

# La búsqueda del significado

Luis Ml. Valdés Villanueva  
(ed.)



*tecno*  
s



Secretariado de publicaciones  
e intercambio científico.  
UNIVERSIDAD DE MURCIA

Uno de los fenómenos más presentes en nuestra experiencia cotidiana es, sin duda, el lenguaje. Casi todas nuestras actividades están llenas de cosas como hablar, escuchar a alguien que habla, leer, escribir, etc. La característica central de todas ellas, lo que las hace *lenguaje*, es que se les adscribe característicamente *significado*. El significado lingüístico es entonces algo con lo que estamos muy familiarizados, quizá más que con cualquier otra cosa. Sin embargo, dista de ser un asunto fácil decir en qué consiste. ¿Qué es lo que hace que unos sonidos o unas manchas de tinta tengan un significado y no otro? ¿Qué es lo que diferencia estos eventos físicos de otros, en apariencia similares, pero que no tienen significado? ¿Cómo es posible que las palabras se refieran a objetos? ¿Cómo puede una cadena de sonidos o una ristra de manchas de tinta decir algo verdadero o falso? Éstas son algunas de las cuestiones que trata la filosofía del lenguaje, cuestiones que, por otra parte, son tan viejas como la filosofía misma.

En este volumen se recogen veintidós artículos de autores como Frege, Russell, Hempel, Tarski, Grice, Quine, Strawson, Kripke, Putnam, Austin, Searle, Davidson, Dummett, Barwise, Perry, Harman, Wilson y Sperber, que han hecho contribuciones extremadamente relevantes a la tarea filosófica central —emprendida desde finales del siglo pasado— de buscar una teoría del significado adecuada.

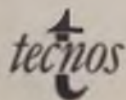
Luis Ml. Valdés Villanueva, autor de la presente selección, ha sido profesor de Lógica y Filosofía del Lenguaje en la Universidad de Valencia y actualmente es catedrático de Lógica de la Universidad de Murcia.

ISBN 84-309-1922-8



9 788430 919222

Filosofía y Ensayo



Secretariado de publicaciones  
e intercambio científico.  
UNIVERSIDAD DE MURCIA

**LA BÚSQUEDA  
DEL SIGNIFICADO**

**LECTURAS DE FILOSOFÍA  
DEL LENGUAJE**



LUIS ML. VALDÉS VILLANUEVA  
(Editor)

# LA BÚSQUEDA DEL SIGNIFICADO

LECTURAS DE FILOSOFÍA  
DEL LENGUAJE



UNIVERSIDAD  
DE MURCIA

Diseño de colección:  
Rafael Celda y Joaquín Gallego

Impresión de cubierta:  
Gráficas Molina

401  
B 876

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de Editorial Tecnos, S.A.

Introducción © LUIS ML. VALDÉS VILLANUEVA, 1991  
© UNIVERSIDAD DE MURCIA, 1991  
© EDITORIAL TECNOS, S.A., 1991  
Telémaco, 43 - 28027 Madrid  
ISBN: 84-309-1922-8  
Depósito Legal: M-1769-1991

---

*Printed in Spain.* Impreso en España por Azalso. Tracia, 17, Madrid

## ÍNDICE

INTRODUCCIÓN .....	Pág. 9
I. TEORÍAS DESCRIPTIVAS DE LA REFERENCIA .....	21
INTRODUCCIÓN .....	23
SOBRE SENTIDO Y REFERENCIA (1892), por Gottlob Frege .....	24
DESCRIPCIONES (1919), por Bertrand Russell .....	46
SOBRE EL REFERIR (1950), por Peter F. Strawson .....	57
NOMBRES PROPIOS Y DESCRIPCIONES (1967), por John R. Searle .....	83
II. TEORÍAS DE LA REFERENCIA DIRECTA .....	95
INTRODUCCIÓN .....	97
IDENTIDAD Y NECESIDAD (1971), por Saul Kripke .....	98
EL SIGNIFICADO DE «SIGNIFICADO» (1975), por Hilary Putnam .....	131
III. SIGNIFICADO Y VERIFICACIÓN .....	195
INTRODUCCIÓN .....	197
PROBLEMAS Y CAMBIOS EN EL CRITERIO EMPIRISTA DE SIGNIFICADO (1950), por Carl G. Hempel .....	199
DOS DOGMAS DEL EMPIRISMO (1953), por Willard V. O. Quine .....	220
SIGNIFICADO Y TRADUCCIÓN (1959), por Willard V. O. Quine .....	244
IV. SIGNIFICADO Y VERDAD .....	271
INTRODUCCIÓN .....	273
LA CONCEPCIÓN SEMÁNTICA DE LA VERDAD Y LOS FUNDAMENTOS DE LA SEMÁNTICA (1944), por Alfred Tarski .....	275
VERDAD Y SIGNIFICADO (1967), por Donald Davidson .....	314
SIGNIFICADO Y VERDAD (1969), por Peter F. Strawson .....	335
INTERPRETACIÓN RADICAL (1973), por Donald Davidson .....	354
¿QUÉ ES UNA TEORÍA DEL SIGNIFICADO? (1975), por Michael A.E. Dummett .....	370
V. ACTOS DE HABLA .....	411
INTRODUCCIÓN .....	413
EMISIONES REALIZATIVAS (1961), por John L. Austin .....	415
¿QUÉ ES UN ACTO DE HABLA? (1965), por John R. Searle .....	431
UNA TAXONOMÍA DE LOS ACTOS ILOCUCIONARIOS (1975), por John R. Searle .....	449
VI. SIGNIFICADO E INTENCIÓN .....	477
INTRODUCCIÓN .....	479
LAS INTENCIONES Y EL SIGNIFICADO DEL HABLANTE (1969), por Paul Grice ..	481
LÓGICA Y CONVERSACIÓN (1975), por Paul Grice .....	511

VII. ALGUNAS PROPUESTAS RECIENTES .....	531
INTRODUCCIÓN .....	533
SITUACIONES Y ACTITUDES (1981), por Jon Barwise y John Perry .....	534
SEMÁNTICA DEL ROL CONCEPTUAL (1982), por Gilbert Harman .....	561
SOBRE LA DEFINICIÓN DE RELEVANCIA (1986), por Deirdre Wilson y Dan Sperber .....	583
ORIGEN DE LOS ARTÍCULOS .....	599



## INTRODUCCIÓN

*A man does not know what he is saying  
until he knows what he is not saying.*

G. K. CHESTERTON

### ¿QUÉ ES LA FILOSOFÍA DEL LENGUAJE?

Uno de los fenómenos que más presencia tiene en nuestra experiencia cotidiana es, sin duda, el lenguaje. Casi todas nuestras actividades están llenas de cosas como hablar, escuchar a alguien que habla, leer, escribir, etc. La característica central de todos esos eventos, lo que los hace *lenguaje*, es que de todos ellos se dice que tienen *significado*. Así pues, parece que el significado lingüístico es algo con lo que nos encontramos muy a menudo, algo con lo que estamos más familiarizados que con cualquier otra cosa. Por eso resulta —a primera vista— chocante que sea tan difícil explicar en qué consiste. Consideremos la oración siguiente:

Oviedo es la capital de España;

si queremos hacernos una idea de cuáles son las dificultades que surgen aquí sólo tenemos que preguntarnos qué quiere decir que esas manchas de tinta negra tienen significado. Ciertamente estamos ante una cosa física lo mismo que la ciudad de *Hull* o la máquina de escribir que tengo sobre mi mesa. Pero cuando digo que esas manchas tienen significado ¿qué es lo quiero decir? Sin duda mediante ellas quiero decir que Oviedo es la capital de España pero ¿qué es lo que hace que tengan el significado que tienen y no otro? ¿Qué diferencia hay entre una ristra de marcas significativa y otra que no lo es? ¿Cómo soy capaz de reconocerla como tal (o de producirla y que tú la reconozcas) aunque no la haya encontrado antes? Parece también que «Oviedo» se refiere a una ciudad y «España» a un país: ¿cómo es posible que unas meras manchas se refieran a ciudades o países o, en general, a objetos? Además, casi todo el mundo estaría de acuerdo en que mediante la oración anterior hacemos un enunciado falso. ¿Cómo puede una ristra de manchas significar algo falso (o verdadero)?

Éstas son algunas de las cuestiones centrales de la filosofía del lenguaje, cuestiones que, por otra parte, son tan viejas como la pro-

pia filosofía. No es sin embargo una tarea fácil delimitar el campo que abarca: todo autor de un texto de filosofía del lenguaje procura escurrir el bulto diciendo de antemano que ésta es un conjunto más o menos cohesionado de opiniones filosóficas sobre el lenguaje que puede que tengan un cierto «aire de familia» pero que, a veces resultan ser tan dispares que no es infrecuente que se contradigan entre sí<sup>1</sup>. Esta situación no es sin embargo algo peculiar de la filosofía del lenguaje (si bien aquí se da la dificultad añadida de separar los problemas que le pertenecen de los que caen dentro del campo de la lingüística o de la ciencia cognitiva); afecta a todas las filosofías regionales y a la filosofía en general. ¿Por qué sucede esto?

En un libro relativamente reciente sobre la filosofía lingüística, L. J. Cohen<sup>2</sup> afirma que «la filosofía es inherentemente autocrítica». Esto es: la filosofía no sólo intenta resolver disputas que, de otro modo, ni siquiera se plantearían, sino que entre sus problemas fundamentales está la investigación de cuál sea la naturaleza misma de la investigación filosófica. La pregunta «¿qué es un problema filosófico?» es ella misma una cuestión filosófica y a pesar de que los filósofos han dedicado bastante tiempo y espacio a discutir este tipo de asuntos no parece que se hayan puesto de acuerdo (o se vayan a poner en un futuro próximo) sobre cuál ha de ser el enunciado correcto de la respuesta. Pero aunque no hayan sido capaces de proporcionar una definición detallada y completa de en qué consiste un problema filosófico, sorprendentemente, sí son capaces de reconocerlo como tal (más o menos, debe admitirse) cuando se encuentran con él. Y ciertamente los filósofos son capaces de reconocer un problema como perteneciente a la filosofía del lenguaje aunque no hayan acertado a trazar sus límites. Pero entonces ¿no podríamos considerar esta, llamémosla «competencia», como un modo *sui generis* de trazar esos límites, de acotar la provincia de la filosofía del lenguaje?

H. P. Grice, P. F. Strawson y J. R. Searle<sup>3</sup> han utilizado, en otro contexto, un argumento que puede aplicarse a este caso. Para ellos el rechazo quineano de la distinción analítico/sintético<sup>4</sup> no sólo va en contra de una tradición filosófica bien establecida (lo cual no sería un

<sup>1</sup> La mayor parte de los libros de texto despachan la cuestión de lo que sea la filosofía del lenguaje en unas pocas líneas triviales o vagas. Creo que esto no se debe a una falta de interés o perspicuidad, sino simplemente a que no es posible caracterizarla de una manera clara y detallada. Así, p. ej., Rosenberg y Travis en una importante antología (*Readings in the Philosophy of Language*, Prentice-Hall, New Jersey, 1961) afirman: «ni nosotros ni nadie, que sepamos, sabe cómo dar tal explicación (de lo que es la filosofía del lenguaje)».

<sup>2</sup> Cfr. L. J. Cohen, *The Dialogue of Reason*, Clarendon Press, Oxford, 1986, p. 1.

<sup>3</sup> Cfr. H. P. Grice y P. F. Strawson, «In defense of a dogma» (1964); recogido en Rosenberg y Travis (1961), *op. cit.*, en nota 1.

<sup>4</sup> Ver W. V. O. Quine, «Dos dogmas del empirismo», pp. 220-243 de este volumen.

argumento de excesivo peso), sino también en contra de nuestra práctica efectiva. Esto es: parece ser un hecho que los usuarios de un término como «analítico» están de acuerdo de manera abrumadora en la aplicación que hacen de él: lo aplican, rehúsan hacerlo y tienen dudas en más o menos *los mismos casos*. Además, ese acuerdo tiene una característica *proyectiva*: no se reduce a situaciones experimentadas previamente, sino que se extiende a casos nuevos. Ahora bien, como Grice y Strawson afirman, «si un par de expresiones opuestas se usan de manera general y habitual aplicadas a los mismos casos, *donde los casos no forman una lista cerrada*, esto es una condición suficiente para decir que hay géneros de casos a los que se aplica la expresión; y no se necesita más para señalar la distinción»<sup>5</sup>.

Esto es: si los hablantes son competentes a la hora de reconocer algo es que *debe de haber algo que reconocer*, independientemente de cuáles sean las razones que los hablantes aleguen o las historias que cuenten sobre su reconocimiento. Tomemos entonces cualquier expresión (y, en particular, la expresión «problema filosófico»); si los que la usan de manera habitual están de acuerdo en hacerlo y en no hacerlo respecto de una clase indefinidamente grande de casos, entonces puede decirse que bajo la expresión en cuestión caen una serie de casos con determinadas características y que las personas que participan de ese acuerdo general son competentes para reconocerlos.

En contra de esto podría argumentarse que todo lo anterior se reduce a decir que un problema filosófico no es más que aquello en lo que un número importante de filósofos están de acuerdo en que es un problema filosófico. Pero esto no es ciertamente lo que se está afirmando. Lo que se dice aquí es que, *al menos*, ésa es una condición que cumplen un buen número de problemas (justamente los que se denominan «problemas filosóficos») y que otros (justamente los no considerados como tales) no cumplen. Por ello —y de acuerdo con un argumento similar al anterior— podríamos afirmar que debe de existir algún conjunto de características subyacentes que determinan qué problemas interesan a los filósofos y que, por lo tanto, constituyen, con toda probabilidad, *problemas filosóficos*. No hay razón alguna para que este argumento no se aplique a la filosofía del lenguaje, por lo que vamos a suponer que existen una serie de rasgos que permiten reconocer lo que es un problema perteneciente a la provincia de la filosofía del lenguaje *aunque no seamos capaces de enunciarlos de modo efectivo*. Al fin y al cabo tampoco lo somos en el caso de otras filosofías regionales —o de la filosofía en general— sin que ello parezca representar un grave problema. Obviamente, esto no quiere

---

<sup>5</sup> H. P. Grice y P. F. Strawson, *op. cit.*, en nota 3.

decir que tanto los problemas filosóficos regionales como los generales no tengan algunas características distintivas. Lo único que se afirma es que de momento, y parece que por mucho tiempo, no vamos a estar en condiciones de proporcionar un conjunto de condiciones necesarias y suficientes que nos permita determinar que algo es un problema filosófico. Es más, puede que toda la empresa de descubrirlos carezca en absoluto de interés.

Lo que haré a continuación será desplegar mi «competencia» para intentar *decir* lo que (al menos parcialmente) *no es* filosofía del lenguaje. Ello nos permitirá, o al menos así lo espero, introducir por vía negativa algunas de esas características distintivas. Dejo a los artículos que componen este libro la tarea de *mostrar* lo que *es* filosofía del lenguaje. No obstante, una muy breve presentación antes de cada uno de los siete grupos en que están divididos intentará situarlos en el marco general de la disciplina. Como la selección que se ha hecho puede requerir que se la justifique, la última parte de esta introducción estará dedicada a ello.

#### LO QUE NO ES FILOSOFÍA DEL LENGUAJE

En primer lugar, la filosofía del lenguaje *no es lingüística*. Es cierto que durante siglos las cuestiones empíricas del análisis gramatical han estado interconectadas con las reflexiones filosóficas sobre el lenguaje. Pero la lingüística moderna hace poco menos que una cuestión de principio el haberse separado de la filosofía y se suele definir a sí misma como el estudio científico (léase empírico) del lenguaje y de los lenguajes, contrapuesto a las actividades más bien *a priori* de los filósofos. Pero este criterio de demarcación es muy vago y, a poco que se lo presione, hará agua por los cuatro costados. Si los lingüistas se interesasen sólo por cosas tales como preparar diccionarios de lenguajes particulares y describir sus características fonéticas, sintácticas, etc., o a cuestiones prácticas tales como la enseñanza de las lenguas, sería relativamente fácil distinguir entre los temas y métodos de ambas disciplinas. Pero ¿no es uno de los temas de la lingüística contemporánea explicar cuestiones como la competencia de un hablante para producir y entender oraciones nuevas? ¿No es verdad que la lingüística teórica se ocupa cada vez menos de reunir datos para centrarse en rasgos universales reconocidos como tales por los hablantes competentes? ¿Y no son éstos problemas que consideraríamos como típicos de la filosofía del lenguaje?

El problema se complica por las interconexiones que existen entre lingüística y filosofía del lenguaje. Veámos un ejemplo. Recientemente J. Katz <sup>6</sup> ha defendido la tesis de que a lo largo de este

<sup>6</sup> Cfr. J. Katz (ed.), *The Philosophy of Linguistics*, Oxford University Press, 1985.

siglo ha habido dos giros lingüísticos en filosofía. En el primero de ellos —que abarcaría hasta comienzos de los cincuenta— el lenguaje se habría convertido en tema central de la investigación filosófica al producirse la ruptura con la tradición idealista del XIX. Frege, Moore, Russell, los miembros del Círculo de Viena, Ryle y Wittgenstein en sus dos períodos, serían los campeones de tal movimiento que no tuvo en cuenta —salvo contadas excepciones<sup>7</sup>— la lingüística científica. El segundo giro sólo se produjo cuando algunos filósofos —con Quine como pionero— empezaron a tomarla en consideración. Así éste, en sus críticas a los conceptos de significado y analiticidad, utilizaba la metodología del estructuralismo americano, particularmente las tesis de Bloomfield acerca de los criterios de sustitución como piedra de toque para decidir acerca de la claridad de los conceptos lingüísticos<sup>8</sup>: la imposibilidad de usar criterios de sustitución que no fueran circulares en el caso de la sinonimia y la analiticidad demostraba que estos conceptos no eran en absoluto respetables. Naturalmente, como el propio Katz apunta, el razonamiento de Quine es válido sólo si se supone que los criterios de sustitución son *el método apropiado* de delimitar conceptos de lingüística (y eso es parte de lo que la metodología, por ejemplo, de Chomsky, pone en cuestión).

El ejemplo aducido pone de manifiesto cómo un problema tan familiar en filosofía del lenguaje como es el del significado o el de la analiticidad puede verse fuertemente afectado por —o incluso depende— posiciones metodológicas mantenidas por la lingüística científica. Pero aunque la filosofía del lenguaje esté emparentada con la lingüística y no sean infrecuentes los trasvases de una disciplina a otra no por ello hemos de concluir que una se reduce a la otra o es parte de ella. Como máximo puede decirse que hay problemas fronterizos, problemas respecto de los cuales tendríamos dudas si se nos preguntase a dónde adscribirlos. Pero la existencia de casos dudosos no es, desde luego, un buen argumento para eliminar una distinción.

La filosofía del lenguaje no es tampoco, como se ha afirmado a veces<sup>9</sup>, *filosofía de la lingüística*. Esta última es una rama de la filo-

<sup>7</sup> Aunque sólo fuese para afirmar que la gramática tradicional era desorientadora. Piénsese, por ejemplo, en la teoría de las descripciones de Russell.

<sup>8</sup> De hecho, y también lo señala Katz, Quine toma en consideración en «Dos dogmas» otras maneras *no lingüísticas* de clarificar los conceptos de analiticidad, sinonimia y significado. Él piensa correctamente que debe utilizarse todo instrumento que ofrezca la posibilidad de clarificar esas nociones, pero cree erróneamente que los criterios de sustitución son el único utensilio que proporciona la lingüística.

<sup>9</sup> A principios de la década de los sesenta Katz y Fodor publicaron un artículo titulado «What is wrong with the philosophy of language?» (*Inquiry*, vol. 5, 1962), en el

sofía de la ciencia, paralela a la filosofía de la matemática o a la filosofía de la física, mientras que la filosofía del lenguaje es un área substantiva de la investigación filosófica. Pero, como en el caso anterior, existen interconexiones, solapamientos y zonas de confluencia entre ambas. Si es cierta la tesis de Katz, anteriormente citada, de los dos giros lingüísticos, hay razones tanto *prácticas* como *teóricas* para pensar que la filosofía de la ciencia lingüística tiene cierto peso sobre cuestiones de filosofía del lenguaje. El desarrollo de la lingüística científica, una vez que ésta se convirtió en objeto de interés por parte de los filósofos del lenguaje, y la proliferación de teorías, llevo aparejado casi inmediatamente el examen de los supuestos filosóficos que incorporaban; el filósofo del lenguaje necesitaba tener criterios para elegir entre la oferta de teorías lingüísticas rivales. El filósofo de la lógica, cuando presenta las lógicas alternativas junto con las premisas filosóficas que están en su base y los problemas que generan, no haría algo excesivamente distinto.

Pero, además de servir de aviso de navegantes, la filosofía de la lingüística resulta interesante para nuestros propósitos por razones estrictamente teóricas. En la obra citada de J. Katz se señala un grupo de ellas pertenecientes a la ontología. Como es sabido, uno de los problemas más importantes de los que se ocupa la metafísica se refiere a la investigación de qué géneros de objetos hay. Si nos centramos en la lógica o en la matemática, la respuesta a la pregunta ¿qué son las proposiciones? o ¿qué son los números? difiere dependiendo de si uno es nominalista (objetos físicos), conceptualista (objetos psíquicos) o realista (objetos abstractos). Pero el mismo enfoque es posible en lingüística pues, podemos preguntarnos, ¿son las oraciones objetos físicos, psíquicos o abstractos? Esto quiere decir que la filosofía de la lingüística podría dar respuesta a la pregunta ontológica general (y, por ende, a la regional) de una manera extremadamente simple. Por ejemplo: si el realismo ha de ser verdadero debe de haber *al menos* un dominio de objetos abstractos. Por consi-

---

que intentaban poner remedio, decían ellos, a la paradójica situación en la que se encontraba la filosofía del lenguaje: a pesar de que su interés se había centrado desde finales del siglo XIX en el estudio de éste, los filósofos no habían sido capaces de formular hasta entonces «una teoría explícita del lenguaje que investigan o un enunciado explícito de los métodos de investigación». La solución consistiría para ellos en convertir la filosofía del lenguaje en un estudio de los conceptos y métodos de la lingüística empírica. Sólo cuatro años más tarde Katz (en *The Philosophy of Language*, Harper Row, Nueva York) confesaba que tanto él como Fodor estaban «completamente equivocados al interpretar la filosofía del lenguaje como una rama de la filosofía de la ciencia... La filosofía del lenguaje es un área de investigación filosófica del conocimiento conceptual, más que una de las diversas ramas de la filosofía contemporánea... Se trata de un área que busca aprehender todo lo que puede aprehenderse sobre el conocimiento conceptual de la manera que tal conocimiento se expresa y comunica en el lenguaje.»

guiente, si las oraciones (o el lenguaje) resultan ser objetos abstractos, entonces puede responderse afirmativamente a la pregunta ¿hay objetos abstractos? Nótese que la oración anterior estaba en forma condicional; entonces, hasta que no esté decidido el *status* de los objetos de la lingüística, hasta que no esté decidido si las oraciones (o el lenguaje) son objetos abstractos, la pregunta ontológica general no podría responderse <sup>10</sup>.

Finalmente, la filosofía del lenguaje no es *filosofía lingüística*. La filosofía del lenguaje, como dice uno de sus más conspicuos representantes, es el nombre de un *tema*. Ciertamente no existe un acuerdo universal sobre a qué denominar «filosofía lingüística». Hay quien piensa que tal disciplina podría caracterizarse como la aplicación de hechos y técnicas aprendidos mediante el estudio del lenguaje a los problemas filosóficos estándares <sup>11</sup>; otros que es lo mismo que «filosofía analítica» <sup>12</sup>; otros, en fin, que no se distingue de la filosofía del lenguaje <sup>13</sup>. Dada la disparidad de opiniones voy a entender por «filosofía lingüística» el nombre de un *método* para habérselas con los problemas filosóficos que, a pesar de su venerable antigüedad, tiene una historia y un desarrollo muy concretos en el pensamiento contemporáneo.

La siguiente caracterización puede servirnos de punto de partida: *Los problemas filosóficos son problemas esencialmente lingüísticos y su solución (o disolución) requiere bien una reforma del lenguaje o una elucidación más adecuada de su funcionamiento*. Entendida de esta manera la filosofía se convierte en análisis del lenguaje y su historia se retrotrae por lo menos hasta la búsqueda de definiciones por parte de Sócrates. Desde entonces ha sido moneda corriente entre los filósofos el pensar que la investigación y el análisis de los conceptos no sólo tienen utilidad para llegar a conocer cómo los hombres describen el mundo, sino también para penetrar en la realidad misma y poder sacar de ese conocimiento consecuencias prácticas. Así, por ejemplo, en *La República* de Platón la búsqueda socrática de definiciones es parte integrante de la empresa de establecer cómo deberían vivir las personas. ¿Qué es entonces lo que añade de nuevo a esto la filosofía lingüística? La novedad consiste en la suposición de que *los problemas filosóficos pueden resolverse a través del análisis puesto*

<sup>10</sup> Obsérvese que el *status* ontológico de las oraciones o del lenguaje tiene una importancia especial en nuestra disciplina. Una de las aspiraciones de algunos de sus practicantes de primera hora era resolver (o disolver) los problemas filosóficos. Piénsese, por ejemplo, en Carnap y sus predicciones acerca de la solución de muchos problemas si utilizásemos el modo formal de hablar.

<sup>11</sup> Cfr. Rosenberg y Travis, *op. cit.*, en nota 1, p. 4.

<sup>12</sup> Cfr. J. J. Acero, *Filosofía y análisis del lenguaje*, Cincel, Madrid, 1985, p. 18.

<sup>13</sup> Cfr. J. Katz, *Linguistic Philosophy*, Allen Unwin, Londres, 1971.

que se derivan o bien de malas comprensiones de nuestro lenguaje o del uso de un lenguaje que no es el adecuado. Esto es: la filosofía lingüística es un *método* de solución de los problemas filosóficos *supuesta la tesis* de que su origen está en alguna anomalía lingüística. Este método, con sus dos tradiciones principales que podemos, convencionalmente, denominar tradición positivista y tradición de la filosofía del lenguaje ordinario<sup>14</sup>, tuvo su época de esplendor durante la primera mitad de este siglo y los nombres de Frege, Russell, Wittgenstein, Carnap, Austin... están unidos a él. Ciertamente todo esto tiene un poco de caricatura (ni siquiera en los mejores tiempos se mantuvo que *todos* los problemas filosóficos fuesen problemas lingüísticos<sup>15</sup>: pero la verdad es que filosofía lingüística y filosofía del lenguaje están, al menos en sus orígenes contemporáneos, indisolublemente unidas: los filósofos antes citados hicieron filosofía del lenguaje aplicando en muchos casos técnicas de la filosofía lingüística: temas como el de la naturaleza de la verdad, la referencia, el significado, la metáfora, los actos de habla, la necesidad lógica y toda una hueste adicional fueron tratados de esta manera. Hoy día el método ha desaparecido prácticamente de la escena aunque los temas permanecen. Los artículos que componen este volumen tienen como tema la filosofía del lenguaje aunque, como el lector podrá comprobar por sí mismo, algunos de ellos utilizan los métodos de la filosofía lingüística.

Se ha mencionado que la práctica de la filosofía lingüística no es algo novedoso en la historia de la filosofía. Pero tampoco lo es la propia distinción —aunque sólo sea implícita— entre filosofía del lenguaje y filosofía lingüística que también se remonta, al menos, hasta Platón: cuando éste en el *Eutifrón* se pregunta qué es la piedad, su investigación puede considerarse como una elucidación del significado del término griego para «piadoso» (filosofía lingüística), mientras que cuando en el *Fedón* se presenta la tesis de que los términos generales adquieren su significado al estar por las Formas correspondientes, se está haciendo teoría del significado (filosofía del lenguaje). No obstante, la moderna filosofía del lenguaje tiene unas características lo suficientemente peculiares como para que se le dedique una atención especial. Ello nos permitirá al mismo tiempo proporcionar alguna justificación para la selección de artículos que componen este volumen.

<sup>14</sup> En esto sigo a Katz, *op. cit.*, en nota anterior, y *Semantic Theory*, Harper Row, Nueva York, 1962.

<sup>15</sup> Piénsese, por ejemplo, en Austin, para quien el lenguaje ordinario es sólo —aunque ya es bastante— «la primera palabra».



## LA FILOSOFÍA DEL LENGUAJE EN EL PENSAMIENTO CONTEMPORÁNEO

La mayor parte de las cuestiones que se debaten en este libro sólo pueden entenderse dentro del contexto de toda una corriente de pensamiento que comienza con la obra de G. Frege y que se vio estimulada por la situación de *impasse* a la que se había visto abocada la filosofía moderna. Creo pues indispensable hacer un breve examen de esa situación aunque antes, siguiendo la táctica de decir lo que *no es* la filosofía del lenguaje, intentaré deshacer algún equívoco.

Una de las cuestiones que invitan aquí a la confusión es el propio título de la disciplina. Muchos filósofos no avisados tienden a pensar que la gramática lógica —si se me permite hablar así— de dicho título es FILOSOFÍA+LENGUAJE y que cualquier combinación de ambos temas, cualquier libro que se considere a sí mismo como filosófico y que hable del lenguaje debe contarse *eo ipso* entre los tratados o monografías de filosofía del lenguaje. Pero no se repara que mediante este procedimiento casi todo filósofo sería filósofo del lenguaje pues es difícil de encontrar en toda la historia de la filosofía un solo autor que no haya tenido el lenguaje entre sus preocupaciones ni un solo libro de filosofía que no dedique alguna de sus páginas a algún tópico relacionado con el lenguaje. A esto podría replicarse que el criterio anterior dejaría de ser trivial si se exigiese que el filósofo o el tratado tuviesen el lenguaje como su preocupación *central*. Pero, aparte de la dificultad de determinar cuando algo es central —algo parecido a la dificultad de determinar cuando algo es central —algo parecido a la dificultad de definir «exactitud» de la que habla Wittgenstein en las *prelaciones* de lo que es filosofía del lenguaje. Equivale a decir (más o menos): «“Filosofía del Lenguaje” es “Filosofía” + “Lenguaje”, donde el último de los sumandos ocupa un lugar central.» ¿Cómo podríamos pues aproximarnos, de una manera más adecuada, a ese «tema» que constituye la filosofía del lenguaje?

Nuestra disciplina se ocupa de un conjunto de problemas que, tradicionalmente, han sido objeto de estudio —si bien no de un modo sistemático— de diversas ramas de la filosofía: la teoría del conocimiento, la metafísica, la psicología filosófica... han sido solamente algunas de ellas. Pero si queremos proporcionar un enfoque integrador de toda esta hueste de cuestiones aparentemente tan dispares, una de las posibilidades que se nos ofrece consiste en considerarlas como estrechamente ligadas a la lógica y defender su carácter preliminar y básico dentro del conjunto de la filosofía: después de todo, el nacimiento de la filosofía del lenguaje moderna está ligado al de la lógica simbólica <sup>16</sup>. G. Frege, el padre fundador de la lógica mo-

<sup>16</sup> En este punto conviene recordar también que Aristóteles, el fundador de la ló-

derna, no fue sólo el autor de la *Conceptografía*, sino que escribió también obras fundamentales sobre las nociones de función, concepto, objeto, sentido, referencia..., que son parte integrante del acervo de cuestiones que trata la filosofía del lenguaje. Su propósito original era el de introducir rigor absoluto en los métodos de prueba de la matemática y para ello consideró necesario construir un lenguaje simbólico en el que pudieran llevarse a cabo las demostraciones con total garantía de que no se deslizaran errores debido al uso incontrolado de la intuición. Ahora bien, para realizar esto comprendió que era necesario efectuar previamente un examen de la estructura formal de los enunciados que componen las demostraciones lo que, a su vez, le llevó a analizar el significado de los enunciados en términos del de sus componentes, esto es: se dio cuenta de la obligatoriedad de realizar un análisis de la *estructura interna* de éstos. Pero este trabajo se salía con mucho de los límites de la lógica formal, al exigir el uso de un combinado de técnicas matemáticas y argumentaciones típicamente filosóficas. Además la obra de Frege afectó de manera substancial a la ubicación de la filosofía del lenguaje dentro del conjunto de las disciplinas filosóficas inaugurando una línea de pensamiento que llega hasta nuestros días y sin cuya toma en consideración es imposible entender las investigaciones de las que son muestra los artículos que componen este libro.

En un libro publicado en la década de los setenta Michael Dummett<sup>17</sup> defiende la tesis de que Frege —sin especiales proclamas por su parte— llevó a cabo una revolución en filosofía similar a la que dos siglos antes había efectuado Descartes. Desde Descartes hasta finales del siglo XIX puede decirse que toda la filosofía está dominada por *problemas epistemológicos*: su preocupación central es la teoría del conocimiento y sus preguntas básicas son del tipo siguiente: ¿cuáles son las capacidades de la mente en sus esfuerzos por lograr conocimiento del mundo exterior? ¿Hasta qué punto son capaces las facultades de la mente de penetrar en la estructura de la realidad? ¿Hasta qué punto son adecuadas las ideas de la mente para representar y descifrar la naturaleza del mundo? ¿Cuáles son los límites y las capacidades de la mente para alcanzar la verdad?<sup>18</sup> Se trata de una dirección de pensamiento —que abarca desde Descartes al idealismo— cuyo tema central es la *oposición* entre sujeto y objeto, entre la

---

gica clásica, no es sólo el autor de los *Primeros y Segundos Analíticos*, sino también de las *Categorías*, del tratado *De Interpretatione* y los *Tópicos*, obras todas ellas que tienen justificados títulos para ser consideradas como de filosofía del lenguaje.

<sup>17</sup> Cfr. Michael Dummett, *Frege. Philosophy of Language*, Duckworth, Londres, 1973, especialmente el capítulo 19. También del mismo autor *The Interpretation of Frege's Philosophy*, Duckworth, Londres, 1981.

<sup>18</sup> Cfr. Milton K. Munitz, *Contemporary Analytic Philosophy*, MacMillan, Nueva York, 1981, p. 4.

*mente que conoce* y el *mundo exterior* <sup>19</sup>. Frege fue el primero en darse cuenta de que este modelo era inadecuado y que había que dejar de pensar en el problema epistemológico como parte central de la filosofía. Así, el cómo adquirimos nuestras ideas, o su propia naturaleza —algo que preocupaba enormemente a los empiristas—, es para él un asunto completamente irrelevante. No es, ciertamente, que nuestras experiencias no sean relevantes: no cabe duda de que no captaríamos los pensamientos que captamos si nuestras experiencias fueran distintas o estuviéramos constituidos de otra manera. Pero Frege tuvo el acierto de señalar que eso no puede constituir el análisis de nuestros conceptos, es más: *depende de él*. La posesión de un concepto es algo que se manifiesta mediante el uso del lenguaje, de modo que el análisis de los conceptos no es algo distinto del análisis del funcionamiento del lenguaje. Las preguntas acerca de *si* es posible que nuestra mente conozca un mundo exterior o de *cómo* es posible esto se vuelven completamente superfluas dado que ya tenemos ese conocimiento. El problema no es, por consiguiente, determinar si es posible lograr conocimiento, sino mostrar las condiciones y procedimientos para lograrlo: cómo pasamos de la ignorancia o la duda a la creencia fundada. Frege pensó que todas estas cuestiones *dependían de una correcta teoría del significado* y, consecuentemente, mantuvo que la primera y principal tarea de la filosofía era la búsqueda de una teoría del significado adecuada: la parte primera y básica de la filosofía pasaba a ser la filosofía del lenguaje <sup>20</sup>.

A partir de este cambio de perspectiva se ha desarrollado la moderna filosofía del lenguaje. Los artículos que componen este volumen son un segmento del diálogo filosófico general <sup>21</sup> que, teniendo como origen los logros de Frege, presenta, como asunto de hecho, unos nombres y temas recurrentes y, creo que puede decirse, bastante bien definidos. No se trata pues de la amalgama filosofía+lenguaje, sino de una disciplina con perfiles propios <sup>22</sup>. No es por tanto un

<sup>19</sup> Cfr. Munitz, *op. cit.*, en nota anterior, p. 5.

<sup>20</sup> Pensemos en la definición de las conectivas de la lógica proposicional mediante el procedimiento de tablas de verdad. Si dudamos acerca de si el análisis clásico es o no adecuado, nuestra duda no ha de interpretarse como siendo acerca de si hemos proporcionado la tabla de verdad correcta: lo es acerca de cuál es el modelo general de análisis adecuado; es una duda acerca de en qué consiste conocer el significado de una conectiva proposicional. Es en este sentido en el que la filosofía del lenguaje es parte primera y básica de la filosofía.

<sup>21</sup> Cfr. L. J. Cohen, *op. cit.*, en nota 2.

<sup>22</sup> Aquí podemos acudir a los libros de texto de filosofía del lenguaje, a las revistas especializadas, a los *simposia*, etc., donde se codifican de manera bastante definida los temas de nuestra disciplina, el diálogo que se mantiene y sus participantes. Obviamente no se pretende que esto sirva como *criterio* de demarcación pero sí, al menos, como *síntoma*. Es además frecuente (y no extraño, dado lo anteriormente expuesto) que

olvido, ni un prejuicio de escuela, que no aparezca texto alguno de Heidegger, Derrida o Habermas <sup>23</sup> —por citar sólo algunas ausencias que pueden resultar chocantes para algunos—, sino fruto del desarrollo de nuestra disciplina y del lugar que ocupa dentro del conjunto del saber filosófico. No se quiere decir con esto que prohibirse el rotular como «filosofía del lenguaje» todo aquello que no case con la concepción que estamos considerando. Sólo me atrevo a sugerir que deberían buscarse denominaciones distintas para asuntos distintos.

En una obra como ésta —pensada para servir como libro de lecturas durante un curso de filosofía del lenguaje— es necesario tomar decisiones sobre qué artículos incluir y excluir. Esas decisiones son a veces teóricas, pero las más de las veces vienen dictadas por constricciones económicas, editoriales, de derechos de edición, etc. En nuestro caso ha sido particularmente doloroso el tener que prescindir de algún representante de la lingüística que, como se ha afirmado, ha estado particularmente presente en el desarrollo de nuestra disciplina. Espero que esta ausencia pueda suplirse sin demasiados problemas.

Una serie de personas e instituciones han cooperado para que este libro viese la luz. Juan José Acero, Pepe Fillol, Alfonso García Suárez, Aurelio Pérez Fustegueras, Daniel Quesada y Enrique Ujaldón tradujeron algunos de los artículos y/o me aconsejaron sobre su inclusión y distribución que, por lo que respecta a sus defectos, es responsabilidad que sólo a mí me compete. Manuel Garrido apoyó el proyecto de modo entusiasta desde el principio. La Universidad de Murcia lo acogió generosamente lo mismo que Editorial Tecnos. Mis alumnos de filosofía del lenguaje de las Universidades de Valencia y Murcia, que me han soportado pacientemente durante estos últimos quince años, me hicieron ver la necesidad de disponer de un libro (más o menos) como éste. Aunque para ellos llegue ya demasiado tarde espero que a sus futuros compañeros les sea de alguna utilidad.

LUIS ML. VALDÉS VILLANUEVA

---

exista un fluido diálogo entre los que practican la filosofía del lenguaje tal como la entendemos aquí y los lingüistas. Quine y Putnam, por citar sólo un ejemplo, han hecho frente a algunas posiciones de Chomsky y éste ha hecho lo mismo con la filosofía del lenguaje de Quine, Grice o Searle. Nada similar ha ocurrido con otros estudiosos de filosofía + lenguaje (cuando han existido intercambios han sido esporádicos y periféricos) por lo que éste es uno más de los apoyos que tendría la no arbitrariedad de esta propuesta de demarcación.

<sup>23</sup> Habermas es uno de los pocos filósofos no pertenecientes a la tradición que estamos considerando que ha utilizado extensivamente nociones típicas de nuestra disciplina (por ejemplo, la de acto de habla). No obstante se ha limitado —sin mucho acierto, pienso— a utilizarlos sin entrar en diálogo alguno con los filósofos del lenguaje.

