeto tendrá su correspondiente modelo en nuestro metalenguaje, en el cual (1) se cumple. La interpretación de *x mantiene R con y* fijará la interpretación de '*x se refiere a y*'. Pero ésta será sólo una relación en *cada modelo admisible*; de ninguna manera servirá para acotar el número de interpretaciones admisibles.» (p. 45; p. 56).

c) «... considerando la opinión de que (1) [*'x se refiere a y* si y sólo si *mantiene R con y*'], si la entendemos como Field quiere que lo hagamos (como si describiese la relación única y determinada entre las palabras y sus referentes), es verdadera. Si... (1) es verdadera, ¿qué la *hace* verdadera? Dado que hay muchas 'correspondencias' entre palabras y cosas, e incluso muchas correspondencias que satisfacen nuestras constricciones, ¿qué es lo que *selecciona* a una particular correspondencia R? No es la corrección empírica de (1), porque esto es asunto de nuestras constricciones operacionales y teóricas. No son... nuestras intenciones (más bien, R toma parte en la determinación de lo que nuestras intenciones significan). Parece que el hecho de que R *sea* la referencia debe ser un hecho *metafísicamente inexplicable*, un tipo de verdad metafísica, primitiva, irracional.» (p. 46; pp.56-57).

d) «... suponiendo un mundo de entidades independientes de la mente y del discurso... hay, como hemos visto, muchas 'correspondencias' diferentes que representan candidatas a relaciones de referencia o posibles relaciones de referencia (de hecho, infinitamente muchas, si hay una cantidad de cosas infinita en el universo). Ni siquiera la exigencia de que (1) [*'x se refiere a y* si y sólo si *x mantiene R con y*'] sea verdadera, bajo la noción de verdad que corresponda a la relación 'real' de referencia metafísicamente seleccionada, logra excluir ninguna de estas candidatas, si, como hemos visto, (1) es ella misma *empíricamente* aceptable (aceptable dados nuestros constreñimientos operacionales y teóricos). Más entonces hay una cantidad infinita de posibles «verdades metafísicas

irracionales» *diferentes*, de la forma «R es la relación de referencia *real* (metafísicamente seleccionada)». Si el que sostiene tal punto de vista concede que podría no ser totalmente *determinada* (la relación metafísicamente seleccionada R puede permitir una pluralidad de interpretaciones admisibles), entonces hasta puede ser concebible que el punto de vista de los constreñimientos operacionales-plus-teóricos sea, después de todo, metafísicamente correcto. Pues, ¿por qué no podría ser un hecho metafísicamente irracional el que la referencia sea la relación: *x* se refiere a *y*, *al menos en un modelo admisible M*? Démonos cuenta de que todas estas infinitamente muchas teorías metafísicas son compatibles con las *mismas* oraciones verdaderas, con la misma «teoría del mundo», y con la misma metodología óptima para descubrir lo que es verdadero.» (pp. 47-48 ; p. 58).

e) «... argumento... dirigido contra la versión «fiscalista» o naturalista del realismo metafísico: supongamos que la interpretación standard I (bajo la cual «gato» se refiere a gatos, «cereza» a cerezas) es idéntica o coexistiva con la relación fiscalista R. R se da entre las ocurrencias de «gato» (o los eventos físicos que se dan en alguien que esté usándolas adecuadamente) y los gatos, etc. La interpretación no-standar J, también descrita, se da coextensiva con cierta relación R' definible en términos de R y de los mundos posibles y permutaciones empleadas al construir J. De forma que R' se da entre las ocurrencias de «gato» (o los eventos físicos que se dan en alguien que utiliza éstas en la forma standar) y las cerezas; etc. Tanto R como R' son «correspondencias»: bajo ambas son «verdaderas» las mismas oraciones. Las acciones que se exigen para la R'-verdad de una oración (es decir, las acciones que tendrán éxito desde el punto de vista de la gente) son las mismas que se exigen para que la oración sea R-verdadera. Si R es «idéntica a la referencia», si R y R', y todas las demás relaciones que asignan extensiones a nuestras palabras en conformidad con aquellos modos que satisfacen nuestros constreñimientos operacionales y teóricos, no son igualmente correctos, y no lo son porque una de ellas -R- es precisamente *la* referencia, entonces este mismo hecho es *inexplicable* bajo una óptica fiscalista.» (pp. 143-144; p. 147).

1.4.1.2.1.3. *Beyond Historicism (1981). En Philosophical Papers. Vol. III. (1983)*

a) «... hay infinitamente muchos modelos admisibles de nuestro lenguaje, i. e., infinitamente muchos modelos que satisfacen todos los constreñimientos operacionales y teóricos. Si se piensa que las entidades de que constan estos modelos, son entidades independientes de la mente y del discurso, entonces la afirmación de que uno de estos modelos es precisamente el único modelo «pretendido» deviene totalmente misteriosa. Cada uno de estos modelos corresponde a una relación de referencia. Así, hay infinitamente muchas relaciones de referencias admisibles, R₁, R₂, R₃,... Alguien que crea que precisamente una de ellas, digamos R₁₇, es realmente la única relación de referencia *real*, la relación de referencia, cree que la palabra «referencia» esta ligada a R₁₇ (y no a R₁, R₂, R₃,...) con *pagamento* metafísico.» (p. 295).

b) «Critizando este argumento Devitt replica que la «verdadera» relación de referencia, R₁₇, es lo que él denomina «conexión causal». Y afirma que la palabra «referencia» se encuentra ligada a R₁₇, no mediante *pagamento* metafísico, sino mediante conexión causal.

Esto es una violación flagrante de la teoría de los tipos. De hecho, no hay relación alguna, que contenga en su extensión un par ordenado uno de cuyos miembros sea esa misma relación. (i.e., que sea tal que algo tenga esa relación con la propia relación)» (pp.295-296).

c) «Pero aún... si sucede que la palabra «referencia» se conecta con la relación R_{17} , mediante la misma R_{17} , también sucederá que la palabra «referencia» se conecta con la relación R_1 mediante la misma R_{17} , con la relación R_2 mediante la misma R_{17} ... Michael Devitt ahora replicará que la palabra «referencia» se conecta con R_1 mediante la misma R_{17} , pero que la palabra «referencia» no se conecta a R_1 mediante R_{17} (no se conecta a R_1 mediante «conexión causal»). Desde el punto de vista de Michael Devitt, lo que importa no es el hecho formal de que R_{17} se encuentre dentro de su propio dominio converso; más bien [se trata] de este hecho formal *plus* el hecho de que R_{17} sea realmente referencia (sea realmente «conexión causal»). Es la creencia en este «hecho» adicional: la, por así decirlo, naturaleza auto-identificante de la conexión entre R_{17} y el nombre «referencia» (o el nombre «conexión causal») lo que yo describía como equivalente a la afirmación de que R_{17} está conectada a la palabra «referencia» (o a la frase «conexión causal») mediante pegamento metafísico. No se trata de que no existan diversas conexiones naturalistas entre la palabra «referencia» y R_{17} ; lo que resulta completamente ininteligible es la idea de que una de ellas declare ella misma tener el honor de hacer que R_{17} sea la relación de referencia independientemente de todos los constreñimientos operacionales y teóricos.» (p.296).

d) «... si alguien dice que la referencia es R_{17} , entonces, el significado de éste enunciado no puede ser explicado informativamente diciendo que la palabra «referencia» refiere a R_{17} ; pues si esta teoría misma es verdadera, entonces el hecho de que la palabra «referencia» refiera a R_{17} es precisamente el hecho de que la palabra «referencia» tenga R_{17} con R_{17} , y lo que no ha sido en absoluto explicado es qué significa decir que el hecho de que la palabra «referencia» tenga R_{17} con R_{17} es el mismo hecho que el hecho de que la palabra «referencia» refiera a R_{17} » (p. 296).

1.4.1.2.1.4. ¿Por qué no hay un mundo prefabricado? (1982) «Why isn't a Ready-Made World?» (1982) *En Phil. Papers. Vol. III.* (1983) (*En Racionalidad y Metafísica. Trad. Cast. 1985*).

a) «Si el materialista no puede definir la referencia, puede, naturalmente, tomarla como noción primitiva. Pero la referencia, como la causalidad, es una noción flexible, relativa a los intereses: el que consideramos que algo se refiere a algo depende del conocimiento de fondo y de nuestra disposición a ser caritativos en la interpretación. Resulta absurdo interpretar una relación tan profundamente humana y tan evasivamente intencional como algo en el mundo y llamar satisfactoria a la teoría metafísica resultante (independientemente de ser materialista o no).» (pp. 43-44).