# I. TEORÍAS DESCRIPTIVAS DE LA REFERENCIA



Los términos singulares, categoría lógica a la que pertenecen, entre otros, nombres propios y descripciones, han atraído tradicionalmente el interés de los filósofos por lo que respecta a la cuestión de cómo se relacionan con el mundo. El dilema era (y es, si se me permite simplificar las cosas) éste: ¿la relación de estos términos con el mundo es directa o indirecta? Gottlob Frege, el fundador de la moderna filosofía del lenguaje, mantuvo en «Sobre sentido y referencia» —como resultado del análisis de ciertos problemas sobre la informatividad de los enunciados de identidad— que esa relación, tanto en el caso de los nombres propios como en el de las descripciones, era indirecta, se producía por medio del sentido. La célebre teoría de las descripciones de Bertrand Russell se enfrenta —aunque sólo en cierta medida— a las posiciones de Frege. Russell defiende que los nombres propios, si son nombres propios *genuinos* (esto es: nombres de particulares como «esto», usados deícticamente para hacer referencia a contenidos de conciencia del hablante, o el pronombre personal de primera persona «yo» usado por éste para referirse a sí mismo), se relacionan con el mundo de modo directo. Pero tanto las descripciones definidas (p. ej., «el autor del *Quijote*») como los nombres propios ordinarios (p. ej., «Cervantes») no tienen referencia directa, es más: estos últimos sólo son realmente modos de abreviar descripciones o, si se quiere, su referencia se logra de manera indirecta, *vía* las descripciones que abrevian. Peter Strawson saltó al estrellato filosófico cuando en 1950 publicó «Sobre el referir» —un duro ataque a las, hasta entonces, indiscutibles tesis de Russell— siguiendo las líneas trazadas por Frege; en particular a Strawson le parecía intolerable la mitología del nombre propio genuino de Russell y consideró que los nombres propios ordinarios son nombres propios genuinos y tienen referencia indirecta. Hay sin embargo un punto en el que Strawson se opone tanto a Frege como a Russell y que anticipa ciertos desarrollos posteriores (p. ej., algunas afirmaciones de la teoría de los actos de habla): estos últimos mantenían que eran los términos los que tenían referencia, mientras que Strawson defenderá la posición de que son los hablantes los que realizan el acto de la referencia por medio del uso de ciertos términos. Por último, John Searle en «Nombres propios y descripciones» presenta algo que quiere ser un compromiso entre las teorías descriptivas y las no descriptivas de la referencia. Frege, afirma él, estaba en lo cierto al mantener que los nombres propios tienen sentido, que su referencia es indirecta, pero ese sentido no es un conjunto de descripciones que singularizan el objeto al que se refiere el nombre, sino un racimo de éstas constitutivamente vago e indeterminado: ésta es justamente la diferencia entre describir y nombrar.

## SOBRE SENTIDO Y REFERENCIA \*

GOTTLOB FREGE

La igualdad <sup>1</sup> induce a la reflexión a través de preguntas relacionadas con ella y que no son fáciles de contestar. ¿Es la igualdad una relación?, ¿es una relación entre objetos?, ¿o bien entre nombres o signos de objetos? Esto último es lo que supuse en mi ideografía. Las razones que parecen hablar en favor de ello son las siguientes:  $a = a$  y  $a = b$  son evidentemente enunciados de diferente valor cognoscitivo:  $a = a$  vale *a priori* y, siguiendo a Kant, puede denominarse analítico, mientras que enunciados de la forma  $a = b$  contienen frecuentemente ampliaciones muy valiosas de nuestro conocimiento y no siempre pueden justificarse *a priori*. El descubrimiento de que cada mañana no se levanta un nuevo sol, sino que siempre es el mismo, fue ciertamente uno de los descubrimientos más trascendentales de la astronomía. Aún ahora, el reconocimiento de un pequeño planeta o de un cometa no es siempre algo evidente. Ahora bien, si en la igualdad quisiéramos ver una relación entre aquello a lo que los nombres «*a*» y «*b*» se refieren, no parecería que  $a = b$  pudiera ser distinto de  $a = a$ , siempre que  $a = b$  fuera cierto. Se habría expresado, en tal caso, una relación de una cosa consigo misma, y además una relación tal que se da en cada cosa respecto de sí misma, pero que ninguna cosa tiene respecto de cualquier otra. Parece que lo que se quiere decir con  $a = b$  es que los signos o nombres «*a*» y «*b*» se refieren a lo mismo, y por lo tanto en la igualdad se trataría precisamente de estos signos; se afirmarían una relación entre ellos. Pero esta relación existiría entre los nombres o signos únicamente en la medida en que éstos denominan o designan algo. Sería una relación inducida por la conexión de cada uno de los dos signos con la misma cosa designada. Esta conexión es arbitraria. No se le puede prohibir a nadie tomar cualquier suceso u objeto producido arbitrariamente, como signo para algo. Con ello, el enunciado  $a = b$  no se referiría entonces ya a la cosa misma, sino tan sólo a nuestro modo de designación; con ella no expresaríamos ningún verdadero conocimiento. Pero esto es justamente lo que quere-

---

\* Versión castellana de Ulises Moulines.

<sup>1</sup> Empleo esta palabra en el sentido de identidad y entiendo « $a = b$ » en el sentido de «*a* es lo mismo que *b*» o «*a* y *b* coinciden».



mos en muchos casos. Si el signo «*a*» sólo se diferencia del signo «*b*» como objeto (en este caso por su forma), y no como signo (es decir, no por el modo como designa algo), entonces el valor cognoscitivo de  $a = a$  sería esencialmente el mismo que el de  $a = b$ , caso de que  $a = b$  fuera verdadero. Una distinción puede darse únicamente en el caso de que la diferencia de signos corresponda a una diferencia en el modo de darse lo designado. Sean *a*, *b*, *c*, las rectas que unen los ángulos de un triángulo con el punto medio de los lados opuestos. El punto de intersección de *a* y *b* es entonces el mismo que el punto de intersección de *b* y *c*. Tenemos, pues, designaciones distintas para el mismo punto, y estos nombres («intersección de *a* y *b*», «intersección de *b* y *c*») indican al mismo tiempo el modo de darse el punto, y de ahí que en el enunciado esté contenido auténtico conocimiento.

Es natural considerar entonces que a un signo (nombre, unión de palabras, signo escrito), además de lo designado, que podría llamarse la referencia del signo, va unido lo que yo quisiera denominar el sentido del signo, en el cual se halla contenido el modo de darse. Según esto, en nuestro ejemplo, la referencia de las expresiones «el punto de intersección de *a* y *b*» y «el punto de intersección de *b* y *c*» sería ciertamente la misma, pero no sería el mismo su sentido. La referencia de «lucero vespertino» y de «lucero matutino» sería la misma, pero el sentido no sería el mismo.

Del presente contexto se desprende que con «signo» y «nombre» he entendido cualquier designación que represente un nombre propio, cuya referencia sea, pues, un objeto determinado (tomada esta palabra en su extensión más amplia), pero no un concepto ni una relación, sobre los cuales se tratará con más detenimiento en otro ensayo \*. La designación de un único objeto puede estar compuesta de varias palabras u otro tipo de signos. Para abreviar, llamaremos nombre propio a cada una de tales designaciones.

El sentido de un nombre propio lo comprende todo aquel que conoce el lenguaje o el conjunto de designaciones al que pertenece <sup>2</sup>; pero con ello, la referencia, caso de que exista, queda sólo parcial-

\* Se refiere a «Sobre concepto y objeto», en G. Frege, *Estudios sobre semántica*, Ariel, Barcelona, 1971, pp. 99-120. (N. del T.)

<sup>2</sup> En el caso de un verdadero nombre propio como «Aristóteles», naturalmente pueden dividirse las opiniones en cuanto a su sentido. Por ejemplo, se podría suponer que este sentido es: el discípulo de Platón y maestro de Alejandro Magno. Quien suponga esto, atribuirá al enunciado «Aristóteles era originario de Estagira» un sentido distinto de aquel para quien el sentido de este nombre fuera: el maestro de Alejandro Magno originario de Estagira. Mientras la referencia siga siendo la misma, pueden tolerarse estas oscilaciones del sentido, a pesar de que deben evitarse en el edificio conceptual de una ciencia demostrativa y de que no deberían aparecer en un lenguaje perfecto.

mente iluminada. Un conocimiento completo de la referencia implicaría que, de cada sentido dado, pudiéramos indicar inmediatamente si le pertenece o no. Esto no lo logramos nunca.

La conexión regular entre el signo, su sentido y su referencia es tal que al signo le corresponde un determinado sentido y a éste, a su vez, una determinada referencia, mientras que a una referencia (a un objeto) no le corresponde solamente un signo. El mismo sentido puede expresarse en diferentes lenguas, e incluso en la misma, de diversas maneras. Naturalmente, hay excepciones a esta situación regular. Es verdad que en un conjunto perfecto de signos, a cada expresión debería corresponderle un sentido determinado; pero las lenguas naturales a menudo no cumplen este requisito, y hay que darse por satisfecho si, sólo en un mismo contexto, tiene la misma palabra siempre el mismo sentido. Quizá puede admitirse que una expresión gramaticalmente correcta que sustituye un nombre propio tiene siempre sentido. Pero con ello no se ha dicho que al sentido le corresponda también una referencia. Las palabras «el cuerpo celeste más alejado de la Tierra» tienen un sentido; pero que tengan también una referencia, es muy dudoso. La expresión «la serie menos convergente» tiene un sentido; pero se demuestra que no tiene referencia, puesto que para cada serie convergente puede encontrarse otra menos convergente, pero que, no obstante, es convergente. Así pues, por el hecho de que se conciba un sentido, no se tiene con seguridad una referencia.

Cuando se usan palabras de la manera habitual aquello de lo que se quiere hablar es su referencia. Pero puede ocurrir también que se quiera hablar de las palabras mismas o de su sentido. Lo primero sucede, por ejemplo, cuando se citan las palabras de otro en estilo directo. Las palabras propias se refieren entonces en primer lugar a las palabras del otro, y tan sólo estas últimas tienen la referencia corriente. Tenemos entonces signos de signos. En el lenguaje escrito se encierran los caracteres, en este caso, entre comillas. Por lo tanto, un carácter que se halla entre comillas no debe ser tomado en su referencia usual.

Si se quiere hablar del sentido de la expresión «A», basta con usar sencillamente la locución «el sentido de la expresión "A"». En el estilo indirecto se habla del sentido, por ejemplo, del discurso de otro. Se ve claramente que, incluso en este modo de hablar, las palabras no tienen su referencia usual, sino que se refieren a lo que habitualmente es su sentido. Para utilizar una expresión breve, vamos a decir: las palabras se usan *indirectamente*, o tienen su referencia *indirecta* en el estilo indirecto. Según esto, distinguimos la referencia habitual de una palabra de su referencia indirecta, y su sentido habitual de su sentido indirecto. La referencia indirecta de una palabra es, pues, su sentido usual. Hay que tener siempre presentes tales excep-

ciones si se quiere concebir correctamente, en cada caso particular, el modo de conexión de signo, sentido y referencia.

De la referencia y del sentido de un signo hay que distinguir la representación a él asociada. Si la referencia de un signo es un objeto sensiblemente perceptible, la representación que yo tengo de él es entonces una imagen interna formada a partir de recuerdos de impresiones sensibles que he tenido, y de actividades que he practicado, tanto internas como externas<sup>3</sup>. Esa imagen está frecuentemente impregnada de sentimientos; la claridad de cada una de sus partes es diversa y vacilante. No siempre, ni siquiera en la misma persona, está unida la misma representación al mismo sentido. La representación es subjetiva: la representación de uno no es la del otro. Por ello se dan múltiples diferencias en las representaciones asociadas al mismo sentido. Un pintor, un jinete y un zoólogo asociarán probablemente representaciones muy distintas al nombre «Bucéfalo». Por eso se diferencia la representación esencialmente del sentido de un signo, el cual puede ser propiedad común de muchos y que, por tanto, no es parte o modo de la mente individual; pues ciertamente no se podrá negar que la Humanidad tiene un tesoro común de pensamientos, que transmite de una generación a otra<sup>4</sup>.

Mientras que, según lo dicho, no existe ninguna objeción para hablar del sentido sin más, en el caso de la representación, en cambio, para ser estrictos, hay que añadir a quién pertenece y en qué momento. Quizás alguien diría: al igual que con la misma palabra uno asocia tal representación, el otro tal otra, también puede uno asociarle tal sentido, el otro tal otro. Sin embargo, la diferencia consiste entonces sólo en el modo de darse la asociación. Esto no impide que ambos conciban el mismo sentido; pero no pueden tener la misma representación. *Si duo idem faciunt, non est idem*. Cuando dos personas se representan lo mismo, cada una tiene, sin embargo, su representación propia. A veces es ciertamente posible constatar diferencias de representaciones y hasta de sensaciones de personas distintas; pero no es posible una comparación exacta, por que no podemos tener juntas representaciones en la misma conciencia.

La referencia de un nombre propio es el objeto mismo que designamos con él; la representación que tenemos entonces es totalmente

---

<sup>3</sup> Junto a las representaciones, podemos poner también las intuiciones o datos sensoriales, en los que las impresiones sensibles y las actividades mismas ocupan el lugar de las huellas que han dejado en el espíritu. Para nuestro propósito, la diferencia es irrelevante, tanto más cuanto que, junto a las sensaciones y actividades, los recuerdos de éstas ayudan a completar la imagen intuitiva. Por intuición o dato sensorial, sin embargo, puede entenderse también un objeto, en la medida en que éste sea sensiblemente perceptible o espacial.

<sup>4</sup> De ahí que sea inútil designar con la palabra «representación» cosas básicamente tan distintas.

subjetiva; entre ambas se halla el sentido, que ciertamente ya no es subjetivo como la representación, pero, con todo, tampoco es el objeto mismo. Quizá sea adecuada la siguiente analogía, para ilustrar estas relaciones. Alguien observa la Luna a través de un telescopio. Compara la Luna con la referencia; es el objeto de observación, que es proporcionado por la imagen real que queda dibujada sobre el cristal del objetivo del interior del telescopio, y por la imagen en la retina del observador. La primera imagen la comparo con el sentido; la segunda, con la representación o intuición. La imagen formada dentro del telescopio es, en verdad, sólo parcial; depende del lugar de observación; pero con todo es objetiva, en la medida en que puede servir a varios observadores. Podría incluso disponerse de modo que pudieran utilizarla varios simultáneamente. Pero, de las imágenes retinianas, cada uno tendría la suya propia. Apenas podría lograrse una congruencia geométrica, debido a la diferente constitución de los ojos, y una coincidencia real estaría excluida. Podría quizá seguir desarrollándose esta analogía, admitiendo que la imagen retiniana de A podría hacerse visible a B; o también que el propio A podría ver su propia imagen retiniana en un espejo. Con esto se mostraría quizá que una representación puede ser tomada ciertamente como objeto, pero que en sí misma no es nunca para el observador lo que es para el que la tiene. Pero seguir en esta dirección nos apartaría demasiado de lo que nos ocupa.

Podemos ahora distinguir tres niveles de diferenciación entre palabras, expresiones o frases enteras. O bien la diferencia se refiere a lo sumo a las representaciones, o bien al sentido pero no a la referencia, o bien, en fin, también a la referencia. Con respecto al primer nivel, hay que hacer notar que, debido a la conexión incierta de las representaciones con las palabras, para uno puede existir una diferencia que otro no encuentra. Las diferencias en la traducción de un escrito original no pasarían de este primer nivel. Entre otras diferencias posibles aquí, están los matices y énfasis con que la poesía [y] la elocuencia tratan de revestir el sentido. Estos matices y énfasis no son objetivos, sino que el oyente o el lector debe dejarse llevar por las alusiones del poeta o del orador. Naturalmente, sin cierto parentesco entre las representaciones humanas, el arte no sería posible; pero nunca puede averiguarse exactamente en qué medida nuestras representaciones corresponden a los propósitos del poeta.

En lo que sigue, no hablaremos ya más de las representaciones e intuiciones; se las ha mencionado aquí únicamente para que la representación que despierta una palabra en un oyente no sea confundida con su sentido o su referencia.

Para posibilitar un modo de expresarnos breve y exacto, vamos a establecer las siguientes locuciones:

Un nombre propio (palabra, signo, fila de signos o expresión) ex-

presa su sentido, se refiere a su referencia o la designa. Con un signo expresamos su sentido y designamos su referencia.

De parte idealista o escéptica, a todo esto quizá se habrá objetado desde hace ya rato lo siguiente: «Hablas aquí sin más de la Luna como de un objeto. ¿Pero cómo sabes tú que el nombre "la Luna" tiene alguna referencia, cómo sabes que hay algo que tenga referencia?» Respondo que nuestro propósito no es hablar de nuestra representación de la Luna, y que tampoco nos conformamos con el sentido, cuando decimos «la Luna», sino que presuponemos una referencia. Sería perder totalmente el sentido si se quisiera suponer que, en el enunciado «la Luna es menor que la Tierra», se está hablando de una representación de la Luna. Si ésta fuera la intención del que habla, utilizaría la expresión «mi representación de la Luna». Desde luego, al hacer aquella presuposición podemos equivocarnos, y tales equivocaciones se dan ciertamente. Pero aquí no tenemos por qué responder a la cuestión de que quizá siempre cometemos tal equivocación; de momento basta con señalar nuestro propósito al hablar o al pensar, para justificar el que hablemos de la referencia de un signo, si bien con la reserva: caso que exista tal.

Hasta aquí sólo se han examinado sentido y referencia de las expresiones, palabras o signos, que hemos llamado nombres propios. Ahora vamos a preguntarnos por el sentido y la referencia de un enunciado asertivo completo. Un tal enunciado contiene un pensamiento<sup>5</sup>. ¿Debe ser considerado este pensamiento como su sentido o como su referencia? Supongamos que el enunciado tiene una referencia. Si sustituimos en él una palabra por otra de la misma referencia, pero de distinto sentido, esto no podrá tener ningún efecto sobre la referencia del enunciado. Sin embargo, vemos que, en tales casos, el pensamiento cambia; pues, por ejemplo, el pensamiento del enunciado «el lucero matutino es un cuerpo iluminado por el sol» es distinto del enunciado «el lucero vespertino es un cuerpo iluminado por el sol». Alguien que no supiera que el lucero vespertino es el lucero matutino podría tomar un pensamiento por verdadero y el otro por falso. El pensamiento no puede, pues, ser la referencia del enunciado; por el contrario, deberemos concebirlo como su sentido. ¿Pero qué hacemos con la referencia? ¿Tenemos derecho a preguntar por ella? ¿Acaso el enunciado entero tiene sólo sentido, pero no referencia? En todo caso, es de esperar que se den tales enunciados, lo mismo que hay partes de un enunciado que tienen sentido, pero no referencia. Y los enunciados que contienen nombres propios sin referencia serán de este tipo. El enunciado «Ulises fue dejado en Ítaca

<sup>5</sup> Por pensamiento no entiendo la actividad subjetiva de pensar, sino su contenido objetivo, que es apto para ser propiedad común de muchos.

profundamente dormido» tiene evidentemente un sentido. Pero, como es dudoso que el nombre «Ulises» que aparece en él tenga una referencia, también es dudoso que lo tenga el enunciado entero. Pero lo que es seguro, no obstante, es que alguien que crea en serio que el enunciado es verdadero o falso, también atribuiría al nombre «Ulises» una referencia, y no sólo un sentido; pues es justamente de la referencia de este nombre de lo que se afirma o se niega el predicado. Quien no admita una referencia no podrá afirmar ni negar de ella un predicado. Pero entonces sería innecesario el llegar hasta la referencia del nombre; uno podría contentarse con el sentido, en el caso de querer quedarse con el pensamiento. Si sólo nos interesásemos por el sentido del enunciado, por el pensamiento, sería innecesario preocuparse de la referencia de una parte del enunciado; pues, con respecto al sentido del enunciado, únicamente es relevante el sentido, no la referencia, de esta parte. El pensamiento sigue siendo el mismo, tanto si el nombre «Ulises» tiene una referencia como si no. Que nos esforcemos por hallar la referencia de una parte del enunciado es señal de que también admitimos y exigimos, en general, una referencia para el enunciado mismo. El pensamiento pierde valor para nosotros tan pronto como vemos que a una de sus partes le falta la referencia. Estamos, pues, bien justificados al no contentarnos con el sentido de un enunciado, y al preguntarnos también por su referencia. Pero ¿por qué queremos que cada nombre propio no tenga únicamente un sentido, sino también una referencia? ¿Por qué no nos basta el pensamiento? Porque, y en la medida en que, nos interesa su valor veritativo. No siempre es éste el caso. Al escuchar un poema épico, por ejemplo, nos cautivan, además de la eufonía del lenguaje, el sentido de los enunciados y las representaciones y sentimientos despertados por ellos. Si nos preguntásemos por su verdad, abandonaríamos el goce estético y nos dedicaríamos a un examen científico. De ahí que nos sea indiferente el que el nombre «Ulises», por ejemplo, se refiera a algo o no, mientras consideremos el poema como obra de arte <sup>6</sup>. Es la búsqueda de la verdad lo que nos incita a avanzar del sentido a la referencia. Hemos visto que a un enunciado hay que buscarle una referencia siempre que interesa la referencia de las partes componentes; y esto es siempre el caso, y sólo entonces, cuando nos preguntamos por los valores veritativos.

Por esto nos vemos impulsados a admitir el *valor veritativo* de un enunciado como su referencia. Por valor veritativo de un enunciado entiendo la circunstancia de que sea verdadero o de que sea falso. No hay más valores veritativos. En aras de la brevedad, al uno lo llamo

<sup>6</sup> Sería de desear que tuviéramos una denominación especial para los signos que sólo han de tener sentido. Si llamásemos a éstos imágenes, las palabras del actor en la escena serían entonces imágenes, y hasta el propio actor sería una imagen.

lo verdadero, al otro lo falso. Cada enunciado asertivo, en el que tenga importancia la referencia de las palabras, debe ser considerado, pues, como un nombre propio, y su referencia, caso de que exista, es o bien lo verdadero o bien lo falso. Estos dos objetos son admitidos, aunque sólo sea tácitamente, por todo aquel que emita juicios, que tenga algo por verdadero, o sea, también por el escéptico. El designar los valores veritativos como objetos puede parecer aquí todavía una ocurrencia arbitraria y quizás un mero juego de palabras, del que no deberían sacarse consecuencias fundamentales. Lo que yo llamo objeto, sólo podrá ser discutido con más precisión teniendo en cuenta el concepto y la relación. Esto quiero reservarlo para otro ensayo\*. Pero, con todo, aquí podría ya quedar claro que en todo juicio<sup>7</sup> —y por muy evidente que éste sea— se ha dado ya el paso del nivel de los pensamientos al nivel de las referencias (de lo objetivo).

Alguno podría verse tentado a considerar la relación del pensamiento con lo verdadero no como la que hay entre el sentido y la referencia, sino como relación del sujeto con el predicado. Verdaderamente puede decirse: «El pensamiento de que 5 es un número primo es verdadero.» Pero si se examina esto más atentamente, se observa que con ello no se dice realmente nada más de lo que se dice en el simple enunciado «5 es un número primo». La afirmación de la verdad radica, en ambos casos, en la forma del enunciado asertivo, y cuando éste no tiene su fuerza habitual, por ejemplo en boca de un actor en escena, el enunciado «el pensamiento de que 5 es un número primo es verdadero» contiene también únicamente un pensamiento, a saber, el mismo pensamiento que el simple «5 es un número primo». De aquí puede desprenderse que la relación del pensamiento con lo verdadero no debe compararse a la del sujeto con el predicado. Efectivamente, sujeto y predicado (entendidos en sentido lógico) son partes del pensamiento; para el conocimiento, se hallan al mismo nivel. Ensamblando sujeto y predicado siempre se consigue únicamente un pensamiento, pero no se pasa nunca de un sentido a su referencia, de un pensamiento a su valor veritativo. Nos movemos en el mismo nivel, no se pasa de un nivel al siguiente. Un valor veritativo no puede ser parte de un pensamiento, como no puede serlo el sol, porque no es un sentido, sino un objeto.

Si es correcta nuestra suposición de que la referencia de un enunciado es su valor veritativo, entonces éste debe permanecer inmodificado cuando una parte del enunciado se sustituye por una expresión de la misma referencia, pero de distinto sentido. Y, de hecho, éste es el caso. Leibniz explica correctamente: «*Eadem sunt, quae sibi*

\* Se refiere de nuevo a «Sobre concepto y objeto», op. cit. *supra*. (N. del T.)

<sup>7</sup> Un juicio no es para mí la mera concepción de un pensamiento, sino el reconocimiento de su verdad.

*mutuo substitui possunt, salva veritate.*» Realmente, ¿qué otra cosa, sino el valor veritativo, podría encontrarse que pertenezca con toda generalidad a cada enunciado en el que interese la referencia de las partes componentes, y que permanezca inmodificado en una sustitución del tipo mencionado?

Ahora bien, si el valor veritativo de un enunciado es su referencia, resulta que, por una parte, todos los enunciados verdaderos tienen la misma referencia, y que, por otra, también todos los enunciados falsos tienen la misma referencia. De ahí que, en la referencia del enunciado, todo lo singular desaparezca. Nunca podemos quedarnos tan sólo con la referencia de un enunciado; pero tampoco el mero pensamiento proporciona ningún conocimiento, sino únicamente el pensamiento junto con su referencia, es decir, su valor veritativo. El juzgar puede ser considerado como el paso de un pensamiento a su valor veritativo. Naturalmente, esto no debe ser tomado como una definición. El juzgar es precisamente algo muy singular e incomparable. También podría decirse que juzgar es distinguir partes dentro de un valor veritativo. Esta distinción ocurre retrocediendo al pensamiento. Cada sentido que pertenezca a un valor veritativo correspondería a su modo propio de descomposición. La palabra «parte» la he utilizado aquí de una manera peculiar. En efecto, la relación del todo a la parte en el enunciado la he transferido a su referencia, al denominar a la referencia de una palabra, parte de la referencia del enunciado cuando esa misma palabra es parte de este enunciado, modo de hablar que naturalmente es impugnabile, porque, en el caso de la referencia, la otra parte no queda determinada por el todo y la parte escogida, y porque la palabra parte se emplea para los cuerpos en un sentido distinto. En su lugar, debería crearse una expresión apropiada.

Vamos ahora a seguir comprobando la suposición de que el valor veritativo de un enunciado es su referencia. Hemos hallado que el valor veritativo de un enunciado permanece inmodificado cuando en éste sustituimos una expresión por otra de igual referencia: pero todavía no hemos considerado el caso en que la expresión a ser sustituida es ella misma un enunciado. Si nuestro punto de vista es correcto, el valor veritativo de un enunciado, que contiene a otro como parte, debe permanecer inmodificado si sustituimos el enunciado componente por otro cuyo valor veritativo es el mismo. Hay que esperar excepciones, cuando el todo o el enunciado componente están en estilo directo o indirecto; pues, como hemos visto, la referencia de las palabras no es entonces la usual. Un enunciado se refiere en el estilo directo a otro enunciado, y en el indirecto, a un pensamiento.

Nos vemos, pues, llevados al estudio de los enunciados subordinados. Éstos aparecen como partes de una estructura enunciativa que, desde el punto de vista lógico, es asimismo un enunciado, a



saber, el enunciado principal. Pero en este punto nos enfrentamos a la pregunta de si también vale para los enunciados subordinados el que su referencia sea un valor veritativo. Del estilo indirecto sabemos ya que ocurre lo contrario. Los gramáticos consideran los enunciados subordinados como representantes de partes del enunciado general, y, según eso, los denominan enunciados nominales, calificativos, adverbiales \*. De aquí podría surgir la suposición de que la referencia de un enunciado subordinado no es un valor veritativo, sino que es análogo a la de un nombre, un calificativo o un adverbio, en resumen, al de una parte del enunciado, cuyo sentido no es un pensamiento, sino sólo una parte del mismo. Únicamente una investigación más detenida puede proporcionar claridad sobre este punto. En ella, no nos atenderemos estrictamente al hilo conductor gramatical, sino que reuniremos lo que es lógicamente similar. Busquemos primero aquellos casos en los que el sentido del enunciado subordinado, como acabamos de suponer, no es un pensamiento autónomo.

A los enunciados nominales abstractos introducidos por «que», pertenece también el estilo indirecto, del cual hemos visto que, en él, las palabras tienen una referencia indirecta, que coincide con lo que habitualmente es su sentido. En este caso, pues, el enunciado subordinado tiene por referencia un pensamiento, no un valor veritativo; por sentido, no un pensamiento, sino el sentido de las palabras «el pensamiento de que...», el cual es sólo parte del pensamiento de toda la estructura enunciativa. Esto sucede después de «decir», «oír», «opinar», «estar convencido», «concluir», y palabras parecidas <sup>8</sup>. La cuestión aparece distinta, y ciertamente bastante complicada, después de palabras como «conocer», «saber», «imaginarse», lo cual será estudiado más adelante.

Que en nuestros casos la referencia del enunciado subordinado es, en realidad, el pensamiento, se ve también por el hecho de que, para la verdad del todo, es indiferente que aquel pensamiento sea verdadero o falso. Compárense, por ejemplo, los dos enunciados: «Copérnico creía que las órbitas de los planetas eran círculos» y «Copérnico creía que la apariencia del movimiento del Sol es producida por el movimiento real de la Tierra». Sin perjuicio de la verdad, puede sustituirse aquí un enunciado subordinado por el otro. El enunciado principal, junto con el subordinado, tiene por sentido úni-

\* Hemos traducido *Nennsatz* por «enunciado nominal», *Beisatz* por «enunciado calificativo» y *Adverbsatz* por «enunciado adverbial». Con esta clasificación más o menos gramatical, Frege quiere distinguir entre enunciados que podrían ser sustituidos por un nombre, por un adjetivo calificativo o por locuciones adverbiales, respectivamente. (*N. del T.*)

<sup>8</sup> En «A mintió al decir que había visto a B», el enunciado subordinado se refiere a un pensamiento, del cual se dice, en primer lugar, que A lo afirmó como verdadero, y, en segundo lugar, que A estaba convencido de su falsedad.

camente un solo pensamiento, y la verdad del todo no implica ni la verdad ni la falsedad del subordinado. En tales casos no está permitido sustituir, en el enunciado subordinado, una expresión por otra que tenga la misma referencia usual, sino solamente por una que tenga la misma referencia indirecta, es decir, el mismo sentido usual. Si alguien quisiera sacar la conclusión: la referencia de un enunciado no es su valor veritativo, «pues entonces podría sustituirse en todas partes por otro enunciado con el mismo valor veritativo», habría demostrado demasiado; con la misma razón podría afirmarse que la referencia de la palabra «lucero matutino» no es Venus; pues no en todas partes podría decirse «Venus» en vez de «lucero matutino». Correctamente sólo puede deducirse que la referencia de un enunciado *no siempre* es su valor veritativo, y que «lucero matutino» no siempre se refiere al planeta Venus, a saber, en el caso en que esa palabra tenga su referencia indirecta. Semejante caso de excepción se presenta en los enunciados subordinados que acabamos de examinar, cuya referencia es un pensamiento.

Cuando se dice «parece que...», lo que se quiere decir es «me parece que...» u «opino que...». Tenemos, pues, el mismo caso. Igualmente ocurre con expresiones como «alegrarse», «lamentar», «aprobar», «censurar», «esperar», «temer». Cuando, hacia el fin de la batalla de Belle-Alliance, Wellington se alegró de que los prusianos vinieran, la razón de su alegría era un convencimiento. Si hubiera estado equivocado, no se habría alegrado menos hasta tanto hubiese durado su ilusión, y antes de adquirir el convencimiento de que venían los prusianos no podía alegrarse de ello, si bien, en realidad, ya se acercaban.

Así como un convencimiento o una creencia es razón de un sentimiento, también puede ser razón de otro convencimiento, como ocurre en la inferencia. En el enunciado: «De la redondez de la Tierra, Colón infirió que, viajando hacia el oeste, podría alcanzar la India», tenemos, como referencia de las partes, dos pensamientos: que la Tierra es redonda, y que Colón puede alcanzar la India viajando hacia el oeste. Nuevamente, aquí importa tan sólo que Colón estaba convencido de lo uno y de lo otro, y que un convencimiento era la razón del otro. Que la Tierra sea realmente redonda y que Colón, viajando hacia el oeste, pudiese realmente alcanzar la India, tal y como él pensaba, es indiferente para la verdad de nuestro enunciado; pero no es indiferente que pongamos, en vez de «la Tierra», el «planeta, que está acompañado de una luna cuyo diámetro es mayor que la cuarta parte de su propio diámetro». También aquí tenemos la referencia indirecta de las palabras.

Éste es el caso también de los enunciados adverbiales de finalidad con «para qué»; pues evidentemente la finalidad es un pensamiento; por eso: referencia indirecta de las palabras, subjuntivo.

El enunciado subordinado con «que» después de «mandar», «pedir», «prohibir» aparecería, en estilo directo, en forma de imperativo. Tal enunciado no tiene referencia, sino sólo un sentido. Una orden, un ruego, no son ciertamente pensamientos, pero, con todo, están al mismo nivel que el pensamiento. De ahí que, en las subordinadas que dependen de «mandar», «pedir», etc., las palabras tienen su referencia indirecta. La referencia de uno de estos enunciados no es, pues, un valor veritativo, sino una orden, un ruego, u otros similares.

Análogamente ocurre, en el caso de la pregunta indirecta, en giros tales como «dudar de que», «no saber que». Es fácil ver que también aquí hay que tomar las palabras en su referencia indirecta. Las interrogativas indirectas con «quién», «qué», «dónde», «cuándo», «cómo», «por medio de qué», etc., a veces se asemejan aparentemente mucho a enunciados adverbiales en los que las palabras tienen su referencia usual. Lingüísticamente, estos casos se diferencian por el modo del verbo. En el caso del subjuntivo tenemos preguntas indirectas y referencia indirecta de las palabras, de modo que un nombre propio no puede ser sustituido en general por otro del mismo objeto.

En los casos considerados hasta aquí, las palabras de los enunciados subordinados tenían su referencia indirecta, y por esto es explicable que también la referencia del enunciado subordinado mismo fuera indirecta; es decir, no un valor veritativo, sino un pensamiento, una orden, un ruego, una pregunta. El enunciado subordinado podría ser concebido como nombre; se podría incluso decir: como nombre propio del pensamiento, la orden, etc., puesto que como tal aparecía en el contexto de la estructura enunciativa.

Llegamos ahora a otros enunciados subordinados, en los que las palabras tienen ciertamente su referencia usual, pero sin que aparezca un pensamiento como sentido, ni un valor veritativo como referencia. Cómo es esto posible, se verá claramente con ejemplos.

«El que descubrió la forma elíptica de las órbitas planetarias murió en la miseria.»

Si en este caso el enunciado subordinado tuviera por sentido un pensamiento, tendría que ser posible expresarlo también en un enunciado principal. Pero esto no puede ser, porque el sujeto gramatical «el que» no tiene ningún sentido independiente, sino que proporciona las relaciones con el segundo miembro del enunciado, «murió en la miseria». De ahí también que el sentido del enunciado subordinado no sea un pensamiento completo y que su referencia no sea un valor veritativo, sino Kepler. Podría objetarse que el sentido del todo contiene, no obstante, un pensamiento como parte, a saber, el de que existió uno que descubrió por primera vez la forma elíptica de las órbitas planetarias; pues quien tuviera por verdadero el todo no

podría negar esta parte. Lo último es indiscutible; pero únicamente debido a que, en caso contrario, la subordinada «el que descubrió la forma elíptica de las órbitas planetarias» no tendría ninguna referencia. Cuando se afirma algo, siempre es evidente la suposición previa de que los nombres propios utilizados, ya sean simples o compuestos, tienen una referencia. Así pues, si se afirma «Kepler murió en la miseria», se presupone con ello que el nombre «Kepler» designa algo; pero por esto, sin embargo, en el sentido del enunciado «Kepler murió en la miseria» no está contenido el pensamiento de que el nombre «Kepler» designa algo. Si éste fuera el caso, la negación no podría ser

«Kepler no murió en la miseria»,

sino

«Kepler no murió en la miseria, o bien el nombre “Kepler” carece de referencia».

Que el nombre «Kepler» designa algo es, por el contrario, suposición tanto de la afirmación

«Kepler murió en la miseria»,

como de la opuesta. Ahora bien, resulta que las lenguas tienen el defecto de que en ellas son posibles expresiones que, por su forma gramatical, están destinadas a designar un objeto, pero que, en casos especiales, no consiguen este objetivo suyo, porque esto depende de la verdad de un enunciado. Por eso depende de la verdad del enunciado

«existió uno que descubrió la forma elíptica de las órbitas planetarias»,

el que la subordinada

«el que descubrió la forma elíptica de las órbitas planetarias»

designe realmente un objeto, o bien que sólo produzca la apariencia de ello, careciendo de hecho de referencia. Y así es cómo llega a parecer como si nuestra subordinada contuviera, como parte de su sentido, el pensamiento de que existió uno que descubrió la forma elíptica de las órbitas planetarias. Si esto fuera correcto, la negación debería ser:

«el que descubrió por primera vez la forma elíptica de las órbi-

tas planetarias, no murió en la miseria, o bien no hubo nadie que descubriese la forma elíptica de las órbitas planetarias».

Esto radica, pues, en una imperfección del lenguaje, de la que, por lo demás, tampoco está completamente libre el lenguaje simbólico del análisis; también en este último caso pueden aparecer filas de signos que producen la ilusión de que se refieren a algo, pero que, por lo menos hasta este momento, todavía carecen de referencia, como por ejemplo, series infinitas divergentes. Esto puede remediarse, por ejemplo, mediante la estipulación especial de que las series infinitas divergentes tienen que referirse al número 0. De un lenguaje lógicamente perfecto (ideografía) hay que exigir que cada expresión, que se haya formado como nombre propio a partir de signos ya introducidos de manera gramaticalmente correcta, designe realmente también un objeto, y que no se introduzca ningún signo como nombre propio sin que antes no se le haya asegurado una referencia. En los tratados de Lógica se previene en contra de la multivocidad de las expresiones como fuente de errores lógicos. Creo que es por lo menos igualmente oportuna la prevención frente a los nombres propios aparentes que no tienen ninguna referencia. La historia de las matemáticas podría narrar todos los errores que han surgido de ahí. Éstos son frecuentes igualmente en el mal uso demagógico, quizá más todavía que las palabras multívocas. Puede servir de ejemplo «la voluntad del pueblo», pues es fácil establecer que, por lo menos, no hay una referencia universalmente aceptada de esta expresión. Por esto no es en absoluto irrelevante taponar de una vez por todas la fuente de esos errores, por lo menos para la ciencia. Objeciones como la antes discutida serán entonces imposibles, porque nunca podrá depender de la verdad de un pensamiento el que un nombre propio tenga una referencia.

Podemos someter a estudio estos enunciados nominales juntamente con una clase de enunciados calificativos \* y adverbiales que están lógicamente emparentados con los primeros.

También algunos enunciados calificativos pueden servir para formar nombres propios compuestos, si bien esto no lo consiguen por sí solos, como en el caso de los nominales. Estos enunciados calificativos deben ser tomados igual que términos calificativos. En vez de «la raíz cuadrada de 4 que es menor que 0», puede decirse también «la raíz cuadrada negativa de 4». Nos hallamos aquí ante el caso en que, a partir de una expresión conceptual, se forma un nombre propio compuesto con la ayuda del artículo determinado en singular; lo

---

\* Traducimos por «enunciado calificativo» el alemán *Beisatz*: se trata de un enunciado (que gramaticalmente será casi siempre de relativo) cuya función es equivalente a un atributo o adjetivo calificativo. (N. del T.)

cual, de todos modos, sólo está permitido cuando cae bajo el concepto un objeto y sólo uno <sup>9</sup>. Las expresiones conceptuales pueden formarse de tal manera que se indiquen sus características por medio de enunciados calificativos, como en nuestro ejemplo por medio del enunciado «que es menor que 0». Es evidente que semejante enunciado calificativo no puede tener, lo mismo que antes el enunciado nominal, un pensamiento por sentido, ni un valor veritativo por referencia, sino que por sentido tiene solamente una parte de un pensamiento que, en algunos casos, puede expresarse también como un único calificativo. También en este caso, lo mismo que en el de los enunciados nominales, falta el sujeto independiente y con él también la posibilidad de volver a dar el sentido de la subordinada en una independiente.

Desde un punto de vista lógico, los lugares, instantes e intervalos son objetos; por lo tanto, la denominación lingüística de un determinado lugar, de un determinado momento o intervalo temporal debe ser considerada como un nombre propio. Enunciados adverbiales de lugar y de tiempo pueden entonces ser utilizados para la formación de un nombre propio semejante, de manera análoga a como lo acabamos de ver para los enunciados nominales y calificativos. Asimismo pueden formarse expresiones de conceptos que se refieren a lugares, etc. También aquí hay que hacer notar que no puede volverse a dar el sentido de estas subordinadas en una principal, porque falta un componente esencial, a saber, la determinación espacial o temporal, a la que sólo se alude por medio de un pronombre relativo a una conjunción <sup>10</sup>.

<sup>9</sup> Según lo anteriormente observado, a una expresión semejante, de hecho se le debería asegurar siempre una referencia por medio de una estipulación especial, por ejemplo, por medio de la definición de que su referencia será el número 0 siempre que bajo el concepto no caiga ningún objeto o más de uno.

<sup>10</sup> Por lo demás, respecto de estos enunciados, son posibles concepciones ligeramente distintas. El sentido del enunciado «después de que Schleswig-Holstein se hubo separado de Dinamarca, se enemistaron Prusia y Austria» podemos volver a darlo bajo la forma «después de la separación de Schleswig-Holstein de Dinamarca, se enemistaron Prusia y Austria». Bajo esta concepción, está suficientemente claro que no debe ser considerado parte de este sentido el pensamiento de que Schleswig-Holstein se separó alguna vez de Dinamarca, sino que esto es la condición necesaria para que la expresión «después de la separación de Schleswig-Holstein de Dinamarca» tenga alguna referencia. Naturalmente, nuestro enunciado puede concebirse de tal manera que con él se diga que Schleswig-Holstein se separó una vez de Dinamarca. Entonces tenemos un caso que deberá ser estudiado más adelante. Para hacer más clara la diferencia, pongámonos en la mente de un chino que, por sus escasos conocimientos de la historia europea, crea que es falso que Schleswig-Holstein se haya separado alguna vez de Dinamarca. Éste considerará que nuestro enunciado, concebido de la primera manera, no es ni verdadero ni falso, y rehusará darle referencia alguna, porque ésta le faltaría a la subordinada. Esta última daría sólo aparentemente una determinación

Incluso en los enunciados condicionales puede reconocerse generalmente, como lo acabamos de ver en el caso de los enunciados nominales, calificativos y adverbiales, un componente que alude indeterminadamente, al que en el enunciado consecuente le corresponde otro igual. Al aludirse los dos, el uno al otro, se unen ambos enunciados en una totalidad que, por lo general, expresa solamente un pensamiento. En el enunciado:

«si un número es menor que 1 y mayor que 0, también su cuadrado es menor que 1 y mayor que 0»,

este componente es «un número» en el antecedente o condicional y «su» en el consecuente. Justamente debido a esa indeterminación, obtiene el sentido la generalidad que se espera de una ley. Pero precisamente así se hace también que el antecedente por sí solo no tenga por sentido ningún pensamiento completo, y que exprese, junto con el consecuente, un pensamiento y uno solo, cuyas partes ya no son pensamientos. En general, es erróneo creer que en un juicio hipotético se interrelacionan dos juicios. Si se dice esto o algo parecido, la palabra «juicio» se usa en el mismo sentido que yo he asociado a la palabra «pensamiento», de modo que yo debería decir: «En un pensamiento hipotético, se interrelacionan dos pensamientos.» Esto podría ser cierto únicamente en el caso en que faltase un componente que aludiese indeterminadamente<sup>11</sup>; pero entonces tampoco se daría generalidad.

Cuando debe aludirse indeterminadamente a un instante en el enunciado condicional o antecedente y en el consecuente, esto ocurre no pocas veces solamente por medio del *tiempo presente* del verbo, que en este caso no connota el presente. Esta forma gramatical es entonces el componente que alude indeterminadamente en el enunciado principal y en el subordinado. «Cuando el Sol se halla en el trópico de Cáncer, tenemos el día más largo en el hemisferio norte», es un ejemplo de este caso. También aquí es imposible expresar el sentido del subordinado en uno principal, ya que este sentido no es un pensamiento completo; pues si dijéramos: «el Sol se halla en el trópico de Cáncer», nos referiríamos con ello a nuestro presente y, de este modo, cambiaríamos el sentido. Todavía menos es el sentido de la principal un pensamiento; tan sólo el todo consistente en la principal y la subordinada es lo que contiene un pensamiento. Por lo

---

temporal. Si, por el contrario, concibe nuestro enunciado del segundo modo, hallaría expresado en él un pensamiento que creería falso, junto a una parte que, para él, carecería de referencia.

<sup>11</sup> A veces falta una alusión lingüística explícita y debe ser deducida de todo el contexto.

demás, también puede aludirse indeterminadamente a varios componentes comunes en el antecedente y el consecuente.

Es obvio que enunciados nominales con «quien», «lo que» y adverbiales con «donde», «cuando», «donde quiera que», «siempre que» frecuentemente deben ser considerados, por el sentido, como enunciados condicionales; por ejemplo: «Quien coge barro, se ensucia.»

También algunos enunciados calificativos representan enunciados condicionales. De este modo, podemos expresar el sentido de nuestro enunciado antes mencionado también en la forma «el cuadrado del número que es menor que 1 y mayor que 0 es menor que 1 y mayor que 0».

La cuestión aparece totalmente distinta cuando el componente común del principal y del subordinado es designado por un nombre propio. En el enunciado:

«Napoleón, que se dio cuenta del peligro para su flanco derecho, dirigió él mismo sus guardias de corps contra la posición enemiga»,

se expresan los dos pensamientos siguientes:

1. Napoleón se dio cuenta del peligro para su flanco derecho;
2. Napoleón dirigió él mismo sus guardias de corps contra la posición enemiga.

Cuándo y dónde ocurrió esto puede saberse ciertamente sólo por el contexto, pero por eso mismo debe considerarse determinado. Si expresamos todo nuestro enunciado como afirmación, afirmamos con él al mismo tiempo los dos enunciados parciales. Si uno de estos enunciados es falso, lo es también el todo. Aquí tenemos el caso en que el subordinado por sí solo tiene por sentido un pensamiento completo (si lo completamos con una indicación temporal y espacial). En consecuencia, la referencia de la subordinada es un valor veritativo. Podemos esperar, pues, que, sin perjuicio de la verdad del todo, pueda ponerse en su lugar un enunciado del mismo valor veritativo. Éste es justamente el caso; debe tenerse en cuenta, tan sólo, que su sujeto ha de ser «Napoleón», por una razón puramente gramatical, puesto que sólo entonces puede ponerse el enunciado en la forma de un enunciado calificativo atribuido a «Napoleón». Pero si se prescinde de la exigencia de que tenga que aparecer en esta forma, y si se admite también la conexión por medio de «y», entonces desaparece esta restricción. Incluso en subordinadas con «aunque» se expresan pensamientos completos. Esta conjunción no tiene propiamente nin-



gún sentido y tampoco modifica el sentido del enunciado, sino que lo aclara sólo de una manera peculiar <sup>12</sup>. En verdad, podríamos sustituir, sin perjuicio de la verdad del todo, el enunciado concesivo por otro del mismo valor veritativo; pero la aclaración parecería entonces ligeramente inadecuada, como si se quisiera cantar una canción de contenido triste de una manera alegre.

En los últimos casos, la verdad del todo incluía la verdad de los enunciados parciales. Caso distinto es aquel en que un enunciado condicional expresa un pensamiento completo, cuando, en vez del componente sólo alusivo, contiene un nombre propio o algo que deba considerarse igual. En el enunciado

«si ahora el sol ya se ha levantado, entonces el cielo está muy nublado»,

el tiempo es el presente, o sea, determinado. También el lugar debe pensarse que está determinado. Aquí puede decirse que se ha postulado una relación entre el valor veritativo del antecedente y del consecuente, o sea, la de que no se da el caso en que el antecedente se refiere a lo verdadero y el consecuente a lo falso. Según esto, nuestro enunciado es verdadero, tanto si el sol todavía no se ha levantado ahora, esté el cielo muy nublado o no, como si el sol se ha levantado ya y el cielo está muy cubierto. Como que, en este caso, sólo interesan los valores veritativos, puede sustituirse cada uno de los enunciados parciales por otro del mismo valor veritativo, sin que cambie el valor veritativo del todo. Naturalmente, la aclaración sería también aquí generalmente inoportuna: el pensamiento parecería ligeramente absurdo; pero esto no tiene nada que ver con su valor veritativo. En todo caso, debe tenerse siempre en cuenta que resuenan pensamientos adicionales, que, sin embargo, en realidad, no están expresados y que por esto no deben ser incluidos en el sentido del enunciado, no interesándonos, por lo tanto, su valor veritativo <sup>13</sup>.

Con esto se habrían discutido los casos simples. Lancemos una mirada retrospectiva hacia lo que hemos averiguado.

El enunciado subordinado, por lo general, no tiene por sentido ningún pensamiento, sino únicamente una parte de alguno y, en consecuencia, no tiene por referencia ningún valor veritativo. La razón consiste, o bien en que, en la subordinada, las palabras tienen su referencia indirecta, de modo que la referencia, y no el sentido de la subordinada, es un pensamiento, o bien en que la subordinada es in-

<sup>12</sup> Análogamente ocurre con «pero», «no obstante».

<sup>13</sup> Podría expresarse el pensamiento de nuestro enunciado también así: «o bien el sol aún no se ha levantado ahora, o el cielo está muy nublado», de donde se infiere cómo debe concebirse este tipo de conexión de enunciados.

completa debido a que hay en ella un componente que sólo alude indeterminadamente, de modo que únicamente junto con la principal puede expresar un pensamiento, y entonces, sin perjuicio de la verdad del todo, puede ser sustituida por otro enunciado del mismo valor veritativo, siempre y cuando no existan impedimentos gramaticales.

Si, después de lo dicho, se examinan todas las subordinadas especiales, se encontrarán pronto algunas que no podrán meterse en esas casillas. Por lo que alcanzo a ver, la razón de ello provendrá de que estas subordinadas tendrán un sentido nada simple. Parece que casi siempre a un pensamiento principal, que expresamos, asociamos pensamientos secundarios que también el oyente, a pesar de que no son expresados, une a nuestras palabras según leyes psicológicas. Y dado que así aparecen por sí mismos asociados a nuestras palabras, casi tanto como el propio pensamiento principal, también nosotros queremos expresar conjuntamente un pensamiento secundario semejante. Por ello se hace más rico el sentido del enunciado, y puede muy bien ocurrir que tengamos más pensamientos simples que enunciados. En algunos casos el enunciado debe ser entendido de este modo, mientras que en otros puede ser dudoso que el pensamiento secundario pertenezca realmente al sentido del enunciado, o bien sólo lo acompañe <sup>14</sup>. Así, podría quizás encontrarse que en el enunciado

«Napoleón, que se dio cuenta del peligro para su flanco derecho, dirigió él mismo sus guardias de corps contra la posición enemiga»,

no se han expresado únicamente los dos pensamientos antes mencionados, sino también el de que el darse cuenta del peligro fue la razón por la cual dirigió sus guardias de corps contra la posición enemiga. De hecho, puede dudarse de si este pensamiento sólo está ligeramente sugerido, o bien está realmente expresado. Se nos plantea la pregunta de si nuestro enunciado sería falso en el caso de que Napoleón hubiese tomado su decisión ya antes de percibir el peligro. Si, a pesar de esto, nuestro enunciado fuera verdadero, entonces nuestro pensamiento secundario no debería considerarse parte del sentido de nuestro enunciado. Probablemente nos decidiríamos por esto último. En el primer caso, la situación estaría bastante embrollada: tendríamos más pensamientos simples que enunciados. Si sustituimos también el enunciado

«Napoleón se dio cuenta del peligro para su flanco derecho»

<sup>14</sup> Esto puede ser de importancia para la cuestión de saber si una afirmación puede ser una mentira, o un juramento, un perjurio.

por otro del mismo valor veritativo, por ejemplo, por

«Napoleón tenía ya más de 45 años de edad»,

se habría alterado entonces no sólo nuestro primer pensamiento, sino también el tercero, y por ello podría también modificarse su valor veritativo —a saber, en el caso en que su edad no hubiese sido la razón de la decisión de dirigir sus guardias de corps contra el enemigo. A partir de esto puede comprenderse por qué no siempre en tales casos pueden reemplazarse mutuamente enunciados del mismo valor veritativo. Pues, justamente entonces, el enunciado, gracias a su conexión con otro, expresa más de lo que expresa por sí solo.

Consideremos ahora algunos casos en los que esto sucede regularmente. En el enunciado

«Bebel se imagina que, por medio de la devolución de Alsacia-Lorena, se podrán acallar los deseos de venganza de Francia»,

se expresan dos pensamientos, de los cuales, no obstante, no pertenece el uno al enunciado del principal y el otro al subordinado, a saber:

1. Bebel cree que, por medio de la devolución de Alsacia-Lorena, se podrán acallar los deseos de venganza de Francia;
2. Por medio de la devolución de Alsacia-Lorena no podrán acallarse los deseos de venganza de Francia.

En la expresión del primer pensamiento, las palabras de la subordinada tienen su referencia indirecta, mientras que esas mismas palabras, en la expresión del segundo pensamiento, tienen su referencia usual. Vemos, pues, que en nuestra estructura enunciativa original, la subordinada debe tomarse como doble, con distintas referencias, de las cuales una es un pensamiento y la otra un valor veritativo. Ahora bien, puesto que el valor veritativo no es toda la referencia del enunciado subordinado, no podemos sustituir sin más éste por otro del mismo valor veritativo. Análogamente ocurre con expresiones como «saber», «reconocer», «es sabido».

Con un enunciado subordinado sobre la razón y el correspondiente enunciado principal, expresamos varios pensamientos, que, sin embargo, no corresponden a cada uno de los enunciados aislados. En el enunciado

«como el hielo es menos denso que el agua, flota en el agua»

tenemos:

1. el hielo es menos denso que el agua;
2. si algo es menos denso que el agua, flota en el agua;
3. el hielo flota en el agua.

No era necesario quizá manifestar explícitamente el tercer pensamiento, al estar contenido en los dos primeros. Por el contrario, ni juntando el primero con el tercero, ni el segundo con el tercero, se obtendría el sentido de nuestro enunciado. Vemos, pues, que en nuestra subordinada

«como el hielo es menos denso que el agua»

se expresa tanto nuestro primer pensamiento como una parte del segundo. De ahí que no podamos, sin más, sustituir nuestra subordinada por otro enunciado del mismo valor veritativo; pues así también quedaría modificado nuestro segundo pensamiento y esto también podría fácilmente afectar su valor veritativo.

Análogamente aparece la cuestión en el enunciado

«si el hierro fuera menos denso que el agua, flotaría en el agua».

Aquí tenemos los dos pensamientos de que el hierro no es menos denso que el agua y de que algo flota en el agua si es menos denso que el agua. Nuevamente, la subordinada expresa un pensamiento y una parte del otro.

Si concebimos el enunciado antes analizado

«después de que Schleswig-Holstein se hubo separado de Dinamarca, se enemistaron Prusia y Austria»,

de forma que en él esté expresado el pensamiento de que Schleswig-Holstein se había separado alguna vez de Dinamarca, tenemos entonces, en primer lugar, este pensamiento, y en segundo lugar, el pensamiento de que en cierto momento, que queda algo más determinado por medio de la subordinada, Prusia y Austria se enemistaron. También en este caso expresa la subordinada no sólo un pensamiento, sino también una parte de otro. De aquí que, en general, no se pueda sustituir por otro enunciado del mismo valor veritativo.

Es difícil agotar todas las posibilidades dadas en el lenguaje; pero, con todo, espero haber hallado, en lo esencial, las razones por las que no siempre se puede sustituir una subordinada por otra del mismo valor veritativo, sin perjuicio de la verdad de la estructura enunciativa entera. Estas razones son

1. que la subordinada no se refiere a ningún valor veritativo, al expresar sólo una parte de un pensamiento;
2. que la subordinada se refiere ciertamente a un valor veritativo, pero no se limita a esto, al comprender su sentido, además de un pensamiento, una parte de otro pensamiento.

El primer caso se da

- a) en la referencia indirecta de las palabras,
- b) cuando una parte del enunciado alude sólo indeterminadamente, en vez de ser un nombre propio.

En el segundo caso, la subordinada puede tomarse como doble, a saber, una vez en su referencia habitual, la otra en su referencia indirecta; o bien, puede ser que el sentido de una parte de la subordinada sea, al mismo tiempo, componente de otro pensamiento, que juntamente con el sentido expresado directamente en la subordinada constituya el sentido total del enunciado principal y de la subordinada.

De todo esto resulta con suficiente probabilidad que los casos en que una subordinada no es sustituible por otra del mismo valor veritativo, no demuestran nada en contra de nuestra idea de que el valor veritativo es la referencia del enunciado, el sentido del cual es un pensamiento.

Volvamos a nuestro punto de partida.

Si, en general, encontramos que el valor cognoscitivo de « $a = a$ » y « $a = b$ » es distinto, esto se explica por el hecho de que, para el valor cognoscitivo, el sentido del enunciado, o sea el pensamiento expresado en él, no entra menos en consideración que su referencia, es decir, su valor veritativo. Ahora bien, si  $a = b$ , la referencia de « $b$ » es ciertamente la misma que la de « $a$ », y, por lo tanto, también el valor veritativo de « $a = b$ » es el mismo que el de « $a = a$ ». Sin embargo, el sentido de « $b$ » puede ser distinto el sentido de « $a$ », y con ello también será el pensamiento expresado en « $a = b$ » distinto del expresado en « $a = a$ »; pero entonces los dos enunciados tampoco tienen el mismo valor cognoscitivo. Si, como hemos hecho más arriba, por «juicio» entendemos el paso del pensamiento a su valor veritativo, también diremos entonces que los juicios son distintos.

## DESCRIPCIONES \*

BERTRAND RUSSELL

Hemos tratado en el capítulo precedente acerca de las palabras *todo* y *alguno*; en este capítulo consideraremos la palabra *el* (*la*) en el singular, y en el próximo capítulo consideraremos la palabra *el* (*la*) en el plural. Puede pensarse que es excesivo dedicar dos capítulos a una palabra, pero para el matemático filósofo es una palabra de enorme importancia: al igual que el gramático de Browning con la enclítica *de* daría la doctrina de esta palabra si estuviese «muerto de cintura para abajo» y no meramente en una prisión.

Hemos tenido ya ocasión de mencionar «funciones descriptivas», esto es: expresiones tales como «el padre de  $x$ » o «el seno de  $x$ ». Estas expresiones han de definirse definiendo primero «descripciones».

Una «descripción» puede ser de dos clases, definida o indefinida (o ambigua). Una descripción indefinida es una expresión de la forma «un tal-y-tal», y una descripción definida es una expresión de la forma «el tal-y-tal». Comencemos con la primera.

«¿Con quién te encontraste?» «Me encontré con un hombre.» «Esto es una descripción muy indefinida.» No estamos por lo tanto separándonos del uso habitual en nuestra terminología. Nuestra pregunta es: ¿Qué asevero realmente cuando asevero «Me encontré con un hombre»? Supongamos por el momento que mi aserción es verdadera, y que de hecho me encontré con Juan. Es claro que lo que asevero *no* es «Me encontré con Juan». Puedo decir «Me encontré con un hombre pero no era Juan»; en ese caso, aunque miento, no me contradigo a mí mismo, como lo haría si cuando digo que me encontré con un hombre quisiese decir realmente que me encontré con Juan. Es claro también que la persona a la que estoy hablando puede entender lo que digo, incluso si es un forastero y no ha oído hablar nunca de Juan.

Pero podemos ir más allá: no solamente Juan sino que tampoco ningún hombre efectivo entra en mi enunciado. Esto resulta obvio cuando el enunciado es falso, puesto que entonces no hay más razón por la que deba suponerse que Juan entra en la proposición que por la que deba suponerse que entra otra persona cualquiera. De hecho

---

\* Versión castellana de Luis Ml. Valdés Villanueva.

el enunciado continuaría siendo significativo, aunque no podría posiblemente ser verdadero, incluso si no hubiese ningún hombre en absoluto. «Me encontré con un unicornio» o «Me encontré con una serpiente marina» son aserciones perfectamente significativas, si sabemos lo que sería ser un unicornio o una serpiente marina, esto es: cuál es la definición de esos monstruos fabulosos. Así pues es solamente lo que llamamos el *concepto* lo que entra en la proposición. En el caso de, por ejemplo, «unicornio», lo único que hay es el concepto: no hay también, en algún lugar entre las sombras, algo irreal que pueda llamarse «un unicornio». Por lo tanto, puesto que es significativo (aunque falso) decir «Me encontré con un unicornio», es claro que esta proposición, correctamente analizada, no contiene como constituyente «un unicornio», aunque contiene el concepto «unicornio».

La cuestión de la «irrealidad» con la que nos enfrentamos en este punto, es una noción muy importante. Desorientados por la gramática, la gran mayoría de los lógicos que han tratado esta cuestión lo han hecho de acuerdo con procedimientos erróneos. Han considerado la forma gramatical como una guía de análisis más segura que lo que de hecho es. Y no han caído en la cuenta de qué diferencias en la forma gramatical son importantes. «Me encontré con Juan» y «Me encontré con un hombre» contarían tradicionalmente como proposiciones de la misma forma, pero de hecho son de formas completamente diferentes: la primera nombra a una persona efectiva, Juan; mientras tanto la segunda incluye una función proposicional y se convierte cuando se explicita en: «La función “Me encontré con  $x$  y  $x$  es humano” es algunas veces verdadera.» (Se recordará que adoptamos la convención de usar «algunas veces» como no implicando más de una vez.) Esta proposición no es obviamente de la forma «Me encontré con  $x$ », que da cuenta de la existencia de la proposición «Me encontré con un unicornio» a pesar del hecho de que no hay tal cosa como «un unicornio».

Debido a la falta del aparato de las funciones proposicionales, muchos lógicos se han visto empujados a la conclusión de que hay objetos irreales. Se argumenta, por ejemplo por parte de Meinong<sup>1</sup>, que podemos hablar sobre «la montaña de oro», «el cuadrado redondo» y así sucesivamente; podemos hacer proposiciones verdaderas de las cuales estas expresiones sean los sujetos; por consiguiente tienen que tener algún género de ser lógico, puesto que de lo contrario las proposiciones en las que ocurren carecerían de significado. En tales teorías, me parece, hay una falta de ese sentido de la realidad que debería preservarse incluso en los estudios más abstractos. La ló-

<sup>1</sup> *Untersuchungen zur Gegenstandstheorie und Psychologie*, 1904.

gica, mantendría yo, no tiene que admitir un unicornio en mayor medida que puede hacerlo la zoología, pues la lógica se interesa por el mundo real tan verdaderamente como la zoología, aunque por sus rasgos más abstractos y generales. Decir que los unicornios tienen una existencia en heráldica o en literatura o en la imaginación es una evasión sobremanera lastimosa y mezquina. Lo que existe en heráldica no es un animal, hecho de carne y hueso, que se mueve y respira por su propia iniciativa. Lo que existe es una figura o una descripción en palabras. Similarmente, mantener que Hamlet, por ejemplo, existe en su propio mundo, a saber, en el mundo de la imaginación de Shakespeare, tan verdaderamente como (pongamos por caso) Napoleón existió en el mundo ordinario, es decir algo deliberadamente confundente o, dicho de otro modo, confuso hasta un grado que es difícilmente creíble. Hay solamente un mundo, el mundo «real»: la imaginación de Shakespeare es parte de él, y los pensamientos que él tuvo al escribir Hamlet son reales. También lo son los pensamientos que tenemos al leer la obra. Pero forma parte de la esencia misma de la ficción el que solamente los pensamientos, sentimientos, etc., de Shakespeare y sus lectores sean reales, y que no haya, adicionalmente, a ellos un Hamlet objetivo. Cuando se han tomado en cuenta todos los sentimientos provocados por Napoleón en los escritores y lectores de historia, no se ha tocado para nada al hombre efectivo; pero en el caso de Hamlet, se ha llegado a sus entrañas. Si nadie pensase sobre Hamlet, no quedaría nada de él; si nadie hubiese pensado sobre Napoleón, pronto se vería lo que alguien (Napoleón) hizo. El sentido de la realidad es vital en lógica, y cualquiera que le haga trampas pretendiendo que Hamlet tiene otro género de realidad está haciendo un flaco servicio al pensamiento. Un sentido robusto de la realidad es muy necesario al construir un análisis correcto de proposiciones sobre unicornios, montañas de oro, cuadrados redondos, y otros pseudo objetos tales.

En obediencia al sentido de la realidad, insistiremos en que, en el análisis de las proposiciones, no ha de admitirse nada «irreal». Pero, después de todo, si no *hay* nada irreal ¿cómo, podemos preguntarnos, si *podríamos* admitir algo irreal? La réplica es que, al tratar de proposiciones, estamos tratando en primera instancia con símbolos, y si atribuimos significación a grupos de símbolos que no tienen significación alguna, caeremos en el error de admitir irrealidades, en el único sentido en que esto es posible, a saber: como objetos descritos. En la proposición «Me encontré con un unicornio», la totalidad de las cuatro palabras juntas constituyen una proposición significativa, y la palabra «unicornio» es significativa por sí misma, en justamente el mismo sentido que la palabra «hombre». Pero las *dos* palabras «un unicornio» no forman un grupo subordinado que tenga significado por sí mismo. Así pues, si atribuimos significado falsamente a esas



dos palabras nos encontramos a cuerdas con «un unicornio» y con el problema de cómo puede haber tal cosa en un mundo donde no hay unicornios. «Un unicornio» es una descripción indefinida que no describe nada. No es una descripción indefinida que describe algo irreal. Una proposición tal como « $x$  es irreal» sólo tiene significado cuando « $x$ » es una descripción definida o indefinida; en ese caso la proposición será verdadera si « $x$ » es una descripción que no describe nada. Pero el que la descripción « $x$ » describa algo o no describa nada, no es en ningún caso un constituyente de la proposición en la que aparece; igualmente «un unicornio», que hemos considerado hace un momento, no es un grupo subordinado que tenga significado por sí mismo. Todo esto resulta del hecho de que, cuando « $x$ » es una descripción, « $x$  es irreal» o « $x$  no existe» no es un sinsentido, sino que siempre son proposiciones significativas y algunas veces verdaderas.

Podemos ahora proceder a definir de manera general el significado de las proposiciones que contienen descripciones ambiguas. Supóngase que queremos hacer un enunciado sobre «un tal-y-tal», donde «tales-y-tales» son aquellos objetos que tienen una cierta propiedad  $\varphi$ , esto es: aquellos objetos  $x$  para los que la función proposicional  $\varphi x$  es verdadera. (Por ejemplo, si tomamos «un hombre» como nuestra instancia de «un tal-y-tal»,  $\varphi x$  será « $x$  es humano».) Digamos ahora que queremos aseverar la propiedad  $\psi$  de «un tal-y-tal», esto es: queremos aseverar que «un tal-y-tal» tiene aquella propiedad que  $x$  tiene cuando  $\psi x$  es verdadera. (Por ejemplo, en el caso de «Me encontré con un hombre»,  $\psi x$  será «Me encontré con  $x$ ».) Ahora bien, la proposición de que «un tal-y-tal» tiene la propiedad  $\psi$  no es una proposición de la forma « $\psi x$ ». Si lo fuera, «un tal-y-tal» tendría que ser idéntico a  $x$  para una  $x$  adecuada; y aunque (en un sentido) esto puede ser verdadero en algunos casos, no es ciertamente verdadero en un caso tal como el de «un unicornio». Es justamente este hecho, que el enunciado de que «un tal-y-tal» tiene la propiedad  $\psi$  no es de la forma  $\psi x$ , lo que hace posible que «un tal-y-tal» sea, en un sentido claramente definible, «irreal». La definición es como sigue:

El enunciado de que «un objeto que tiene la propiedad  $\varphi$  tiene la propiedad  $\psi$ »

significa:

«La aserción conjunta de  $\varphi x$  y  $\psi x$  no es siempre falsa».

Dentro de los límites de la lógica, esta es la misma proposición que podría expresarse por «algunos  $\varphi$  son  $\psi$ »; pero retóricamente hay una diferencia, puesto que en un caso hay una sugerencia de sin-

gularidad, y en el otro caso de pluralidad. Este no es, sin embargo, el punto importante. El punto importante es que las proposiciones que son verbalmente sobre «un tal-y-tal», cuando se las analiza correctamente, se encuentra que no contienen ningún constituyente representado por esta expresión. Y esta es la razón por la que tales proposiciones pueden ser significativas incluso cuando no hay tal cosa como un tal-y-tal.

La definición de *existencia*, tal como se aplica a las descripciones ambiguas, resulta de lo que se dijo al final del capítulo precedente. Decimos que «existen hombres» o que «existe un hombre» si la función proposicional « $x$  es humano» es algunas veces verdadera; y generalmente «un tal-y-tal» existe si « $x$  es tal-y-tal» es algunas veces verdadera. Podemos poner esto en otro lenguaje. La proposición «Sócrates es un hombre» es sin duda *equivalente* a «Sócrates es humano», pero no es exactamente la misma proposición. El *es* de «Sócrates es humano» expresa la relación de sujeto y predicado; el *es* de «Sócrates es un hombre» expresa identidad. Es una desgracia para el género humano el que haya elegido emplear la misma palabra «es» para esas dos ideas enteramente diferentes —una desgracia a la que, desde luego, un lenguaje lógico simbólico pone remedio—. La identidad en «Sócrates es un hombre» es identidad entre un objeto nombrado (aceptando «Sócrates» como un nombre, sujeto a puntualizaciones explicadas más adelante) y un objeto descrito ambiguamente. Un objeto descrito ambiguamente «existirá» cuando al menos una tal proposición es verdadera, esto es: cuando hay al menos una proposición verdadera de la forma « $x$  es un tal-y-tal», donde « $x$ » es un nombre. Es característico de las descripciones ambiguas (como opuesto a las definidas) que pueda haber un número cualesquiera de proposiciones verdaderas de la forma anterior —Sócrates es un hombre, Platón es un hombre, etc.—. Así, «existe un hombre» se sigue de Sócrates, de Platón o de cualquier otro. Respecto de las descripciones definidas, por otro lado, la forma de proposición correspondiente, a saber: « $x$  es el tal-y-tal» (donde « $x$ » es un nombre), puede solamente ser verdadera para un valor de  $x$  como máximo. Esto nos lleva al tema de las descripciones definidas, que han de definirse de una manera análoga a la empleada para las descripciones ambiguas, pero bastante más complicada.

Llegamos ahora al tema principal del presente capítulo, a saber: la definición de la palabra *el* (o *la*), en el singular. Un punto muy importante sobre la definición de «un tal-y-tal» se aplica igualmente a «el tal-y-tal»; la definición que ha de buscarse es una definición de proposiciones en las que aparece esta frase, no una definición de la frase misma aisladamente. En el caso de «un tal-y-tal», esto es bastante obvio: nadie podría suponer que «un hombre» era un objeto definido, que pudiese ser definido por sí mismo. Sócrates es un hom-

bre, Platón es un hombre, Aristóteles es un hombre, pero no podemos inferir que «un hombre» significa lo mismo que significa «Sócrates», y también lo mismo que significa «Platón» y también lo mismo que significa «Aristóteles», puesto que estos tres nombres tienen diferentes significados. No obstante, cuando hemos enumerado todos los hombres que hay en el mundo, no queda nada de lo que podamos decir: «Esto es un hombre, y no solamente eso, sino que es *el* “un hombre”, la entidad quintaesencial que es precisamente un hombre indefinido sin ser nadie en particular». Resulta, por supuesto, completamente claro, que cualquier cosa que hay en el mundo es definida: si esto es un hombre, esto es un hombre determinado y no otro cualquiera. Así pues no puede encontrarse en el mundo una entidad tal que «un hombre», como algo opuesto al hombre específico. Y de acuerdo con esto es natural que no definamos «un hombre» mismo, sino las proposiciones en las que aparece.

En el caso de «el tal-y-tal» esto es igualmente verdadero, aunque a primera vista menos obvio. Podemos mostrar que esto tiene que ser el caso, mediante una consideración de la diferencia entre un *nombre* y una *descripción definida*. Tómese la proposición «Scott es el autor de *Waverley*». Tenemos aquí un nombre «Scott», y una descripción «el autor de *Waverley*», que se asevera que se aplican a la misma persona. La distinción entre un nombre y todos los demás símbolos puede explicarse como sigue:

Un nombre es un símbolo simple cuyo significado es algo que puede aparecer solamente como sujeto, esto es: algo del género que, en el Capítulo XIII, definimos como un «individuo» o un «particular». Y un símbolo «simple» es aquel que no tiene parte alguna que sea símbolo. Así «Scott» es un símbolo simple, puesto que, aunque tiene partes (a saber: letras separadas), esas partes no son símbolos. Por otro lado, «el autor de *Waverley*» no es un símbolo simple, puesto que las palabras separadas que componen la expresión son partes que son símbolos. Si, como puede ser el caso, cualquier cosa que *parece* ser un «individuo» es en realidad susceptible de análisis posterior, tendremos que contentarnos con lo que puede llamarse «individuos relativos», que serán términos que, a través de todo el contexto en cuestión, no son analizados jamás y jamás aparecen de otra manera que como sujetos. Y en ese caso habremos de contentarnos, correspondientemente, con «nombres relativos». Desde el punto de vista de nuestro presente problema, a saber: la definición de las descripciones, este problema, el de si estos son nombres absolutos o solamente nombres relativos, puede ignorarse, puesto que atañe a diferentes estadios en la jerarquía de «tipos», mientras que hemos de comparar parejas tales como «Scott» y «el autor de *Waverley*», en las que ambos se aplican al mismo objeto, y no plantean el problema de los tipos. Por lo tanto, podemos por el momento tratar a los nombres

como si fueran absolutos; nada de lo que tendremos que decir depende de esta suposición, pero su expresión en palabras se abreviará un tanto mediante ella.

Tenemos, entonces, dos cosas para comparar: (1) un *nombre*, que es un símbolo simple, que designa directamente un individuo que es su significado y que tiene este significado por sí mismo, independientemente de los significados de todas las demás palabras; (2) una *descripción* que consta de varias palabras cuyos significados están ya fijados, y a partir de los cuales resulta cualquier cosa que haya de considerarse como el «significado» de la descripción.

Una proposición que contiene una descripción no es idéntica a aquello en lo que la proposición se convierte cuando la descripción se substituye por un nombre, incluso si el nombre nombra el mismo objeto que describe la descripción. «Scott es el autor de *Waverley*» es obviamente una proposición diferente de «Scott es Scott»: la primera es un hecho de la historia de la literatura, la segunda es una prograssada trivial. Y si ponemos cualquier otro distinto de Scott en lugar de «el autor de *Waverley*», nuestra proposición se convertirá en falsa y por lo tanto no es ya ciertamente la misma proposición. Pero, podría decirse, nuestra proposición es esencialmente de la misma forma que (pongamos por caso) «Scott es Sir Walter», en la que se dice que dos nombres se aplican a la misma persona. La réplica es que, si «Scott es Sir Walter» significa realmente «la persona llamada “Scott” es la persona llamada “Sir Walter”», entonces los nombres están siendo usados como descripciones; esto es: el individuo, en vez de ser nombrado, está siendo descrito como la persona que tiene ese nombre. Éste es un modo en que se usan los nombres frecuentemente en la práctica y, como regla general, no habrá nada en la fraseología que muestre si están siendo usados de esta manera o *como* nombres. Cuando se usa un nombre directamente, para indicar meramente aquello de lo que estamos hablando, no es parte del *hecho* aseverado, o de la falsedad si resulta que nuestra aserción es falsa: es meramente parte del simbolismo en el cual expresamos nuestro pensamiento. Lo que queremos expresar es algo que podría (por ejemplo) traducirse a un lenguaje foráneo; se trata de algo para lo cual las palabras efectivas son un vehículo, pero de lo que no son parte alguna. Por otra parte, cuando hacemos una proposición sobre «la persona llamada “Scott”», el nombre efectivo «Scott» entra en lo que estamos aseverando y no meramente en el lenguaje usado al hacer la aserción. Nuestra proposición será ahora una proposición diferente si lo substituímos por «la persona llamada “Sir Walter”». Pero en la medida en que estamos usando nombres *como* nombres, el que digamos «Scott» o el que digamos «Sir Walter» es tan irrelevante para lo que estamos aseverando como el hecho de que hablemos en inglés o en francés. Así pues, en la medida en que los nombres son usados *como* nombres «Scott es

Sir Walter» es la misma proposición trivial que «Scott es Scott». Esto completa la demostración de que «Scott es el autor de *Waverley*» no es la misma proposición que la que resulta de substituir «el autor de *Waverley*» por un nombre, sea cual fuere el nombre por el que se substituya.

Cuando usamos una variable, y hablamos de una función proposicional, digamos  $\varphi x$ , el proceso de aplicar enunciados generales sobre  $x$  a casos particulares consistirá en sustituir la letra « $x$ » por un nombre, suponiendo que  $\varphi$  es una función que tiene individuos para sus argumentos. Supóngase, por ejemplo, que  $\varphi x$  es «siempre verdadera»; sea, pongamos por caso, la «ley de identidad»,  $x = x$ . Entonces podemos substituir « $x$ » por cualquier nombre queelijamos, y obtendremos una proposición verdadera. Suponiendo por el momento que «Sócrates», «Platón» y «Aristóteles» son nombres (una suposición muy temeraria), podemos inferir a partir de la ley de identidad que Sócrates es Sócrates, Platón es Platón y Aristóteles es Aristóteles. Pero cometeríamos una falacia si intentásemos inferir, sin premisas adicionales, que el autor de *Waverley* es el autor de *Waverley*. Esto resulta de lo que acabamos de probar, a saber: que si en una proposición substituímos «el autor de *Waverley*» por un nombre, la proposición que obtenemos es diferente. Es decir, aplicando el resultado a nuestro presente caso: Si « $x$ » es un nombre, « $x = x$ » no es la misma proposición que «el autor de *Waverley* es el autor de *Waverley*», sin importar qué nombre pueda ser « $x$ ». Así pues, del hecho de que todas las proposiciones de la forma « $x = x$ » son verdaderas no podemos inferir sin más ayuda, que el autor de *Waverley* es el autor de *Waverley*. De hecho, las proposiciones de la forma «el tal-y-tal es el tal-y-tal» no son siempre verdaderas: es necesario que el tal-y-tal exista (un término que se explicará en breve). Es falso que el actual Rey de Francia es el actual Rey de Francia, o que el cuadrado redondo es el cuadrado redondo. Cuando substituímos un nombre por una descripción, las funciones proposicionales que son «siempre verdaderas» pueden convertirse en falsas, si la descripción no describe nada. No hay ningún misterio en esto tan pronto como nos damos cuenta (lo cual fue probado en el párrafo precedente) que cuando substituímos una descripción el resultado no es un valor de la función proposicional en cuestión.

Estamos ahora en posición de definir las proposiciones en las que aparece una descripción definida. La única cosa que distingue «el tal-y-tal» de «un tal-y-tal» es la implicación de unicidad. No podemos hablar de «el habitante de Londres», puesto que habitar en Londres es un atributo que no es único. No podemos hablar sobre «el actual Rey de Francia», puesto que no hay ninguno; pero podemos hablar sobre «el actual Rey de Inglaterra». Así pues, proposiciones sobre «el tal-y-tal» implican siempre las proposiciones correspondientes sobre

«un tal-y-tal», con el *addendum* de que no hay más que un tal-y-tal. Una proposición tal como «Scott es el autor de *Waverley*» podría no ser verdadera si *Waverley* no se hubiese escrito jamás, o si varias personas lo hubieran escrito; y no más podría serlo cualquier otra proposición resultante de una función proposicional  $x$  mediante la substitución de « $x$ » por «el autor de *Waverley*». Podemos decir que «el autor de *Waverley*» significa «el valor de  $x$  para el que “ $x$  escribió *Waverley*” es verdadera». Así, por ejemplo, la proposición «el autor de *Waverley* era escocés» incluye:

- (1) « $x$  escribió *Waverley*» no es siempre falsa;
- (2) «si  $x$  e  $y$  escribieron *Waverley*,  $x$  e  $y$  son idénticos» es siempre verdadera;
- (3) «si  $x$  escribió *Waverley*,  $x$  era escocés» es siempre verdadera.

Estas tres proposiciones, traducidas al lenguaje ordinario, enuncian:

- (1) al menos una persona escribió *Waverley*;
- (2) como máximo una persona escribió *Waverley*;
- (3) quienquiera que escribió *Waverley* era escocés.

Todas estas tres proposiciones están implicadas por «el autor de *Waverley* era escocés». Conversamente, las tres juntas (pero no dos de ellas) implican que el autor de *Waverley* era escocés. Por lo tanto las tres juntas pueden considerarse como definiendo lo que se quiere decir mediante la proposición «el autor de *Waverley* era escocés».