1.4.1.2.1.5. *Introduction to Realism and Reason. Philosophical Papers, vol. III.* (1983).

a) «...hay un conjunto de cosas -llamémosle el conjunto de los gatos - tal que cuando un hombre usa la palabra «gato», la palabra está por este conjunto (desde una perspectiva de ojo de Dios), y [hay] un conjunto diferente de cosas -llamémosle el conjunto de los gatos* -, tal que cuando una mujer usa la palabra «gato», al palabra está por este conjunto (desde

la perspectiva de ojo de Dios); hay una relación entre eventos -llamémosla la relación de causación- tal que, cuando un hombre usa la palabra «causa», la palabra está por esa relación (en una perspectiva de ojo de Dios), y [hay] una relación diferente, tal que cuando una mujer usa la palabra «causa», la palabra está por *esa* relación (desde una perspectiva de ojo de Dios); y así sucesivamente.» (p.X).

b) «... un filósofo puede decir, «Cuando el niño llega a `asociar` la palabra *mesa* con ciertas percepciones, imágenes, etc., la `asocia` no en el sentido semántico (*mesa* no refiere a las impresiones visuales que desencadenan la preferencia *hay una mesa en frente de mí*), sino en un sentido causal. La causa de que [el niño] albergue ciertas creencias, es parcialmente el hecho de que se producen ciertas impresiones visuales. Pero estas impresiones visuales, «representaciones mentales», o lo que sea, son a su vez causadas por ciertos eventos externos. Normalmente, son causadas por la presencia, de hecho, de una mesa. Así, indirectamente, la palabra *mesa*, llega a asociarse con las mesas externas». (pp. X -XI).

c) «Para ver por qué esta respuesta no es solución al problema [de la fijación de la referencia], imaginemos que esto lo está diciendo, en primer lugar una mujer filósofo, y, a continuación un hombre filósofo. Cuando la mujer filósofo dice esto, está señalando que la creencia del niño *Hay una mesa en frente de mí* está en una cierta relación -la relación de efecto*-con ciertas impresiones visuales y que éstas están en la relación *efecto** con ciertos eventos externos. De hecho, son causadas* por la presencia* de una mesa*. Así, indirectamente, la palabra *mesa* ha llegado a asociarse con mesas*. Cuando un hombre filósofo dice esto, está señalando que las mismas impresiones visuales son causadas por la presencia de una mesa. Así, la palabra *mesa* ha llegado a asociarse con mesas. Por supuesto, ambos tienen razón. La palabra *mesa* está «indirectamente asociada» con mesas* (del modo indicado por la mujer), y también «indirectamente asociada» con mesas (en el modo indicado por el hombre). No se sigue que haya una cosa tal como la correspondencia (la Única la metafísicamente seleccionada correspondencia) entre palabras y cosas.» (p. XI).

1.4.1.2.2. El Argumento de la «Petitio Principii».

1.4.1.2.2.1. Reason Truth and History (1981) (Trad. Cast. 1988).

a) «...¿por qué no decir que son precisamente nuestras *intenciones*, implícitas o explícitas, las que fijan la referencia de nuestros términos?» (p.42; p. 52).

b) «El problema es que las nociones de «intención» y «estado mental» poseen cierta ambigüedad. Llamemos *puro* a un estado mental si su presencia o ausencia depende sólo de lo que acaece dentro del hablante. El que yo tenga o no un dolor sólo depende de lo que acaece «dentro de mí», pero que sepa o no que la nieve es blanca depende no sólo de que algo acaezca o no «dentro» de mí (creer o confiar en que la nieve es blanca), sino que también depende de que la nieve sea o no blanca, y, de este modo, de algo que acaece «fuera» de mi cuerpo y de mi mente. Así pues, el *dolor* es un estado mental puro, pero el *conocimiento* es un estado mental *impuro*...» (p. 42; p.53).

c) «¿Y la creencia? Hemos definido la creencia *puesta entre paréntesis* («mundo nociónal») de forma que tener una creencia puesta entre paréntesis de que (hay agua en la mesa), o tener un mundo nociónal que incluya el haber agua sobre la mesa, es un estado mental puro. Pero... *el creer que hay agua en la mesa* (sin poner nada entre «paréntesis»

presupone que la palabra «agua» se *refiere realmente al agua*, y esto depende de la naturaleza real de ciertos «paradigmas», de las relaciones causales directas que se mantengan con estos paradigmas, etc.... En resumen, creer que hay agua en la mesa es un estado mental *impuro*.» (p.42; p. 53).

d) «A la intención le ocurre lo mismo que a la creencia. Los estados mentales intencionales puros -por ejemplo, el acto intencional de que el término «agua» se refiera al agua en *el mundo nocional* de uno - no fijan la referencia en el mundo real. Los estados mentales intencionales impuros -el acto intencional de que el término «agua» se refiera al agua real- *presuponen* la capacidad de referirse al agua (real).» (p. 43; p. 53).

e) «...*creer presupone la capacidad de referir*. Y exactamente del mismo modo, los actos intencionales ya presuponen la capacidad de referir. Las intenciones no son acontecimientos mentales que *causan* la referencia de las palabras; las intenciones (en el sentido ordinario e impuro) contienen la referencia como un *componente* integral. Explicar la referencia en términos de intención (impura) es circular. Y el problema de cómo los estados mentales intencionales *puros*, de creencia, etc., puedan (en el marco causal apropiado) constituir o causar la referencia es precisamente lo que hemos hallado demasiado enigmático.» (p.43; p.54).

f) «Incluso si hay una relación fisicalista determinada R (sea o no definible en el lenguaje de la ciencia con un número finito de palabras) que *sea* precisamente la referencia (independientemente de cómo la describamos), *este* mismo hecho no puede ser consecuencia de nuestras intenciones referenciales; más bien, tal y como hemos apuntado repetidas veces toma parte en la determinación de lo que nuestras mismas intenciones referenciales significan. La opinión de Kripke según la cual «agua es H₂O» es verdadera en todos los mundos posibles, podría ser acertada incluso si la referencia en el mundo real se fijase sólo por costreñimientos operacionales y teóricos; su punto de vista presupone la noción de referencia, no nos dice si la referencia está determinada o qué es la referencia.» (p.46; p.58).

g) «El externalista replicará ahora que la palabra «electrón» no está conectada con los manuales mediante una cadena causal del tipo *apropiado*. (Pero, ¿cómo podemos tener intenciones que determinen qué cadenas causales son del «tipo apropiado» salvo que ya seamos capaces de *referirnos* a las cosas?. (p.51; p.61).

h) «Por otra parte, si digo «la palabra `caballo` se refiere a aquellos objetos que tienen una propiedad conectada con mi emisión, en ciertas ocasiones, de la proferencia «Hay un caballo delante de mi» mediante una *cadena causal del tipo apropiado*», entonces tropiezo con la dificultad de que, si soy capaz de especificar cuál es el tipo de cadena causal, debo ser capaz de referirme de *antemano* a las clases de cosas y propiedades que constituyen este tipo de cadena causal. Pero, ¿cómo obtuve tal capacidad? » (p.66; p.75).

1.4.2. Argumentos ontológicos

1.4.2.1. Contra la Naturalización de la Referencia.

Why there isn't a Ready-Made World (1982). En Phil. Papers. Vol. III. (1983) (Trad. Cast. en Racionalidad y Metafísica. (1985).

a) «El metafísico materialista a menudo toma las relaciones causales como ejemplo de estructura dada. Los hechos tienen causas; los objetos tienen «poderes causales». Y orgullosamente proclama su realismo sobre estas cosas... en el sentido del realista metafísico. Bien,... pero mi pregunta ahora no es si ese tipo de realismo está justificado, sino si es realmente compatible con el materialismo. ¿Es la *casualidad* una relación física?» (p. 211; p. 20).

b) «... si «A causa B» simplemente quisiera decir: «siempre que ocurre un hecho del tipo A, entonces un hecho del tipo B le sigue en el tiempo», entonces 'causa' sería físicamente definible. Se han hecho muchos intentos para definir la causalidad de este modo buscando una definición que se aplicara a leyes auténticamente causales y no a secuencias que consideramos causales o no causales por alguna otra razón. Muy pocos filósofos creen hoy en día que esto sea posible.» (p. 212; p.22).

c) «Pero supongamos, sin embargo, que «causa» (en este sentido) es algo físicamente definible. Una causa... es condición *suficiente* para su efecto; siempre que se da una causa, el efecto *ha de* seguirse de ella (al menos en un mundo determinístico). Siguiendo a Mill, llamemos a una causa de ese tipo *causa total*... Mill señaló que, en el lenguaje ordinario, la palabra «causa» raramente (si alguna vez lo hace) significa «causa total».... La puntualización de Mill es que nosotros consideramos ciertas partes de la causa total como «telón de fondo», mientras que nos referimos sólo a aquella parte que nos interesa como *la causa*» (pp.212-213; pp.22-23).

d) «...Cuando se dice que la presencia de ciertas partículas de H₂O en nuestro medio ambiente «causa» que usemos la palabra agua como lo hacemos, ciertamente no se quiere decir que la presencia de H₂O sea la «causa total». En su sentido ordinario «causa» se puede parafrasear a menudo por una locución que tenga que ver con *explicación*; la presencia de H₂O en nuestro medio ambiente, nuestra dependencia del H₂O para la vida, etc., son «parte de» la explicación del hecho de que tengamos una palabra que usamos del modo como la palabra agua.» (p.213; p.23).