Podemos simplificar de alguna manera estas tres proposiciones. La primera y la segunda juntas son equivalentes a: «Hay un término  $c$  tal que “ $x$  escribió *Waverley*” es verdadera cuando  $x$  es  $c$  y falsa cuando  $x$  no es  $c$ .» En otras palabras: «Existe un término  $c$  tal que “ $x$  escribió *Waverley*” es siempre equivalente a “ $x$  es  $c$ ”.» (Dos proposiciones son «equivalentes» cuando las dos son verdaderas o las dos son falsas.) Tenemos aquí, para empezar, dos funciones de  $x$ , « $x$  escribió *Waverley*» y « $x$  es  $c$ », y formamos una función de  $c$  considerando la equivalencia de esas dos funciones de  $x$  para todos los valores de  $x$ ; procedemos entonces a aseverar que la función resultante de  $c$  es «algunas veces verdadera», esto es: que es verdadera para al menos un valor de  $c$ . (Obviamente no puede ser verdadera para más de un valor de  $c$ .) Estas dos condiciones juntas se definen como dando el significado de «el autor de *Waverley* existe».

Podemos definir ahora «el término que satisface la función  $\phi x$  existe». Ésta es la forma general de lo que lo anterior es un caso particular. «El autor de *Waverley*» es «el término que satisface la función “ $x$  escribió *Waverley*”». Y «el tal-y-tal» incluirá siempre referencia a

alguna función proposicional, a saber: aquella que define la propiedad que hace de una cosa un tal-y-tal. Nuestra definición es como sigue:

«El término que satisface la función  $\varphi x$  existe» significa:  
«Hay un término  $c$  tal que  $\varphi x$  es siempre equivalente a “ $x$  es  $c$ ”.»

Para definir «el autor de *Waverley* era escocés», tenemos todavía que tomar en cuenta la tercera de nuestras tres proposiciones a saber: «Quienquiera que escribiese *Waverley* era escocés.» Esto se satisfará añadiendo meramente que el  $c$  en cuestión ha de ser escocés. Así «el autor de *Waverley* era escocés» es:

«Hay un término  $c$  tal que (1) “ $x$  escribió *Waverley*” es siempre equivalente a “ $x$  es  $c$ ”, (2)  $c$  es escocés.»

Y de manera general: «el término que satisface  $\varphi x$  satisface  $\psi x$ » se define de modo que signifique:

«Hay un término  $c$  tal que (1)  $\varphi x$  es siempre equivalente a “ $x$  es  $c$ ”, (2)  $\psi c$  es verdadera.»

Ésta es la definición de las proposiciones en las que aparecen las descripciones.

Es posible tener un amplio conocimiento por lo que respecta a un término descrito, esto es, conocer muchas proposiciones concernientes a «el tal-y-tal», sin conocer efectivamente lo que es el tal-y-tal, esto es, sin conocer ninguna proposición de la forma « $x$  es el tal-y-tal», donde « $x$ » es un nombre. En una novela policíaca las proposiciones sobre «el hombre que cometió el crimen» se acumulan, con la esperanza de que al final bastarán para demostrar que fue A el que cometió el crimen. Podemos incluso llegar a decir que, en todo conocimiento tal que pueda ser expresado en palabras —con la excepción de «esto» y «aquello» y un puñado de otras palabras cuyo significado varía en ocasiones diferentes— ningún nombre, en sentido estricto, aparece como tal, sino que aquello que parece ser semejante a los nombres es realmente una descripción. Podemos significativamente preguntar si Homero existió, algo que no podríamos hacer si «Homero» fuese un nombre. La proposición «el tal-y-tal existe» es significativa, ya sea verdadera o falsa; pero si  $a$  es el tal-y-tal (donde « $a$ » es un nombre), las palabras « $a$  existe» carecen de significado. Es solamente de las descripciones —definidas o indefinidas— de las que puede aseverarse significativamente la existencia; pues si « $a$ » es un nombre, *tiene* que nombrar algo: lo que no nombra nada no es un nombre y,

por lo tanto, si intentó ser un nombre, es un símbolo desprovisto de significado, mientras que una descripción, como «el actual Rey de Francia», no se torna en incapaz de aparecer significativamente meramente sobre la base de que no describe nada, por la razón de que ella es un símbolo *complejo*, cuyo significado se deriva del de sus símbolos constituyentes. Así, cuando preguntamos si Homero existió, estamos usando la palabra «Homero» como una descripción abreviada: podemos reemplazarla por, pongamos por caso, «el autor de la *Iliada* y la *Odisea*». La misma consideración se aplica a casi todos los usos de lo que tienen el aspecto de asemejarse a nombres propios.

Cuando las descripciones aparecen en proposiciones, es necesario distinguir entre lo que pueden llamarse ocurrencias «primarias» y «secundarias». La distinción abstracta es como sigue. Una descripción tiene una ocurrencia «primaria» cuando la proposición en la que aparece resulta de substituir « $x$ » por la descripción en alguna función proposicional  $\phi x$ ; una descripción tiene una ocurrencia «secundaria» cuando el resultado de substituir  $x$  por la descripción en  $\phi x$  da solamente una *parte* de la proposición en cuestión. Un ejemplo hará esto más claro. Considérese «el actual Rey de Francia es calvo». Aquí «el actual Rey de Francia» tiene una ocurrencia primaria, y la proposición es falsa. Toda proposición en la que una descripción que no describe nada tiene una ocurrencia primaria es falsa. Pero considérese ahora «el actual Rey de Francia no es calvo». Esto es ambiguo. Si tomamos primero « $x$  es calvo», substituímos « $x$ » por «el actual Rey de Francia», y a continuación negamos el resultado, la ocurrencia de «el actual Rey de Francia» es secundaria y nuestra proposición es verdadera; pero si tomamos « $x$  no es calvo» y substituímos « $x$ » por «el actual Rey de Francia», entonces «el actual Rey de Francia» tiene una ocurrencia primaria y la proposición es falsa. La confusión entre ocurrencias primarias y secundarias es una fuente habitual de falacias por lo que se refiere a las descripciones.

Las descripciones aparecen principalmente en matemáticas en la forma de *funciones descriptivas*, esto es: «el término que tiene la relación  $R$  con  $y$ », o «el  $R$  de  $y$ » como podemos decir en analogía con «el padre de  $y$ » y frases similares. Decir, por ejemplo, «el padre de  $y$  es rico» es decir que la siguiente función proposicional de  $c$ : « $c$  es rico, y « $x$  engendró a  $y$ » es siempre equivalente a « $x$  es  $c$ » es «algunas veces verdadera», esto es, es verdadera para al menos un valor de  $c$ . Obviamente no puede ser verdadera para más de un valor.

La teoría de las descripciones, brevemente bosquejada en el presente capítulo, es de la mayor importancia tanto para la lógica como para la teoría del conocimiento. Pero para propósitos de la matemática, las partes más filosóficas de la teoría no son esenciales, y por lo tanto han sido omitidas en la anterior exposición, que se ha limitado a los requisitos matemáticos más escuetos.



## SOBRE EL REFERIR \*

P. F. STRAWSON

### I

Muy comúnmente usamos ciertos géneros de expresiones para mencionar o hacer referencia a alguna persona individual, a un objeto singular, a un evento particular, a un lugar o a un proceso, en el curso de lo que normalmente describiríamos como hacer un enunciado sobre esa persona, objeto, lugar, evento o proceso. Denominaré este modo de usar las expresiones, «uso referencial singularizador». Las clases de expresiones más comúnmente usadas de este modo son: pronombres demostrativos en singular («éste» y «ése»); nombres propios (por ejemplo, «Venecia», «Napoleón», «Juan»), pronombres personales e impersonales en singular («él», «ella», «yo», «tú», «ello») y frases que comienzan con el artículo determinado seguidas de un sustantivo, adjetivado o no, en singular (por ejemplo, «la mesa», «el hombre viejo», «el rey de Francia»). Cualquier expresión de alguno de estos géneros puede aparecer como sujeto de lo que tradicionalmente se consideraría como una oración en singular de sujeto-predicado y, si tal fuera el caso, ejemplificaría el uso que deseo discutir.

No quiero decir que las expresiones que pertenecen a estos géneros no tengan nunca otro uso que el que intento discutir. Por el contrario, resulta obvio que lo tienen. Es obvio que cualquiera que emitiera la oración «la ballena es un mamífero» estaría usando la expresión «la ballena» de una manera muy diferente de como la usaría alguien que tuviese la ocasión de emitir seriamente la oración «La ballena arremetió contra el barco». Obviamente, en la primera oración *no* se está mencionando una ballena particular, mientras que en la segunda oración es evidente que se la *está* mencionando. Es más, si hubiera dicho «Napoleón fue el más grande soldado francés», estaría usando la palabra «Napoleón» para mencionar a cierto individuo, pero no estaría usando la frase «el más grande soldado francés», para mencionar a un individuo, sino para decir algo sobre un individuo mencionado anteriormente. Resultaría natural decir que al usar esta

---

\* Versión castellana de Luis Ml. Valdés.

oración estaba hablando *sobre* Napoleón y que lo que estaba *diciendo* sobre él era que fue el más grande soldado francés. Pero, desde luego, *podríamos* usar la expresión «el más grande soldado francés» para mencionar a un individuo; por ejemplo, diciendo «el más grande soldado francés murió en el exilio». Por consiguiente, resulta obvio que al menos algunas expresiones pertenecientes a los géneros que he mencionado *pueden* tener usos diferentes del que deseo discutir. Tampoco pretendo decir que en una oración dada nunca hay más de una expresión usada en el sentido que propongo discutir. Por el contrario, es obvio que puede haber más de una. Por ejemplo, resultaría natural decir que, al usar seriamente la oración «La ballena arremetió contra el barco», se estaba diciendo algo sobre cierta ballena y sobre cierto barco, que se estaba usando cada una de las expresiones «la ballena» y «el barco» para mencionar un objeto particular; o, dicho de otra manera, que se estaba usando ambas expresiones en el sentido referencial singularizador. En general, sin embargo, limitaré mi atención a casos donde una expresión, usada en este sentido, aparece como sujeto gramatical de una oración.

Pienso que es correcto decir que la teoría de las descripciones de Russell, que trata de la última de las cuatro clases de expresiones que he mencionado más arriba (esto es, de expresiones de la forma «el tal-y-tal»), es aún ampliamente aceptada entre los lógicos por entender que proporciona una explicación correcta del uso de tales expresiones en el lenguaje ordinario. Deseo mostrar, en primer lugar, que esta teoría, entendida de tal manera, contiene algunos errores fundamentales.

¿A qué pregunta o preguntas sobre frases de la forma «el tal-y-tal» pretendía responder la teoría de las descripciones? Creo que al menos una de las preguntas puede ser ilustrada del siguiente modo. Supóngase que alguien emitiera ahora la oración «El rey de Francia es sabio». Nadie diría que la oración que se ha emitido carezca de significado. Todo el mundo estaría de acuerdo en que es significativa. Pero todo el mundo sabe también que en la actualidad no hay rey de Francia. Una de las preguntas que la teoría de las descripciones se proponía responder era ésta: ¿cómo puede una oración como «El rey de Francia es sabio» ser significativa, incluso si no existe nada que responda a la descripción que contiene, esto es, en este caso, nada que responda a la descripción «el rey de Francia»? Y una de las razones por las que Russell pensó que era importante dar una respuesta correcta a esta pregunta, fue el hecho de que considerase importante mostrar que era errónea otra respuesta que podría darse. La respuesta que pensó que era errónea, y a la que deseaba dar una alternativa, podría presentarse como la conclusión de uno de los argumentos falaces siguientes. Llamemos a la oración «El rey de Francia es sabio» la oración O. El primer argumento es entonces el siguiente:

1) La frase «el rey de Francia» es el sujeto de la oración O.  
Consecuentemente 2) si O es una oración significativa, O es una oración *sobre* el rey de Francia.

Pero 3) si no existe en ningún sentido un rey de Francia, la oración no es acerca de nada y, por consiguiente, no es acerca del rey de Francia.

Consecuentemente 4) puesto que O es significativa, debe existir (o subsistir) en algún sentido (en algún mundo) el rey de Francia.

Y el segundo argumento es como sigue:

1) Si O es significativa, es verdadera o falsa.

2) O es verdadera si el rey de Francia es sabio y falsa si el rey de Francia no es sabio.

3) Pero el enunciado de que el rey de Francia es sabio y el enunciado de que el rey de Francia no es sabio son igualmente verdaderos sólo si hay (en algún sentido, en algún mundo) algo que sea el rey de Francia.

Por tanto 4) puesto que O es significativa, entonces se sigue la misma conclusión anterior.

Estos son, obviamente, malos razonamientos y, como era de esperar, Russell los rechaza. El postular un mundo de entidades extrañas al cual pertenece el rey de Francia es una ofensa, dice Russell, contra «ese sentido de la realidad que debe preservarse incluso en los estudios más abstractos». El hecho de que Russell rechace estos razonamientos es, *sin embargo*, menos interesante que el alcance que, al rechazar su conclusión, concede al más importante de sus principios. Me referiré a la frase «el rey de Francia» como la frase D. Creo que las razones de Russell para rechazar estos dos argumentos pueden resumirse de la siguiente manera. El error surge, dice, de creer que D, que ciertamente es el sujeto *gramatical* de O, es también el sujeto *lógico* de O. Pero D no es el sujeto lógico de O. De hecho, aunque O tiene gramaticalmente un sujeto singular y un predicado, no es, lógicamente, una oración de sujeto-predicado. La proposición que expresa es un género complejo de proposición *existencial* parte de la cual podría ser descrita como una proposición «existencial singularizadora». Para mostrar la forma lógica de la proposición tendríamos que volver a escribir la oración de una forma gramatical lógicamente apropiada, de tal modo que la engañosa similitud de O con la oración que expresa una proposición de sujeto-predicado desaparezca y quedemos a salvo de argumentos tan malos como los bosquejados más arriba. Antes de volver a ocuparnos de los detalles del análisis de Russell de la oración O, observemos lo que su respuesta, tal como hasta aquí la he presentado, parece implicar. Su respuesta pa-

rece implicar que, en el caso de una oración que es similar a O en que 1) tiene gramaticalmente la forma de sujeto-predicado y 2) su sujeto gramatical no se refiere a nada, la única alternativa a su carencia de significado es que no tenga realmente (esto es, lógicamente) la forma sujeto-predicado, sino alguna otra forma completamente diferente. Y esto, a su vez, parece implicar que si existen algunas oraciones que sean genuinamente de la forma sujeto-predicado, entonces el hecho de que sean significativas, de que tengan significado, garantiza que *hay* algo a lo que el sujeto lógico (y gramatical) se refiere. Además, la respuesta de Russell parece implicar que existen tales oraciones. Pues, si bien es verdad que podemos ser inducidos a pensar erróneamente que O es de la forma sujeto-predicado por su similitud gramatical con otras oraciones, entonces con toda seguridad debe haber otras oraciones gramaticalmente similares a O que *son* de la forma sujeto-predicado. Para mostrar no sólo que la respuesta de Russell parece implicar estas conclusiones, sino también que aceptó al menos la primera de ellas, es suficiente considerar lo que dice sobre la clase de expresiones que él llama «nombres lógicamente propios» y que pone en contraposición con expresiones como D, que llama «descripciones definidas». De los nombres lógicamente propios Russell dice, o sus afirmaciones implican, lo siguiente:

1) Que ellos y sólo ellos pueden aparecer como sujeto de las oraciones que son genuinamente de la forma sujeto-predicado.

2) Que una expresión que intente ser un nombre lógicamente propio *carece de significado*, a menos que haya algún objeto singular al que la expresión represente, puesto que el *significado* de tal expresión es, precisamente, el objeto individual que designa. En consecuencia, la expresión, para que sea un nombre, *tiene que designar* algo.

Es fácil ver que si alguien cree estas dos proposiciones, entonces la única manera que tiene de salvar la significación de O es negar que sea una oración, lógicamente hablando, de la forma sujeto-predicado. Puede decirse de modo general que Russell sólo reconoce dos modos en los que pueden ser significativas las oraciones que parecen ser, según su estructura gramatical, acerca de alguna persona particular, objeto o evento individual:

1) El primero ha de ser aquel en que la forma gramatical de la oración sea desorientadora respecto de su forma lógica, y que haya de ser analizable al igual que O, como un género especial de oración existencial.

2) El segundo ha de ser aquel en que el sujeto gramatical de la oración sea un nombre lógicamente propio, cuyo significado es la cosa individual que designa.

Creo que Russell está totalmente equivocado en esto, y que las oraciones que son significativas y comienzan con una expresión unida de modo referencial singularizador no pertenecen a ninguna de estas dos clases. Las expresiones usadas de modo referencial singularizador no son nunca nombres lógicamente propios ni descripciones, si lo que se quiere decir al llamarlas «descripciones» es que deben analizarse de acuerdo con el modelo proporcionado por la teoría de las descripciones de Russell.

No existen nombres lógicamente propios ni descripciones (en este sentido).

Consideremos ahora los detalles del análisis de Russell. De acuerdo con él, cualquiera que asevere O aseveraría que:

- 1) Hay un rey en Francia.
- 2) No hay más que un rey en Francia.
- 3) No hay nada que sea rey de Francia y no sea sabio.

Es fácil darse cuenta de cómo llegó Russell a este análisis y de cómo esto le permite dar respuesta a la pregunta con que comenzábamos, a saber ¿cómo puede ser significativa la oración O cuando no hay rey de Francia? El modo en que llegó a este análisis fue claramente preguntándose cuáles serían las circunstancias en que diríamos que cualquiera que emitiera la oración O habría hecho una aserción verdadera. Y parece totalmente claro, y no tengo deseos de ponerlo en cuestión, que las oraciones anteriores 1)-3) describen las circunstancias que son al menos condiciones *necesarias* para cualquiera que haga una aserción verdadera al emitir la oración O. Pero, como espero mostrar, decir esto no es en absoluto lo mismo que decir que Russell ha dado una explicación correcta del uso de la oración O, o ni siquiera que ha dado una explicación que, aunque incompleta, es correcta hasta cierto punto; y ciertamente no es lo mismo que decir que el modelo de traducción presentado es un modelo correcto para todas (o para algunas) oraciones singulares que comienzan con una frase de la forma «el tal-y-tal».

Resulta también fácil advertir cómo este análisis capacita a Russell para responder a la pregunta de cómo la oración O puede ser significativa, incluso si no existe rey de Francia. Pues, si este análisis es correcto, cualquiera que emita hoy la oración O aseveraría a la vez tres proposiciones, una de las cuales (a saber, que hay un rey de Francia) sería falsa; y puesto que la conjunción de tres proposiciones, de las cuales una es falsa, es ella misma falsa, la aserción total sería significativa, pero falsa. De esta manera no podría aplicarse a tal aserción ninguno de los malos argumentos a favor de las entidades subsistentes.

## II

Quiero ahora plantear ciertas distinciones como paso previo dirigido a mostrar que la solución de Russell a este problema es errónea, y hacia el logro de la solución correcta. Con este propósito, y para abreviar, denominaré, en el resto de esta sección, a una expresión que tenga un uso referencial singularizador «una expresión», y a una oración que empiece con tal expresión «una oración». Las distinciones que plantearé son solamente aproximadas y expeditivas, y, sin duda, podrían presentarse casos difíciles que requieran su refinamiento. Con todo, creo que servirían para mis propósitos. Las distinciones se plantean entre:

- A<sub>1</sub>) una oración,
- A<sub>2</sub>) un uso de una oración,
- A<sub>3</sub>) una emisión de una oración,

y correspondientemente, entre:

- B<sub>1</sub>) una expresión,
- B<sub>2</sub>) un uso de una expresión,
- B<sub>3</sub>) una emisión de una expresión.

Consideremos de nuevo la oración «el rey de Francia es sabio». Es fácil imaginar que esta oración fue emitida en diversos momentos desde, digamos, el comienzo del siglo XVII en adelante, durante los reinados de los sucesivos monarcas franceses y también es igualmente fácil imaginar que fuese emitida en los períodos subsiguientes en que Francia no era una monarquía. Adviértase que me resultaba natural hablar de que «la oración», o de que «esta oración» fue emitida en diversos momentos durante este período; o, en otras palabras, que sería natural y correcto hablar de *una y la misma* oración que es emitida en todas estas diversas ocasiones. Usaré la expresión A<sub>1</sub>, «una oración», en el sentido en que sería correcto hablar de una y la misma oración que es emitida en todas esas diversas ocasiones. Existen, sin embargo, diferencias obvias entre las distintas *ocasiones de uso* de esta oración. Por ejemplo, si una persona la emitiera en el reinado de Luis XIV y otra la emitiese en el de Luis XV sería natural decir (suponer) que estaban hablando, respectivamente, de diferentes personas; y podría sostenerse que la primera persona, al usar la oración, hacía una aserción verdadera, mientras que la segunda, usando la misma oración, hacía una aserción falsa. Si, por otra parte, dos personas diferentes emitiesen simultáneamente la oración (por ejemplo, si una la escribiera y otra la dijera de viva voz) durante el reinado de Luis XIV, sería natural decir (suponer) que ambos esta-

ban hablando de la misma persona, y en ese caso, al usar la oración, o **lon** dos *tienen que* haber hecho una aserción verdadera o ambos *tienen que* haberla hecho falsa. Esto aclara lo que quiero decir con *un uso* de una oración. Las dos personas que emitieron la oración, una **on** el reinado de Luis XV y otra en el de Luis XIV, hicieron un uso *diferente* de la misma oración; mientras que las dos personas que **emitieron** la oración simultáneamente en el reinado de Luis XIV hicieron el mismo uso <sup>1</sup> de la misma oración. Obviamente, tanto en el **caso** de esta oración como en el de muchas otras, no podemos hablar **de** que *la oración* será verdadera o falsa, sino sólo de que se usa para **hacer** una aserción verdadera o falsa, o (si se prefiere) para expresar una proposición verdadera o falsa. Y es igualmente obvio que no podemos hablar de que *la oración* sea *acerca de* una persona particular, puesto que la misma oración puede usarse en momentos diferentes para hablar acerca de personas particulares completamente diferentes; sólo podemos hablar de *un uso* de la oración para hablar acerca **de** una persona particular. Por último, quedará suficientemente claro **»** qué me refiero cuando hablo de una emisión de una oración, si digo que las dos personas que la emitieron simultáneamente en el reinado de Luis XIV hicieron dos emisiones diferentes de la misma oración, aunque el mismo uso de ella.

Si en vez de considerar toda la oración «El rey de Francia es sabio» tomamos en cuenta aquella parte constituida por la expresión «el rey de Francia», resulta obvio que podemos hacer distinciones análogas, aunque no idénticas, entre 1) la expresión, 2) un uso de la expresión y 3) una emisión de la expresión. Las distinciones no serán idénticas; obviamente, no podemos decir correctamente que la expresión «el rey de Francia» se use para expresar una proposición verdadera o falsa, ya que en general sólo las oraciones pueden usarse verdadera o falsamente, y de modo similar, únicamente al usar una oración, y no al usar una expresión aislada, se puede hablar acerca de una persona particular. En cambio, diremos en este caso que la expresión *se usa para mencionar o hacer referencia* a una persona particular cuando se usa la oración para hablar acerca de ella. Pero, como es obvio en este caso y en muchos otros, no puede decirse que la *expresión* B<sub>1</sub> mencione, o haga referencia a algo, del mismo modo que no puede decirse que *la oración* sea verdadera o falsa. La misma expresión puede tener diferentes usos mencionadores, al igual que la misma oración puede usarse para hacer enunciados con distintos va-

<sup>1</sup> Esta «usanza» de uso es, desde luego, diferente de: *a*) la usanza corriente en que «uso» (de una palabra, frase, oración, particular) = (en líneas generales) «reglas de uso» = (en líneas generales) «significado»; y *b*) mi propia usanza en la frase «uso referencial singularizador de expresiones», en la que «uso» = (en líneas generales) «modo de usar».

lores de verdad. «Mencionar» o «hacer referencia» no son cosas que haga una expresión; son cosas que alguien puede hacer al usar una expresión. Mencionar o hacer referencia a algo es una característica de un *uso* de una expresión, del mismo modo que «ser acerca de» algo y verdad-o-falsedad son características de un *uso* de una oración.

Un ejemplo muy diferente puede ayudarnos a esclarecer más estas distinciones. Consideremos otro caso de una expresión que tiene un uso referencial singularizador, a saber, la expresión «yo», y consideremos la oración «Yo tengo calor». Innumerables personas pueden usar esta misma oración, pero es lógicamente imposible que dos personas diferentes hagan el mismo *uso* de esta oración o, si se prefiere, que la usen para expresar la misma proposición. La expresión «yo» puede usarse correctamente por (y sólo por) cualquier persona para referirse a sí misma. Decir esto es decir algo acerca de la expresión «yo»; es, en cierto sentido, dar su significado. Esta es la clase de cosas que pueden decirse acerca de las *expresiones*. Pero no tiene sentido decir de la *expresión* «yo» que se refiere a una persona particular. Esta es la clase de cosas que solamente puede decirse de un uso particular de la expresión.

Permítaseme usar «tipo» como abreviatura de «oración o expresión». No estoy diciendo, entonces, que haya oraciones y expresiones (tipos), y usos y emisiones de ellas, del mismo modo que hay barcos y zapatos y lacre. Estoy diciendo que no podemos decir *las mismas cosas* acerca de los tipos, usos de los tipos y emisiones de los tipos. Y el hecho es que hablamos acerca de los tipos; y esta confusión es fácil que surja al no advertir las diferencias entre lo que puede decirse acerca de éstos y lo que sólo puede decirse acerca de los *usos* de tipos. Tenemos tendencia a imaginar que hablamos de oraciones y expresiones cuando estamos hablando de usos de oraciones y expresiones.

Esto es lo que hace Russell. En su contra, de manera general, diré esto. El significado (al menos en un sentido importante) es una función de la oración o expresión; mencionar y hacer referencia, y verdad y falsedad, son funciones del uso de la oración o expresión. Dar el significado de una expresión (en el sentido en que estoy usando la palabra) es dar *directrices generales* para su uso para hacer referencia a, o mencionar, objetos o personas particulares; dar el significado de una oración es dar las *directrices generales* para su uso al hacer aserciones verdaderas o falsas. No es hablar acerca de una ocasión particular cualquiera del uso de la oración o expresión. El significado de una expresión no puede identificarse con el objeto al que nos referimos mediante su uso en una ocasión particular. El significado de una oración no puede identificarse con la aserción que hacemos mediante su uso, en una ocasión particular. Hablar sobre el sig-



nificado de una expresión u oración no es hablar sobre su uso en una ocasión particular, sino sobre las reglas, hábitos y convenciones que gobiernan su uso correcto, en todas las ocasiones, para hacer referencia o aseverar. De esta manera, la cuestión de si una oración o expresión es significativa o no nada tiene que ver con la cuestión de si la oración, emitida en una ocasión particular, está siendo usada o no, en una ocasión, para realizar una aserción verdadera-o-falsa, o si la expresión está siendo usada, en esta ocasión, para hacer referencia a algo o mencionarlo.

El error de Russell tuvo su origen en pensar que hacer referencia o mencionar, si es que ocurren de alguna manera, deben ser significar. No distinguió  $B_1$  de  $B_2$ ; confundió expresiones con su uso en un contexto particular y confundió, así, significar con mencionar, con hacer referencia. Si hablo de mi pañuelo, puedo, quizás, sacar de mi bolsillo el objeto al que me estoy refiriendo, pero no puedo sacar de mi bolsillo el significado de la expresión «mi pañuelo». Puesto que Russell confundió significar con mencionar, pensó que si existiesen expresiones que tuviesen un uso referencial singularizador, que fueran lo que parecen ser (esto es, sujetos lógicos) y no otra cosa disfrazada, su significado tenía que ser el objeto particular al que se referían en su uso. De ahí la enojosa mitología del nombre lógicamente propio. Pero si alguien me pregunta por el significado de la expresión «esto» —en tiempos candidato favorito de Russell para este *status*— no le presentaría el objeto al que acabo de referirme con el uso de la expresión, añadiendo que el significado de la palabra cambia cada vez que se la usa. Ni tampoco le presentaría todos los objetos a que me he referido, o podría haberme referido, con su uso. Explicaría e ilustraría las convenciones que rigen el uso de la expresión. En eso consiste dar el significado de la expresión, lo cual es muy diferente a dar (en cualquier sentido de dar) el objeto al cual se refiere, ya que la expresión misma no se refiere a nada, aunque pueda usarse en diferentes ocasiones para referirse a innumerables cosas. De hecho, existe en inglés un sentido de la palabra «mean \*» que se aproxima al de «indicar, mencionar, o hacer referencia a»; por ejemplo, cuando alguien dice (algo molesto) «I mean you», o cuando señalo y digo «That's the one I mean». Pero *the one I meant* es completamente diferente del significado de la expresión que he usado para hablar de él. En este sentido especial de «mean» son las personas las que significan, y no las expresiones. Las personas usan expresiones para referirse a cosas particulares. Pero el significado de una expresión no es el

\* Dado que ni el verbo castellano «significar», ni la expresión «querer decir» —que serían las traducciones naturales de *to mean*— tienen el sentido a que Strawson alude, preferí dejar en inglés en este caso las expresiones en las que aparece con dicho sentido. (N. del T.)

conjunto de cosas o la cosa singular a la que podemos referirnos correctamente con su uso; el significado es el conjunto de reglas, hábitos, convenciones, para su uso al hacer referencia.

Sucede lo mismo con las oraciones, incluso de modo más obvio. Todo el mundo sabe que la oración «La mesa está llena de libros» es significativa, y todo el mundo sabe también lo que significa. Pero, si pregunto «¿Acerca de qué objeto es esa oración?», estoy planteando una pregunta absurda, una pregunta que no puede plantearse acerca de la oración, sino sólo acerca de algún uso de la oración: y en este caso la oración no se ha usado para hablar acerca de algo, sólo se la ha tomado como un ejemplo. Al saber lo que significa se sabe como podría usarse correctamente para hablar sobre algo: así, conocer el significado no tiene nada que ver con conocer un uso particular de la oración para hablar acerca de algo. Del mismo modo, si pregunto «¿es la oración verdadera o falsa?» planteo una pregunta absurda, que no es menos absurda si añado: «Debe ser lo uno o lo otro, puesto que es significativa». La pregunta es absurda, puesto que la *oración* no es ni verdadera ni falsa, así como tampoco es *acerca de* algún objeto. Desde luego, el hecho de que sea significativa es lo mismo que el hecho de que *pueda* usarse correctamente para hablar de algo y que, al usarla así alguien hará una aserción verdadera o falsa. Y añadiré que se usará para hacer una aserción verdadera o falsa *sólo* si la persona que la usa *está* hablando de algo. Si, cuando la emite, no habla de nada, entonces su uso no es genuino, sino un uso espurio o pseudo-uso; no hace una aserción verdadera o falsa, aunque pueda pensar que la está haciendo. Y esto señala el camino hacia la respuesta correcta al problema al que la teoría de las descripciones da una respuesta fatalmente incorrecta. Lo importante es que la pregunta sobre si la oración es o no significativa es totalmente independiente de la pregunta que puede plantearse acerca de un uso particular de ella, a saber, la pregunta sobre si es un uso genuino o espurio, de si se usa para hablar de algo, o para fabular, o como un ejemplo filosófico. La pregunta sobre si la oración es o no significativa es la pregunta de si existen en ese lenguaje hábitos, convenciones, o reglas tales que la oración podría usarse lógicamente para hablar de algo; y por ello es totalmente independiente de la pregunta sobre si ha sido usada de tal manera en una ocasión particular.

### III

Consideremos de nuevo la oración «El rey de Francia es sabio» y las cosas verdaderas y falsas que Russell dice sobre ella.

Hay, al menos, dos cosas verdaderas que Russell diría sobre esta oración:

- 1) La primera de ellas es que es significativa; que si alguien la emitiese ahora estaría emitiendo una oración significativa.
- 2) La segunda es que si alguien la emitiese ahora haría una aseveración verdadera sólo si, actualmente, existiera de hecho uno y sólo un rey de Francia, y si fuera sabio.

¿Cuáles son las cosas falsas que Russell diría sobre la oración?  
Estas:

- 1) Que cualquiera que la emitiese en la actualidad estaría haciendo una aseveración verdadera o falsa.
- 2) Que parte de lo que estaría aseverando sería que en la actualidad existía uno y sólo un rey de Francia.

Ya he dado algunas razones para pensar que estos dos enunciados son incorrectos. Supongamos ahora que alguien te dijese, efectivamente, y con un aspecto completo de seriedad: «El rey de Francia es sabio». ¿Dirías: «Eso no es verdad»? Creo con toda seguridad que no lo dirías. Pero supongamos que la misma persona continuase preguntándote si pensabas que lo que acababa de decir era verdadero o falso, si estabas de acuerdo o no con lo que acaba de decir. Creo que te inclinarías a decir, aunque con alguna vacilación, que ninguna de las dos cosas; que la cuestión de si su enunciado era verdadero o falso simplemente *no se plantea*, puesto que no existe una persona tal que sea el rey de Francia. Podrías, si hablara completamente en serio (si tuviera el aspecto de andar un tanto extraviado por los siglos), decir algo semejante a: «Me temo que debes estar en un error. Francia no es una monarquía. No hay rey de Francia». Esto da pie a la cuestión de que, si una persona emitiese seriamente la oración, tal emisión sería en algún sentido una *evidencia* de que *creta* que había un rey de Francia. No sería una evidencia a favor de su creencia, en el sentido de que si alguien coge su impermeable esto es evidencia de su creencia de que está lloviendo. Pero tampoco sería una evidencia de su creencia en el sentido de que el hecho de que una persona diga «está lloviendo» es evidencia a favor de su creencia de que está lloviendo. Podríamos plantearlo del siguiente modo. Decir «El rey de Francia es sabio» es *implicar*, en algún sentido de «implicar», que hay un rey de Francia. Pero este es un sentido muy especial y extraño de «implicar». «Implica», en este sentido, no equivale, ciertamente, a «entraña» (o implica lógicamente). Y esto resulta del hecho de que, cuando, como respuesta a su enunciado, decimos (como podríamos hacerlo) «No hay ningún rey de Francia», *no* diríamos ciertamente que estábamos *contradiendo* el enunciado de que el rey de Francia es sabio. No estamos diciendo, por cierto, que es falso. Estamos más bien dando una razón para decir que la cuestión de si es verdadero o falso no se plantea.

Y es aquí donde la distinción que tracé antes puede ayudarnos. La oración «El rey de Francia es sabio» es ciertamente significativa pero esto no quiere decir que cualquier uso particular de ella sea verdadero o falso. La usamos verdadera o falsamente cuando la usamos para hablar acerca de alguien, cuando al usar la expresión «el rey de Francia» mencionamos en efecto a alguien. El hecho de que la oración y la expresión, respectivamente, sean significativas es precisamente el hecho de que la oración *podría* ser usada, en determinadas circunstancias, para decir algo verdadero o falso, y de que la expresión *podría* ser usada, en ciertas circunstancias, para mencionar a una persona particular; y conocer su significado es conocer qué clase de circunstancias son éstas. De esta manera, cuando emitimos la oración sin mencionar de hecho a nadie mediante el uso de la frase «El rey de Francia», la oración no deja de ser significativa; simplemente *no logramos* decir nada verdadero o falso porque no logramos mencionar a nadie mediante este uso particular de esta frase perfectamente significativa. Es, si se quiere, un uso espurio de la oración, aunque podamos (o no) creer erróneamente que se trata de un uso genuino.

Tales usos espurios <sup>2</sup> son muy familiares. El novelar y el mundo de ficción sofisticados <sup>3</sup> dependen de ellos. Si empezase diciendo «El rey de Francia es sabio», y continuase «y vive en un castillo de oro y tiene cien esposas», etc., un oyente me entendería perfectamente bien, sin suponer *o bien* que estaba hablando acerca de una persona particular, *o* que estaba haciendo un enunciado falso en el sentido de que existía una persona tal como la descrita por mis palabras. (Vale la pena añadir que allí donde el uso de oraciones y expresiones pertenece abiertamente al mundo de ficción, el sentido de las palabras «acerca de» puede cambiar. Como dijo Moore, es perfectamente natural y correcto decir que algunos de los enunciados de los *Pickwick Papers* son acerca del Sr. Pickwick. Pero donde el uso de oraciones y expresiones no pertenecen abiertamente al mundo de la ficción, este uso de «acerca de» parece menos correcto, esto es, no sería *en general* correcto decir que un enunciado era acerca del Sr. X. o de tal-y-tal, a menos que existiese tal persona o cosa. Sólo donde el novelar corre el riesgo de ser tomado en serio podríamos responder a la pregunta «¿acerca de quién está hablando?» con «No habla acerca de nadie»; pero al decir esto no estamos diciendo que lo que se decía era falso o sin sentido.)

<sup>2</sup> La elección de la palabra «espurio» me parece ahora desafortunada, al menos para algunos usos no estándar. Preferiría llamar ahora a algunos de ellos usos «secundarios».

<sup>3</sup> La fórmula no sofisticada comienza con «Érase una vez...».

Sin embargo, dejando aparte los usos que pertenecen abiertamente al mundo de ficción, he dicho que usar una expresión como «el rey de Francia» al comienzo de una oración era implicar, en algún sentido de «implicar», que había un rey de Francia. Cuando una persona usa esta expresión no *asevera*, ni lo que dice *entraña*, una proposición existencial singularizadora. Pero una de las funciones convencionales del artículo determinado es la de actuar como *señal* de que se hace una referencia singularizadora —una señal y no una aserción disfrazada—. Cuando empezamos una oración con «el tal-y-tal» el uso de «el» muestra, pero no enuncia, que nos referimos, o intentamos referirnos, a un individuo particular de la especie «tal-y-tal». *Cuál* sea el individuo particular es cuestión a determinar por el contexto, tiempo, lugar y otras características de la situación de emisión. Ahora bien, siempre que una persona usa una expresión, existe la presunción de que ella piensa que la está usando correctamente: así, cuando usa la expresión «el tal-y-tal» en un sentido referencial singularizador existe la presunción de que esa persona piensa que hay *algún* individuo de esa especie, y que el contexto de uso determinará de modo suficiente cuál es el individuo que tiene presente. Usar la palabra «el» en este sentido es implicar (en el sentido relevante de «implicar») que se cumplen las condiciones existenciales descritas por Russell. Pero usar «el» en este sentido no es enunciar que esas condiciones se cumplen. Si doy comienzo a una oración con una expresión de la forma «el tal-y-tal» y a continuación se me impide decir más, no he formulado enunciado alguno, pero puedo haber mencionado a algo o a alguien.

La aserción existencial singularizadora, que Russell supone que forma parte de cualquier aserción en la que se haga un uso referencial singularizador de una expresión de la forma «el tal-y-tal» es, observa Russell, un compuesto de dos aserciones. Decir que hay un  $\varphi$  es decir algo compatible con que haya varios  $\varphi$ ; decir que no hay más que un  $\varphi$  es decir algo compatible con que no haya ninguno. Decir que hay un  $\varphi$  y sólo uno es combinar estas dos aserciones. Hasta aquí me he centrado principalmente en la pretendida aserción de existencia, y me he ocupado en menor medida de la pretendida aserción de unicidad. Un ejemplo que ponga el énfasis sobre el último aspecto servirá para revelar más claramente el sentido de «implicada» en el que una aserción existencial singularizadora está implicada, aunque no *entrañada*, por el uso de expresiones de modo referencial singularizador. Consideremos la oración «La mesa está llena de libros». Es completamente cierto que en cualquier uso normal de esta oración, la expresión «la mesa» se usaría para hacer una referencia singularizadora, esto es, para referirse a una mesa particular. Es éste un uso muy estricto del artículo determinado, en el sentido en que Russell habla, en la página 30 de los *Principia Mathematica*, de

usar el artículo «*estrictamente*, de manera que implique unicidad». En la misma página Russell dice que una frase de la forma «el tal-y-tal» usada estrictamente, «sólo tendrá aplicación en el caso de que haya un tal-y-tal y no más». Ahora bien, es absolutamente falso que la frase «la mesa», en la oración «La mesa está llena de libros», cuando es usada normalmente, «tenga solamente aplicación en el caso que haya solamente una mesa y no más». Además es tautológicamente verdadero que, en tal uso, la frase tendrá aplicación sólo en el caso en que haya una mesa y no más *a la que se está haciendo referencia*, y que se entenderá que tiene aplicación sólo en el caso de que haya una mesa —que es a la que se entiende que se hace referencia mediante el uso de la frase— y no más. Usar la oración no es aseverar, sino implicar (en el sentido especial ya discutido) que hay sólo una cosa que es, *a la vez*, del género especificado (esto es, la mesa) y *a la que está haciendo referencia* el hablante. Obviamente, no es aseverar esto. Hacer referencia no es tampoco decir que se está haciendo referencia. Decir que hay *una u otra mesa* a la que se está haciendo referencia no es lo mismo que hacer referencia a una mesa particular. No tendríamos ningún uso para frases tales como «el individuo al que me refería» a menos que hubiese algo que se considerase como hacer referencia. (No tendría sentido decir que has señalado si no hubiese nada que se considerase como señalar.) Así, una vez más, llego a la conclusión de que hacer referencia o mencionar una cosa particular no puede disolverse en un género cualquiera de aserción. Hacer referencia no es aseverar, aunque hagamos referencia con el objeto de aseverar algo.

Permítaseme ahora tomar como ejemplo de uso referencial singularizador una expresión que no sea de la forma «el tal-y-tal». Supongamos que tiendo mis manos hacia alguien, poniéndolas cuidadosamente en forma de copa, y diciéndole a la vez que lo hago: «Esto es un rojo hermoso». El, al mirar mis manos y no ver nada en ellas, puede decir: «¿Qué es?, ¿de qué estás hablando?» O quizás: «¿Pero si no hay nada en tus manos!» Desde luego, sería absurdo decir que al emitir «¿Pero si no tienes nada en tus manos!», estuviera *negando* o *contradiendo* lo que dije. De esta manera, «esto» no es una descripción disfrazada en el sentido de Russell. Ni es un nombre lógicamente propio. Porque es necesario saber lo que significa la oración para reaccionar de esa manera ante su emisión. Precisamente a causa de que el significado de la palabra «esto» es independiente de cualquier referencia particular que pueda hacerse al usarla, aunque no independiente del modo en que puede usarse para hacer referencia, puedo, como en este ejemplo, usarla para  *fingir*  que estoy haciendo referencia a algo.

La moraleja general de todo esto es que la comunicación es un asunto de aserción explícita o disfrazada en bastante menor medida

de lo que los lógicos suelen suponer. La aplicación particular de esta moraleja general, en la que estoy interesado, es su aplicación al caso en que se hace referencia singularizadora. Una parte de la significación de las expresiones del género que estoy discutiendo consiste en que pueden usarse, en una gran variedad de contextos, para hacer referencias singularizadoras. No forma parte de su significación aseverar que están siendo usadas de tal modo o que satisfacen las condiciones de uso. Así, la distinción realmente importante que necesito trazar es entre:

- 1) utilizar una expresión para hacer una referencia singularizadora, y
- 2) aseverar que hay un individuo y sólo uno que tiene ciertas características (por ejemplo, que es de un cierto género o que está en cierta relación con el hablante, o ambas cosas).

Esta es, en otras palabras, la distinción entre:

- 1) oraciones que contienen una expresión usada para indicar o mencionar o hacer referencia a una persona o cosa particular, y
- 2) oraciones existenciales singularizadoras.

Lo que Russell hace es ir asimilando progresivamente más y más las oraciones de la clase 1) a las oraciones de la clase 2) y, por consiguiente, se ve envuelto en insuperables dificultades respecto a los sujetos lógicos y, en general, respecto a los valores de las variables individuales: dificultades que le llevaron por último a la teoría lógicamente desastrosa de los nombres desarrollada en *Inquiry into Meaning and Truth* y en *Human Knowledge*. Este punto de vista sobre el significado de las expresiones que son sujetos lógicos, y que constituye el principal incentivo de la teoría de las descripciones, hace imposible, al mismo tiempo, que Russell encuentre alguna vez unos sustitutos satisfactorios para estas expresiones que, comenzando con frases sustantivas, va degradando progresivamente de la condición de sujetos lógicos<sup>4</sup>. La raíz del problema no es simplemente, como a veces se dice, la fascinación que provoca la relación entre un nombre y su portador. Ni siquiera los nombres llegan al *standard* imposible, que se ha establecido. Se trata más bien de la combinación de dos concepciones erróneas más radicales: en primer lugar, el no lograr captar la importancia de la distinción (ver Sección II) entre lo que puede decirse de una expresión y lo que puede decirse de un uso par-

<sup>4</sup> Y esto a pesar de la señal de peligro que la frase «forma gramaticalmente *desorientadora*» constituye.

particular de ella; en segundo lugar, el no lograr reconocer que el uso referencial singularizador de las expresiones, inocuo y necesario, es distinto de, pero complementario del uso predicativo o adscriptivo de las mismas. Las expresiones que de hecho pueden aparecer como sujetos lógicos singulares son expresiones de la clase que enumeré al principio (demostrativos, frases sustantivas, nombres propios, pronombres): decir esto es lo mismo que decir que esas expresiones, junto con su contexto (en el sentido más amplio), son las que se usan para hacer referencias singularizadoras. La finalidad de las convenciones que rigen los usos de tales expresiones es, junto con la situación de emisión, el garantizar la unicidad de la referencia. Pero hacer esto basta y sobra. Mientras hacemos referencia no alcanzamos, ni podemos llegar a alcanzar, el grado de completa explicitud en el que la función referencial ya no se realiza. La referencia singularizadora que se ha hecho, si es que se hace, es asunto del uso particular en el contexto particular; la significación de la expresión usada es el conjunto de reglas y convenciones que permiten llevar a cabo tales referencias. De ahí que usando expresiones significativas podamos fingir referirnos a algo, en el mundo de la fábula o en el de la ficción, o pensar erróneamente que nos estamos refiriendo a algo cuando, en realidad, no nos estamos refiriendo a nada <sup>5</sup>.

Esto muestra la necesidad de que distingamos dos géneros (entre muchos otros) de convenciones o reglas lingüísticas: reglas para hacer referencia y reglas para atribuir y adscribir, así como la necesidad de que investiguemos las primeras. Si reconocemos esta distinción de uso como lo que es, estamos entonces en vías de resolver un buen número de viejos rompecabezas lógicos y metafísicos.

Mis dos últimas secciones tratarán, aunque sólo a grandes rasgos, de estas cuestiones.

#### IV

Uno de los principales propósitos para los que usamos el lenguaje es el de enunciar hechos acerca de cosas, personas y eventos. Si queremos cumplir este propósito debemos disponer de algún medio de anticiparnos a la pregunta «¿De qué (de quién, de cuál) está usted hablando?», y también a la de «¿Qué estás diciendo de ello (de él, de ella)?».

<sup>5</sup> Esta oración me parece ahora que está expuesta a numerosas objeciones, principalmente a causa del uso restrictivo no explícito de la expresión «hacer referencia». Podría ser más exactamente expresada de la manera siguiente: «De aquí que podamos, usando expresiones significativas, hacer referencias de manera secundaria, como en el fabular o en el mundo de ficción, o pensar erróneamente que nos estamos refiriendo a algo primariamente cuando no estamos refiriéndonos a nada de ese modo».



La función de anticiparse a la primera pregunta es la función referencial (o identificativa). La función de anticiparse a la segunda es la función atributiva (o descriptiva, o clasificativa, o adscriptiva). En la oración castellana convencional que se usa para enunciar, o afirmar que se enuncia, un hecho acerca de una cosa, persona o suceso individuales, la realización de estas dos funciones puede asignarse, de manera tosca y aproximada, a expresiones separables<sup>6</sup>. En tal oración, esta asignación de expresiones a sus papeles separados corresponde a la clasificación gramatical convencional de sujeto y predicado. No hay nada sacrosanto en el empleo de expresiones separables para estas dos tareas. Se podrían emplear, y se emplean, otros métodos. Existe, por ejemplo, el método de emitir una única palabra o una frase atributiva ante la presencia visible del objeto al que se hace referencia; o el método análogo ejemplificado por el cartel con las palabras «peligroso para camiones» sobre un puente, o por el atar a una calabaza el cartel en el que se lea «primer premio». También es posible imaginar un juego muy complicado en que nunca se usase una expresión de modo referencial singularizador, sino que sólo se emitiesen oraciones existenciales singularizadoras, intentando que el oyente fuera capaz de identificar aquello de que se está hablando por medio de una acumulación de cláusulas de relativo. (Esta descripción de los propósitos del juego muestra en qué sentido sería un juego: éste no es el uso corriente que hacemos de las oraciones existenciales). Hay dos aspectos que requieren ser subrayados. El primero es que la necesidad de realizar estas dos tareas para enunciar hechos particulares no requiere ninguna explicación trascendental: prestarle atención es elucidar, en parte, el significado de la frase, «enunciar un hecho». El segundo es que incluso esta elucidación se hace en términos derivados de la gramática de la oración singular convencional; que incluso la distinción lingüística, abiertamente funcional, entre los papeles identificativos y atributivos que las palabras pueden desempeñar en el lenguaje está motivada por el hecho de que el lenguaje ordinario nos ofrece expresiones separables a las que pueden ser asignadas, de manera plausible y aproximada, diferentes funciones. Y esta distinción funcional ha proyectado largas sombras filosóficas. Las distinciones entre particular y universal, entre substancia y cualidad, son ejemplos de sombras pseudomateriales proyectadas por la gramática de la oración convencional, en la que expresiones separables juegan papeles distinguibles<sup>7</sup>.

<sup>6</sup> Omito las oraciones relacionales puesto que no exigen, en principio, una modificación de lo que digo, sino una complicación de detalle.

<sup>7</sup> Lo que se dice o implica en las dos últimas oraciones de este párrafo no me parece ya verdadero a menos que se puntualice considerablemente.

Usar una expresión separada para realizar la primera de esas tareas es usar una expresión de modo referencial singularizador. Ahora quisiera decir algo en general acerca de las convenciones de uso para las expresiones usadas de este modo y contrastarlas con las del uso adscriptivo. Procedo, pues, a una breve ilustración de estas observaciones generales, y a efectuar algunas aplicaciones posteriores de ellas.

Lo que, en general, se requiere para hacer una referencia singularizadora es, obviamente, algún o algunos dispositivos para mostrar tanto *que* se intenta hacer una referencia singularizadora como *de qué* referencia singularizadora se trata; algún dispositivo que exija y capacite al oyente o al lector para identificar aquello de que se habla. Para garantizar este resultado, el contexto de emisión es de una importancia casi imposible de exagerar; y por «contexto» quiero decir, por lo menos, el tiempo, lugar, situación, identidad del hablante, los temas que constituyen el centro inmediato de interés y las historias personales del hablante y de aquéllos a los que se dirige. Además del contexto hay, desde luego, convenciones: convenciones lingüísticas. Pero excepto en el caso de los nombres propios genuinos, de los que diré algo más posteriormente, el cumplimiento de las condiciones contextuales, enunciadas con más o menos precisión, se exige *convencionalmente* (o *lógicamente*, en un sentido más amplio de la palabra) para el uso referencial correcto de las expresiones, en un sentido en el que no vale para los usos adscriptivos correctos. Para la correcta aplicación de una expresión en su uso adscriptivo a una cierta cosa se requiere, simplemente, que ésta sea de una cierta clase, que tenga ciertas características. El requisito para la aplicación correcta de una expresión en su uso referencial a cierta cosa es algo que es adicional a cualquier requisito derivado del significado adscriptivo que la expresión pueda tener; a saber: es el requisito de que la cosa esté en cierta relación con el hablante y con el contexto de emisión. Llamemos a esto el requisito contextual. De esta manera, por ejemplo, en el caso límite de la palabra «yo» el requisito contextual es que la cosa sea idéntica con el hablante; pero en el caso de la mayoría de las expresiones que tienen un uso referencial, este requisito no puede especificarse de un modo tan preciso. Una diferencia posterior, y completamente general, entre las convenciones para hacer referencia y las convenciones para describir, es aquella con la que ya nos hemos encontrado, a saber: que el cumplimiento de las condiciones para un uso adscriptivo correcto de una expresión es parte de lo que se enuncia por medio de tal uso; pero el cumplimiento de las convenciones para un uso referencial correcto de una expresión no es nunca parte de lo que se enuncia, aunque es implicado (en el sentido relevante de «implicado») por tal uso.

Las condiciones para hacer referencia han sido descuidadas o mal

interpretadas por los lógicos. Las razones de este descuido no son difíciles de ver, aunque sean difíciles de enunciar brevemente. Dos de ellas son, a grandes rasgos: 1) la preocupación de la mayoría de los lógicos por las definiciones; 2) la preocupación de algunos lógicos por los sistemas formales. 1) Una definición, en el sentido más familiar, es una especificación de las condiciones del uso adscriptivo o clasificatorio correcto de una expresión. Las definiciones no tienen en cuenta requisitos contextuales. De este modo, mientras que la búsqueda del significado, o la búsqueda del análisis de una expresión, se concibe como la búsqueda de una definición, el descuido o la mala interpretación de las convenciones diferentes de las adscriptivas, es inevitable. Quizá sería mejor decir (puesto que no quiero legislar sobre «significado» o «análisis») que los lógicos no han logrado advertir que los problemas del uso son más amplios que los problemas del análisis y del significado. 2) La influencia de la preocupación por las matemáticas y la lógica formal se ve más claramente (para no tomar ejemplos más recientes) en los casos de Leibniz y Russell. El constructor de cálculos no interesado en, ni obligado a, hacer enunciados fácticos, se acerca a la lógica aplicada con un prejuicio. Es natural que suponga que los tipos de convención, con cuya adecuación en un campo está familiarizado, deban ser realmente adecuados, ¡ojalá se pudiera ver cómo!, en un campo totalmente diferente, el de los enunciados/de hecho. De este modo vemos a Leibniz pugnando desesperadamente por hacer de la unicidad de las referencias singularizadoras una cuestión de lógica en sentido estricto, y a Russell pugnando desesperadamente para hacer lo mismo, aunque de manera diferente, tanto con la implicación de unicidad como con la de existencia.

Debe quedar claro que la distinción que intento trazar es principalmente una distinción entre los diferentes papeles que las expresiones pueden jugar en el lenguaje, y no una distinción entre diferentes grupos de expresiones, ya que algunas de ellas pueden aparecer en uno u otro papel. Algunos de los géneros de palabras sobre las que hablaré tienen predominantemente, si no exclusivamente, un papel referencial. Esto es completamente verdadero, por lo que respecta a los pronombres y los nombres propios ordinarios. Algunos pueden aparecer como todos o partes de expresiones que tienen un uso predominantemente referencial, y como todos o partes de expresiones que tienen predominantemente un uso adscriptivo o clasificatorio. Los casos más obvios son los nombres comunes, o los nombres comunes precedidos de adjetivos, incluidos los adjetivos o los adjetivos verbales (participiales); y, menos obviamente, los adjetivos o los adjetivos verbales (participiales) solos. Las expresiones que pueden tener un uso referencial difieren también entre sí, al menos en alguno de los tres modos siguientes, que no son mutuamente independientes.

1) Difieren en el grado en que la referencia que se hace con su uso depende del contexto de emisión. Palabras como «yo» y «ello» están situadas en un extremo de esta escala —el extremo de la máxima dependencia—, y frases como «el autor de *Waverley*» y «el decimotavo rey de Francia» están situadas en el otro.

2) Difieren en el grado de «significado descriptivo» que poseen: por «significado descriptivo» entiendo «la limitación convencional, en la aplicación a cosas de una cierta clase general, o que poseen ciertas características generales». En un extremo de esta escala están los nombres propios que usamos más comúnmente en el discurso ordinario; podemos llamar «Horacio» a hombres, perros y motocicletas. El puro nombre no tiene significado descriptivo (excepto el que pueda adquirir *como resultado* de alguno de sus usos como nombre). Una palabra como «el» tiene un significado descriptivo mínimo, pero tiene alguno. Frases substantivas como «la tabla redonda» tienen el máximo significado descriptivo. Una posición intermedia muy interesante es la ocupada por nombres propios «impuros» como «La Tabla Redonda», expresiones substantivas a las que les han salido letras mayúsculas.

3) Finalmente, pueden dividirse en las dos clases siguientes: a) aquellas cuyo uso referencial correcto está gobernado por algunas convenciones *generales* referenciales-*cum*-adscriptivas; b) aquellas cuyo uso referencial correcto no está gobernado por convenciones generales, sean de la clase contextual o adscriptiva, sino por convenciones *ad hoc* para cada uso particular (aunque no para cada emisión particular). A la primera clase pertenecen tanto los pronombres (que tienen el mínimo significado descriptivo) como las frases substantivas (que tienen el máximo). A la segunda clase pertenecen, hablando en líneas generales, la clase más familiar de los nombres propios. Desconocer el nombre de una persona no es desconocer el lenguaje. Esta es la razón de por qué no hablamos del significado de los nombres propios. (Pero esto no quiere decir que no tengan significado). De nuevo las frases tales como «el Viejo Pretendiente» \* ocupan una posición intermedia. De esta manera, sólo podemos hacer referencia a un viejo pretendiente, pero saber de qué viejo pretendiente se trata no es conocer una convención general, sino una convención *ad hoc*.

En el caso de las frases de la forma «el tal-y-tal» usadas referencialmente, el uso de «el», junto con la posición de la frase en la oración (esto es, al comienzo o a continuación de un verbo transitivo o de una preposición) actúa como señal *de que* se hace una referencia

---

\* Las frases «*The Old Pretender*» y «*The Young Pretender*» se usan para designar, respectivamente, al hijo y nieto de Juan II de Inglaterra, que mantuvieron sucesivamente sus derechos al trono británico contra la Casa de Hannover. (N. del T.)

singularizadora; y el nombre, o nombre y adjetivo, que le sigue, junto con el contexto de emisión, muestran *qué* referencia singularizadora me está haciendo. En general, la diferencia funcional entre los nombres comunes y adjetivos reside en que los primeros se usan, natural y comúnmente, de modo referencial, mientras que los últimos no se usan, común y naturalmente, de este modo, excepto cuando califican nombres, aunque pueden usarse, y de hecho se usan, solos. Y, desde luego, esta diferencia funcional no es independiente de la fuerza descriptiva peculiar de cada palabra. En general, debemos esperar que la fuerza descriptiva de los nombres sea tal que resulten ser los instrumentos más eficientes para la tarea de mostrar qué referencia singular se intenta hacer cuando tal referencia es señalada; y también debemos esperar que la fuerza descriptiva de las palabras que usamos natural y comúnmente para hacer referencias singulares refleje nuestro interés en las características relevantes, relativamente permanentes y de conducta, de las cosas. Estas dos expectativas no son independientes una de otra y, si prestamos atención a las diferencias entre la clase más corriente de los nombres comunes y la clase más corriente de adjetivos, encontraremos que ambas se satisfacen. Estas diferencias pertenecen al género que Locke presenta, de una manera original, cuando habla de que nuestras ideas de substancias son *colecciones* de ideas simples; cuando dice que «las facultades constituyen una gran parte de nuestras ideas de substancias», y cuando pasa a contrastar la identidad de la esencia real y nominal, en el caso de las ideas simples, con la carencia de identidad y el carácter fluctuante de la esencia nominal, en el caso de las substancias. La «substancia» misma es el gravoso tributo que Locke paga por su confuso conocimiento de la diferencia existente en la función lingüística predominante, que persiste incluso cuando el nombre ha sido expandido en una cadena más o menos indefinida de adjetivos. Russell repite el error de Locke con una diferencia cuando, admitiendo la inferencia de la sintaxis a la realidad, hasta el punto de percibir que puede desembarazarse de esta incógnita metafísica sólo si puede purificar totalmente el lenguaje de la función referencial, traza su programa para «abolir particulares»; un programa, de hecho, para abolir la distinción de uso lógico que me esfuerzo en subrayar.

El requisito contextual para el uso referencial de los pronombres puede enunciarse en algunos casos con la mayor precisión (por ejemplo, «yo» y «tú») y sólo con la mayor vaguedad en otros («ello» y «éste»). No me propongo decir nada más acerca de los pronombres, excepto señalar un síntoma adicional del fracaso en reconocer como tal el uso referencial singularizador; a saber: el hecho de que ciertos lógicos hayan tratado de elucidar la naturaleza de las variables presentando oraciones como «él está enfermo», «esto es verde», como ejemplos de algo parecido a lo que en el habla ordinaria se llama una

*función oracional*. Es cierto, desde luego, que la palabra «él» puede usarse en diferentes ocasiones para referirse a diferentes personas o animales; lo mismo pueden hacer la palabra «Juan» y la frase «el gato». Lo que disuade a tales lógicos de tratar estas dos expresiones como cuasi-variables es, en el primer caso, la persistente superstición, ya casi desaparecida, de que un nombre está lógicamente ligado a un individuo singular, y, en el segundo caso, el significado descriptivo de la palabra «gato». Pero «él», que tiene un amplio rango de aplicaciones y mínima fuerza descriptiva, sólo adquiere un uso como palabra referencial. Este hecho, junto con el fracaso en otorgar a las expresiones, usadas referencialmente, el lugar que les corresponde dentro de la lógica (el lugar dejado por el mítico «nombre lógicamente propio») es lo que explica el erróneo intento de elucidar la naturaleza de las variables haciendo referencia a palabras tales como «él», «ella», «ello».

Se dice a veces de los nombres propios ordinarios que son esencialmente palabras, cada una de las cuales se usa para referirse a un individuo. Esto es evidentemente falso. Muchos nombres personales ordinarios —los nombres *par excellence*— se usan correctamente para referirse a muchas personas. Un nombre personal ordinario es, a grandes rasgos, una palabra usada referencialmente, cuyo uso *no* está dictado por ningún significado descriptivo que pueda tener la palabra, *ni* está prescrito por ninguna regla general para su uso como expresión referencial (o como parte de una expresión referencial), tal como encontramos en el caso de palabras como «yo», «esto» y «el», sino que está gobernado por convenciones *ad hoc* para cada conjunto particular de aplicaciones de la palabra a una persona dada. La cuestión importante es que la corrección de tales aplicaciones no se sigue de ninguna regla o convención *general* para el uso de la palabra en cuanto tal. (El límite de lo absurdo y de la circularidad obvia se alcanza en el intento de tratar a los nombres como descripciones disfrazadas en el sentido de Russell, pues lo que es implicado, en el sentido especial, aunque no entrañado, por el hecho de que ahora haga referencia a alguien al nombrarlo, es simplemente la existencia de alguien a quien *ahora estoy haciendo referencia*, a quien *me refiero convencionalmente* mediante ese nombre). Sin embargo, incluso esta característica de los nombres es sólo un síntoma del propósito para el cual se emplean. Actualmente, nuestra elección de los nombres es parcialmente arbitraria y parcialmente dependiente de prácticas legales y sociales. Sería perfectamente posible tener un *sistema* completo de nombres basados, por ejemplo, en las fechas de nacimiento o en una clasificación minuciosa de las diferencias fisiológicas y anatómicas. Pero el éxito de tal sistema dependería enteramente de la adecuación de las asignaciones de nombres que resultarían con el propósito de hacer referencias singularizadoras; y esto dependería, a

nu vez, de la multiplicidad de las clasificaciones usadas y del grado en que atraviesan, fortuitamente, las agrupaciones sociales normales. Si se dan ambas cosas en un grado suficiente, la selectividad proporcionada por el contexto se encargaría del resto, al igual que ocurre con nuestros hábitos actuales de nombrar. Si tuviéramos tal sistema, podríamos utilizar palabras-nombre tanto descriptivamente (como hacemos en la actualidad, hasta cierto punto y en sentido diferente, con algunos nombres famosos) como referencialmente. Pero es mediante criterios derivados de la consideración de los requisitos de la tarea referencial, como valoraríamos la adecuación de cualquier sistema de nombrar. Desde el punto de vista del nombrar, ningún tipo de clasificación será mejor o peor que cualquier otra, simplemente por el tipo de clasificación —natal o anatómica— de que se trate.

Ya he mencionado la clase de los cuasi-nombres, de las frases substantivas a las que les han salido letras mayúsculas y de las que son ejemplos frases como «la Gloriosa», «la Gran Guerra», «la Anunciación», «la Tabla Redonda». Mientras que el significado descriptivo de las palabras que siguen al artículo determinado es todavía relevante para su papel referencial, las letras mayúsculas son un signo de esa selectividad extralógica en su uso referencial, que es característica de los nombres puros. Tales frases se encuentran impresas o escritas cuando algún miembro de una clase de eventos o cosas es de un interés excepcional en cierta sociedad. Estas frases son nombres embrionarios. Una frase puede, por razones obvias, entrar en, o salir de, esta clase (por ejemplo, «la Gran Guerra»).

## V

Quiero concluir considerando, muy brevemente, tres problemas adicionales que conciernen a los usos referenciales.

a) *Referencias indefinidas*. No todos los usos referenciales de expresiones singulares se anticipan a la pregunta «¿De qué (de quién, de cuál) estás hablando?». Existen algunos usos que invitan a formular esta pregunta, mientras que otros niegan la intención o capacidad para responderla. Ejemplos de esto son comienzos de oraciones tales como «Una persona me dijo que...», «Alguien me dijo que...». La doctrina ortodoxa (russelliana) sostiene que tales oraciones son existenciales, pero no existenciales singularizadoras. Esto parece ser erróneo por diversas razones. Es ridículo sugerir que parte de lo que se asevera es que la clase de los hombres o personas no es vacía. Esto está *implicado* ciertamente en el sentido de implicación que ya nos es familiar; pero esta implicación es también en la misma medida una implicación de la unicidad del objeto particular de referencia, como

cuando empiezo una oración con una frase como «la mesa». La diferencia en el uso de los artículos determinados e indeterminados es, a grandes rasgos, como sigue. Usamos «el» cuando se ha hecho una referencia previa y cuando «el» señala que se está haciendo la misma referencia; o cuando, ante la ausencia de una referencia indefinida previa, se espera que el contexto (incluyendo el conocimiento que se supone tiene el oyente) capacite al oyente para decir qué referencia definida se está haciendo. Usamos «un» cuando estas condiciones no se cumplen, o cuando, aunque podría hacerse una referencia definida, deseamos mantener desconocida la identidad del individuo a quien, o al que, nos estamos refiriendo. Este es el uso *elusivo* de frases como «una cierta persona» o «alguien» que podrían desarrollarse, no como «alguien, pero no sabrías (o yo no sé) quién», sino como «alguien, pero no te digo quién».

b) *Enunciados de identificación*. Me refiero, bajo esta etiqueta, a enunciados como los siguientes:

1a) Éste es el hombre que cruzó a nado el canal dos veces en el mismo día.

2a) Napoleón fue el hombre que ordenó la ejecución del duque de Enghien.

El problema de esos enunciados reside en que sus predicados gramaticales no parece que se usen de un modo exactamente adscriptivo como sucede con los predicados gramaticales de los enunciados:

1b) Este hombre cruzó a nado el canal dos veces en un día.

2b) Napoleón ordenó la ejecución del duque de Enghien.

Pero sí, para evitar borrar las diferencias entre 1a y 1b, y 2a y 2b, se dice que las frases que forman los complementos gramaticales de 1a y 2a se usan de un modo referencial, nos quedaríamos perplejos sobre lo que se dice en esas oraciones. Parece entonces que estamos haciendo referencia dos veces a la misma persona y, una de dos, o no decimos nada de ella, y de este modo no hacemos ningún enunciado, o lo identificamos consigo mismo produciendo así una identidad trivial.

El espantajo de trivialidad puede descartarse. Sólo se plantea para aquellos que piensan que el objeto al que hacemos referencia mediante el uso de una expresión es su significado, y de este modo piensan que el sujeto y el complemento de estas oraciones tienen el mismo significado, porque podrían usarse para hacer referencia a la misma persona.

Creo que las diferencias entre las oraciones del grupo a y las del



grupo *b* pueden entenderse mejor si consideramos las diferencias entre las circunstancias en las que diríamos *1a* y aquéllas en las que diríamos *1b*. Diríamos *1a* en lugar de *1b* si supiésemos o creyésemos que nuestro oyente sabía o creía que *alguien* había cruzado el canal a nado dos veces en un día. Decimos *1a* cuando consideramos que nuestro oyente está en la posición del que puede preguntar «¿Quién cruzó el canal a nado dos veces en un día?» (y al preguntar eso no está diciendo que alguien lo hiciera, aunque su pregunta implica —en el sentido relevante— que alguien lo hizo). Estas oraciones son especie de respuestas a tales preguntas. Es mejor llamarlas «enunciados de identificación» que «identidades». La oración *1a* no asevera ni más ni menos que la oración *1b*. Se trata solamente de que decimos *1a* a quien consideramos que conoce ciertas cosas que consideramos que desconoce la persona a la que decimos *1b*.

Esta es, en lo más esencial, la solución al problema de Russell acerca de las «frases denotativas» unidas por «es»; uno de los problemas cuya solución afirma ser mérito de la teoría de las descripciones.

c) *La lógica de sujetos y predicados*. Muchas de las cosas que he dicho acerca del uso referencial singularizador de las expresiones pueden extenderse, con las modificaciones adecuadas, al uso referencial no singularizador de expresiones, esto es, a algunos usos de expresiones que contienen «los», «todos los», «algunos», «algunos de los», etc., seguidas de un nombre, adejetivado o no, en plural; a algunos usos de expresiones que contienen «ellos», «de ellos», «éstos», «ésos» y a conjunciones de nombres. Las expresiones de la primera clase tienen un interés especial. En general, la crítica ortodoxa moderna, inspirada por la lógica matemática, de doctrinas tradicionales como la del cuadrado de oposición y de algunas formas del silogismo tradicionalmente reconocidas como válidas descansa en el fracaso habitual en reconocer el sentido especial en que las aserciones existenciales pueden estar implicadas por el uso referencial de las expresiones. Se dice que a las proposiciones universales del cuádruple esquema se les debe dar o una interpretación existencial negativa (por ejemplo, para *A* «no existen  $X_s$  que no sean  $Y_s$ »), o deben interpretarse como conjunciones de enunciados existenciales negativos y afirmativos de la forma, por ejemplo (para *A*), «no existen  $X_s$  que no sean  $Y_s$  y existen  $X_s$ ». A las formas *I* y *O* normalmente se les da una interpretación positivamente existencial. Entonces se ve que cualquiera de las opciones anteriores que se escoja lleva aparejado el rechazo de algunas leyes tradicionales. El dilema, sin embargo, es aparente. Si no interpretamos las proposiciones del esquema ni como positiva, ni como negativamente existenciales, por una parte, y ni como positiva y negativamente existenciales, por otra, sino como oraciones tales que *la cuestión de si se las usa para hacer aserciones verdaderas o*

*falsas no se plantea, excepto cuando la condición existencial se cumple por el término sujeto, entonces son válidas todas las reglas tradicionales. Y esta interpretación está mucho más cerca de la mayoría de los usos corrientes de expresiones que comienzan con «todo» y «alguno» que cualquier alternativa russelliana, ya que esas expresiones se usan corrientemente de modo referencial. Si a una persona que toma las cosas al pie de la letra y que no tiene hijos se le pregunta si todos sus hijos están durmiendo, ciertamente no responderá «Sí», dado que no tiene ninguno, pero tampoco responderá «No» sobre esta misma base. Ya que no tiene hijos la cuestión no se plantea. Decir esto no es decir que no pueda usar la oración «Todos mis hijos están dormidos» con la intención de engañarte haciéndote pensar que los tengo. Ni supone un debilitamiento de mis tesis el conceder que las frases singulares de la forma «el tal-y-tal» puedan a veces usarse con un propósito similar. Ni las reglas aristotélicas ni las russellianas dan cuenta de la lógica exacta de cualquier expresión del lenguaje ordinario, porque el lenguaje ordinario no tiene lógica exacta.*

## NOMBRES PROPIOS Y DESCRIPCIONES \*

J. R. SEARLE

El *status* de los nombres propios ha confundido siempre a los filósofos desde los días de Platón. Parece no haber duda de que adjetivos y nombres comunes como «rojo» y «mesa» tienen sentido o significado, y, obviamente, lo mismo parece sostenerse para las llamadas descripciones definidas como «la flor roja» o «el hombre que está al lado de la mesa». Pero ¿qué decir sobre los nombres propios como «Winston Churchill» y «San Francisco»? ¿Tienen sentido del mismo modo que adjetivos, nombres comunes y descripciones definidas? En la historia de la filosofía las respuestas a esta pregunta han sido cruciales para responder a la cuestión general de cómo las palabras se relacionan con el mundo.

### TEORÍAS DEL SENTIDO Y DEL NO-SENTIDO DE LOS NOMBRES PROPIOS

De acuerdo con una teoría largamente sostenida, los nombres propios simplemente representan objetos sin tener ningún otro sentido o significado que el de representar objetos. Una formulación temprana de esta teoría está en el *Teeteto* de Platón, y las más sofisticadas versiones modernas de esta idea están en el *Tractatus Logico-Philosophicus* de Wittgenstein y en *Philosophy of Logical Atomism* de Russell. Según Wittgenstein, el significado de un nombre propio es simplemente el objeto por el que está. Tal vez la formulación más famosa de esta teoría del no-sentido de los nombres propios es la afirmación de Mill de que los nombres propios tienen denotación pero no connotación. Para Mill un nombre común como «caballo» tiene tanto una connotación como una denotación; connota aquellas propiedades que serían especificadas en una definición de la palabra «caballo» y denota todos los caballos. Pero un nombre propio sólo denota a su portador.

La anterior es una famosa y atractiva teoría de los nombres propios pero hay ciertas notorias dificultades en ella. Una es que algunas

---

\* Versión castellana de José Fillol y Enrique Ujaldón.

veces uno encuentra los nombres propios en enunciados de identidad, enunciados de la forma «*a* es idéntico a *b*». Como Frege señaló, si los nombres propios simplemente representan objetos y nada más, ¿cómo podrían tales enunciados transmitir siquiera información fáctica? Si interpretamos tales enunciados como siendo solamente sobre el referente de los nombres, parece entonces que deben ser triviales puesto que si son verdaderos dicen sólo que un objeto es idéntico consigo mismo. Si por otra parte interpretamos los enunciados como dando información acerca de los nombres, entonces parece que deben ser arbitrarios, puesto que podemos asignar cualquier nombre que deseemos a un objeto. La solución de Frege fue argüir que, además de los nombres y los objetos a los que refieren, debemos distinguir un tercer elemento, el sentido (*Sinn*) del nombre en virtud del cual y sólo en virtud del cual se refiere al objeto. En el enunciado «La estrella de la tarde es idéntica a la estrella de la mañana», las expresiones «la estrella de la tarde» y «la estrella de la mañana» tienen la misma referencia pero diferentes sentidos. El sentido proporciona el modo de presentación (*Art des Gegebenseins*) del objeto; el objeto está, por así decirlo, iluminado desde un lado (*einseitig beleuchtet*) por el sentido de la expresión, y es gracias a que las dos expresiones tienen diferentes sentidos por lo que el enunciado puede transmitirnos información fáctica. Lo que el enunciado transmite es que uno y el mismo objeto tiene dos conjuntos diferentes de propiedades especificadas por los dos sentidos diferentes de los dos nombres, y así un enunciado tal puede ser un enunciado de hecho y no una mera trivialidad o una decisión verbal arbitraria. Todos los nombres propios, para Frege, tenían sentido del mismo modo que lo tienen las expresiones «la estrella de la mañana» y «la estrella de la tarde».

Esto presenta una imagen completamente diferente de los nombres propios respecto de la teoría clásica del no-sentido. Según la teoría clásica, los nombres, si son realmente nombres, tienen necesariamente una referencia y no tienen en absoluto sentido. Según la teoría fregeana, esencialmente tienen un sentido y sólo contingentemente referencia. Hacen referencia si y sólo si hay un objeto que satisface su sentido. En la primera teoría los nombres propios son *sui generis*, y realmente para Platón (en el *Teeteto*) y Wittgenstein (en el *Tractatus*) son el vínculo conector especial entre las palabras y el mundo; en la segunda teoría los nombres propios son sólo una especie de descripciones definidas disfrazadas: cada una es equivalente en significado a una descripción definida, a la descripción definida que da una formulación explícita de su sentido. De acuerdo con la primera teoría, nombrar es anterior a describir; de acuerdo con la segunda, describir es anterior a nombrar, porque un nombre sólo nombra describiendo el objeto que nombra.

## PROS Y CONTRAS DE LAS TEORÍAS

Hay un franco conflicto entre la teoría clásica del no-sentido y la teoría fregeana del sentido y la referencia. Parte del encanto del problema presentado por este conflicto es que cada uno tiene interesantes argumentos en su favor.

El sentido común parece inclinarnos hacia la teoría del no-sentido, al menos por lo que respecta a la mayoría de los nombres propios no son equivalentes a descripciones definidas porque, por ejemplo, llamar a un objeto por su nombre no es un modo de describirlo. Nombrar es una preparación para describir, no un modo de describir. Además, no tenemos definiciones de la mayoría de los nombres propios; las entradas de los diccionarios para los nombres propios usualmente ofrecen enunciados de hechos contingentes que describen el objeto referido por el nombre. Las descripciones no son equivalentes definicionales del nombre, porque son sólo contingentemente verdaderas del portador. Pero el nombre no es «verdadero de» el portador, es su nombre.

No sólo no tenemos equivalentes definicionales para los nombres propios, sino que tampoco está en absoluto claro cómo podríamos intentar obtener definiciones de nombres propios si las deseásemos. Si, por ejemplo, intentásemos presentar una descripción completa del objeto como el sentido del nombre, sobrevendrían consecuencias extrañas —por ejemplo, cualquier enunciado verdadero acerca del objeto que usase el nombre como sujeto sería analítico y cualquiera falso sería autocontradictorio—. El significado de un nombre (y tal vez la identidad del objeto) cambiaría cada vez que hubiese un cambio en el objeto, y el mismo nombre tendría diferentes significados para los diferentes usuarios del nombre.

Tales consideraciones del sentido común pesan en alguna medida a favor de la teoría del no sentido; sin embargo, ésta también presenta serias dificultades. Primero, como ha sido mostrado, no puede dar cuenta de la ocurrencia de nombres propios en enunciados informativos de identidad. Segundo, de modo similar es incapaz de explicar la ocurrencia de los nombres propios en enunciados existenciales. En enunciados tales como «Hay un lugar como África» y «Cerberos no existe», los nombres propios no puede decirse que tengan referencia, pues ningún sujeto de un enunciado existencial puede tenerla. Si lo hiciese, la precondition de que tuviese un valor de verdad garantizaría su verdad si fuese afirmativo y su falsedad si fuese negativo. (Éste es otro modo de decir que «existe» no es un predicado.) Todo enunciado existencial afirmativo establece en efecto que un cierto concepto o predicado es instanciado. (Tal como Frege lo expresó, la existencia es un concepto de segundo orden.) Un enunciado existencial

afirmativo no se refiere a un objeto y establece que existe; más bien expresa un concepto y establece que ese concepto es instanciado.

Así, si un nombre propio aparece en un enunciado existencial, parece que debe tener algún contenido conceptual o descriptivo. Pero, si tiene un contenido descriptivo, entonces parece que la teoría de Frege debe ser correcta, porque ¿qué podría ser ese contenido descriptivo salvo el sentido del nombre propio? De este modo, la ocurrencia de los nombres propios en enunciados existenciales añade otra grave dificultad para los teóricos del no-sentido. Sin embargo, lo peor aún no ha llegado.

¿Qué explicación puede dar el teórico del no-sentido de la existencia del objeto al que se refiere un nombre propio? Si uno conviene con el Wittgenstein del *Tractatus* en que el significado de un nombre propio es literalmente el objeto por el que está, entonces parece que la existencia de esos objetos que son nombrados por nombres propios genuinos no puede ser un hecho contingente ordinario. La razón de esto es que cambios tales en el mundo como la destrucción de algunos objetos no pueden destruir el significado de las palabras, porque cualquier cambio en el mundo debe no obstante ser describable en palabras. Pero esto parece forzarnos a la posición de que hay una clase de objetos en el mundo cuya existencia es de algún modo necesaria, aquellos objetos que son los significados de los nombres propios reales. Realmente, parece, si aceptamos esta idea, que no podría tener ningún sentido afirmar o negar la existencia de los objetos nombrados por nombres propios genuinos. Como Platón observó, no podemos decir de un elemento que existe o que no existe (*Teeteto* 201 D - 202 A).

Hay al menos dos modos de tratar este problema de la existencia del referente, un modo metafísico y un modo lingüístico. En el *Tractatus*, Wittgenstein adopta una conclusión metafísica, diciendo que los objetos forman la sustancia del mundo (2.021). Su existencia no puede ser ni aseverada ni negada. En su libro *An Introduction to Wittgenstein's Tractatus*, G.E.M. Anscombe adopta una salida lingüística al problema. Dice simplemente que debemos distinguir los nombres propios genuinos de los aparentes. Sólo las expresiones que tienen portadores son nombres propios genuinos. Es, en efecto, una verdad necesaria que un nombre propio genuino tiene un portador, pero esto no nos fuerza a aceptar la idea de que haya una clase de objetos que tengan una existencia necesaria; más bien conduce a la idea de que, el que una expresión sea un nombre propio o no, depende de si su pretendido referente existe realmente o no (o al menos ha existido o existirá). Así pues, si descubrimos que César nunca existió, probaríamos en efecto que «César» no era un nombre propio genuino.