e) «Cuando se dice que una palabra se refiere a x sólo en el caso de que (el uso de) la palabra esté relacionada con x por «una cadena causal del tipo apropiado», la noción de cadena causal implicada es, más bien, la noción de cadena *explicativa*. Incluso si la noción de «causa total» fuera físicamente definible, no sería posible *usarla* ni en la vida cotidiana ni en la filosofía; la noción que el materialista realmente usa cuando emplea «cadena causal», etc., en sus explicaciones filosóficas, es la noción intuitiva de explicación.» (p.213; p.24).

f) «Pero desde luego, esta noción [*explicación*] no es definible físicamente. Para comprobar que no lo es, obsérvese primero que «explicar» (y «causado» cuando tiene la fuerza de «explica por qué x ocurrió») son nociones abstractas. Incluso cuando imaginamos un mundo en el que hay cosas o propiedades no físicas, podemos concebirlas como causantes de otras cosas que ocurren.» (p.213; p.24).

g) «Una definición de «causado» (en este sentido `explicativo`) que fuera demasiado de «primer orden», demasiado ligada a las magnitudes concretas que son las «magnitudes fundamentales» de la física en nuestro mundo, haría *conceptualmente imposible* que un espíritu inmaterial (o cualquier hecho que implique magnitudes que no son «físicas en

nuestro mundo) pudiera ser una causa. Por esto, la anterior definición humeneana de *causa total* - A es la causa (total) de B si y sólo si un hecho del tipo A es seguido siempre en el tiempo por un hecho del tipo B - no contiene ningún término físico *específico* (excepto «tiempo»): esta definición es lo suficientemente abstracta como para poder aplicarla a mundos posibles diferentes del nuestro (aunque incluso así sea un intento fallido). ¿Podría haber una definición igualmente abstracta (y más afortunada) de «causa» en el sentido explicativo?» (p.214; pp. 24-25).

h) «...lo que para un hombre... es «condición de fondo», puede ser «causa» para otro. Qué es «causa» o «explicación» y qué no lo es depende del conocimiento previo y de nuestros motivos para hacer la pregunta. Ninguna relación puramente *formal* entre sucesos será sensible a esta relatividad de los argumentos explicativos con respecto al conocimiento de fondo y a los intereses.» (p. 214; p. 25).

i) «Richard Boyd (1980) ha sugerido que el intento de *definir* la causalidad era erróneo en su totalidad: los fisicalistas deberían tomar una noción como primitiva. Con ello puede querer decir... que... una definición de «causa» (o cualquier otra cosa) en el formalismo habitual de la matemática o la física (que contiene *nombres* sólo para numerosos reales enumerables, etc.) no resulta razonable. Si es así, ello no sería un argumento contra... que toda propiedad y relación *físicas* sean definibles en una *extensión infinitaria* de la física, en un lenguaje que admita nombres y oraciones infinitamente largas. (De hecho, si una relación o propiedad *no* es físicamente definible ni siquiera en este sentido amplio, ¿qué queremos decir al llamarla «física»? Pero podría haber querido decir que «causa» ha de tomarse como una noción literalmente irreductible, cuya indefinibilidad física no es debida a accidentes sintácticos como serían los límites en la longitud de las fórmulas. ¿puede, sin embargo, considerarse materialista un filósofo que acepta la existencia de un fenómeno de *causalidad* irreductible?» (p.215; pp. 26-27).

j) «Cuando Boyd,..., dice que una cierta microestructura es un «poder causal» ... no quiere decir que la microestructura en cuestión sea la *causa total* de los sucesos explicados ... Los «poderes causales» son propiedades que *explican* algo, dadas unas condiciones de fondo y ciertos cánones de importancia y revelancia...

... sea como sea, importancia y revelancia son atributos del pensamiento y el razonamiento, no de la naturaleza. Proyectarlos en «el mundo real» del realista, en lo que Kant llamó el mundo *nouménico*, es mezclar el Idealismo Objetivo (o quizás el aristotelismo medieval) con el materialismo de una forma totalmente incoherente.» (p. 215; pp. 27-28).

k) «Si ciertos sucesos explican *intrínsecamente* a otros, si hay importancias, relevancias, cánones sobre lo que son las condiciones «normales», etc., dados en el mundo mismo independientemente de la mente, entonces el mundo es, en muchos aspectos, *como* una mente, o está investido de algo muy similar a la razón. Y si *eso* es verdad, el materialismo *no puede* ser verdad. Se puede intentar revivir el proyecto de la metafísica especulativa,..., pero no se debería hacer pasar *este* tipo de metafísica por *física* (futura).» (pp. 215-216; p. 28).

l) «David Lewis ... ha propuesto analizar «causa» usando este tipo de condicional contrafáctico. La idea es que «A causa B» puede analizarse en términos de «si A *no* hubiera ocurrido, B *no* habría sucedido».» (p.217; p.30).

m) «Lewis ... da *condiciones de verdad* para los condicionales contrafácticos. Postula que existen de hecho «otros mundos posibles» (como en la ciencia-ficción) y que hay una «métrica de semejanza» que determina en qué medida están «cerca» o son «semejantes» dos mundos posibles cualesquiera. Un condicional contrafáctico, «Si X hubiera sucedido, entonces Y habría tenido lugar», es verdadero exactamente cuando Y es, *de hecho* verdad en todos los «mundos paralelos» *más cercanos* al mundo en que X es efectivamente verdadero.» (p. 218; pp. 31-32).

n) «...una teoría que precisa de una ontología de mundos paralelos y de una «métrica de semejanza» desde luego no tiene una ontología *materialista*. Lo que es más importante, no tiene una ontología *coherente*: no sólo la idea de la existencia real de mundos paralelos es un tanto fantástica, sino que la idea de una métrica de semejanza «intrínseca»... sensible a lo que consideramos condiciones relevantes o ... normales, que pondera qué tipo de rasgos *nosotros* consideramos semejanzas o diferencias entre estados de cosas, es una idea que implica de nuevo que el mundo es como una mente, o que está imbuido de algo muy similar a la razón. Y si *esto* es verdad, ha de tener una *explicación* (apropiadamente metafísica). Es difícil que el idealismo sea ni siquiera *un poquitín* verdadero. («Todo es física, salvo que hay esta métrica de semejanza» simplemente carece de sentido). (p. 218; p.32).

1.4.2.2. *Contra el Esencialismo*

1.4.2.2.1. *Reason Truth and History* (1981). (Vers. Cast. 1988).

a) «... «del mismo tipo que» es una expresión que no tiene sentido si no es desde un sistema categorial que señale qué propiedades no cuentan como semejanzas. Después de todo, cualquier cosa es del mismo tipo que cualquier otra de *varias* maneras. Todo este alambicado relato acerca de cómo nos referimos a algunas cosas en virtud del hecho de que estén conectadas con nosotros por cadenas causales del tipo apropiado, y aún a otras en virtud del hecho de que «son del mismo tipo que» las cosas conectadas con nosotros por cadenas causales del tipo apropiado, e incluso por «descripción», no es tan falso como ocioso. Lo que hace que los caballos con los que no he interactuado sean del «mismo tipo» que aquellos con los que sí lo he hecho, es que tanto los primeros como los últimos son *caballos*. La formulación del problema por parte del realista metafísico lo disfraza una vez más, como si hubiera que empezar con todos esos objetos en sí mismos, adquiriendo entonces un tipo peculiar de lazo con algunos pocos (los caballos con los que tengo conexión real, vía cadena causal del tipo apropiado), topándome entonces con el problema de conseguir que mi palabra («caballo») cubra no sólo aquellos objetos que he «enlazado», sino también aquellos otros que no puedo enlazar, bien porque están demasiado lejos en el espacio-tiempo, bien por cualquier otra causa. Y la «solución» del realista metafísico- es decir que la palabra cubre *automáticamente* no sólo los objetos que he enlazado, sino también los objetos que son *del mismo tipo* -del mismo tipo *en sí mismos*. Pero entonces, lo que se afirma, después de todo, es que el mundo consta de Objetos que se Auto-identifican, pues afirmar que es el mundo, y no los sujetos pensantes, el que clasifica las cosas en tipos, significa precisamente esto. » (p. 53; pp. 62-63).

1.4.2.2.2 «*Why there isn't a Ready - Made World* (1982). En *Phil. Papers. Vol. III.* (1983). (Trad. Cast. en *Racionalidad y Metafísica.* (1985).

a) «los empiristas tardíos ... negaron que hubiera en absoluto lugar para la noción de esencia. He aquí una manera típica de argumentar en este sentido: «Supongamos un trozo de arcilla con el que hemos modelado una estatua ... el trozo de arcilla y la estatua son *una* cosa, no dos, ... solamente tiene sentido hablar de una «propiedad esencial» de algo *relativamente a una descripción*. Relativamente a la descripción «esa estatua» una forma determinada es una propiedad esencial del objeto; relativamente a la descripción «ese trozo de arcilla», la forma *no* es una propiedad esencial. La pregunta: «¿cuáles son las propiedades esenciales de una cosa en sí misma?» es una cuestión sin sentido.» (p. 206; pp. 10-11).

b) «La negación de las esencias es también una negociación de la estructura intrínseca: un electrón en mi cuerpo tiene una cierta carga eléctrica, pero desde el punto de vista descripto, es un error pensar que el tener esa carga sea una propiedad intrínseca del objeto (excepto de forma relativa a la descripción «electron») en una forma en que no lo es el ser una parte de mi cuerpo. En resumen, se piensa normalmente (o se pensaba hasta hace poco tiempo) que una cosa no está «intrínsecamente» relacionada con ninguna de sus propiedades (o relaciones) más de lo que está con cualquier de sus otras propiedades o relaciones.» (p. 206; p. 11).

c) Según Kripke, la estatua y el trozo de arcilla son dos objetos, no uno. El hecho de que el trozo de arcilla tenga una propiedad modal, a saber la de «ser una cosa que podría haber sido esférica» que no tiene la estatua ... muestra ya que los objetos no pueden ser idénticos, desde el punto de vista de Kripke. Ahora bien, esto suena muy raro la primera vez que se oye. ¿si pongo la estatua en una balanza, he puesto en ella *dos objetos* ? Si el trozo de arcilla pesa diez quilos y la estatua otros diez, ¿por qué marca la balanza diez en vez de veinte, si los dos objetos están en ella?» (pp. 218-219; p. 34).

d) Al igual que puedo aceptar que no soy el mismo objeto que el agregado de moléculas de mi cuerpo sin que con ello niegue que ahora mismo *consisto* en esas moléculas (la diferencia entre objetos está en las diferentes afirmaciones que son verdaderas de cada uno, no en su distinción física), se puede estar de acuerdo con Kripke en que la estatua no es el mismo objeto que el trozo de arcilla sin negar por ello que la arcilla sea la materia de la estatua. De nuevo, la diferencia entre objetos reside en las diferentes afirmaciones verdaderas que puedan hacerse sobre ellos y no en su distinción física.» (p. 219; p. 34).

e) «Pero ahora empieza a aparecer como si los objetos, individuados propiamente, *tuvieran* esencias, *tuvieran algunas propiedades* en una forma especial, ¿Las ideas de Kripke pueden servir de ayuda al materialista metafísico? (Como es bien sabido, Kripke personalmente es notablemente contrario al materialismo).» (p.219; p. 34).

f) «Un mundo materialista cuya ontología incluya «mundos posibles» puede introducir objetos intensionales apropiados identificandolos con funciones de mundos posibles en regiones espacio-temporales... Con esto tendríamos que, de hecho, la «estatua» y el «trozo de arcilla» son «objetos distintos (... construcciones lógicas diferentes) incluso si ocupan la misma región espacio-temporal en el mundo real, puesto que hay otros mundos posibles en que no ocupan el mismo ámbito espacio-temporal.» (pp. 219-220; p. 34).

g) «Las funciones de este tipo [de mundos posibles en regiones espacio-temporales] se usan habitualmente en la semántica moderna para representar *conceptos*. Nadie dudaría de que el *concepto* «esa estatua» es un *concepto* distinto del *concepto* «ese pedazo de arcilla»; la cuestión es la de si hay algún *individuo* en el mundo real al que se le aplique *esencialmente* uno de estos conceptos y los otros sólo accidentalmente. La región espacio-temporal misma no es ese individuo, y es difícil ver cómo podría encontrar un materialista un individuo así en su ontología.» (p. 220; pp. 34-35).

h) «Sin duda se pueden conseguir tantos objetos como se quiera apoyándose en los mundos posibles y la teoría de conjuntos (de hecho...se obtienen *demasiados*). Consideremos la afirmación metafísica de que los pensamientos están en algún tipo de conexión intrínseca con los objetos externos. Si los sucesos que tienen lugar en mi cerebro están en una región espacio-temporal que tiene una conexión conjuntista con alguna entidad abstracta que implica ciertos objetos externos, entonces esa misma región espacio-temporal tendrá similares conexiones conjuntistas con otras entidades abstractas que implican otros objetos externos. Desde luego, el materialista puede decir que los «pensamientos» implican *intrínsecamente* ciertos objetos externos *identificándolos* con una determinada entidad abstracta y no con otras. Pero si se supone que esta identificación es un rasgo de la realidad misma, ha de haber entonces *esencias* en el mundo en un sentido que no cabe esperar explicar por pura teoría de conjuntos.» (p. 220; p. 35).

i) «La dificultad estriba en que Kripke individúa los objetos por sus *propiedades modales*, por medio de lo que (esencialmente) *podrían y no podrían ser*. La ontología de Kripke presupone el esencialismo: no puede usarse para fundarlo. Y las propiedades modales no son, a primera vista, elementos del mobiliario materialista del mundo.» (p. 220; p.36).

j) «...yo mismo he hablado de «propiedades esenciales» en otros lugares... Dije que...no hay mundo posible en el que agua no sea H_2O . Una vez que hemos descubierto lo que el agua es en el mundo real, hemos descubierto su *naturaleza*: ¿no es esto esencialismo?».

Es un tipo de esencialismo, pero no de la especie que podría ayudar al materialista, porque lo que he dicho es que nuestra *intención* ha sido desde antiguo que un líquido sólo se *consideraría* «agua» si tiene la misma composición que los ejemplos paradigmáticos de agua (o que la mayoría de ellos)... incluso antes de que *conociéramos* definitivamente la composición que los ejemplos paradigmáticos de agua (o que la mayoría de ellos)... incluso antes de que *conociéramos* definitivamente la composición del agua. Si estoy en lo cierto en este punto, *dadas esas intenciones referenciales*, siempre ha sido imposible que un líquido fuera agua si no era H_2O , incluso si iniciásemos investigaciones empíricas para encontrarlo. Así... la «esencia» del agua en *este* sentido es producto de nuestro uso de la palabra, del tipo de intenciones referenciales que tenemos: este tipo de esencia «no está dada en el mundo» del modo que lo exige *una teoría esencialista de la referencia* para poder despegar del suelo.» (pp. 220-221; p. 36).

k) «Concluyo que sin importar cómo se tomen las teorías de Kripke (o las mías), ya se interpreten metafísicamente, como teorías acerca de «esencias» objetivas que en alguna forma *están ahí fuera*, ya se interpreten como teorías acerca de nuestras prácticas e intenciones referenciales, no sirven de ayuda al materialista. Interpretadas metafísicamente son bastante realistas, pero su realismo no es de tipo materialista; en interpretación

semántica *presuponen* la noción de referencia y no pueden usarse como apoyo de una explicación metafísica de la referencia como correspondencia intrínseca entre pensamientos y cosas.» (p.221; p.37).

1.4.2.2.3. *Introduction to Realism and Reason. Phil. Papers (1983).*

a) «... la idea de que lo «no-psicológico» fija la referencia -i. e., que la *naturaleza* misma determina aquello sobre lo que versan nuestras palabras- es totalmente ininteligible. Para empezar, pensar que una relación signíca *está dada en la naturaleza* es retroceder al esencialismo medieval, a la idea de que ahí fuera hay «objetos y especies que se auto-identifican». Una idea tal tenía sentido en el contexto de la concepción medieval del mundo, que poseía no sólo una ontología muy elaborada (esencia y existencia, forma sustancial, etc.), sino también una psicología muy elaborada (p.e., la distinción del Aquinate entre la «imaginación pasiva» y la «imaginación productiva», lo mismo que sus «fantasmada», «especies inteligibles», etc.), y una muy elaborada correspondencia entre las dos (Dios lo ha dispuesto de tal manera que la «especie inteligible» producida por la imaginación productiva actuando sobre los fantasmata tendría una correspondencia pre-establecida con las formas sustanciales). En el contexto de una concepción del mundo propia del siglo veinte, por contraste, decir en el tono más intimidatorio de voz de uno «yo creo que las conexiones causales determinan a qué corresponden nuestras palabras» es solamente decir que uno cree en un *no se sabe qué* que resuelven nuestro problema *no se sabe cómo*.» (p. XII).