Una consecuencia de la doctrina wittgensteniana, aceptada y elucidada por Russell en su período del atomismo lógico, es que lo que

En el lenguaje ordinario consideramos que son nombres propios no lo son en absoluto porque la existencia de sus portadores es un hecho contingente y de ningún modo se sigue del *status* de las expresiones del lenguaje. Esto es, de algún modo, una consecuencia irónica y desafortunada, porque la teoría del no sentido, que comienza como una explicación de sentido común de los nombres propios ordinarios, cuando se la sigue por este camino termina con la sorprendente e implausible conclusión de que los nombres propios ordinarios no son realmente nombres propios (no son «nombres lógicamente propios»), sino descripciones definidas disfrazadas.

La posición de Anscombe evita la incómoda conclusión de que ningún nombre propio ordinario es un nombre propio genuino; sin embargo, todavía hay convincentes objeciones en su contra. Mantiene que es un criterio para que una expresión sea un nombre propio que tenga un portador. Esto tiene la extraña consecuencia de que la pertenencia a una categoría sintáctica resulta contingente dependiendo de la relación no sintáctica nombre-portador. Además, tiene la extraña consecuencia de que muchos nombres propios obvios, tales como «Cerbero» y «Zeus», no son realmente nombres propios genuinos, porque Cerbero y Zeus no existen. Su posición es simplemente la decisión arbitraria de usar la expresión «nombre propio genuino» de tal manera que sólo las expresiones que tengan portadores son nombres propios genuinos. Pero, si aceptamos la teoría del no sentido y rechazamos la teoría wittgensteniana de la sustancia, parece que estuviésemos abocados a alguna decisión arbitraria e insatisfactoria. En su trabajo posterior Wittgenstein rechazó su anterior análisis de los nombres propios, diciendo que confundía el portador de un nombre con el significado del nombre: «Cuando el Sr. N. N. muere, se dice que muere el portador del nombre, no el significado» (*Investigaciones filosóficas*, § 40).

De este modo, al menos inicialmente, el sentido común parece favorecer la teoría del no sentido, pero es incapaz de explicar la ocurrencia de los nombres propios en enunciados informativos de identidad y en enunciados existenciales. Además, para los teóricos del no sentido la naturaleza de la existencia de los referentes de los nombres propios origina serios problemas. La teoría del sentido parece implausible al pretender que los nombres propios sean simplemente descripciones definidas abreviadas, pero al menos tiene el mérito de explicar los problemas acerca de enunciados de identidad y enunciados existenciales.

### UNA PROPUESTA DE SOLUCIÓN

La antinomia planteada por las dos teorías opuestas de los nombres propios admite una posible solución. Podemos replantearnos la

cuestión «¿Tienen los nombres propios sentido?» como «¿Los nombres propios implican predicados descriptivos?» o simplemente como «¿Hay proposiciones que contengan un nombre propio como sujeto y una expresión descriptiva como predicado analítico?». Pero esta cuestión tiene una forma más débil y otra más fuerte: la más débil, «¿Hay tales enunciados analíticos?», y la más fuerte, «¿Hay enunciados analíticos donde el sujeto sea un nombre propio y el predicado una descripción que sea suficientemente específica para identificar uno y sólo un objeto (a partir de aquí llamada una descripción identificadora)?». Es característico de un nombre propio que sea usado para referirse al mismo objeto en diferentes ocasiones. El uso del mismo nombre en ocasiones diferentes presupone que el objeto es el mismo; una condición necesaria de identidad de la referencia es la identidad del objeto al que se refiere. Pero presuponer que el objeto es el mismo a su vez presupone un criterio de identidad, esto es, presupone una habilidad por parte del hablante para responder a la pregunta: «¿En virtud de qué el objeto en el tiempo  $t$ , al que se hace referencia mediante el nombre  $N$ , es idéntico con el objeto al que se hace referencia mediante el mismo nombre en el tiempo  $t'$ ?». Para decirlo de un modo más simple, «¿El objeto en el tiempo  $t$  es el mismo *qué* que objeto en la ocasión  $t'$ ?» El hueco indicado por el *qué* tiene que ser llenado por un término general descriptivo; es la misma montaña, la misma persona, el mismo río, etc. El término general aporta un criterio de identidad en cada caso. Esto nos da una respuesta afirmativa a la pregunta más débil. Algún término general está analíticamente vinculado a un nombre propio: el Everest es una montaña, el Misisipí es un río, De Gaulle es una persona. Cualquier cosa que no sea una montaña no podría ser el Everest, etc.; para asegurar la continuidad de la referencia necesitamos un criterio de identidad, y el término general asociado con el nombre proporciona el criterio. Incluso si alguien quisiese sostener que De Gaulle podría convertirse en un árbol o en un caballo y todavía ser De Gaulle, debe haber algún criterio de identidad. De Gaulle no podría convertirse en cualquier cosa y seguir siendo todavía De Gaulle, y decir esto es decir que algún término o rango de términos está analíticamente relacionado con el nombre «De Gaulle».

Una tentación es decir que, si continuamos llamando a un objeto «Everest», la propiedad de ser llamado «Everest» es suficiente para garantizar que es el mismo. Pero el núcleo del análisis anterior es que estamos justificados al llamarle «Everest» sólo si podemos dar una razón para suponer que es idéntico a lo que solíamos llamar «Everest», y dar como razón que se llama «Everest» sería circular. En este sentido, al menos, los nombres propios tienen en efecto «connotaciones».

Pero la respuesta «sí» a la cuestión más débil no implica la misma



respuesta a la más fuerte, y es esa forma más fuerte la que es crucial para decidir si un nombre propio tiene un sentido o no, como Frege usaba la palabra. Según Frege, el sentido de un nombre propio contiene el «modo de presentación» que identifica el referente, y desde luego un predicado descriptivo simple no nos provee de un modo de presentación (una descripción identificadora). Que Sócrates era un hombre puede ser analíticamente verdadero, pero el predicado «hombre» no es una descripción identificadora específica de Sócrates.

Consideremos ahora la formulación más fuerte de la cuestión. Al menos dos consideraciones nos inclinan a decir que puede haber algo parecido a una respuesta afirmativa. Primero, aprendemos a usar los nombres propios y se los enseñamos a otros sólo por ostensión o descripción y ambos métodos conectan el nombre al objeto sólo en virtud de especificar características del objeto para distinguirlo de otros objetos. Segundo, cualquiera que use un nombre propio debe estar preparado para responder a la pregunta «¿Acerca de quién o qué estás hablando?», y las respuestas a esta pregunta, cuando son adecuadas, tomarán la forma o de descripciones verbales identificadoras o de presentaciones ostensivas del objeto. Ambas consideraciones sugieren una estrecha conexión entre la capacidad de usar el nombre y un conocimiento de las características del objeto suficientes para distinguirlo de otros objetos.

Pero ¿cuán estrecha es la conexión? Supongamos que pedimos a los usuarios de un nombre propio, por ejemplo «Aristóteles», que establezcan lo que consideran como los hechos establecidos y esenciales acerca de él. Sus respuestas constituirían un conjunto de descripciones identificadoras, y la totalidad de ellas sería la descripción identificadora. Por ejemplo, Aristóteles era un griego; un filósofo; el tutor de Alejandro Magno; el autor de la *Ética a Nicómaco*, de la *Metafísica*, del *De Interpretatione*; y el fundador de la escuela conocida como el Liceo de Atenas. Aunque ningún elemento particular de estas descripciones está analíticamente ligado con el nombre «Aristóteles», algún subconjunto indefinido de ellas lo está. Un erudito clásico podría descubrir que Aristóteles nunca fue el tutor de Alejandro o que nunca escribió la *Metafísica*; pero si un erudito clásico afirmase haber descubierto que Aristóteles no escribió ninguna de las obras que se le atribuyen, que nunca tuvo ninguna relación con Platón o Alejandro, que nunca estuvo cerca de Atenas, y ni siquiera fue un filósofo sino que de hecho era un oscuro pescadero veneciano del Renacimiento tardío, entonces el «descubrimiento» sería un mal chiste. El conjunto original de los enunciados sobre Aristóteles constituye el respaldo descriptivo del nombre en virtud del cual y sólo en virtud del cual podemos enseñar y usar el nombre. Tiene sentido el negar alguno de los miembros del conjunto de descripciones del por-

tador del nombre, pero negar todos ellos es hacer desaparecer totalmente las precondiciones para usar el nombre.

Para replantear este punto supongamos que tenemos medios independientes para identificar un objeto. ¿Cuáles son, entonces, las condiciones bajo las cuales podría decirse del objeto «Éste es Aristóteles»? Lo que se está sosteniendo es que las condiciones, los poderes descriptivos del enunciado, son que un número suficiente pero hasta ahora inespecificado de estos enunciados son verdaderos del objeto. En suma, si se prueba que no es verdadera de algún objeto independientemente localizado *ninguna* de las descripciones que los usuarios del nombre de ese objeto creen que son verdaderas de ese objeto, entonces no hay ningún objeto idéntico al portador del nombre. Es una condición necesaria para que un objeto sea Aristóteles que satisfaga al menos alguna de estas descripciones. Éste es otro modo de decir que la disyunción de estas descripciones está analíticamente relacionada con el nombre «Aristóteles» —lo cual es una respuesta casi afirmativa a la pregunta «¿Tienen los nombres propios sentido?» en su formulación más fuerte. Debería advertirse entre paréntesis que la descripción «llamado Aristóteles», aunque tiene un *status* peculiar, no es más crucial que cualquier otra descripción, porque es un hecho contingente que Aristóteles fuese y sea llamado Aristóteles.

La respuesta, entonces, a la cuestión «¿Tienen los nombres propios sentido?» —si esta pregunta quiere decir si los nombres propios son usados o no para describir o especificar características de objetos— es «no». Pero, si se pregunta si los nombres propios están o no lógicamente conectados con características del objeto al cual se refieren, la respuesta es «sí, de un modo laxo».

Algunos filósofos suponen que es una objeción a esta clase de explicación el que la misma palabra es algunas veces usada como un nombre para más de un objeto, pero esto es un hecho totalmente irrelevante y no una objeción a esta explicación. El que diferentes objetos sean llamados «John Smith» no es más relevante para la cuestión «¿Tienen los nombres propios sentido?» que el hecho de que tanto los asientos de los parques como las casas de finanzas sean llamados «bancos» lo sea para la cuestión «¿Tienen sentido los términos generales?». Ambos «banco» y «John Smith» sufren de homonimia, pero uno no demuestra que una palabra no tiene significado señalando que tiene varios.

Esta solución es un compromiso entre Mill y Frege. Mill tenía razón al pensar que los nombres propios no implicaban ninguna descripción particular, que no tienen definiciones, pero Frege estaba en lo cierto suponiendo que cualquier término singular debería tener un modo de presentación y por tanto, de algún modo, un sentido. Su error estuvo en tomar la descripción identificadora que puede ser sustituida por el nombre como una definición.

Este análisis de los nombres propios nos capacita para explicar todas las dificultades surgidas en este artículo. ¿Cómo puede un nombre propio aparecer en un enunciado existencial? Un enunciado tal como «Aristóteles nunca existió» establece que un número suficiente, pero hasta ahora inespecificado de respaldos descriptivos de «Aristóteles» son falsos. Sobre cuáles se afirme que son falsos no está todavía claro, pues el respaldo descriptivo de Aristóteles no es aún preciso. Suponiendo que de las proposiciones de las que se cree que son verdaderas de Aristóteles la mitad fueran de un hombre y la mitad de otro, ¿Diríamos que Aristóteles nunca existió? La cuestión no está decidida de antemano. Esto significa que, en la medida en que «Aristóteles» tenga un sentido, es mucho menos preciso que el de una descripción definida.

Similarmente, es fácil explicar los enunciados de identidad usando nombres propios. «El Everest es el Jomolungma» establece que el respaldo descriptivo de ambos nombres es verdadero del mismo objeto. Si el respaldo descriptivo de los dos nombres, para la persona que hace el enunciado, es el mismo, o si uno contiene el otro, el enunciado es analítico; si no, es sintético. El instinto de Frege iba por buen camino al inferir del hecho de que hacemos enunciados de identidad fácticamente informativos usando nombres propios que éstos deben tener un sentido, pero estaba equivocado al suponer que este sentido es tan claro como una descripción definida. Su famoso ejemplo de «la estrella de la mañana — la estrella de la tarde», lo extravió aquí, porque, aunque el sentido de estos nombres está claro, estas expresiones no son nombres propios paradigmáticos, sino que están en la línea fronteriza entre las descripciones definidas y los nombres propios.

### LA IMPRECIÓN DE LOS NOMBRES PROPIOS

Hemos visto que, en la medida en que pueda decirse que los nombres propios tienen un sentido, éste es impreciso. ¿Es la imprecisión en cuanto a qué características constituyen las condiciones necesarias y suficientes para aplicar un nombre propio algo que se deba a un mero accidente, un producto de la dejadez lingüística? ¿O ésta se deriva de las funciones que los nombres propios realizan? Pedir los criterios para aplicar el nombre «Aristóteles» es preguntar de modo formal qué es Aristóteles; es pedir un conjunto de criterios de identidad para el objeto Aristóteles. «¿Qué es Aristóteles?» y «¿Cuáles son los criterios para aplicar el nombre “Aristóteles”?» formulan la misma pregunta, la primera en el modo material y la segunda en el modo formal del lenguaje. Así, si antes de usar el nombre llegásemos a un acuerdo sobre las características precisas que constituyen la identidad de

Aristóteles, nuestras reglas para usar el nombre serían precisas. Pero esta precisión sólo se lograría a costa de que cualquier uso del nombre entrañase algunas cualidades específicas. En efecto, el nombre mismo llegaría a ser a su vez lógicamente equivalente a este conjunto de descripciones, pero, si este fuera el caso, estaríamos en la posición de ser capaces de referirnos a un objeto sólo por medio de su descripción. De hecho, esto es precisamente lo que la institución de los nombres propios nos permite evitar y lo que distingue los nombres propios de las descripciones. Si los criterios para los nombres propios fueran en todos los casos muy rígidos y específicos, entonces un nombre propio no sería más que una abreviatura de estos criterios y funcionaría exactamente como una descripción definida elaborada. Pero la peculiaridad e inmensa conveniencia pragmática de los nombres propios en nuestro lenguaje reside precisamente en el hecho de que nos capacitan para referirnos públicamente a objetos sin estar forzados a plantear cuestiones y llegar a un acuerdo con respecto a que características descriptivas constituyen exactamente la identidad del objeto. Ellas funcionan no como descripciones, sino como perchas en las que colgar descripciones. Así pues, la laxitud de los criterios para los nombres propios es una condición necesaria para aislar la función referencial de la función descriptiva del lenguaje.

Para establecer de modo diferente el mismo punto, supóngase que preguntamos: «¿Por qué tenemos nombres propios?» Obviamente, para referirnos a individuos. «Sí, mas también las descripciones definidas podrían hacer eso.» Pero sólo a costa de especificar las condiciones de identidad cada vez que se hace la referencia. Supongamos que decidimos abandonar «Aristóteles» y en sustitución decir «el maestro de Alejandro». Entonces es una verdad lógica que el hombre al que se refiere es el maestro de Alejandro —pero es un hecho contingente que Aristóteles enseñara a Alejandro— (aunque es un hecho necesario que Aristóteles tiene la suma lógica —disyunción inclusiva— de las propiedades comúnmente atribuidas a él).

Podemos clarificar algunos de los puntos expuestos en este artículo comparando los nombres propios paradigmáticos con nombres propios degenerados como «el Banco de Inglaterra». Para los últimos casos límite de nombres propios, parece que el sentido es dado tan directamente como en una descripción definida; las condiciones del uso del nombre son obvias. Por otra parte, un nombre propio puede adquirir un uso rígido sin tener la forma verbal de una descripción: Dios es justo, omnipotente, omnisciente *por definición* para los creyentes. Para nosotros «Homero» simplemente significa «el autor de la *Iliada* y la *Odisea*». La forma puede confundirnos a menudo; el Sacro Imperio Romano no fue ni sacro, ni romano, ni un imperio, pero fue sin duda el Sacro Imperio Romano. De nuevo, puede ser

convencional llamar a las chicas solamente «Marta»; pero, si llamo a mi hijo «Marta», puedo estar generando confusión pero no mintiendo. Y, desde luego, no todos los nombres propios paradigmáticos son iguales con respecto a la naturaleza de su «contenido descriptivo». Habrá, por ejemplo, una diferencia entre los nombres de la gente viva donde la capacidad del usuario del nombre para reconocer a la persona puede ser una importante «descripción identificadora», y los nombres de los personajes históricos. Pero el hecho esencial que hay que tener en cuenta cuando se tratan estos problemas es que tenemos la institución de los nombres propios para realizar el acto de habla de la referencia. La existencia de estas expresiones deriva de nuestra necesidad de separar el hacer referencia de las funciones descriptivas del lenguaje. Pero la referencia nunca aparece en completo aislamiento de la descripción, porque sin descripción alguna la referencia sería enteramente imposible.

## II. TEORÍAS DE LA REFERENCIA DIREC



Las teorías de Frege y Russell tienen —a pesar de sus diferencias sustanciales— una amplia zona de confluencia, en particular por lo que respecta al tratamiento de los nombres propios ordinarios. Ambos afirman que éstos tienen un respaldo descriptivo consistente en un conjunto de descripciones sucesivamente analizables cuando aparecen en ellas nombres propios ordinarios. Para poner un ejemplo: si un nombre como «Platón» se analiza en términos de la descripción «el maestro de Aristóteles», ésta, puesto que contiene el nombre propio «Aristóteles», se analizaría a su vez en términos de otra, pongamos por caso: «el maestro del maestro de Alejandro», que, a su vez, resultaría nuevamente analizable, y así sucesivamente. Esta concepción fue mantenida prácticamente sin discusión hasta comienzos de los años setenta en que fue seriamente puesta en tela de juicio por las investigaciones de, entre otros, Saul Kripke e Hilary Putnam. Lo que se presentaba ahora, una vuelta, en cierto sentido, a las tesis de Aristóteles y John Stuart Mill, se dio en llamar «teoría de la referencia directa» y consiste en la afirmación de que la relación entre términos singulares —como, por ejemplo, los nombres propios— y el mundo no está mediada por ningún concepto descriptivo. No obstante, debe precisarse un poco este rótulo que, en algunos aspectos, resulta desorientador. Pues la teoría de la referencia directa no afirma que la relación entre nombres y *nominata* no pueda estar mediada en modo alguno, sino que se limita a la afirmación más cauta de que no es el sentido de una expresión, su modo de presentación en términos fregeanos, lo que lleva a su referente: el nombre propio «Platón» no es entonces abreviatura alguna para un conjunto de descripciones analizables en, por así decirlo, una cascada de nuevas descripciones, cada una de las cuales contiene un nombre propio. «Identidad y necesidad» es uno de los lugares clásicos donde Kripke expone, *inter alia*, lo esencial de su teoría de la referencia directa, presenta su célebre distinción entre «designadores accidentales» (p. ej., «el inventor de las lentes bifocales») y «designadores rígidos» (p. ej., el nombre propio ordinario «Benjamin Franklin») y extrae la conclusión de que los designadores rígidos seleccionan al mismo individuo a través de todos los mundos posibles. En «El significado de "significado"» Hilary Putnam da una nueva vuelta de tuerca a la cuestión al defender una tesis similar a la de Kripke para nombres de géneros naturales (p. ej., «agua», «oro», «tigre», etc.) desafiando así la opinión recibida de acuerdo con la cual dichos nombres eran el paradigma de los términos descriptivos.



## IDENTIDAD Y NECESIDAD \*

SAUL KRIPKE

Un problema que con frecuencia ha surgido en la filosofía contemporánea es el siguiente: «¿Cómo son posibles los enunciados *contingentes* de identidad?» Esta pregunta se formula por analogía al modo como Kant formuló su pregunta: «¿Cómo son posibles los juicios sintéticos *a priori*?» En ambos casos generalmente se ha dado algo por sentado; en el caso de Kant, que los juicios sintéticos *a priori* eran posibles y, en el caso de la literatura filosófica contemporánea, que los enunciados contingentes de identidad son posibles. No pretendo ocuparme de la pregunta kantiana excepto para mencionar esta analogía. Después de que un libro más bien voluminoso fue escrito con el fin de responder a la pregunta acerca de cómo eran posibles los juicios sintéticos *a priori*, otros vinieron posteriormente a sostener que la solución al problema era que los juicios sintéticos *a priori* eran, por supuesto, imposibles y que un libro que tratase de mostrar otra cosa estaba escrito en vano. No discutiré quién tenía razón acerca de la posibilidad de los juicios sintéticos *a priori*. Pero en el caso de los enunciados contingentes de identidad, la mayoría de los filósofos han solido pensar que la idea de un enunciado contingente de identidad desemboca en algo así como la siguiente paradoja. Puede darse un argumento como el siguiente contra la posibilidad de los enunciados contingentes de identidad <sup>1</sup>:

\* Versión castellana de Margarita M. Valdés.

<sup>1</sup> Este trabajo fue presentado oralmente, sin texto escrito, en la Universidad de Nueva York en la serie de conferencias sobre identidad que se recogen en el volumen editado por Milton K. Munitz bajo el título de *Identity and Individuation*. La conferencia se grabó en cinta magnetofónica y el presente artículo representa una transcripción de dicha cinta, con leves modificaciones que no intentan cambiar el estilo del original. Si el lector imagina las oraciones de este trabajo como pronunciadas, extemporáneamente, con las pausas y énfasis adecuados, esto puede facilitar su comprensión. Sin embargo, puede haber todavía pasajes que resulten difíciles de seguir; el tiempo permitido hacía necesaria una presentación condensada del argumento. (Una versión más amplia de algunas de estas tesis, aún algo comprimidas y que también no son más que la transcripción de algunas observaciones orales, aparecerá en otro lugar. [Véase, S. Kripke, «Naming and Necessity», en D. Davidson y G. Harman (eds.), *Semantics of Natural Language*, D. Reidel, Dordrecht, 1972, N. del T.]) Ocasionalmente tuvieron que ser suprimidas algunas reservas, ampliaciones y autocomplacencias de mis observaciones, especialmente en la discusión de la identificación teórica y el problema

Primero, la ley de la substitutividad de los idénticos dice que, para cualesquiera objetos  $x$  e  $y$ , si  $x$  es idéntico a  $y$ , entonces, si  $x$  tiene una cierta propiedad  $F$ , también la tiene  $y$ :

$$(1) \quad (x) (y) [(x = y) \supset (Fx \supset Fy)]$$

Por otro lado, con toda seguridad cualquier objeto es necesariamente idéntico a sí mismo:

$$(2) \quad (x) \Box (x = x)$$

Pero

$$(3) \quad (x) (y) (x = y) \supset [\Box (x = x) \supset \Box (x = y)]$$

es una instancia de substitución de (1), esto es, de la ley de substitutividad. A partir de (2) y (3), podemos concluir que, para toda  $x$  e  $y$ , si  $x$  es igual a  $y$ , entonces, es necesario que  $x$  sea igual a  $y$ :

$$(4) \quad (x) (y) ((x = y) \supset \Box (x = y))$$

Esto es así porque la cláusula  $\Box (x = x)$  del condicional, se elimina dado que se sabe que es verdadera.

Éste es un argumento que ha sido expresado muchas veces en la filosofía reciente. Su conclusión, sin embargo, ha sido considerada con frecuencia como sumamente paradójica. Por ejemplo, David Wiggins, en su artículo «Identity Statements», dice:

Ahora bien, existen indudablemente enunciados de identidad contingentes. Sea  $a = b$  uno de ellos. A partir de su simple verdad y de (5) [= (4) anteriormente] podemos derivar « $\Box (a = b)$ ». Pero, ¿cómo puede haber entonces algún enunciado contingente de identidad? <sup>2</sup>.

Dice entonces que son posibles cinco diferentes reacciones a este argumento y las rechaza todas, pero reacciona él a su vez. No quiero discutir todas las posibles reacciones a este argumento, sino solamente mencionar la segunda de las que rechaza Wiggins. Ésta dice así:

mente-cuerpo. Las notas a pie de página, que fueron añadidas al original, habrían resultado aún más pesadas si esto no se hubiera hecho.

<sup>2</sup> R. J. Butler (ed.), *Analytical Philosophy, Second Series*, Basil Blackwell, Oxford, 1965, p. 41.

Podríamos aceptar el resultado y alegar que siempre y cuando  $a$  y  $b$  sean nombres propios no hay nada que resulte inapropiado. La consecuencia de esto es que ningún enunciado de identidad contingente puede hacerse por medio de nombres propios.

Y luego pasa a decir que le desagrada esta solución, así como a muchos otros filósofos les ha desagradado también, en tanto que otros más la han defendido.

¿Qué es lo que hace que el enunciado (4) parezca sorprendente? Dice que, para cualesquiera objetos  $x$  e  $y$ , si  $x$  es  $y$ , entonces es necesario que  $x$  sea  $y$ . He mencionado ya que alguien podría objetar este argumento sobre la base de que la premisa (2) es ya falsa, que no es el caso que todo sea necesariamente idéntico a sí mismo. Bueno, por ejemplo, ¿soy yo necesariamente idéntico a mí mismo? Alguien podría argüir que en algunas situaciones que podemos imaginar yo no habría ni siquiera existido y, por lo tanto, el enunciado «Saul Kripke es Saul Kripke» habría sido falso o no sería el caso que yo fuese idéntico a mí mismo. Quizá no sería ni verdadero ni falso, en un mundo semejante, decir que Saul Kripke es idéntico a sí mismo. Bueno, esto pudiera ser así, pero realmente depende de la posición filosófica que uno adopte con respecto a un asunto que no discutiré, esto es, lo que se haya de decir acerca de los valores de verdad de enunciados que mencionan objetos que no existen en el mundo actual, ni en ningún mundo posible o situación contrafáctica. Interpretemos aquí la necesidad de manera débil. Podemos tener por necesarios a aquellos enunciados tales que, siempre que los objetos mencionados en ellos existan, los enunciados serán verdaderos. Si quisiéramos ser muy cuidadosos sobre este asunto, tendríamos que entrar en la cuestión de la existencia como predicado y preguntar si el enunciado puede reformularse en la forma: Para cualquier  $x$ , es necesario que, si  $x$  existe, entonces  $x$  es idéntico a sí mismo. No entraré aquí en esta forma particular de sutileza porque no es algo que vaya a ser relevante para mi tema principal. Tampoco voy a considerar realmente la fórmula (4). Cualquiera que crea la fórmula (2) está comprometido, en mi opinión, con la fórmula (4). Si  $x$  e  $y$  son las mismas cosas y si de alguna manera podemos hablar, digamos, en el habla ordinaria, acerca de propiedades modales de un objeto, si podemos hablar de modalidad *de re* y de que un objeto tiene *necesariamente* ciertas propiedades en cuanto tal, entonces la fórmula (1), pienso yo, tiene que ser verdadera. Si  $x$  e  $y$  son el mismo objeto, en donde  $x$  es una propiedad cualquiera, incluso una propiedad que suponga operadores modales, y si  $x$  tuviese una cierta propiedad  $F$ , entonces  $y$  tiene que tener la misma propiedad  $F$ ; y esto es así, aunque la propiedad  $F$  sea ella misma de una forma tal que tenga alguna otra propiedad  $G$  necesariamente, en

particular la de ser necesariamente idéntico a un cierto objeto. Bueno, no discutiré propiamente la fórmula (4) porque en sí misma no afirma que sea necesario ningún enunciado verdadero de identidad en particular. De ninguna manera dice nada acerca de *enunciados*. Dice que para todo *objeto x* y para todo *objeto y*, si *x* e *y* son el mismo objeto, entonces es necesario que *x* e *y* sean el mismo objeto. Y considero que esto, si pensamos acerca de ello (de cualquier manera, si alguien no piensa así, no argumentaré en favor de esto aquí), equivale realmente a algo que difiere muy poco del enunciado (2). Dado que *x*, por la definición de identidad, es el único objeto idéntico a *x*, «(y) (y = x  $\supset$  Fy)» me parece nada más una forma un tanto locuaz de decir “Fx” y, así, (x) (y) (y = x  $\supset$  Fx) dice lo mismo que (x)Fx, no importa lo que sea *F* en particular, aun cuando *F* represente la propiedad de identidad necesaria con *x*. De esta manera, si *x* tiene esta propiedad (de identidad necesaria con *x*), trivialmente todo lo que sea idéntico a *x* la tiene, como afirma (4). Pero es claro que a partir del enunciado (4) uno puede ser capaz de deducir que varios enunciados particulares de identidad tienen que ser necesarios y se supone, entonces, que esto es una consecuencia muy paradójica.

Dice Wiggins: «Ahora bien, existen indudablemente enunciados de identidad contingentes.» Un ejemplo de un enunciado de identidad contingente es el enunciado de que el primer director general de Correos de los Estados Unidos es idéntico al inventor de los lentes bifocales, o que ambos son idénticos al hombre a quien el *Saturday Evening Post* reclama como su fundador (a propósito, creo que falsamente). Ahora bien, algunos de estos enunciados son francamente contingentes. Es un hecho francamente contingente que uno y el mismo hombre haya tanto inventado los lentes bifocales, como desempeñado el trabajo de director general de Correos de los Estados Unidos. ¿Cómo podemos reconciliar esto con la verdad del enunciado (4)? Bueno, éste también es un asunto en el que no quiero entrar en detalle salvo para mostrarme sumamente dogmático al respecto. Creo que esta cuestión quedó muy bien resuelta por Bertrand Russell en su noción del alcance de una descripción. De acuerdo con Russell, uno puede decir con propiedad, por ejemplo, que el autor de *Hamlet* pudo no haber escrito *Hamlet* o, más aún, que el autor de *Hamlet* pudo no haber sido el autor de *Hamlet*. Ahora bien, aquí, desde luego, no negamos la necesidad de la identidad de un objeto consigo mismo, sino que decimos que es verdad, con respecto a un determinado hombre, que de hecho fue la única persona que escribió *Hamlet* y, en segundo lugar, que el hombre que de hecho fue el hombre que escribió *Hamlet* pudo no haber escrito *Hamlet*. En otras palabras, si Shakespeare hubiera decidido no escribir tragedias, pudiera no haber escrito *Hamlet*. En estas circunstancias, el hombre que de hecho escribió *Hamlet* no hubiera escrito *Hamlet*. Russell destaca

esto diciendo que la primera ocurrencia de la descripción «El autor de *Hamlet*» en semejante enunciado, tiene alcance amplio<sup>3</sup>. Esto es, decimos: «El autor de *Hamlet* tiene la siguiente propiedad: que pudo no haber escrito *Hamlet*.» No aseveramos que el siguiente enunciado pudiera haber sido el caso: que el autor de *Hamlet* no escribió *Hamlet*, puesto que esto no es verdadero. Eso sería equivalente a decir que pudo haber sucedido que alguien escribió *Hamlet* y, sin embargo, no escribió *Hamlet*, lo cual sería una contradicción. Ahora bien, haciendo a un lado los detalles de la formulación particular que hace Russell de este asunto, la cual depende de su teoría de las descripciones, ésta parece ser la distinción que cualquier teoría de las descripciones tiene que hacer. Por ejemplo, si alguien se encontrara al rector de Harvard y lo tomara por un profesor ayudante, podría decir: «Tomé al rector de Harvard por un profesor ayudante.» Con esto no quiere decir que tomó como verdadera la proposición «El rector de Harvard es un profesor ayudante.» Pudo haber querido decir esto, por ejemplo, si hubiera creído que se había instaurado algún tipo de sistema democrático en Harvard, de tal manera que su rector hubiese decidido desempeñar el trabajo de un profesor ayudante. Pero probablemente no es esto lo que quiere decir. Lo que quiere decir más bien, como lo señala Russell, es: «Alguien es rector de Harvard y yo pensé que era un profesor ayudante.» En uno de los ejemplos de Russell alguien dice: «Pensé que tu yate era más grande de lo que es», y el otro replica: «No, mi yate no es más grande de lo que es.»

Suponiendo que la noción de modalidad *de re*, y por lo tanto la de cuantificar en contextos modales, tiene algún sentido, tenemos efectivamente una solución plenamente adecuada al problema de evitar las paradojas si sustituimos descripciones por los cuantificadores universales en (4), ya que la única consecuencia que podríamos obtener<sup>4</sup> en el caso de los lentes bifocales, por ejemplo, es que hay un hombre el cual sucede que fue tanto el que inventó los lentes bifocales como el primer director general de Correos de los Estados Unidos

<sup>3</sup> La segunda ocurrencia de la descripción tiene alcance reducido.

<sup>4</sup> En la teoría de Russell,  $F(\lambda xGx)$  se sigue de  $(x)Fx$  y  $(\exists!x)Gx$ , siempre y cuando la descripción en  $F(\lambda xGx)$  tenga como su alcance el contexto completo (en la terminología de Russell en 1905, tenga una «ocurrencia primaria»). Sólo entonces  $F(\lambda xGx)$  es «acerca» de la denotación de « $\lambda xGx$ ». Aplicando esta regla a (14) obtenemos los resultados indicados en el texto. Nótese que en la forma antigua  $\Box(\lambda xGx = \lambda xHx)$ , si una o ambas descripciones tienen «ocurrencias primarias», la fórmula no afirma la necesidad de  $\lambda xGx = \lambda xHx$ ; si ambas tienen una ocurrencia secundaria, entonces sí la afirma. Así pues, en un lenguaje que no tenga indicadores explícitos de alcance, las descripciones tendrán que interpretarse con el alcance más pequeño posible; sólo entonces  $\sim A$  será la negación de  $A$ ,  $\Box A$  la necesidad [necessitation] de  $A$ , y así sucesivamente.

y el cual es necesariamente idéntico a sí mismo. Hay un objeto  $x$  tal que  $x$  inventó los lentes bifocales y, como cuestión de hecho contingente, hay un objeto  $y$ , tal que  $y$  es el primer director general de Correos de los Estados Unidos y, finalmente, es necesario que  $x$  es  $y$ . ¿Qué son aquí  $x$  e  $y$ ? Aquí, tanto  $x$  como  $y$  son Benjamin Franklin y, ciertamente, puede ser necesario que Benjamin Franklin sea idéntico a sí mismo, de modo que no hay problemas en el caso de las descripciones si aceptamos la noción russelliana de alcance<sup>5</sup>. Quiero simplemente dejar aquí esta cuestión, de manera dogmática, y pasar a considerar el problema acerca de los nombres planteado por Wiggins. Wiggins dice que podría aceptar el resultado y alegar que siempre y cuando  $a$  y  $b$  sean nombres propios no hay nada que resulte inapropiado. Aunque luego rechaza esto.

Ahora bien, ¿cuál es el problema especial acerca de los nombres propios? Al menos en el caso de que uno no esté familiarizado con la literatura filosófica sobre este tema, se tiende a pensar, ingenuamente, con respecto a los nombres propios, algo así como lo siguiente: Primero, si alguien dice «Cicerón era un orador», entonces usa el nombre «Cicerón» en ese enunciado, simplemente para señalar a un determinado objeto y luego atribuir una cierta propiedad al objeto,

<sup>5</sup> Una distinción más antigua que tenía el mismo propósito era, desde luego, la distinción medieval entre *de dicto* y *de re*. Muchos lógicos, especialmente Smullyan, han señalado que la distinción de alcance de Russell elimina las paradojas modales.

Para evitar malentendidos, permítaseme enfatizar que, por supuesto, no estoy afirmando que la noción russelliana de alcance resuelve el problema de Quine del «esencialismo»; lo que sí muestra, especialmente en conjunción con algunos enfoques modernos de teoría de los modelos de la lógica modal, es que la lógica modal cuantificada no tiene que negar la verdad de todas las instancias de  $(x)(y)(x = y \cdot \supset \cdot Fx \supset Fy)$ , ni todas las instancias de  $\langle (x)(Gx \supset Ga) \rangle$  (en donde « $a$ » ha de reemplazarse por una descripción definida no vacía cuyo alcance es todo « $Ga$ »), a fin de evitar que se convierta en una verdad necesaria, que uno y el mismo hombre inventó los lentes bifocales y dirigió el primer departamento de Correos. La definición contextual de Russell de las descripciones no tiene que adoptarse a fin de asegurar estos resultados; pero otras teorías lógicas, la fregeana u otras, las cuales consideran a las descripciones como primitivas, deben expresar de alguna manera los mismos hechos lógicos. Frege mostró que un contexto simple, no iterado, que contenga una descripción definida con alcance reducido la cual no puede interpretarse como si fuese «acerca» de la denotación de la descripción, sí puede interpretarse como acerca de su «sentido». Algunos lógicos se han interesado en el problema acerca de las condiciones en las cuales, en un contexto intensional, una descripción con alcance reducido es equivalente a la misma con alcance amplio. Una de las virtudes del tratamiento russelliano de las descripciones en la lógica modal es que entonces la respuesta (en términos generales, que la descripción sea un «designador rígido» en el sentido de esta conferencia) frecuentemente se sigue de otros postulados para la lógica modal cuantificada; no se necesitan postulados especiales como en el tratamiento de Hintikka. Incluso si las descripciones se toman como primitivas, la postulación especial de cuándo es irrelevante el alcance, puede deducirse frecuentemente a partir de axiomas más básicos.

esto es, en este caso, atribuye a un determinado hombre la propiedad de haber sido un orador. Si alguien más usa otro nombre, digamos el de «Tulio», sigue hablando acerca del mismo hombre. Uno atribuye la misma propiedad al mismo hombre si dice «Tulio es un orador». El hecho o situación objetiva representado por el enunciado es, por así decirlo, el mismo, sea que uno diga «Cicerón es un orador» o «Tulio es un orador». Parecería, por tanto, que la función de los nombres es *simplemente* la de hacer referencia y no la de describir a los objetos así nombrados mediante propiedades tales como «ser el inventor de los lentes bifocales» o «ser el primer director general de Correos». Parecería que la ley de Leibniz y la ley (1) valdrían no solamente en la forma universalmente cuantificada, sino también en la forma «si  $a = b$  y  $Fa$ , entonces  $Fb$ », siempre que « $a$ » y « $b$ » estén en lugar de nombres y « $F$ » esté en lugar de un predicado que exprese una propiedad genuina del objeto:

$$(a = b \cdot Fa) \supset Fb$$

Podemos recorrer el mismo argumento una vez más para obtener la conclusión «si  $a = b$ , entonces necesariamente  $a = b$ », en donde  $a$  y  $b$  están en lugar de cualquier nombre. De este modo, podríamos aventurar esta conclusión: que siempre que « $a$ » y « $b$ » sean nombres propios, si  $a$  es  $b$ , es necesario que  $a$  sea  $b$ . Los enunciados de identidad entre nombres propios tienen que ser necesarios si es que han de ser de alguna manera verdaderos. Esta tesis ha sido defendida por Ruth Barcan Marcus, por ejemplo, en uno de sus trabajos acerca de la interpretación filosófica de la lógica modal <sup>6</sup>. De acuerdo con esta tesis, por ejemplo, siempre que uno haga un enunciado correcto de identidad entre dos nombres, tal como, por ejemplo, que Cicerón es Tulio, su enunciado tiene que ser necesario si es que es verdadero. Pero esta conclusión *parece* francamente falsa. (Yo, como otros filósofos, tengo la costumbre de decir menos de lo que en realidad quiero decir, conforme a la cual «parece francamente falsa» querría decir «es francamente falsa». De hecho creo que la tesis es verdadera, aunque no en la forma defendida por la Sra. Marcus.) De cualquiera manera, parece francamente falsa. El profesor Quine dio un ejemplo en su réplica a la profesora Marcus en el simposio: «Creo que veo problemas, de todas maneras, en la distinción entre nombres propios y descripciones tal como la traza la profesora Marcus. El paradigma de la asignación de nombres propios es el etiquetar. Podemos etiquetar

<sup>6</sup> «Modalities and Intensional Languages», *Boston Studies in the Philosophy of Science*, vol. I, Humanities Press, Nueva York, 1963, pp. 71 ss. Véase también los «Comments» de Quine y la discusión subsecuente.

una hermosa tarde al planeta Venus con el nombre propio "Hesperus". Podemos también etiquetar al mismo planeta, un día, antes de la salida del sol, con el nombre propio "Phosphorus".» (Quine piensa que algo así como esto fue realmente hecho alguna vez.) «Cuando por fin descubrimos que hemos etiquetado al mismo planeta dos veces, nuestro descubrimiento es empírico y no por el hecho de que los nombres propios sean descripciones.» De acuerdo con lo que se nos cuenta, originalmente se pensó que el planeta Venus visto en la mañana era una estrella y era llamada «la Estrella Matutina», o (para deshacernos de cualquier cuestión relativa al uso de una descripción) era llamada «Phosphorus». Uno y el mismo planeta, cuando se le veía en la tarde, era considerado como otra estrella, la Estrella de la Tarde, y era llamada «Hesperus». Más tarde, los astrónomos descubrieron que Phosphorus y Hesperus eran una y la misma cosa. Seguramente que ningún acopio de raciocinios *a priori* que hubiesen elaborado podría concebiblemente haber hecho posible que dedujeran que Phosphorus es Hesperus. De hecho, dada la información que tenían, pudo haber resultado a la inversa. Por lo tanto, se argumenta, el enunciado «Hesperus es Phosphorus» tiene que ser una verdad contingente ordinaria, empírica, que podría haber resultado de otra manera y, de este modo, la tesis de que los enunciados verdaderos de identidad son necesarios, tiene que ser falsa. Otro ejemplo dado por Quine en *Word and Object* está tomado del profesor Schrödinger, el famoso pionero de la mecánica cuántica: Una determinada montaña puede verse tanto desde el Tibet como desde el Nepal. Cuando se la veía desde una dirección se le llamaba «Everest»; cuando se la veía desde otra dirección se le llamaba «Gaurisanker» y luego, más tarde, se hizo el descubrimiento empírico de que el Gaurisanker es el Everest. (Quine dice además, que colige que el ejemplo, de hecho, es geográficamente incorrecto. Creo que uno no debiera confiar en los físicos en cuanto a la información geográfica.)

Una posible reacción a este argumento es, desde luego, negar que nombres como «Cicerón», «Tulio», «Gaurisanker» y «Everest» son realmente nombres propios. «Mira —alguien podría decir (alguien lo dijo, su nombre era Bertrand Russell)— sólo por el hecho de que enunciados como "Hesperus es Phosphorus" y "el Gaurisanker es el Everest" son contingentes, podemos ver que los nombres en cuestión no son en realidad puramente referenciales.» No estás simplemente etiquetando un objeto, en la terminología de la Sra. Marcus, realmente lo estás describiendo. ¿A qué equivale el hecho contingente de que Hesperus es Phosphorus? Bueno, equivale al hecho de que *la* estrella que está en una determinada posición celeste en la tarde es *la* estrella que está en una determinada posición celeste en la mañana. De igual manera, el hecho contingente de que el Gaurisanker es el Everest, equivale al hecho de que la montaña vista desde tal y cual



ángulo en Nepal es la montaña vista desde tal y cual otro ángulo en el Tibet. Por lo tanto, nombres tales como «Hesperus» y «Phosphorus» no pueden ser más que abreviaciones de descripciones. El término «Phosphorus» *tiene* que significar: «la estrella vista...» o (seamos precavidos porque de hecho resultó que no era una estrella) «*el cuerpo celeste* visto desde tal y cual posición a tal y cual hora de la mañana», y el nombre «Hesperus» tiene que significar: «el cuerpo celeste visto en tal y cual posición a tal y cual hora de la tarde». De esta manera, Russell concluye, si queremos reservar el término «nombre» para cosas que realmente sólo nombran un objeto sin describirlo, los únicos nombres propios genuinos que podemos tener son los nombres de nuestros propios datos sensoriales inmediatos, de los objetos «que se nos hacen presentes de manera inmediata». Los únicos nombres de esa naturaleza que aparecen en el lenguaje son demostrativos tales como «esto» y «eso»; y resulta fácil ver que este requisito de la necesidad de la identidad, entendido en el sentido de que deja exentas de toda duda imaginable a las identidades entre nombres, puede garantizarse, realmente, sólo para los nombres demostrativos de los datos inmediatos de los sentidos, ya que sólo en esos casos un enunciado de identidad entre dos nombres diferentes puede tener inmunidad general a la duda cartesiana. Hay algunas otras cosas que Russell admite algunas veces como objetos de aprehensión inmediata, tales como el propio yo, pero no es necesario entrar aquí en detalles.

Otros filósofos (por ejemplo la Sra. Marcus en su réplica, por lo menos en la discusión verbal tal como yo la recuerdo; no sé si esto se publicó, de manera que tal vez no deberíamos «etiquetarle» esto a ella <sup>7</sup>), han dicho: «Si los nombres son en realidad simplemente etiquetas, etiquetas genuinas, entonces un buen diccionario debería ser capaz de decirnos que son nombres del mismo objeto.» Tienes un objeto *a* y un objeto *b* con los nombres «Juan» y «José»; entonces, de acuerdo con la Sra. Marcus, un diccionario debería ser capaz de decirte si «Juan» y «José» son nombres del mismo objeto o si no lo son. Desde luego, yo no sé lo que los diccionarios ideales deberían hacer, pero los nombres propios ordinarios no parecen satisfacer este requisito. *Puedes* ciertamente, en el caso de los nombres propios ordinarios, hacer descubrimientos totalmente empíricos de que, digamos, Hesperus es Phosphorus, aunque previamente hayamos pensado de distinta manera. Podemos estar en duda acerca de si el Gaurisanker es el Everest o de si Cicerón es en efecto Tulio. Aún ahora, podríamos concebiblemente descubrir que nos equivocamos al suponer que Hesperus es Phosphorus. Tal vez los astrónomos cometieron un error. Así pues, parece que esta tesis está equivocada y que, si por un

<sup>7</sup> Sí deberíamos. Véase su observación en la p. 115, *op. cit.*, en la discusión a continuación de los artículos.

nombre entendemos no una noción artificial de nombre, tal como la de Russell, sino un nombre propio en el sentido ordinario, entonces puede haber enunciados contingentes de identidad en los que se usan nombres propios y la tesis contraria parece francamente equivocada.

En la filosofía reciente, un gran número de otros enunciados de identidad han sido destacados como ejemplos de enunciados contingentes de identidad, diferentes, tal vez, de cualquiera de los tipos de enunciados que antes mencioné. Uno de ellos, es por ejemplo, el enunciado «El calor es el movimiento de las moléculas». Primero, se supone que la ciencia descubrió esto. Se supone que los científicos empíricos descubrieron en sus investigaciones (y supongo yo que lo hicieron) que el fenómeno externo al que llamamos «calor» es, en realidad, la agitación molecular. Otro ejemplo de un descubrimiento semejante, es que el agua es  $H_2O$  y algunos ejemplos más son que el oro es un elemento con tal y cual número atómico, que la luz es un haz de fotones, y así sucesivamente. Estos son enunciados de identidad, en algún sentido de «enunciados de identidad». En segundo lugar, se piensa, son enunciados claramente contingentes de identidad sólo porque fueron descubrimientos científicos. Después de todo, pudo haber resultado que el calor no fuese el movimiento de las moléculas; había otras teorías alternativas del calor propuestas, por ejemplo, la teoría calórica. Si estas teorías del calor hubiesen sido correctas, entonces el calor no hubiera sido el movimiento de las moléculas, sino alguna sustancia llamada «calórica», que empapa al objeto caliente; y fue una cosa natural de la ciencia, y no debida a ninguna necesidad lógica, que una teoría resultó ser la correcta y la otra resultó estar equivocada.

De manera que aquí, una vez más, tenemos aparentemente otro ejemplo de un enunciado contingente de identidad. Se supone que es un ejemplo muy importante por su conexión con el problema mente-cuerpo. Hay muchos filósofos que han querido ser materialistas y ser materialistas de una manera particular que hoy se conoce con el nombre de «la teoría de la identidad». De acuerdo con esta teoría, un determinado estado mental, tal como que una persona tenga dolor, es idéntico a un determinado estado del cerebro (o, tal vez, de su cuerpo entero de acuerdo a algunos defensores de la teoría); sea como fuere, idéntico a determinado estado material o neuronal de su cerebro o de su cuerpo. Y así, conforme a esta teoría, que yo tenga dolor en este instante, si es que lo tuviera, sería idéntico a que mi cuerpo o mi cerebro estuvieran en un estado determinado. Otros han objetado que esto no puede ser porque, después de todo, podemos imaginar que mi dolor exista aun cuando el estado de mi cuerpo no existiera. Podemos tal vez imaginar que yo no estuviese de ningún modo encarnado y, sin embargo, que tuviese dolor o, de manera inversa, podríamos imaginar que mi cuerpo existiera y estuviera exactamente

en el mismo estado aun cuando no hubiera dolor. De hecho, concebiblemente mi cuerpo podría estar en este estado aunque no hubiera ninguna mente «detrás de él», por así decirlo. La réplica común consiste en conceder que todas estas cosas podrían haber sucedido, pero argumentar que no son pertinentes para la cuestión relativa a la identidad del estado mental con el estado físico. Esta identidad, se dice, es simplemente otra identificación contingente científica, similar a la identificación del calor con el movimiento molecular o del agua con el  $H_2O$ . Así como podemos imaginar al calor sin ningún movimiento molecular, así también podemos imaginar un estado mental sin ningún estado cerebral que le corresponda. Pero, así como el primer hecho no daña la identificación del calor con el movimiento de las moléculas, así tampoco el segundo hecho daña la identificación de un estado mental con el estado cerebral correspondiente. De esta manera, muchos filósofos recientes han sostenido que es muy importante para el entendimiento teórico del problema mente-cuerpo el que pueda haber enunciados contingentes de identidad de esta forma.

Para expresar por fin lo que *yo* pienso, a diferencia de lo que parece ser el caso, o de lo que otros piensan, pienso que en ambos casos, tanto en el de los nombres como en el de las identificaciones teóricas, los enunciados de identidad son necesarios y no contingentes. Esto es, son necesarios si es que son *verdaderos*; desde luego que los enunciados falsos de identidad no son necesarios. ¿Cómo puede uno defender semejante tesis? Tal vez no cuento con una respuesta completa a esta pregunta, aunque estoy convencido de que la tesis es verdadera. Pero para iniciar una respuesta, permítaseme hacer algunas distinciones que quiero usar. La primera es entre un *designador rígido* y un *designador no rígido*. ¿Qué significan estos términos? Puedo dar como ejemplo de un designador no rígido una expresión tal como «el inventor de los lentes bifocales». Supongamos que fue Benjamin Franklin quien inventó los lentes bifocales, de manera que la expresión «el inventor de los lentes bifocales» designa o se refiere a un hombre determinado, esto es, a Benjamin Franklin. Sin embargo, podemos imaginar fácilmente que el mundo pudo haber sido distinto, que en otras circunstancias diferentes alguien más hubiera llegado a esta invención antes de lo que llegó Benjamin Franklin y, en ese caso, *él* hubiera sido el inventor de los lentes bifocales. Así, en este sentido, la expresión «el inventor de los lentes bifocales» es no rígida. Bajo ciertas circunstancias un hombre hubiera sido el inventor de los lentes bifocales, bajo otras circunstancias otro hombre lo hubiera sido. A modo de comparación, considérese la expresión «la raíz cuadrada de 25». Independientemente de los hechos empíricos, podemos dar una prueba aritmética de que la raíz cuadrada de 25 es, de hecho, el número 5 y, por haber probado esto matemáticamente, lo que hemos probado es necesario. Si pensamos que los números son

de alguna manera entidades, y vamos a suponerlo por lo menos para el propósito de esta conferencia, entonces la expresión «la raíz cuadrada de 25» necesariamente designa a un número determinado, esto es, al 5. A una expresión semejante la llamo «designador rígido». Algunos filósofos consideran que cualquiera que use simplemente la noción de designador rígido o no rígido, muestra que ha caído en una cierta confusión o que no ha prestado atención a ciertos hechos. ¿Qué quiero decir con «designador rígido»? Quiero decir un término que designa al mismo objeto en todos los mundos posibles. Para hacernos de una confusión, que por cierto, no es mía, no estoy usando «podría haber designado a un objeto diferente» para referirme al hecho de que el lenguaje podría haber sido usado de manera diferente. Por ejemplo, la expresión «el inventor de los lentes bifocales» podría haber sido usada por habitantes de este planeta para referirse siempre al hombre que corrompió a Hadleyburg. Este habría sido el caso, en primer lugar, si la gente de este planeta no hubiera hablado castellano, sino algún lenguaje diferente que coincidiera fonéticamente con el castellano y, en segundo lugar, si en ese lenguaje la expresión «el inventor de los lentes bifocales» significara «el hombre que corrompió a Hadleyburg». Entonces, desde luego, se referiría en su lenguaje a quienquiera que de hecho hubiese corrompido a Hadleyburg en esa situación contrafáctica. No es eso lo que yo quiero decir. Lo que quiero expresar al decir que una descripción pudiera haberse referido a algo diferente, es que en *nuestro* lenguaje tal como *nosotros* lo usamos al describir una situación contrafáctica, pudiera haber habido un objeto diferente que satisficiera las condiciones descriptivas que *nosotros* damos para la referencia. Así, por ejemplo, cuando hablamos acerca de otro mundo posible o situación contrafáctica, usamos la frase «el inventor de los lentes bifocales» para referirnos a quienquiera que hubiera inventado los lentes bifocales en esa situación contrafáctica, no a la persona a quien la gente *en* esa situación contrafáctica hubiera llamado «el inventor de los lentes bifocales». Ellos podrían haber hablado un lenguaje diferente que coincidiera fonéticamente con el castellano, en el cual «el inventor de los lentes bifocales» fuese usado de una manera diferente. No es éste el asunto que me interesa aquí. Para el caso, podrían haber sido sordos y mudos, o podría no haber habido gente alguna. (Aun en el caso en que no hubiera habido gente, podría haber habido un inventor de los lentes bifocales; Dios o Satanás servirían para este caso.)

En segundo lugar, al hablar acerca de la noción de un designador rígido, no quiero implicar que el objeto referido tenga que existir en todo mundo posible, esto es, que tenga que existir necesariamente. Algunas cosas, tal vez entidades matemáticas tales como los enteros positivos, si es que existen, existen necesariamente. Algunas personas han sostenido que Dios no sólo existe, sino que existe necesaria-

mente; otras, que existe contingentemente; otras, que contingentemente Él no existe y otras que necesariamente Él no existe<sup>8</sup>; se han intentado estas cuatro opiniones. Sea como fuere, cuando uso la noción de designador rígido no implico que el objeto referido tenga que existir necesariamente. Todo lo que quiero decir es que en cualquier mundo posible donde el objeto en cuestión *exista*, en cualquier situación en la que el objeto *existiera*, usamos el designador en cuestión para designar a ese objeto. En una situación en la que el objeto no exista, entonces debemos decir que el designador no tiene referente y que el objeto en cuestión así designado no existe.

Como dije, muchos filósofos encontrarán la noción de designador rígido objetable *per se*. La objeción que la gente hace puede expresarse de la siguiente manera: Mira, estás hablando de situaciones que son contrafácticas, esto es, estás hablando acerca de otros mundos posibles. Ahora bien, después de todo, esos mundos están completamente desarticulados del mundo actual, el cual no es simplemente otro mundo posible, sino que es el mundo actual. Así pues, antes de que hables, digamos, de un objeto tal como Richard Nixon, en algún otro mundo posible, tienes que decir qué objeto *sería* Richard Nixon en este otro mundo posible. Hablemos acerca de una situación en la cual, como *tú* dirías, Richard Nixon hubiese sido un miembro de la SDS\*. El miembro de la SDS acerca del cual hablas, es ciertamente muy diferente de Nixon en muchas de sus propiedades. Antes de que podamos siquiera decir si este hombre hubiera sido Richard Nixon o no, tenemos que establecer criterios de identidad a través de mundos posibles. Aquí están esos otros mundos posibles. Hay toda clase de objetos en ellos, con propiedades diferentes a aquellas de cualquier objeto actual; algunos de ellos se asemejan a Nixon en algunos aspectos, algunos de ellos se le asemejan en otros aspectos. Bueno, ¿cuál de estos objetos es Nixon? Tiene uno que dar un criterio de identidad, y esto muestra cómo la noción misma de designador rígido cae en un círculo. Supongamos que designamos a un cierto número como el número de los planetas; luego, si esa es nuestra manera favorita, por así decirlo, de designar a este número, entonces, en cualquier otro mundo posible tendremos que identificar cualquiera que sea el número de los planetas con el número 9, el cual es el número de los planetas en el mundo actual. De esta manera, argumentan varios filósofos, por ejemplo, Quine implícitamente y explícitamente mu-

<sup>8</sup> Si no hay una deidad y, especialmente, si la inexistencia de una deidad es necesaria, es dudoso que podamos usar «Él» para referirnos a una deidad. El uso dentro del texto debe tomarse en un sentido no literal.

\* Las iniciales SDS corresponden a *Students for a Democratic Society* (Asociación de Estudiantes por una Sociedad Democrática); ésta surgió en el medio estudiantil liberal americano en los años sesenta. (*N. del T.*)

chos otros detrás de él, no podemos realmente preguntar si un designador es rígido o no porque necesitamos primero un criterio de identidad a través de mundos posibles. Se ha sostenido incluso una tesis extrema, de acuerdo a la cual, puesto que los mundos posibles están tan desarticulados del nuestro, no podemos realmente decir que ningún objeto que contengan sea el *mismo* que un objeto que exista ahora, sino sólo porque hay algunos objetos que se asemejan más o menos a cosas del mundo actual. Por lo tanto, no deberíamos hablar realmente de lo que hubiera sido verdadero de Nixon en otro mundo posible, sino solamente de qué «contrapartidas» (el término que usa David Lewis)<sup>9</sup> de Nixon hubiera habido. Algunas personas en otros mundos posibles tienen perros a los que llaman «Checkers»; otros favorecen a la ABM\* pero no tienen ningún perro llamado «Checkers». Hay varias personas que se parecen a Nixon más o menos, pero realmente no se puede decir de ninguna de ellas que sea Nixon, son solamente *contrapartidas* de Nixon y uno escoge cuál es la mejor contrapartida fijándose en cuál se asemeja más a Nixon, de acuerdo con los criterios favoritos de cada uno. Estas ideas se hallan ampliamente difundidas, tanto entre los defensores de la lógica modal con cuantificación, como entre sus detractores.

Me parece que este tipo de discurso de alguna manera tomó la metáfora de los mundos posibles con demasiada seriedad. Es como si un «mundo posible» fuese como un país extranjero o un planeta distante, por allá, a lo lejos. Es como si viéramos con dificultad a través de un telescopio varios actores en ese planeta distante. La tesis de David Lewis, efectivamente parece ser la más razonable si uno toma esta figura literalmente. Nadie a lo lejos, en otro planeta, puede ser estrictamente idéntico a alguien de aquí. Pero aun si contamos con medios de transporte maravillosos que nos permitan llevar a una y la misma persona de un planeta a otro, necesitamos realmente criterios epistemológicos de identidad para ser capaces de decir si alguien, en ese planeta distante, es idéntico a alguien de aquí.

Me parece que todo esto es una manera totalmente desviada de considerar las cosas. A lo que equivale es a la tesis de que las situaciones contrafácticas tienen que describirse en términos puramente cualitativos. De ser así, no podemos decir, por ejemplo, «Si por lo menos Nixon le hubiera dado un soborno suficiente al senador X, hubiera logrado meter a Carswell», porque esto se refiere a cierta gente, Nixon y Carswell, y habla acerca de qué cosas hubieran sido

<sup>9</sup> David K. Lewis, «Counterpart Theory and Quantified Modal Logic», *Journal of Philosophy*, n.º 65, 1968, pp. 113 ss.

\* Las iniciales ABM corresponden a *Anti Ballistic Missiles*, programa armamentista norteamericano que el Congreso de EE.UU. discutió en los años sesenta. (*N. del T.*)

verdaderas de aquéllas en una situación contrafáctica. En lugar de esto tenemos que decir: «Si un hombre que se hace la raya del pelo de tal y cual manera y sostiene tales y cuales opiniones políticas, hubiera sobornado a un hombre que era senador, el cual tenía tales y cuales otras cualidades, entonces hubiera sido confirmado un hombre que era juez en el Sur y tenía otras muchas cualidades que lo asemejaban a Carswell.» En otras palabras, tenemos que describir las situaciones contrafácticas en términos puramente cualitativos y luego preguntar: «Dado que la situación contiene gente o cosas con tales y cuales cualidades, ¿cuál de esta gente es (o es contrapartida de) Nixon, cuál es Carswell, y así sucesivamente? Esto me parece equivocado. ¿Quién nos va a prohibir decir: «Nixon hubiera logrado meter a Carswell si hubiera hecho ciertas cosas»? Hablamos de *Nixon* y preguntamos qué hubiera sido verdadero de *él* en una situación contrafáctica determinada. Podemos decir que, si Nixon hubiera hecho tal y cual cosa, hubiera perdido las elecciones y Humphrey hubiera ganado. Aquellos a los que me opongo argüirían: «Si, pero ¿cómo descubre usted si el hombre acerca del cual habla es de hecho Nixon?» Efectivamente, sería muy difícil descubrirlo si uno estuviera mirando la situación completa a través de un telescopio, pero no es esto lo que hacemos en este caso. Los mundos posibles no son algo a lo que se aplica una pregunta epistemológica tal como ésta y, si la frase «mundos posibles» es lo que hace pensar a alguien que sí se aplica, debería simplemente deshacerse de esta frase y usar alguna otra expresión, digamos «situación contrafáctica», que pudiera dar lugar a menos confusiones. Si decimos: «Si Nixon hubiera sobornado a tal y cual senador, hubiera logrado meter a Carswell», lo que está *dado* en la mera descripción de esa situación, es que es una situación en la que hablamos de Nixon, de Carswell y de tal y cual senador; y me parece que no hay más objeciones al hecho de *estipular* que hablamos acerca de determinada gente, que al de estipular que hablamos de ciertas *cualidades*. Los defensores de la otra tesis consideran que hablar de ciertas cualidades es inobjetable. No dicen: «¿Cómo sabemos que esta cualidad (en otro mundo posible) es la de rojez?», sin embargo, les parece objetable hablar de ciertas *personas*. No veo yo ninguna razón para poner objeciones en un caso y no en el otro. Y creo que esto realmente proviene de una idea de los mundos posibles como si fuesen algo que existiera por allá, pero muy lejos de aquí, y que sólo pudieran verse a través de un telescopio especial. La tesis de David Lewis es aún más objetable. De acuerdo con Lewis, cuando decimos: «Bajo ciertas condiciones Nixon hubiera logrado meter a Carswell», lo que realmente queremos decir es: «Algún hombre, distinto a Nixon, pero que se asemeja bastante a él, hubiera logrado meter a un juez distinto a Carswell, pero que se le asemeja bastante.» Tal vez esto sea así, es decir, que algún hombre que se asemeja mucho a

Nixon hubiera logrado meter a algún hombre que se pareciera mucho a Carswell. Pero *esto* no consolaría ni a Nixon ni a Carswell, ni haría que Nixon patalease y dijera: «Yo debería de haber hecho tal y cual cosa para lograr meter a Carswell.» La cuestión es la de si *el propio* Nixon, en determinadas circunstancias, hubiera logrado meter a *Carswell*, y considero que la objeción simplemente está basada en una concepción equivocada.

En cambio, podemos perfectamente bien hablar acerca de designadores rígidos y no rígidos. Más aún, tenemos una prueba simple, intuitiva, en favor de ellos. Podemos decir, por ejemplo, que el número de los planetas podría haber sido un número diferente del número que de hecho es. Por ejemplo, podría haber habido solamente siete planetas. Podemos decir que el inventor de los lentes bifocales podría haber sido alguien distinto del hombre que *de hecho* inventó los lentes bifocales<sup>10</sup>. Por otro lado, no podemos decir que la raíz cuadrada de 81 pudiera haber sido diferente del número que de hecho es, ya que dicho número simplemente tiene que ser el 9. Si aplicamos esta prueba intuitiva a los nombres propios, tales como, por ejemplo, «Richard Nixon», intuitivamente parecen resultar ser designadores rígidos. Primero, aun cuando hablemos acerca de la situación contrafáctica en la que suponemos que Nixon hizo cosas diferentes, asumimos que seguimos hablando del propio Nixon. Decimos: «Si Nixon hubiese sobornado a determinado senador, hubiera logrado meter a Carswell», y asumimos que con «Nixon» y «Carswell»

<sup>10</sup> Algunos filósofos piensan que las descripciones definidas, en inglés, son ambiguas, que algunas veces «The inventor of bifocal» [el inventor de los lentes bifocales] designa rígidamente al hombre que de hecho inventó los lentes bifocales. De manera tentativa me inclino por el rechazo de esta tesis, interpretada como una tesis acerca de la lengua inglesa (por oposición a un lenguaje posible hipotético), pero no discutiré este asunto aquí.

Lo que quisiera hacer notar es que esta supuesta ambigüedad, en contra de algunas opiniones, no puede reemplazar la noción russelliana de alcance de una descripción. Considérese la oración «El número de planetas podría haber sido necesariamente par». Obviamente esta oración puede interpretarse de manera que exprese una verdad; de haber habido ocho planetas, el número de planetas habría sido necesariamente par. Pero sin las distinciones de alcance, tanto una interpretación «referencial» (rígida) como una interpretación no rígida de la descripción harían el enunciado falso. (Ya que el número de planetas es nueve, la interpretación rígida equivale a la falsedad de que nueve podría haber sido necesariamente par.)

La interpretación «rígida» es equivalente a la ocurrencia primaria russelliana; la no rígida, al alcance más reducido —algunos, siguiendo a Donnellan, quizá vagamente, han llamado a esta interpretación el uso «atributivo»—. Se ignora entonces la posibilidad de alcances intermedios. En el caso presente, la interpretación de  $\diamond \square$  (el número de planetas es par) que se tiene en el punto de mira, no convierte al alcance de la descripción  $\square$  (el número de planetas es par) ni en el más amplio, ni en el más reducido posibles.



nos seguimos refiriendo exactamente a la misma gente que en el mundo actual. Y parece que no podemos decir: «Nixon podría haber sido un hombre diferente del hombre que era de hecho», a menos, por supuesto, de que nuestra intención sea decirlo de manera metafórica: él podría haber sido una *clase* diferente de persona (si es que se cree en el libre albedrío y en que la gente no es corrupta de manera inherente). Podría pensarse que el enunciado es verdadero en ese sentido, pero Nixon no podría haber sido, en el otro sentido literal, una persona diferente de la persona que, de hecho, es, aun cuando el trigésimo séptimo presidente de los Estados Unidos pudiera haber sido Humphrey. De manera que, la frase «el trigésimo presidente de los Estados Unidos» es no rígida, en tanto que parecería que «Nixon» es rígido.

Permítaseme hacer otra distinción antes de regresar al asunto de los enunciados de identidad. Esta distinción es muy fundamental a la vez que difícil de practicar. En discusiones recientes, muchos filósofos que han reflexionado sobre la significatividad de distintas categorías de verdad, las han considerado idénticas. Algunos de los que las han identificado son defensores vociferantes de ellas, en tanto que otros, tales como Quine, dicen que todas son idénticamente carentes de significado. Sin embargo, generalmente no se las distingue. Éstas son categorías tales como «analítico», «necesario» y «a priori» e, incluso, algunas veces, «cierto». No hablaré acerca de todas ellas sino solamente acerca de las nociones de aprioricidad y necesidad. Con frecuencia se sostiene que éstas son sinónimas. (Muchos filósofos no deberían probablemente ser descritos como sosteniendo que son sinónimas; simplemente las *usan* de manera intercambiable.) Quisiera distinguirlas. ¿Qué queremos decir cuando llamamos a un enunciado *necesario*? Queremos decir simplemente, primero, que, el enunciado en cuestión es verdadero y, segundo, que no podría haber sido de otra manera. Cuando decimos que algo es *contingentemente* verdadero, queremos decir que, aunque de hecho es el caso, podría haber sido el caso que las cosas hubieran sido de otra manera. Si quisiéramos asignar esta distinción a alguna rama de la filosofía, deberíamos asignarla a la metafísica. Hay, por el contrario, la noción de una *verdad a priori*. Una *verdad a priori* se supone que es tal que puede *conocerse* como verdadera independientemente de toda experiencia. Nótese que esto no dice nada, en y por sí mismo, acerca de todos los mundos posibles, a menos que se ponga esto en la definición. Todo lo que dice es que puede conocerse como verdadera del mundo actual, independientemente de toda experiencia. A partir de que sepamos, independientemente de toda experiencia, que algo es verdadero del mundo actual, podría seguirse, mediante un argumento filosófico, que tiene que saberse que es también verdadero de todos los mundos posibles. Pero si esto ha de ser establecido, se necesita que

algún argumento filosófico lo establezca. Ahora bien, si fuéramos a asignar *esta* noción a una rama de la filosofía, pertenecería a la epistemología y no a la metafísica. Tiene que ver con la manera como podemos saber que ciertas cosas son de hecho verdaderas. Ahora bien, puede ser el caso, por supuesto, que todo lo que sea necesario sea algo que *pueda* ser conocido *a priori*. (A propósito, nótese que la noción de verdad *a priori* definida de esta manera, contiene *otra* modalidad: *puede* conocerse independientemente de toda experiencia. Esto es un poco complicado porque hay aquí una doble modalidad.) No tendré tiempo para explorar aquí estas nociones en todo su detalle, pero podemos percatarnos desde el inicio que estas dos nociones de ninguna manera son trivialmente la misma. Si son coextensas, esto requiere de algún argumento filosófico que lo establezca. Como dije, pertenecen a distintos dominios de la filosofía; uno de ellos tiene algo que ver con el *conocimiento*, de lo que puede conocerse de ciertas maneras acerca del mundo *actual*; la otra tiene que ver con la *metafísica*, *cómo pudo haber sido* el mundo, dado que el mundo es como es, ¿pudo haber sido, de algún modo, diferente? Ahora bien, yo sostengo que, de hecho, ninguna de las dos clases de enunciados está contenida en la otra. Pero, de lo único que tenemos que hablar aquí es de esto: ¿Todo lo que es necesario es cognoscible *a priori*, o conocido *a priori*? Considérese el ejemplo siguiente: la conjetura de Goldbach. Ésta dice que todo número par es la suma de dos primos. Éste es un enunciado matemático y si es de alguna manera verdadero, tiene que ser necesario. Ciertamente uno no puede decir que, aunque de hecho todo número par sea la suma de dos primos, podría haber habido algún número extra que fuese par y no fuese la suma de dos primos. ¿Qué significaría esto? Por otro lado, la respuesta a la pregunta de si todo número par *es* de hecho la suma de dos primos, no se conoce y no contamos en el presente con un método para decirlo. Así pues, ciertamente no conocemos *a priori*, o ni siquiera *a posteriori*, que todo número par sea la suma de dos primos. (Bueno, tal vez tenemos alguna prueba en favor de esto por el hecho de que no han sido encontrados contraejemplos). De cualquier manera, ciertamente no conocemos *a priori* que todo número par sea, de hecho, la suma de dos primos. Pero, por supuesto, la definición dice simplemente «*puede* conocerse independientemente de la experiencia» y alguien podría decir que si es verdadera, *podríamos* conocerla independientemente de la experiencia. Es difícil entender exactamente lo que significa esta afirmación. Podría ser así. Una cosa que podría significar es que si fuese verdadera podríamos *probarla*. Esta afirmación es ciertamente equivocada si se aplica en general a los enunciados de la matemática y si tenemos que trabajar dentro de un sistema fijo. Esto es lo que Gödel probó. Y aun si lo que queremos decir es una «prueba intuitiva en general», podría simplemente ser el

caso (este punto de vista es al menos tan claro y tan probable como el contrario) que aunque el enunciado fuese verdadero, simplemente no hubiese ninguna manera de cómo la mente humana pudiera probarlo nunca. Por supuesto, una manera de cómo una mente *infinita* podría ser capaz de probarlo sería examinando cada uno de los números naturales, uno por uno, y comprobándolo. En este sentido, por supuesto, puede, tal vez, conocerse *a priori*, pero sólo por una mente infinita y entonces eso nos lleva a otras cuestiones complicadas. No quiero discutir los problemas acerca de la concebibilidad de realizar un número infinito de actos, tal como sería examinar todos los números uno por uno. Existe una amplia literatura filosófica sobre esto: unos han declarado que es lógicamente imposible, otros, que es lógicamente posible, y otros que no saben. La cuestión principal es que no es trivial que sólo porque un enunciado sea necesario pueda ser conocido *a priori*. Se requieren considerables aclaraciones antes de decidir que puede conocerse de esta manera. Y así, esto muestra que aun si todo lo necesario es *a priori* en algún sentido, esto no debe tomarse como una cuestión trivial de definición. Es una tesis filosófica independiente que requiere de algún trabajo.

Otro ejemplo que uno pudiera dar, está en relación con el problema del esencialismo. He aquí un atril. Una pregunta que con frecuencia ha surgido en filosofía es: ¿Cuáles son sus propiedades esenciales? ¿Qué propiedades, aparte de algunas triviales como la identidad consigo mismo, son tales que este objeto tiene que tenerlas si existe de alguna manera <sup>11</sup>, y si el objeto no las tuviera, no sería este objeto <sup>12</sup>. Por ejemplo, podría ser una propiedad esencial de

<sup>11</sup> Esta definición es la formulación común de la noción de propiedad esencial, pero tiene que hacerse una excepción para la existencia misma; de acuerdo con la definición dada, la existencia sería trivialmente esencial. Deberíamos considerar la existencia como esencial a un objeto sólo si el objeto necesariamente existe. Tal vez haya otras propiedades rebuscadas, que supongan la existencia, para las cuales la definición es igualmente objetable. (Agradezco esta observación a Michael Slote.)

<sup>12</sup> Las dos cláusulas de la oración anotada dan definiciones equivalentes de la noción de propiedad esencial, ya que  $\Box(\exists x)(x = a \supset Fa)$  es equivalente a  $\Box(x)(\sim Fx \supset x \neq a)$ . La segunda formulación, sin embargo, resulta seductoramente poderosa a favor de las teorías de «identificación a través de mundos posibles», pues sugiere que consideremos «un objeto *b* en otro mundo posible» y pongamos a prueba si es identificable con *a* preguntando si carece de alguna de las propiedades esenciales de *a*. Permítaseme, por tanto, enfatizar, que aunque una propiedad esencial es (trivialmente) una propiedad sin la cual un objeto no puede ser *a*, de esto no se sigue, de ninguna manera, que las propiedades esenciales puramente cualitativas de *a*, conjuntamente, formen una condición suficiente para ser *a*, ni que *ninguna* cualidad puramente cualitativa, sea suficiente para que un objeto sea *a*. Más aún, incluso si pudieran existir condiciones cualitativas necesarias y suficientes para que exista un objeto que sea Nixon, de todas maneras habría poca justificación para exigir una descripción puramente cualitativa de todas las situaciones contrafácticas. Podemos preguntar si Nixon podría haber sido miembro del partido Demócrata sin ocuparnos de estas sutilezas.

Este atril el estar hecho de madera, y no de hielo. Tomemos solamente el enunciado más débil, que no está hecho de hielo; éste lo establecerá de manera tan vehemente, tal vez tan dramática, como sea necesario. Suponiendo que este atril esté realmente hecho de madera, ¿podría haber sido hecho este mismísimo atril desde el principio de su existencia, de hielo, digamos, de agua congelada del Támesis? Uno tiene la impresión de que *no* podría; aunque de hecho uno podría haber elaborado ciertamente un atril con aguas del Támesis, congelarlo mediante algún procedimiento y colocarlo exactamente aquí en lugar de esta cosa. Si alguien hubiera hecho eso, habría hecho, por supuesto, un objeto *diferente*; no habría sido *este mismísimo atril* y, así, no tendríamos un caso en el que este mismo atril, aquí presente, había sido hecho de hielo o había sido hecho de agua del Támesis. La cuestión de si podría posteriormente, digamos dentro de un minuto, convertirse en hielo es otra cosa. De manera que parecería, si un ejemplo como este es correcto, y esto es lo que han sostenido los defensores del esencialismo, que este atril no podría haber sido hecho de hielo; esto es, que en cualquier situación contrafáctica de la cual pudiésemos decir que este atril de alguna manera existió, tendríamos que decir también, que no estaba hecho de agua del Támesis congelada. Por supuesto que algunos han rechazado cualquier noción semejante de propiedad esencial por considerarla carente de significado. Generalmente se ha rechazado porque han sostenido (y creo que esto es lo que diría, por ejemplo, Quine) que depende de la noción de identidad a través de mundos posibles y que ésta, en sí misma, carece de significado. Como ya rechacé esta tesis, no volveré a tratarla. Podemos hablar acerca de *este mismísimo objeto* y de si podría haber tenido ciertas propiedades que de hecho no tiene; por ejemplo, podría haber estado en otro cuarto distinto al que en realidad está, aun en este preciso momento, pero no podría haber sido hecho desde el principio de agua congelada.

Si la tesis esencialista es correcta, sólo puede ser correcta si distinguimos con precisión entre las nociones de verdad *a priori* y verdad *a posteriori*, por un lado y, por el otro, entre verdad necesaria y verdad contingente, ya que, aunque sea necesario el enunciado de que esta mesa, si de alguna manera existe, no está hecha de hielo, no es algo que ciertamente conozcamos *a priori*. Lo que sabemos es que, en primer lugar, los atriles generalmente no están hechos de hielo, por lo general están hechos de madera. Esto parece madera; no se siente frío y probablemente si fuera hecho de hielo sí se sentiría. Por lo tanto, concluyo, probablemente esto no está hecho de hielo. Aquí todo mi juicio es *a posteriori*. Podría descubrir que alguien me estuviese haciendo una ingeniosa jugarreta y que, en realidad, este atril fuese de hielo, pero lo que digo es lo siguiente: dado que en realidad no está hecho de hielo sino de madera, uno no puede imaginar que

bajo ciertas circunstancias podría haber sido hecho de hielo. Así, tenemos que decir que aunque no podemos saber *a priori* si esta mesa está hecha de hielo o no, dado que no está hecha de hielo, *necesariamente* no está hecha de hielo. En otras palabras, si *P* es el enunciado de que el atril no está hecho de hielo, uno conoce por un análisis filosófico *a priori* algún condicional de la forma «si  $\bar{P}$ , entonces necesariamente *P*». Si la mesa no está hecha de hielo, necesariamente no está hecha de hielo. Por otro lado, entonces, conocemos mediante una investigación empírica que *P*, el antecedente del condicional, es verdadero, que esta mesa no está hecha de hielo. Podemos concluir por *modus ponens*:

$$\begin{array}{r} P \supset \Box P \\ P \\ \hline \Box P \end{array}$$

La conclusión. « $\Box P$ », es que es necesario que la mesa no esté hecha de hielo y esta conclusión es conocida *a posteriori*, ya que una de las premisas en las que se basa es *a posteriori*. De esta manera, la noción de propiedades esenciales puede mantenerse siempre y cuando se distingan las nociones de verdad *a priori* y verdad necesaria, y yo la mantengo.

Regresemos al asunto de las identidades. Con respecto al enunciado «Hesperus es Phosphorus» o al enunciado «Cicerón es Tulio», uno puede descubrir todos estos enunciados mediante una investigación empírica y podría resultar que nos equivocásemos en nuestras creencias empíricas. Así, generalmente se alega, tienen que ser, entonces, contingentes tales enunciados. Algunos han aceptado el otro lado de la moneda y han sostenido: «En razón de este argumento acerca de la necesidad, los enunciados de identidad entre nombres tienen que ser cognoscibles *a priori*, de manera que, posiblemente, sólo una categoría muy especial de nombres realmente funciona como nombres; las otras cosas son nombres espurios, descripciones disfrazadas o algo por el estilo. Sin embargo, una cierta clase muy reducida de enunciados de identidad son conocidos *a priori* y son éstos los que contienen a los nombres genuinos.» Si se acepta la distinción que hice, no necesita uno pasar a ninguna de las dos conclusiones. Uno puede sostener que ciertos enunciados de identidad entre nombres, aunque con frecuencia conocidos *a posteriori* y probablemente no cognoscibles *a priori*, son de hecho necesarios si es que son verdaderos. Así pues, nos queda lugar para sostener esto; aunque, desde luego, que nos quede lugar para sostener esto, no significa que debamos sostenerlo. Veamos pues cuáles son las pruebas a favor. En primer lugar, recuérdese la observación que hice acerca de que los nom-

bres propios parecen ser designadores rígidos, como cuando usamos el nombre «Nixon» para hablar acerca de un hombre determinado, aun en situaciones contrafácticas. Si decimos «Si Nixon no hubiera escrito la carta a Saxbe, tal vez hubiera logrado meter a Carswell» hablamos en este enunciado acerca de Nixon, Saxbe y Carswell, precisamente los mismos hombres que hay en el mundo actual, y de lo que les hubiera pasado a ellos en ciertas circunstancias contrafácticas. Si los nombres son designadores rígidos, entonces no puede haber duda acerca de si las identidades son necesarias, porque «*a*» y «*b*» serán designadores rígidos de un cierto hombre o cosa *x*. Entonces, aun en todo mundo posible, «*a*» y «*b*» se referirán ambos a este mismo objeto *x* y a ningún otro; de esta manera, no habrá ninguna situación en la que *a* pudiera no haber sido *b*. Esa tendría que ser una situación en la cual, el objeto al que también estamos llamando ahora «*x*» no hubiera sido idéntico a sí mismo. Por lo tanto, no podría uno tener una situación en la cual Cicerón no hubiese sido Tulio o en la que Hesperus no hubiese sido Phosphorus<sup>13</sup>.