2. EN DEFENSA DEL REALISMO INTERNO

2.1. Caracterización del Realismo Interno

2.1.1. *Meaning and Moral Sciences. (1978)*

a) «Según un cierto modo de concebirlo, el Realismo es una teoría empírica... (que) explica el hecho de que las teorías científicas tienden a la 'convergencia', en el sentido de que las teorías anteriores son, muchas veces, casos límites de las posteriores (por lo que es posible considerar que los términos teóricos conservan su referencia a través de muchos cambios de teoría)... [También] explica el hecho de que el uso del lenguaje contribuye a alcanzar nuestros objetivos, obtener satisfacción, o lo que sea.» (p.123).

b) «La explicación realista (en este sentido)... [sostiene] que no es que el lenguaje refleje el mundo, sino que los hablantes reflejan el mundo, i.e., su entorno, *construyendo una representación simbólica del mundo*... una 'correspondencia' entre palabras y conjuntos de cosas (formalmente, una relación de 'satisfacción', en el sentido de Tarski) que pueda ser vista como un *modelo explicativo* de la conducta colectiva del hablante... Me referiré al realismo en este sentido... como realismo interno.» (p.123).

2.1.2. *Quantum Mechanics and the Observer. (1981). En Phil. Papers, vol III (1983).*

a) «... como apunta Dummett, los reduccionistas sólo renuncian a la teoría de verdad como correspondencia respecto a aquellos enunciados que tratan de reducir; para los

enunciados pertenecientes a la clase *reductora* retienen típicamente la tesis de que: (1) verdad y justificación son independientes; (2) la verdad es determinada y bivalente; (3) hay, de todos modos, en el límite ideal, sólo una descripción verdadera y completa. En suma, *el reduccionismo es una forma de idealismo subjetivo* (cuando la clase reductora es la de los enunciados sobre datos sensibles); mientras que el 'verificacionismo', o el 'no-realismo', defendido por Dummett y yo mismo, no deniega la realidad de cualquier objeto del discurso ordinario o científico, no construye alguno de estos objetos con una construcción a partir de otros, sino que consiste más bien, en una renuncia a los tres supuestos (ya citados) sobre la propia verdad. Si ahí aquí una especie de idealismo, es un *idealismo trascendental* y no un idealismo subjetivo.» (p.226).

2.1.3. Reason, Truth and History. (1981) (Trad. Cast. 1988)

a) «... lo característico de tal concepción [la perspectiva «internalista»] es sostener que sólo tiene sentido formular la pregunta ¿de qué objetos consta el mundo? desde dentro de una teoría o descripción... » (p.49; p.59)

b) «No existe un punto de vista como el del Ojo Divino que podamos imaginar o conocer con provecho. Sólo existen diversos puntos de vista de personas reales que reflejan aquéllos propósitos e intereses a los que subordinan sus descripciones y teorías.» (p.50; p. 59).

c) «Desde una perspectiva «internalista», los signos tampoco corresponden intrínsecamente a objetos con independencia de quién o cómo los emplee. Pero un signo empleado de un modo determinado por una comunidad de usuarios puede corresponder a determinados objetos dentro del *esquema conceptual de esos usuarios*. Los «objetos» no existen independientemente de los esquemas conceptuales. Desmenuzamos el mundo en objetos cuando introducimos uno u otro esquema descriptivo, y, puesto que tanto los objetos como los símbolos son internos al esquema descriptivo, es posible indicar cómo se emparejan.» (p. 52; p. 61).

d) «Yo diría que el mundo sí consiste en Objetos que se Auto-Identifican, en un sentido, -pero en un sentido no asequible al externalista-... Si... los propios objetos son tanto contruidos como descubiertos, son tanto producto de nuestra invención conceptual como del factor «objetivo» de la experiencia, el factor independiente de nuestra voluntad, entonces los objetos pertenecen intrínsecamente a ciertas etiquetas; porque esas etiquetas son las herramientas que usamos para construir una visión del mundo en que tales objetos ocupan un lugar preeminente. Pero este tipo de objetos que se Auto-Identifican no es independiente de la Mente; y lo que el externalista quiere es concebir el mundo mismo como si consistiese de objetos que son independientes al mismo tiempo que se Auto-Identifican. Y esto es lo que no se puede hacer.» (p. 53; p.63).

e) «El internalismo no niega que haya *inputs* experienciales en el conocimiento; el conocimiento no es un relato que no tenga otra constricción que la coherencia *interna*; lo que niega es que existan *inputs que no estén configurados en alguna medida por nuestros conceptos*, por el vocabulario que utilizamos para dar cuenta de ellos y describirlos, o *inputs que admitan una sola descripción, independiente de toda opción conceptual.*» (p. 54; p.64).

f) «... nuestras concepciones de coherencia y aceptabilidad están profundamente entrelazadas en nuestra psicología. Depende de nuestra biología y de nuestra cultura, y no

están en absoluto «libres de valores». Pero son **nuestras** concepciones y lo son de algo real. Definen un tipo de *objetividad para nosotros*, si bien ésta no es la objetividad metafísica del punto de vista del Ojo Divino. Objetividad y racionalidad humanas es lo que tenemos.» (p. 55; p.64).

g) «Sugiero que la mejor lectura de Kant consiste en considerarlo como el primer autor que propuso lo que he denominado la perspectiva «internalista» o «realista» con respecto a la verdad, pese a que Kant en ningún momento afirme explícitamente que lo este haciendo.» (p.60; p.69).

h) «[Para Kant] todo lo que decimos sobre un objeto tiene la forma: éste es tal como nos afecta de tal o cual modo. Nada de lo que afirmamos acerca de un objeto describe el objeto tal y como es «en sí mismo», independientemente de su afecto sobre *nosotros*, sobre seres con nuestra naturaleza racional y con nuestra constitución biológica... No podemos suponer ningún tipo de similitud entre nuestra idea de un objeto y cualquiera que sea la realidad independiente-de-la-mente responsable de nuestra experiencia del objeto.» (p.61; p.70).

i) «Kant no duda que hay *alguna* realidad independiente de la mente; para él, es virtualmente un postulado de la razón (la existencia de) cosas en sí mismas, objetos nouménicos o *noumenos*; colectivamente, el *mundo nouménico*. Pero no podemos formarnos una concepción real de tales cosas nouménicas: la noción de mundo nouménico es más bien un tipo de límite del pensamiento que un concepto claro... (... quizás Kant esté en lo cierto, quizás no podamos dejar de pensar que hay, *de algún modo*, una «base» independiente de la mente para nuestra existencia, aún cuando las tentativas de hablar de ella conduzcan de inmediato al sinsentido) (pp. 61-62; pp. 70-1).

2.1.4. «Why there ins't a Ready-Made World (1882). En *Phil. Papers. vol. III* (1983). (Trad. Cast. en *Racionalidad y Metafísica*. (1985))

a) «Una dirección... sería una especie de pragmatismo... el realismo «interno». Un realismo que *reconoce* una diferencia entre «p» y «yo pienso que p», entre tener razón y pensar meramente que se tiene razón, sin colocar esta objetividad ni en una correspondencia trascendente ni en un mero contexto.» (p.45)

2.1.5. «Why Reason can't be Naturalized?» (1982). En *Phil. Papers. vol III* (19873) (Trad. Cast. en *Racionalidad y Metafísica*. (1985).

a) «No tenemos una noción de la «existencia» de las cosas ni de la «verdad» de las afirmaciones que sea independiente de nuestras interpretaciones y de los procedimientos prácticos que dan sentido al hablar de «existencia y verdad» en el campo de esas interpretaciones. ¿«Existen» los *campos* como cosas físicamente reales? Sí, los campos existen relativamente a un esquema para describir y explicar fenómenos físicos; relativamente a otro esquema, hay partículas, partículas virtuales, partículas «fantasma», partículas... ¿Es verdad que existen objetos *marrones* ? Sí, en relación a la versión del mundo del sentido común: aunque no se pueda dar una condición necesaria y suficiente... para considerar marrón a un objeto, en forma de una fórmula finita y cerrada del lenguaje de la física. ¿Existen las disposiciones? Sí, en nuestra forma habitual de hablar... Tenemos muchos modos legítimos de hablar aunque diferentes e irreducibles entre sí, y en cada uno de ellos hay afirmaciones de «existencia» verdaderas» (pp. 230-231; pp. 53-54).

2.1.6. *Representation and Reality*. (1988)

a) «Mi punto de vista -que he llamado «realismo interno» en *Razón, Verdad e Historia* (hubiera sido mejor llamarlo simplemente «realismo pragmático»)... es que la misma situación puede ser descrita de muchos modos diferentes, en dependencia de cómo usemos las palabras. La situación no determina por sí misma el *modo* de usar palabras como «objeto», «entidad», y «existe». Lo erróneo en la noción de objeto existiendo «independientemente» de esquemas conceptuales es que no hay usos canónicos aparte de las elecciones conceptuales, ni siquiera de las nociones lógicas.» (p.114).

b) «La esencia del «realismo interno» es que la verdad no trasciende el uso. Enunciados diferentes -incluso, en algunos casos, enunciados incompatibles entre sí según los canones de la Lógica y Semántica clásicas- pueden ser verdaderos de la misma situación, si las palabras -a veces, las propias palabras lógicas- son usadas de modos diferentes.» (pp.115-116).

c) «Desde mi perspectiva del «realismo interno»,...no hay tal totalidad como *Todos los Objetos que Hay*, más acá o más allá de la ciencia. «Objeto» tiene en sí mismo varios usos, y, así como inventamos creativamente nuevos usos de palabras, nos encontramos con que podemos hablar de «objetos» que no eran «valores de variable» alguna en cualquier lenguaje previamente hablado por nosotros. Lo que pudiera parecer una formulación inocente del problema -«Aquí están los objetos de referencia. Aquí están los hablantes usando palabras. ¿Cómo podemos describir la relación entre los hablantes y los objetos?»- deja de ser inocente cuando lo que se busca no es un «procesador de lenguaje natural» que funcione en algún contexto restringido, sino una «Teoría de la Referencia». Desde el punto de vista internamente realista, el problema real carece de sentido.» (p.120).

2.2. La verdad como justificación racional idealizada vs. la verdad como correspondencia trascendente y vs. la verdad como justificación concreta inmanente.

2.2.1. *Vagueness and alternative Logic*. (1980). *En Phil. Papers, vol. III* (1983).

a) «En mi opinión, la verdad es justificación idealizada (la verdad es lo que sería justificado en condiciones óptimas, donde las concepciones dependen en particular de la aserción (realizada), del contexto, y de los intereses, de maneras complejas)... Una perspectiva que hace depender la verdad de lo que pueda ser justificado y que permite que no todas las verdades sean traducibles a algún determinado lenguaje ideal, no es realismo metafísico, pero preserva una intuición central del realismo: que la verdad es una propiedad sustantiva de las aserciones, una propiedad que depende de hechos que, en general, trascienden la memoria y experiencias presentes del hablante.» (p.280).

b) «Sobre lo que muchos filósofos han llegado a estar de acuerdo, a pesar de varios desacuerdos en sus posiciones, es que: (1) La verdad es una propiedad sustancial de las aserciones, una propiedad que depende a la vez de lo que significa la aserción y del mundo, y no simplemente una noción sincategoremática que funciona para elevar el nivel del lenguaje (i.e., las explicaciones puramente «desentrecorolladoras» son inadecuadas; (2) la verdad no puede ser simplemente justificación (relativa a la memoria y a la experiencia presente del hablante; (3) la concepción de la verdad como «copia» es incoherente. Si ello

es así, hay entonces un problema filosófico significativo (no resuelto por el trabajo de Tarski, por más que sea valioso como contribución a la lógica técnica) conectado con la noción de verdad.» (p. 281).

2.2.2. *Reference and Truth (1981). En Phil. Papers, vol III. (1983)*

a) «Yo trato la verdad como *idealización* de la justificación. La verdad no puede ser simplemente justificación... Se supone que la verdad es una propiedad de un enunciado que (éste) no puede perder, mientras que la justificación puede perderse (de hecho, la justificación es a la vez temporal y relativa a una persona); la justificación es una cuestión de grado, mientras que la verdad no lo es (o no lo es del mismo modo).» (p.84).

b) «Un enunciado es verdadero, en mi opinión, si puede ser justificado bajo condiciones epistémicas ideales para muchos tipos de enunciados,... algo similar a los «planos sin rozamiento» en Física; no podemos alcanzar (esas condiciones) respecto a muchas clases de enunciados, y si alguna vez podemos, no podemos tener la certeza de haberlas alcanzado, más allá de la posibilidad teórica de que nuestra mente funcionase de modo diferente a como lo hace en la actualidad. Pero tampoco podemos alcanzar realmente los planos sin rozamiento, y, sin embargo, hablar de planos sin rozamiento tiene un «valor en efectivo», en la medida en que podemos aproximarnos a ellos hasta un alto grado de precisión. Del mismo modo, podemos aproximarnos a las condiciones epistémicas ideales respecto de muchos tipos de enunciados, hasta un alto grado y con un alto grado de certeza, y esto es lo que confiere un «valor en efectivo» al discurso acerca de lo que podría ser una justificación en tales condiciones.» (pp.84-85).

c) «... los dos puntos claves de una teoría idealizada de la verdad son: (1) que la verdad es independiente de la justificación aquí y ahora, pero no independiente de toda posibilidad de justificación. Pretender que un enunciado es verdadero es pretender que puede ser justificado; (2) que se espera que la verdad sea estable, o «convergente»; si un enunciado o su negociación pueden ser indistintamente justificados, siendo las condiciones tan ideales como cabría esperar, no tendría sentido pensar que este enunciado *tiene* valor de verdad.» (p.85).

d) «Los realistas tradicionales han pretendido que la verdad trasciende incluso la justificabilidad idealizada (... pudiéramos estar «realmente» engañados por un malvado demonio, pudiéramos ser realmente cerebros en una cubeta, etc.). Yo repliqué a estos argumentos que una teoría epistémicamente ideal tendría que poseer necesariamente modelos (p. e., gracias al teorema de Löwenheim-Skolem) y, de hecho, modelos que satisfagan todas las constricciones operacionales y teóricas (siendo así modelos «pretendidos»). Concluí que el realismo metafísico -la opinión de que la verdad rebasa incluso la justificación idealizada- es incoherente. Por otro lado, identificar la verdad con justificación temporal, en tanto opuesta a la justificación idealizada, es sacrificar el principio de que alguno de los *enunciados* que están *ahora* justificados podría llegar a ser no-verdadero; y considero que este principio es una parte central de nuestra visión empírica del mundo. De este modo, he recuperado la distinción kantiana entre realismo metafísico y realismo empírico, y rechazo el primero, mientras que acepto el segundo (realismo «interno»)» (p. 85).

e) «Estoy de acuerdo con Dummett en rechazar la teoría de la verdad como correspondencia. Pero no estoy de acuerdo con la opinión de Dummett que las condiciones de justificación de las oraciones cambian en la medida en que lo hace nuestro cuerpo total de conocimientos, y no podemos considerarlas como fijadas de una vez por todas». (p. 85).

f) «No sólo podemos averiguar que los enunciados considerados ahora como justificados son falsos, sino que incluso podemos averiguar que (determinados) procedimientos tenidos ahora por justificatorios no son tales, y que hay otros procedimientos de justificación que son mejores... Así como la naturaleza objetiva del entorno contribuye a fijar la referencia de los términos, también contribuyen a fijar las condiciones de verdad objetivas de las oraciones, si bien no al modo realista metafísico.» (pp.85-86)

2.2.3. *Reason Truth and History* (1981) (Trad. Cast. 1988)

a) «Desde una perspectiva internalista, la «verdad» es una especie de aceptabilidad racional (idealizada) -una especie de coherencia ideal de nuestras creencias entre sí y con nuestras experiencias, *considerándolas como experiencias representadas en nuestro sistema de creencias*- y no una correspondencia con «estados de cosas» independientes de la mente o del discurso.» (pp.49-50; p.59).

b) «... las dos ideas clave de la teoría de la verdad-idealización son las siguientes: (1) la verdad es independiente de la justificación aquí y ahora, no independiente de *toda* justificación. Afirmar que un enunciado es verdadero es afirmar que podría ser justificado. (2) Es de esperar que la verdad sea estable o «convergente»; si tanto un enunciado como su negación pueden ser «justificados» no tiene sentido pensar que tal enunciado *posee* un valor de verdad, por mucho que las condiciones fueran tan ideales como uno soñase alcanzar.» (p. 56; p.65).

c) «En realidad, Kant no *afirma* que está abandonando la teoría de la verdad como correspondencia... afirma que la verdad es «la correspondencia de un juicio con un objeto», (pero no) lo que un realista metafísico quiere decir... En opinión de Kant, cualquier juicio acerca de objetos externos o internos afirma que el mundo nouménico *como un todo*, es de tal modo que ésta es la descripción que construiría un ser racional (... con nuestra naturaleza racional), dada la información disponible de que dispone un ser, con nuestros órganos sensoriales (... con nuestra naturaleza sensible)... En *este* sentido, el juicio adscribe una facultad ('power')... al mundo nouménico *como un todo*... No hay una *correspondencia uno-a-uno entre las cosas para nosotros y las cosas en sí mismas*. Kant, no sólo renuncia a la idea de similitud entre nuestras ideas y las cosas en sí mismas; renuncia incluso a la idea de un isomorfismo abstracto. Y esto significa que, en su filosofía, no hay una teoría de la verdad como correspondencia.» (p. 63; p.72).

d) «Kant cree que tenemos conocimiento *objetivo*... El uso del término «conocimiento» y «objetivo»... comporta... que, *a pesar de todo*, hay una noción de verdad...; no... la correspondencia con la forma en que las cosas son en sí mismas... (sino que)..., un fragmento de conocimiento (un «enunciado verdadero») es un enunciado que aceptaría un ser racional a partir de una cantidad suficiente de experiencia de la clase que los seres con nuestra naturaleza pueden obtener efectivamente... la Verdad es bondad última de ajuste.» (p.64; pp.72-73).

e) «Un no-realista «interno» concibe los enunciados condicionales como enunciados que entendemos, en buena medida, aprehendiendo sus condiciones de *justificación* (como entendemos los demás enunciados). Esto no significa que el realista «interno» abandone la distinción entre verdad y justificación, sino que aprehendemos la verdad (*justificación idealizada*) tal y como aprehendemos cualquier otro concepto; por medio de una comprensión (en gran medida implícita) de los factores que hacen racionalmente aceptable decir que algo es verdadero.» (pp.112-113; p.127).

f) «...la posición no-realista... asume una noción *objetiva* de aceptabilidad racional. El no-realista rechaza la opinión de que la verdad es una correspondencia con un mundo «prefabricado». Y ésto es lo que le convierte en un no-realista (metafísico). Pero rechazar la teoría de la verdad como correspondencia no es del todo lo mismo que considerar *subjetivas* tanto la verdad como la aceptabilidad.» (p.123; p. 127).