Aparte de la identificación de la necesidad con la aprioricidad, ¿qué es lo que ha hecho que la gente piense de la otra manera? Hay dos cosas que han hecho pensar a la gente de esa otra manera<sup>14</sup>. Algunas personas tienen la tendencia a considerar los enunciados de identidad como enunciados metalingüísticos, a identificar el enunciado «Hesperus es Phosphorus» con el enunciado metalingüístico «“Hesperus” y “Phosphorus” son nombres del mismo cuerpo celeste»; y eso, por supuesto, podría haber sido falso. Podríamos haber usado los términos «Hesperus» y «Phosphorus» como nombres de dos cuerpos celestes diferentes. Pero, desde luego, esto no tiene

<sup>13</sup> Por tanto, estoy de acuerdo con Quine en que «Hesperus es Phosphorus» es (o puede ser) un descubrimiento empírico; con Marcus, en que es necesario. De acuerdo con el punto de vista presente, tanto Quine como Marcus se equivocan al identificar las cuestiones epistemológicas y metafísicas.

<sup>14</sup> Las dos pretendidas confusiones, especialmente la segunda, se relacionan con la confusión de la cuestión metafísica de la necesidad de «Hesperus es Phosphorus» con la cuestión epistemológica de su aprioricidad; ya que si Hesperus es identificado por su posición en el cielo en la tarde y Phosphorus por su posición en la mañana, un investigador bien puede saber, previamente a la investigación empírica que Hesperus es Phosphorus si, y sólo si, uno y el mismo cuerpo ocupa la posición *x* en la tarde y la posición *y* en la mañana. La equivalencia material a priori de los dos enunciados, sin embargo, no implica su equivalencia estricta (necesaria). (La misma observación se aplica al caso que aparece más adelante del calor y el movimiento molecular.) Observaciones similares se aplican, hasta cierto punto, a la relación entre «Hesperus es Phosphorus» y «“Hesperus” y “Phosphorus” nombran la misma cosa». Una confusión que también opera es, desde luego, la confusión entre lo que *nosotros* diríamos de una situación contrafáctica y cómo la describiría la gente en esa situación; también esta confusión está probablemente relacionada con la existente entre aprioricidad y necesidad.

nada que ver con la necesidad de la identidad. En el mismo sentido « $2 + 2 = 4$ » podría haber sido falso. Las expresiones « $2 + 2$ » y « $4$ » podrían haber sido usadas para referirse a dos números diferentes. Se puede imaginar un lenguaje, por ejemplo, en el que « $+$ », « $2$ » y « $=$ » fuesen usados de manera estándar, pero en el que « $4$ » fuese usado como el nombre, digamos, de la raíz cuadrada de menos 1, o como deberíamos llamarla, « $i$ ». Entonces « $2 + 2 = 4$ » sería falso, ya que 2 más 2 no es igual a la raíz cuadrada de menos 1. Pero no es esto lo que queremos. No queremos decir simplemente que un determinado enunciado, el cual de hecho usamos para expresar algo verdadero, podría haber expresado algo falso. Queremos usar el enunciado de la manera como *nosotros* lo usamos y ver si podría haber sido falso. Hagámoslo pues. ¿Qué idea tiene la gente? Dicen: «Mira, Hesperus podría no haber sido Phosphorus. Sucede que un cierto planeta era visto en la mañana y era visto en la tarde, y a continuación simplemente resultó, como una cuestión de hecho empírica, que eran uno y el mismo planeta. Si las cosas hubieran resultado de otra manera, hubieran sido dos planetas diferentes, o dos cuerpos celestes diferentes; entonces, ¿cómo puedes decir que un enunciado semejante es necesario?»

Ahora bien, hay dos cosas que dicha gente puede querer decir. Primero, pueden querer decir que no sabemos *a priori* si Hesperus es Phosphorus. Esto ya lo he concedido. Segundo, pueden querer decir que realmente pueden imaginar circunstancias que ellos llamarían circunstancias en las que Hesperus no habría sido Phosphorus. Pensemos qué sería una circunstancia semejante, suponiendo que usamos aquí esos términos como *nombres* de un planeta. Por ejemplo, podría haber sido el caso que Venus efectivamente saliera en la mañana exactamente en la misma posición en la que lo vemos, pero que, por otro lado, Venus no estuviera en la posición que de hecho ocupa Venus en la tarde y Marte estuviera en su lugar. Todo esto es contrafáctico porque, de hecho, Venus está ahí. Ahora bien, también uno puede imaginar que en ese otro mundo posible contrafáctico, la Tierra hubiese estado habitada por gente y que ésta hubiera usado los nombres «Phosphorus» para Venus en la mañana y «Hesperus» para Marte en la tarde. Bueno, todo esto está muy bien, pero ¿sería ésta una situación en la cual Hesperus no fuera Phosphorus? Desde luego que sería una situación en la que la gente habría sido capaz de *decir* con verdad «Hesperus no es Phosphorus»; pero se supone que nosotros tenemos que describir las cosas en nuestro lenguaje, no en el de ellos. Así pues, describámoslas en nuestro lenguaje. Bueno, ¿cómo podría suceder realmente que Venus no estuviera en esa posición en la tarde? Digamos, por ejemplo, que hay algún cometa que pasa todas las tardes y da un pequeño tirón a las cosas. (Sería una manera científica muy simple de imaginarlo; aunque realmente no tan sim-

ple, es algo bastante difícil de imaginar en realidad.) Sucede simplemente que pasa todas las tardes y las cosas se desplazan un poco. Marte se desplaza exactamente a la misma posición en la que está Venus y luego el cometa vuelve a colocar a las cosas en sus posiciones normales por la mañana. Pensando en este planeta que ahora llamamos Phosphorus, ¿qué cosa diríamos? Bueno, podemos decir que el cometa pasa y da un tirón a Phosphorus, de manera que no está en la posición normalmente ocupada por Phosphorus en la tarde. Si decimos esto, y realmente usamos «Phosphorus» como el nombre de un planeta, entonces tenemos que decir que, en esas circunstancias, Phosphorus no estaría en la tarde en la posición en la que de hecho lo vimos, o alternativamente, que Hesperus no estaría en la tarde en la posición en la que de hecho lo vimos. Podríamos decir que en esas circunstancias, no habríamos llamado «Hesperus» a Hesperus porque Hesperus habría estado en una posición diferente. Pero, esto no bastaría para hacer diferente a Phosphorus de Hesperus; lo que sucedería entonces, en cambio, es que Hesperus habría estado en una posición diferente a la posición en la que de hecho está y, tal vez, no en una posición tal que la gente lo hubiera llamado «Hesperus». Sin embargo, no sería esta una situación en la que Phosphorus no hubiera sido Hesperus.

Tomemos otro ejemplo que puede resultar más claro. Supongamos que alguien usa «Tulio» para referirse al orador romano que denunció a Catilina y usa el nombre «Cicerón» para referirse al hombre cuyas obras tuvo que estudiar en su tercer año de latín en la escuela secundaria. Por supuesto que puede no saber de antemano que precisamente el mismo hombre que denunció a Catilina escribió esas obras; ese es un enunciado contingente. Pero el hecho de que éste sea un enunciado contingente, no debería hacernos pensar que el enunciado de que Cicerón es Tulio, en el caso de ser verdadero, y de hecho es verdadero, es contingente. Supongamos, por ejemplo, que Cicerón realmente denunció a Catilina y que pensó que esta proeza política era tan importante que no debería molestarse en escribir ninguna obra literaria. ¿Diríamos que estas serían circunstancias en las que él no habría sido Cicerón? Me parece que la respuesta es no; que diríamos, en cambio, que en esas circunstancias Cicerón no habría escrito ninguna obra literaria. No es una propiedad necesaria de Cicerón —a la manera como la sombra sigue al hombre— que hubiese escrito ciertas obras; podemos fácilmente imaginar una situación en la que Shakespeare no hubiese escrito las obras de Shakespeare, u otra en la que Cicerón no hubiese escrito las obras de Cicerón. Lo que puede ser el caso es que nosotros *fijemos la referencia* del término «Cicerón» mediante el uso de una frase descriptiva tal como «el autor de estas obras». Pero una vez que tenemos fijada esta referencia, entonces usamos el nombre «Cicerón» *rígidamente* para designar



al hombre que de hecho hemos identificado mediante su calidad de autor de estas obras. No lo usamos para designar a quienquiera que hubiese escrito estas obras en lugar de Cicerón, si es que alguien más las escribió. Podría haber sucedido que el hombre que escribió estas obras no fuese el hombre que denunció a Catilina. Casio pudo haber escrito estas obras; pero no diríamos entonces que Cicerón habría sido Casio, a menos de que estemos hablando en una forma muy laxa y metafórica. Diríamos que Cicerón, al cual podemos haber identificado y llegado a conocer mediante sus obras, no las habría escrito y que alguien más, digamos Casio, las habría escrito en su lugar.

Tales ejemplos no constituyen un fundamento para pensar que los enunciados de identidad son contingentes. Considerarlos como tal fundamento es malinterpretar la relación entre un *nombre* y una *descripción usada para fijar su referencia*, esto es, considerarlos como si fuesen *sinónimos*. Aunque fijemos la referencia de un nombre tal como «Cicerón» como el hombre que escribió tales y cuales obras, al hablar de situaciones contrafácticas, cuando hablamos de Cicerón, no hablamos entonces de quienquiera que en tal situación contrafáctica *hubiese* escrito tales y cuales obras, sino más bien de Cicerón, a quien hemos identificado mediante la propiedad contingente de que es el hombre que de hecho, esto es, en el mundo actual, escribió ciertas obras <sup>15</sup>.

Espero que esto esté razonablemente claro dentro de un ámbito reducido. Ahora bien, de hecho he estado presuponiendo algo que realmente no creo que sea, en general, verdadero. Supongamos que sí fijamos la referencia de un nombre mediante una descripción. Aun si lo hacemos, no hacemos entonces al nombre *sinónimo* de la descripción, sino que, por el contrario, usamos el nombre *rígidamente*

<sup>15</sup> Si con respecto al atril, alguien contraargumenta que después de todo *podía* haber resultado haber estado hecho de hielo y, por lo tanto, podía haber sido hecho de hielo, yo respondería que lo que en realidad quiere decir es que *un atril* podía haberse visto exactamente como éste y haber estado colocado en la misma posición que éste y, sin embargo, haber sido hecho de hielo. En suma, yo podría haber estado en la *misma situación epistemológica* en relación con *un atril hecho de hielo* a la que de hecho estoy en relación con *este* atril. En el texto principal he argumentado que la misma respuesta debería darse a las protestas de que Hesperus podía haber resultado ser distinto de Phosphorus o Cicerón de Tulio. Aquí, entonces, la noción de «contrapartida» hace valer sus méritos; ya que no es esta mesa, sino una «contrapartida» epistémica, la que sería recortada del hielo; no son diferentes Hesperus-Phosphorus-Venus, sino dos contrapartidas distintas, en dos de las funciones que desempeña de hecho Venus (la de Estrella de la Tarde y la de Estrella de la Mañana). Precisamente por este hecho, no es *esta mesa* la que podía haber sido hecha de hielo. Los enunciados acerca de las propiedades modales de *esta mesa* nunca se refieren a contrapartidas. Sin embargo, si alguien confunde los problemas epistemológicos con los metafísicos, se encaminará a la teoría de las contrapartidas defendida por Lewis y otros.

para referirnos al objeto así nombrado, incluso al hablar acerca de situaciones contrafácticas en donde la cosa nombrada no hubiese satisfecho la descripción en cuestión. Ahora bien, esto es lo que pienso que de hecho es verdadero para aquellos casos de nombrar en los cuales la referencia se fija mediante una descripción. Pero, de hecho, también pienso, en oposición a los teóricos más recientes, que la referencia de los nombres rara vez o casi nunca se fija mediante una descripción. Y con esto no quiero decir simplemente lo que dice Searle: «Lo que fija la referencia no es una sola descripción sino más bien un haz o familia de propiedades.» Quiero decir que las propiedades en este sentido no son usadas *de ninguna manera*; pero no tengo tiempo para entrar aquí en este asunto. Así pues, supongamos que por lo menos la mitad de las tesis predominantes sobre el nombrar son verdaderas, que la referencia se fija mediante descripciones. Aun si esto fuera verdadero, el nombre no sería sinónimo de la descripción, sino que sería usado para *nombrar* a un objeto que hemos seleccionado por el hecho contingente de que satisface una determinada descripción y, de esta manera, aun cuando podamos imaginar un caso en el que el hombre que escribió estas obras no hubiese sido el hombre que denunció a Catilina, no deberíamos decir que es un caso en el cual Cicerón no hubiera sido Tulio. Deberíamos decir que es un caso en el que Cicerón no escribió esas obras, sino que lo hizo Casio. La identidad de Cicerón y Tulio se sostiene aún.

Volvamos al caso del calor y el movimiento de las moléculas. ¡Seguramente que este es un caso de identidad contingente! La filosofía reciente ha enfatizado esto una y otra vez. Así pues, si este es un caso de identidad contingente, imaginemos en qué circunstancias sería falso. Ahora bien, con respecto a este enunciado, sostengo que las circunstancias que aparentemente piensan los filósofos como circunstancias en las que habría sido falso, no son en realidad tales. En primer lugar, por supuesto, se argumenta que «el calor es el movimiento de las moléculas» es un juicio *a posteriori*: La investigación científica podría haber arrojado distintos resultados. Como dije anteriormente, esto no muestra nada en contra de la tesis de que es necesario —por lo menos, si es que estoy en lo correcto—. Pero, seguramente, aquí la gente tenía en mente circunstancias muy específicas en las cuales ellos pensaban, el juicio de que el calor es el movimiento de las moléculas habría sido falso. ¿Cuáles eran esas circunstancias? Pueden destilarse a partir del hecho de que descubrimos empíricamente que el calor es el movimiento de las moléculas. ¿Cómo fue esto? ¿Qué fue lo que primero descubrimos cuando encontramos que el calor es el movimiento de las moléculas? Hay un determinado fenómeno externo que podemos sentir mediante el sentido del tacto y que produce una sensación que llamamos «la sensación de calor». Posteriormente descubrimos que el fenómeno externo que produce

esta sensación, la cual sentimos mediante nuestro sentido del tacto es, de hecho, la agitación molecular en la cosa que tocamos, un grado muy alto de agitación molecular. De esta manera, podría pensarse que para imaginar una situación en la cual el calor no hubiese sido el movimiento de las moléculas, necesitamos sólo imaginar una situación en la cual hubiésemos tenido exactamente la misma sensación pero que hubiese sido producida por algo distinto al movimiento de las moléculas. De manera similar, si quisiéramos imaginar una situación en la que la luz no fuese un haz de fotones, podríamos imaginar una situación en la que fuésemos sensibles a algo distinto, pero exactamente de la misma manera, que produjese lo que llamamos experiencias visuales aunque no mediante la acción de un haz de fotones. Para fortalecer este argumento, o para mirar el otro lado de la moneda, podríamos considerar también una situación en la que *estemos* interesados en el movimiento de las moléculas, pero en la cual dicho movimiento no nos produjese una sensación de calor. Y podría haber sucedido que nosotros, o al menos, las criaturas que habitasen este planeta, estuvieran constituidas de tal manera, digamos, que un incremento en el movimiento de las moléculas no nos produjese esta sensación, sino que, por el contrario, una reducción de la velocidad de las moléculas produjese en nosotros exactamente la misma sensación. Pensaríamos que esta sería una situación en la cual el calor no sería el movimiento de las moléculas o, más precisamente, en la cual la temperatura no sería la energía cinética molecular media.

Sin embargo, creo que esto no sería así. Pensemos nuevamente en la situación; pensemos primero en ella en el mundo actual. Imaginen ahora mismo el mundo invadido por unos marcianos quienes efectivamente tienen la misma sensación que llamamos «sensación de calor» cuando tocan el hielo, el cual tiene un movimiento molecular reducido, y que no tienen la sensación de calor —de hecho, tal vez, justo la contraria— cuando ponen sus manos cerca del fuego, el cual causa una gran cantidad de agitación molecular. ¿Diríamos, «¡Ah!, esto arroja algunas dudas sobre que el calor sea el movimiento de las moléculas, puesto que hay otra gente que no obtiene la misma sensación»? Obviamente no, y nadie pensaría semejante cosa. Diríamos, en cambio, que los marcianos, de alguna manera, sienten la misma sensación que nosotros tenemos cuando tocamos lo caliente al tocar lo frío y que no tienen una sensación de calor cuando tocan lo caliente. Pero ahora pensemos en una situación contrafáctica<sup>16</sup>. Supongamos que la Tierra hubiese sido habitada desde su

---

<sup>16</sup> ¿No es también contrafáctica la situación que acabo de describir? Por lo menos, bien podría serlo si tales marcianos de hecho nunca nos invadieran. Estrictamente hablando, la distinción que quisiera trazar contrasta la manera como *hablaríamos* en una

mismo principio por tales criaturas. Primero, imaginemos que ninguna criatura en absoluto la habita: no hay entonces nadie para sentir ninguna sensación de calor. Pero no diríamos que en esas circunstancias sería necesariamente el caso que el calor no existe; diríamos que el calor podría haber existido, por ejemplo, si hubiese fuegos que calentasen el aire.

Supongamos que las leyes de la física no fueran demasiado diferentes: Los fuegos sí calientan el aire. Entonces, habría habido calor aunque no hubiese, por ahí, criaturas para sentirlo. Ahora supongamos que tiene lugar la evolución y se crea la vida y que hay, por ahí, algunas criaturas. Pero éstas no son como nosotros; son más parecidas a los marcianos. ¿Diríamos ahora que el calor se ha convertido de repente en frío, dada la manera como lo sienten las criaturas de este planeta? No; creo que deberíamos describir esta situación como una en la cual, aunque las criaturas de este planeta tuviesen nuestra sensación de calor, no la obtendrían al ser expuestas al calor. La obtendrían cuando fuesen expuestas al frío. Esto es algo perfectamente imaginable; podemos imaginar esto de la misma manera como podemos imaginar a nuestro planeta invadido por criaturas de este tipo. Piénsese esto en dos etapas. Primero hay una etapa en la que no hay ninguna criatura en absoluto; uno puede imaginar que sigue habiendo tanto frío como calor en el planeta, aunque no haya nadie, por ahí, para sentirlo. Luego, a través de un proceso evolutivo, el planeta llega a estar habitado por seres con una estructura neuronal diferente a la nuestra. Entonces, estas criaturas podrían haber sido tales que fuesen insensibles al calor, que no lo sintiesen de la manera como nosotros lo sentimos, pero, por otro lado, que sintiesen lo frío de manera muy similar a como nosotros sentimos lo caliente. Aun así, el calor sería el calor y el frío sería el frío y, en particular, entonces, esto no iría en contra del decir que en esta situación contrafáctica el calor seguiría *siendo* el movimiento molecular, que seguiría siendo eso que es producido por los fuegos, y así sucesivamente, de la misma manera como habría sido si no hubiese ninguna criatura sobre el planeta. De manera similar, podríamos imaginar que el planeta estuviese habitado por criaturas que tuviesen sensaciones visuales cuando hubiese ondas sonoras en el aire. No deberíamos entonces decir: «En esas circunstancias, el sonido habría sido luz.» Deberíamos decir, en cam-

---

situación (posiblemente contrafáctica), si se diera, con la manera como *hablamos* de hecho *de* una situación contrafáctica, sabiendo que no se da; esto es, la distinción entre el lenguaje que hubiéramos usado en una situación y el lenguaje que *usamos* de hecho para describirla. (Considérese la descripción: «Supongamos que todos hablamos en alemán.» Esta descripción está hecha en castellano.) Podemos hacer más vivido el primer caso imaginado que la situación contrafáctica se da de hecho.

bio: «El planeta estaba habitado por criaturas que eran de alguna manera visualmente sensibles al sonido y, tal vez, incluso visualmente sensibles a la luz.» Si esto es correcto, puede seguir siendo y seguirá siendo una verdad necesaria que el calor es el movimiento de las moléculas y que la luz es un haz de fotones.

Para enunciar brevemente la tesis: usamos tanto el término «calor» como la expresión «el movimiento de las moléculas» como designadores rígidos de un cierto fenómeno externo. Puesto que el calor es de hecho el movimiento de las moléculas y los designadores son rígidos, por el argumento que he dado aquí, va a ser *necesario* que el calor sea el movimiento de las moléculas. Lo que nos produce la ilusión de contingencia es el hecho de que hemos identificado al calor por el hecho contingente de que hay criaturas sobre este planeta (a saber, nosotros) que son sensibles a él de una manera determinada, esto es, que son sensibles al movimiento molecular o al calor —éstas son una y la misma cosa—. Y esto es contingente. Así, usamos la descripción «aquello que causa tal y cual sensación, o aquello que sentimos de tal y cual manera» para identificar el calor. Pero al usar este hecho, usamos una propiedad contingente del calor, de la misma manera como usamos para identificar a Cicerón la propiedad contingente de haber escrito tales y cuales obras. Usamos, entonces, los términos «calor» en un caso y «Cicerón» en el otro, *rígidamente*, para designar a los objetos a los que representan. Y, desde luego, el término «el movimiento de las moléculas» es rígido, siempre representa al movimiento de las moléculas, nunca a otro fenómeno. De esta manera, como dijo el arzobispo Butler, «todo es lo que es y no otra cosa». Por tanto, «el calor es el movimiento de las moléculas» será necesario, no contingente, y uno tiene solamente la *ilusión* de contingencia, de la misma manera como uno podía tener la ilusión de contingencia al pensar que esta mesa podría haber sido hecha de hielo. Podríamos pensar que uno puede imaginar esto; pero si lo intentamos, podemos percatarnos al reflexionar, que lo que realmente imaginamos es simplemente que habría otro atril, aquí, en esta misma posición, que estaría hecho realmente de hielo. El hecho de que podamos identificar este atril por ser el objeto que vemos y tocamos en tal y cual posición es otra cosa.

Ahora bien, ¿cómo se relaciona todo esto con el problema mente-cuerpo? Generalmente se sostiene que este es un enunciado contingente de identidad exactamente igual a: «El calor es el movimiento de las moléculas.» Eso no puede ser. No puede ser un enunciado contingente de identidad exactamente igual a «el calor es el movimiento de las moléculas» porque, si estoy en lo correcto, «El calor es el movimiento de las moléculas» no es un enunciado contingente de identidad. Examinemos, por ejemplo, este enunciado: «El que yo tenga dolor en tal y cual momento es que yo esté en tal y cual

«estado cerebral en tal y cual momento» o «El dolor en general es tal y cual estado neuronal (cerebral)».

Se sostiene que este enunciado es contingente sobre las siguientes bases. Primero, podemos imaginar que el estado cerebral existe aunque no haya ningún dolor en absoluto. Es solamente un hecho científico el que siempre que estamos en un determinado estado cerebral, tenemos dolor. Segundo, se podría imaginar a una criatura que tiene dolor pero que no está en ningún estado cerebral especificado, quizá que no tenga ningún cerebro. La gente piensa incluso, por lo menos *prima facie*, aunque puede estar equivocada, que puede imaginarse criaturas completamente desencarnadas, o por lo menos, ciertamente, criaturas que no tienen cuerpos de ninguna manera semejantes al nuestro. Así pues, parece que podemos imaginar circunstancias definidas en las que esta relación habría sido falsa. Ahora bien, nótese que si estas circunstancias son efectivamente circunstancias, no podemos tratarlas simplemente diciendo que esto es nada más una ilusión, esto es, algo que en apariencia podemos imaginar, pero que en realidad no podemos, de la misma manera como pensábamos erróneamente que podíamos imaginar una situación en la que el calor no fuese el movimiento de las moléculas. Porque, aunque podemos decir que seleccionamos el calor de manera contingente, por la propiedad contingente de que nos afecta de tal y cual forma, no podemos decir, de la misma manera, que seleccionamos contingentemente el dolor por el hecho de que nos afecta de tal y cual forma. Desde esta perspectiva, habría un estado cerebral y nosotros lo seleccionaríamos por el hecho contingente de que nos afecta como dolor. Ahora bien, eso podría ser verdadero del estado cerebral, pero no puede ser verdadero del dolor. La experiencia misma tiene que ser *esta experiencia* y no se puede decir que ser un dolor sea una propiedad contingente del dolor que ahora tengo <sup>17</sup>. En realidad parecería

<sup>17</sup> Las teorías de la identidad más populares defendidas hoy en día explícitamente no satisfacen este sencillo requisito; ya que estas teorías generalmente sostienen que un estado mental es un estado cerebral y que lo que convierte al estado cerebral en un estado mental es su «papel causal», el hecho de que tiende a producir determinada conducta (así como las intenciones producen acciones, o el dolor, conducta de dolor) y de ser producido por determinados estímulos (por ejemplo, el dolor por pinchazos de alfiler). Si se consideran contingentes las relaciones entre el estado cerebral y sus causas y efectos, *ser-tal-y-cual-estado-mental* es una propiedad contingente del estado cerebral. Supongamos que *X* es un dolor. Quienes defienden una teoría de la identidad del tipo causal, sostienen: (1) que *X* es un estado cerebral; (2) que el hecho de que *X* es un dolor ha de analizarse (en términos generales) como el hecho de que *X* es producido por determinados estímulos y produce determinada conducta. El hecho mencionado en (2) es considerado, desde luego, contingente; el estado cerebral *X* bien podría existir y no tender a producir la conducta apropiada en ausencia de otras condiciones. Así pues, (1) y (2) afirman que un determinado dolor *X* podría haber existido y, sin

que ambos términos, «mi dolor» y «el que yo esté en tal y cual estado cerebral» son, antes que nada, designadores rígidos. Eso es, siempre que algo es tal y cual dolor, es esencialmente ese preciso objeto, a saber: tal y cual dolor; y siempre que algo es tal y cual estado cerebral, es esencialmente ese preciso objeto, a saber: tal y cual estado cerebral. De esta manera, ambos son designadores rígidos. Uno no puede decir que este dolor podría haber sido otra cosa, algún otro estado. Los dos son designadores rígidos.

Segundo, la manera de seleccionarlos en que pensaríamos —a saber, el dolor por ser una experiencia de cierto tipo y el estado cerebral por ser el estado de algún objeto material de tal y cual configuración molecular—, ambas son maneras que seleccionan a sus objetos esencialmente y no accidentalmente, esto es, los seleccionan mediante propiedades esenciales. Siempre que las moléculas *estén* en esta configuración *tenemos* tal y cual estado cerebral. Siempre que sientas *esto*, tienes un dolor. Así pues, parece que el defensor de la teoría de la identidad se encuentra en problemas, ya que, dado que tenemos dos designadores rígidos, el enunciado de identidad en cuestión es necesario. Por el hecho de que estos designadores seleccionan a sus objetos esencialmente, no podemos decir que cuando parece que imaginas que el enunciado de identidad es falso, se trata realmente de una ilusión, semejante a la ilusión que uno tiene en el caso del calor y el movimiento molecular, porque esta última ilusión dependía del hecho de que seleccionamos al calor mediante una propiedad contingente. Así pues, hay poco espacio para maniobrar, tal vez ninguno<sup>18</sup>. El defensor de la teoría de la identidad, quien sostiene que

---

embargo, no haber sido un dolor. Esto me parece en sí mismo evidentemente absurdo. Imagínense un dolor, ¿sería posible que *éste mismo* pudiera haber existido y sin embargo no haber sido un dolor?

Si  $X = Y$ , entonces  $X$  e  $Y$  comparten todas las propiedades, incluyendo las propiedades modales. Si  $X$  es un dolor e  $Y$  el estado cerebral correspondiente, entonces *ser un dolor* es una propiedad esencial de  $X$  y *ser un estado cerebral* es una propiedad esencial de  $Y$ . Si la relación de correspondencia es, de hecho, la identidad, entonces tiene que ser *necesario* de  $Y$  que corresponda a un dolor y *necesario* de  $X$  que corresponda a un estado cerebral, en realidad a este estado cerebral particular  $Y$ . Ambas afirmaciones parecen falsas; *parece* claramente posible que  $X$  hubiera existido sin el estado cerebral correspondiente, o que el estado cerebral hubiera existido sin haber sido sentido como dolor. Los partidarios de la teoría de la identidad no pueden, en contra de su casi universal práctica presente, aceptar estas intuiciones; tienen que negarlas y explicarlas de manera que se deshagan de ellas. Esto, de ninguna manera, es una cosa que se pueda hacer fácilmente.

<sup>18</sup> Puede resultar de alguna ayuda reformular brevemente el argumento. Si «dolor» y «estímulo de la fibra C» son designadores rígidos de fenómenos, quien los identifique tiene que considerar la identidad necesaria. ¿Cómo puede reconciliarse esta necesidad, con el hecho obvio de que el estímulo de la fibra C podría haber resultado no estar correlacionado de ninguna manera con el dolor? Podríamos tratar de

el dolor es el estado cerebral, tiene que sostener también que necesariamente es el estado cerebral. Por tanto, éste no puede conceder, uno que tiene que negar, que podría haber habido situaciones en las que uno hubiese tenido dolor pero no el estado cerebral correspondiente. Ahora bien, esto está muy lejos de negarse en las discusiones sobre la teoría de la identidad; de hecho, tanto el materialista como su oponente lo conceden de entrada. Dice: «Por supuesto, *podía* haber sido el caso que tuviésemos dolores sin los estados cerebrales. No trata de una identidad contingente.» Pero eso no puede ser. Tiene que sostener que somos presa de una ilusión al pensar que podemos imaginar que podía haber habido dolores sin estados cerebrales. Y el único modelo en el que puedo pensar de lo que podría ser la ilusión, o al menos, el modelo dado por la analogía que sugieren los propios materialistas, a saber, el calor y el movimiento molecular, simplemente no funciona en este caso. De manera que el materialista se encuentra con un desafío muy peliagudo. Tiene que mostrar que estas cosas que pensamos que podemos entender que son posibles, en realidad no lo son. Tiene que mostrar que estas cosas que podemos imaginar, no son realmente cosas que podemos imaginar. Y para esto se requiere algún argumento filosófico muy distinto a la clase de argumento que ha sido dado para el caso del calor y el movimiento molecular; éste tendría que ser un argumento más profundo y más sutil que cualquiera que yo pueda desentrañar, y más fino que cualquiera de los que han aparecido en cualquier literatura materialista que yo

responder por analogía con el caso del calor y el movimiento molecular; esta última identidad también es necesaria, sin embargo, alguien puede creer que, antes de que la investigación científica hubiese mostrado otra cosa, el movimiento molecular podría haber resultado no ser calor. La respuesta es, por supuesto, que lo que realmente es posible, es que la gente (o algunos seres racionales sensibles) podían haber estado en *la misma situación epistémica* en la que nosotros de hecho estamos, e identificar un fenómeno de la misma manera que nosotros identificamos el calor, a saber, sintiéndolo mediante la sensación que llamamos «la sensación de calor», sin que dicho fenómeno fuese el movimiento molecular. Más aún, los seres podrían no haber sido sensibles al movimiento molecular (es decir, al calor) mediante ningún mecanismo neuronal. Es imposible explicar de la misma manera la posibilidad obvia de que la estimulación de la fibra C no hubiese sido dolor. Aquí, también, tendríamos que suponer que podíamos haber estado en la misma situación epistemológica, e identificar algo de la misma manera que identifica el dolor, sin que correspondiese a la estimulación de la fibra C. Pero, la manera como identificamos al dolor es sintiéndolo; y si la estimulación de la fibra C podría haber ocurrido sin que sintiésemos ningún dolor, entonces habría ocurrido la estimulación de la fibra C sin que hubiera *habido* ningún dolor, lo cual es contrario a la necesidad de la identidad. El problema es que, aunque «calor» es un designador rígido, el calor se selecciona por la propiedad contingente de ser sentido de una manera determinada; el dolor, por otro lado, es seleccionado por una propiedad esencial (de hecho necesaria y suficiente). Que una sensación sea *sentida* como dolor es que *sea* dolor.



haya leído. Así, la conclusión de esta investigación sería que la herramienta analítica que usamos va en contra de la tesis de la identidad y, de esta manera, va en contra de la tesis general de que los estados mentales son simplemente estados físicos<sup>19</sup>.

El próximo tema tendría que ser mi propia solución al problema mente-cuerpo, pero esa no la tengo.

<sup>19</sup> Todos los argumentos en contra de la teoría de la identidad que se apoyan en la necesidad de la identidad, o en la noción de propiedad esencial, están, por supuesto, inspirados en el argumento de Descartes en favor de su dualismo. Los argumentos anteriores, los cuales fueron superficialmente refutados mediante las analogías del calor y el movimiento molecular, y del inventor de los lentes bifocales que también fue director general de Correos, tenían dicha inspiración, como también la tiene mi argumento presente. R. Albritton y M. Slote me han informado que ellos también, independientemente, han intentado dar argumentos esencialistas en contra de la teoría de la identidad, y probablemente otros también han intentado hacer semejante cosa.

El argumento cartesiano más simple puede tal vez reformularse de la siguiente manera: Sea «A» un *nombre* (designador rígido) del cuerpo de Descartes. Entonces, Descartes argumenta que ya que él podía haber existido aun cuando A no existiese,  $\Diamond$  (Descartes  $\neq$  A), por lo tanto, Descartes  $\neq$  A. Aquellos que lo han acusado de cometer una falacia modal han olvidado que «A» es rígido. Su argumento es válido, y su conclusión es correcta, siempre y cuando su (tal vez dudosa) premisa se acepte. Por otro lado, siempre que se considere que Descartes dejó de existir a raíz de su muerte, «Descartes  $\neq$  A» puede establecerse sin hacer uso de un argumento modal; pues de ser así, sin duda que A sobrevivió a Descartes cuando A fue un cadáver. Así pues, A tenía una propiedad (existir en un momento dado) que Descartes no tenía. El mismo argumento puede mostrar que una estatua no es el pedazo de piedra o el cúmulo de moléculas de las cuales se compone. La mera no-identidad, entonces, puede ser una conclusión débil. [Véase D. Wiggins, *Philosophical Review*, vol. 77 (1968), pp. 90 ss.] Sin embargo, el argumento modal cartesiano, con toda seguridad, puede desplegarse para sostener también conclusiones relevantes más fuertes.

## EL SIGNIFICADO DE «SIGNIFICADO» \*

HILARY PUTNAM

El lenguaje es la primera gran área de la capacidad cognitiva del hombre de la que estamos empezando a disponer de una descripción que no es esquemática en demasía. Gracias al trabajo de los lingüistas transformacionales contemporáneos <sup>1</sup>, se está en el proceso de elaborar una muy sutil descripción de algunos lenguajes al menos. Ciertos rasgos de estos lenguajes parecen ser *universales*. Los rasgos de esa índole que resulten ser «específicos de la especie» —«no susceptibles de explicación mediante algunas consideraciones generales que apelarían a la utilidad funcional o a la simplicidad de sistemas arbitrarios al servicio de las funciones del lenguaje»— pueden arrojar alguna luz sobre la estructura de la mente. Aunque es extremadamente difícil decir en qué medida queda así iluminada una estructura universal del *lenguaje*, en tanto que algo opuesto a una estructura universal de las estrategias de aprendizaje general que son innatas <sup>2</sup>, el hecho mismo de que pueda tener lugar esta discusión testimonia la riqueza y la generalidad del material descriptivo que los lingüistas están comenzando a proporcionarnos; y testimonia también la popularidad del análisis, en tanto en cuanto los rasgos que parece que son candidatos a considerarse rasgos lingüísticos «específicos de la especie» no son, en ningún sentido, rasgos superficiales o fenomenológicos del lenguaje, sino pertenecientes al nivel de la estructura profunda.

El inconveniente más serio de todo este análisis, en la medida en que resulta importante para el filósofo, es que no se ocupa del significado de las palabras. El análisis de la estructura profunda de las formas lingüísticas nos proporciona una descripción de la *sintaxis* de las lenguas naturales incomparablemente más poderosa que nunca.

\* Versión castellana de Juan José Acero.

<sup>1</sup> Las personas que han contribuido a esta área son en la actualidad demasiado numerosas como para dar una relación de todas ellas: las primeras fueron, naturalmente, Zellig Harris y Noam Chomsky.

<sup>2</sup> Para una discusión de este tema, véase H. Putnam, «Innateness Hypothesis and Explanatory Models in Linguistics», *Synthese*, 17 (1967), 12-22 (reimpreso en *Philosophical Papers*, vol. 2: *Mind, Language and Reality*, Cambridge University Press, 1975); Noam Chomsky, *Problems of Knowledge and Freedom*, Nueva York, 1971, cap. I.

Pero la dimensión del lenguaje asociada a la palabra «significado», pese al actual aluvión de esfuerzos heroicos, cuando no mal encaminados, sigue en la oscuridad en igual medida que siempre.

En este ensayo deseo explorar por qué ha ocurrido esto. En mi opinión, la razón por la que la denominada semántica está en mucho peor estado que la teoría sintáctica es que el concepto *precientífico* sobre el que se basa la semántica —el concepto precientífico de *significado*— se encuentra mucho peor perfilado que el concepto precientífico de la sintaxis. Como sucede corrientemente en filosofía, las dudas escépticas en torno a tal concepto no ayudan para nada a esclarecer o a mejorar la situación más de lo que lo hacen las afirmaciones dogmáticas de los filósofos conservadores de que todo anda bien en el mejor de los mundos posibles, que es éste. La razón por la que el concepto precientífico de significado esté mal perfilado no se aclara recurriendo al argumento escéptico o nominalista de que los significados no existen. De hecho, el resultado de nuestra discusión será que los significados no existen del modo en que tendemos a pensar que lo hacen. Pero tampoco los electrones existen como Bohr pensó que lo hacían. Existe toda la diferencia del mundo entre esta afirmación y la de que los significados (o los electrones) no existen.

Voy a hablar casi enteramente del significado de palabras más que del significado de oraciones, porque tengo la impresión de que nuestro concepto de significado-de-una-palabra es más defectuoso que nuestro concepto de significado-de-una-oración. Pero haré algunos breves comentarios sobre los argumentos de filósofos como Donald Davidson, que insisten en que el concepto de significado-de-una-palabra *debe* ser secundario y en que el estudio del significado de las oraciones debe ser primario. Puesto que considero que las teorías tradicionales del significado se alimentan de mitos (nótese que el tema del «significado» es el tema filosófico por excelencia del cual no hay otra cosa sino «teoría»; literalmente: nada a lo que se le pueda poner el cartel de, o ridiculizar mediante las palabras, «el punto de vista del sentido común»), será necesario que discuta y trate de desenredar un cierto número de temas a propósito de los cuales las opiniones heredadas son, creo yo, falsas. El lector me prestará su máxima ayuda en la labor de aclarar estos puntos si amablemente asume que no hay *nada* claro de antemano.

## SIGNIFICADO Y EXTENSIÓN

Desde la Edad Media, al menos, los que han escrito de teoría del significado han pretendido descubrir una ambigüedad en el concepto ordinario de significado y han introducido un par de términos —*extensión e intensidad*, o *Sinn (sentido)* y *Bedeutung (referencia)*, o

lo que sea— para deshacer la ambigüedad de la noción. La *extensión* de un término, en el parloteo lógico común, es simplemente el conjunto de cosas de las que el término es verdadero. Así, «conejo», en su sentido más común en español, es verdadero de todos y solos los conejos, de modo que la «extensión» de «conejo» es precisamente el conjunto de los conejos. Incluso esta noción —y en esta nebulosa materia es la *menos* problemática— tiene, no obstante, sus problemas. Aparte de los que hereda de la noción, emparentada con ella, de *verdad*, el ejemplo precedente de «conejo» en su sentido más frecuente en español ilustra uno de esos problemas: estrictamente hablando, «conejo» no es un término, sino un par ordenado que consta de un término y de un «sentido» (u ocasión de uso, o algo más que distinga a un término en un sentido del mismo término usado en un sentido diferente) que tiene una extensión. Otro problema es éste: un «conjunto», en sentido matemático, es un objeto «de-sí-o-no»; cualquier objeto dado pertenece definitivamente a  $S$  o definitivamente