2.2.4. *Why Reason can't be Naturalized?» (1982) En Phil Papers vol III (1983)*

(Trad. Cast. en *Racionalidad y Metafísica*. (1985)

a) «La verdad, en el único sentido en el cual tenemos una noción vital y efectiva de ella, es aceptabilidad racional (o más bien, aceptabilidad racional bajo condiciones epistémicas suficientemente buenas). Qué condiciones son epistémicamente mejores o peores depende del tipo de discurso, al igual que la propia aceptabilidad racional.» (p.231; p.54).

b) «... verdad y aceptabilidad racional -el que una afirmación sea correcta y el que alguien esté en situación de hacerla- son relativas al tipo de lenguaje que nosotros estemos usando y a la clase de contexto en que estemos... Esto no significa... que una aserción sea correcta *siempre* que aquellos que utilizan el lenguaje en cuestión acepten que es correcta en este contexto. Hay que *equilibrar* dos hechos, ambos puestos de manifiesto por filósofos de tipos muy diferentes: (a) hablar de qué es «correcto» en cualquier área sólo tiene sentido en el contexto de una *tradición heredada*; pero (b) las tradiciones mismas pueden ser *criticadas*.» (p. 234; p.60).

c) «Lo que estoy diciendo es que los «cánones» aceptados por una cultura o subcultura, ya sea explícita o implícitamente, no pueden *definir* lo que es la razón, incluso dentro de un contexto, porque *presuponen* la razón (razonabilidad) en su interpretación. De un lado, no hay noción de razonabilidad, ..., *sin* culturas, prácticas, procedimientos; de otro lado, las culturas, prácticas, y procedimientos que heredamos no son un algoritmo que haya que seguir ciegamente... la razón, en este sentido, es, al mismo tiempo, immanente (no se puede encontrar fuera de los juegos lingüísticos e instituciones concretas) y trascendente (una idea regulativa para criticar la conducta de *toda* actividad e institución.» (p.234; p.60).

d) «Los filósofos que pierden de vista la immanencia de la razón... se pierden en fantasías filosóficas características. «El lenguaje ideal», «La Lógica Inductiva», «el criterio empirista de significación», ... «la idea absoluta»... Los filósofos que pierden de vista la trascendencia de la razón se convierte en relativistas culturales (o históricos). » (pp.234-235; pp 60-61).

e) «Consideremos ahora la posición del relativista cultural que dice: «cuando digo que algo es *verdad*, quiero decir que es correcto de acuerdo con las normas de *mi* cultura». Si añade: «Cuando un miembro de una cultura diferente dice que algo es verdad, lo que quiere decir (lo sepa o no) es que lo es en conformidad con las normas de *su* cultura», entonces está exactamente en la misma posición que un solipsista metodológico.» (p.

f) «En general, si R.R. [un relativista cultural americano] entiende cada emisión *p* que él usa como queriendo decir «es verdad según normas de la cultura americana que *p*», entonces debe entender sus propias afirmaciones hermenéuticas, las expresiones que usa para interpretar otras, del mismo modo, no importa con cuantos calificativos del tipo «de acuerdo con las normas de la cultura alemana», notas a pie de página, glosas, comentarios, o cualquier otra cosa, las acompañe. Las otras culturas se convierten en, por así decir, construcciones lógicas sacadas de los procedimientos y prácticas de la cultura americana. Si ahora intenta añadir «la situación es la inversa desde el punto de vista de la otra cultura», se encuentra con el mismo panorama con el que se encontraba el solipsista metodológico: la pretensión trascendental de una situación *simétrica* no puede entenderse si la doctrina relativista es correcta. Y decir, como suelen hacerlo los relativistas, que la otra cultura tiene conceptos «incommensurables» no resuelven nada. Esto es sólo la afirmación trascendental expresada en una jerga peculiar.» (pp.237-238; pp. 65-66).

g) «... no se puede reconocer a los otros en este sentido trascendental, que implica reconocer que la situación es *realmente* simétrica, si se piensa que en la realidad los otros son construcciones obtenidas a partir de los datos sensoriales propios. Tampoco es posible reconocer a los otros, en este sentido, si se piensa que la *única* noción de verdad que se puede entender es «verdad-según-lo-determinan-las-normas-de-esa-cultura»». (p. 238; p. 66).

h) «... el relativista cultural puede llegar a ser un imperialista cultural. Puede decir, «bien, entonces, la verdad -la única noción de verdad que entiendo- viene definida por las normas de *mi* cultura»... Tal punto de vista ya no es relativista. Postula una noción *objetiva* de verdad, aunque esta noción se entienda como un producto de nuestra cultura, que ha de ser definida por nuestros criterios culturales (supongo que el imperialista cultural es uno de *nosotros*). En este sentido, así como el solipsismo consistente llega a ser indistinguible del realismo. (como dijo Wittgenstein en el *Tractatus*), el relativismo cultural es un *tipo* especial de realismo.

Es realismo en cuanto que acepta una diferencia objetiva entre lo que es verdad y lo que meramente se piensa como verdad (otra cuestión es la de si es capaz de *explicar* consistentemente esa diferencia).

No se trata de realismo *metafísico* o trascendental puesto que la verdad no puede ir más allá de la asertabilidad correcta, al contrario de lo que afirmaba el realismo metafísico. Pero la noción de asertabilidad correcta se fija por «criterios», en un sentido positivista: algo es afirmable correctamente sólo si las normas de la cultura especifican que lo es; estas normas operan al modo de una *definición operacional* de asertabilidad correcta.» (pp. 238-239; pp. 66-67).

i) «Este punto de vista es, sin embargo, auto-refutable, al menos en nuestra cultura...; la argumentación se apoya en el hecho de que nuestra cultura, a diferencia de culturas totalitarias o teocráticas, no tiene «normas» que decidan las *cuestiones filosóficas* (algunos filósofos han pensado que las tienen; pero han tenido que postular una «gramática profunda» accesible sólo a *ellos*, y no descriptible para la investigación lingüística o antropológica ordinaria). Así, la afirmación filosófica:

Una afirmación es verdadera (se puede afirmar correctamente) sólo si se puede afirmar según las normas de la cultura moderna europea y americana,

no es ni afirmable ni refutable en la medida en que exige el asentimiento de todos los que no se han desviado de las normas de la cultura moderna europea y americana. Así, si esta afirmación es verdadera, se sigue que no es verdadera (no se puede afirmar correctamente). Por lo tanto, no es verdadera, Q.E.D. (Creo que *todas* las teorías que identifican verdad o asertabilidad correcta con aquello con lo que la gente esta de acuerdo, o con lo que estaría de acuerdo a la larga, o con lo que la gente educada e inteligente estaría de acuerdo, o...estaría de acuerdo a la larga, son contingentemente refutables en el mismo modo.» (p. 239; pp. 67-68).

j) Es verdad que hablamos un lenguaje público, que heredamos interpretaciones, que hablar de verdad o falsedad sólo cobra sentido en el contexto de una «tradición heredada» como dice Wittgenstein; pero también es verdad que rehacemos nuestro lenguaje, que construimos nuevas versiones a partir de las antiguas, y que tenemos que usar la razón para hacerlo, y, por decirlo todo, incluso para entender o aplicar las normas que no cambiamos ni criticamos. Las definiciones consensuales de la razón no funcionan, porque el consenso entre los adultos *presupone* la razón en vez de definirla.» (p. 240; p. 69).

k) «Si se abandonan por completo las nociones de justificación, aceptabilidad racional, asertabilidad con pruebas suficientes, asertabilidad correcta, y nociones similares, entonces «verdadero» también desaparece, excepto como un mero mecanismo para el ascenso semántico, es decir, como un puro mecanismo para pasar de un lenguaje a otro. La mera introducción de un predicado de verdad tarskiano no puede definir, para un lenguaje, una noción de *corrección* que no esté previamente definida. Rechazar las nociones de justificación y asertabilidad correcta mientras se *mantiene una noción realista metafísica de verdad*, podría, por otro lado, no sólo resultar peculiar (¿qué razón habría para considerar la verdad, en el sentido de la «correspondencia», como algo *más claro* que la asertabilidad?) sino incoherente: las nociones que los metafísicos naturalistas usan para explicar la verdad y la referencia, por ejemplo, la noción de causalidad (explicación), y la noción de *tipo apropiado* de cadena causal, dependen de nociones que presuponen la noción de razonabilidad.» (pp.245-246; p.79).

l) «Pero si *toda* noción de corrección, tanto epistémica como (metafísicamente) realista se elimina, ¿qué son, entonces, nuestras afirmaciones sino meros sonidos?, ¿qué son nuestros pensamientos, sino *meras* subvocalizaciones? La eliminación de lo «normativo» es un intento de suicidio mental.

De poco sirven las nociones de «decisión que acepto» y «método que conduce a decisiones que acepto». Si el *único* tipo de corrección comprensible que tiene una afirmación es «alcanzarla por un método que ofrece decisiones que *yo* acepto», entonces quedo sometido a un solipismo del momento actual. Quedo sometido al solipismo, porque esto *es* metodológicamente, un sustituto solipista de la asertabilidad («decisiones que yo acepto»), y vimos antes que el solipismo metodológico sólo es consciente si es un solipismo real. Y al solipismo del momento actual, porque ésta es una noción temporal (un sustituto de la asertabilidad con pruebas suficientes en *un tiempo*, no de toda la asertabilidad en las mejores de las condiciones) y si la *única* clase de corrección que tienen mis subvocalizaciones actuales es mi asertabilidad *actual* (se defina como se defina); si no hay noción de decisión *límite*, aunque sea difusa, entonces mis subvocalizaciones no tratan, en ningún sentido,

sobre algo que vaya más allá del momento actual. (Incluso el pensamiento «hay un futuro» es «correcto» desde el punto de vista, sólo en el sentido de ser *afirmable del momento actual*.» (p. 246; pp. 79-80).

2.2.5. *Introduction to the Phil Papers. vol III (1983).*

a) «La fórmula «verdad es justificación» es engañosa de varios modos, razón por la que yo la he evitado en mis propios escritos, a pesar de la inspiración que he recibido a partir de la obra de M. Dummett. Y ello porque en primer lugar sugiere algo que Dummett sin duda cree, y yo no: que es posible especificar de modo efectivo las condiciones de justificación para todas las oraciones de un lenguaje natural. En segundo lugar, sugiere algo acerca de lo cual resultan ambiguos los escritos de Dummett: que hay tal cosa como justificación *conclusiva*, incluso en el caso de las oraciones empíricas.» (pp. XVI-XVII).

b) «Mi propia opinión (para la que he sugerido el nombre de «realismo interno») es que la verdad debe ser identificada con justificación en el sentido de justificación *idealizada*. A veces, ésta parece ser también la opinión de Dummett (en ocasiones habla de una «brecha entre justificación y verdad»); y otras veces, escribe como si las oraciones del lenguaje ordinario acerca de objetos materiales más allá de la ciencia teórica pudieran ser verificadas conclusivamente.» (p. XVII).

c) «Considérese la oración «Hay una silla en mi despacho en este mismo momento». Bajo condiciones epistémicas suficientemente buenas cualquier persona normal podría verificar esto, donde las condiciones epistémicas suficientemente buenas podrían, por ejemplo, consistir en que uno goce de buena visión, esté ahora mi despacho iluminado, no haber ingerido algún agente alucinógeno, etc. ¿Cómo sé que estas condiciones para este tipo de juicio son mejores que... carecer de buena visión, o que se está mirando a la habitación con un telescopio desde una gran distancia, o haber tomado LSD? Parcialmente, sabiendo cómo opera un discurso de este tipo (de qué «juego de lenguaje» se trata, en el sentido de Wittgenstein), y parcialmente, disponiendo de un montón de información empírica. No hay una regla general simple ni un método universal para saber qué condiciones son mejores o peores para justificar un juicio empírico arbitrario.» (p.XVII).

d) «Según esta mi opinión,... la «verdad» (justificación idealizada) es tan vaga, «interesada», y sensible al contexto como somos nosotros *mismos*. Las «condiciones de verdad» para una oración arbitraria no son *obtenibles* en el sentido de Dummett. Yo rechazo las «teorías del significado». (Por supuesto, es posible obtener las condiciones de verdad *superficiales* de las oraciones de un lenguaje -si el lenguaje ha sido canónicamente formalizado- p.e., «*la nieve es blanca*» es verdadera si y sólo si *la nieve es blanca*, e incluso construir una teoría de las que todas sean teoremas. Esto es lo que Donald Davidson llama una «teoría del significado» para el lenguaje. Donde difiero de Davidson es que yo no atribuyo en absoluto significación filosófica a las condiciones superficiales de verdad en ausencia de una explicación de en qué consista el *conocimiento* que de ellas tenga un hablante nativo).» (p.XVII).

e) «(Pero) nosotros aprendemos [estas condiciones] precisamente del modo como piensa Dummett, al menos en el caso de la parte menos teórica del lenguaje, adquiriendo una práctica... (más) no un conocimiento que pudiera ser aplicado como si fuera un

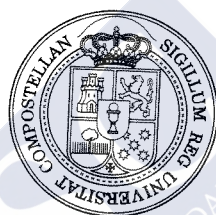
algoritmo. Aprendemos que, en ciertas circunstancias, se supone que aceptaremos (normalmente) «Hay una silla enfrente de mí». Pero se espera que utilicemos el 'coco'. Podemos rehusar aceptar «Hay una silla enfrente de mí, incluso cuando nos parece exactamente como si hubiera una silla enfrente de nosotros, si hacemos caso omiso de los gritos de advertencia de nuestra inteligencia general. La imposibilidad (al menos, en la práctica) de formalizar las condiciones de asertabilidad para oraciones arbitrarias es precisamente la imposibilidad de formalizar la inteligencia general misma.» (pp.XVII-XVIII)

f) «Qué, sin embargo, [la asertabilidad garantizada idealizada (verdad)] sea una noción revelante; que haya condiciones epistémicas mejores o peores para muchos juicios, y que sea una cuestión de hecho la de cuál sería el veredicto si las condiciones fueran suficientemente buenas, un veredicto hacia el cual convergería nuestra opinión, si fuéramos razonables, es el corazón de mi propio realismo. Se trata de un realismo que pretende ser un tipo humano de realismo, una creencia en que es una cuestión de hecho el que algo sea correctamente afirmable para nosotros, en tanto opuesto a lo que sería correctamente afirmable desde la perspectiva del Ojo Divino, tan cara al realista Metafísico clásico.» (p.XVIII).

Luis Villegas
Septiembre, 1989

OBRAS DE HILARY PUTNAM CONSULTADAS

- 1.- (1978) *Meaning and Moral Sciences*. Routledge & Kegan Paul. London
- 2.- (1980 a) «Vagueness and Alternative Logic». En *Realism and Reason*. Phil. Papers. Vol. III. (1983): 271-286.
- 3.- (1980 b) «Models and Reality». *Journal of Symbolic Logic* XLV: 464-81. Rep. en *Realism and Reason*. Phil. Papers. Vol. III. (1983): 1-25
- 4.- (1981 a) *Reason, Truth and History*. Cambridge U.P. Cambridge. (trad. cast. *Razón, verdad e Historia*. Tecnos Madrid. 1988)
- 5.- (1981 b) «Beyond Historicism». En *Realism and Reason*. Phil. Papers vol. III (1983): 287-303
- 6.- (1981 c) «Quantum Mechanics and the Observer». *ERTKENNTNIS* XVI: 193-219. Rep. en *Realism and Reason*. Phil Papers vol III(1983): 248-270)
- 7.- (1981 d) «Referenza/verità». Enciclopedia. G. Einaudi Ed. Firenze. vol. XI: 725-41. Rep. «Reference and Truth» en *Realism and Reason*. Phil. Papers. vol. III (1983): 69-86.
- 8.- (1982 a) «Why there isn't a Ready-made World?» *SYNTHESE* 51/2: 141-167. Rep. en *Realism and Reason*. Phil. Papers vol III (1983): 205-228. (trd. cast. «por qué no hay un mundo prefabricado». en *Racionalidad y Metafísica*. (1985): 9-49)
- 9.- (1982 b) «Why Reason can't be Naturalized». *SYNTHESE* 52/1: 3-23. Rep. en *Realism and Reason*. Phil. Papers. vol III (1983): 229-247. (Trad. cast. «por qué no puede naturalizarse la Razón». En *Racionalidad y Metafísica*. (1985): 51-81
- 10.- (1983) *Realism and Reason*. Philosophical papers vol. III. Cambridge U.P. Cambridge
- 11.- (1985) *Racionalidad y Metafísica*. Revista Teorema. Madrid.
- 12.- (1988) *Representation and Reality*. M.I.T. Press. Cambridge Mass.





CURSOS E CONGRESOS
DA
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA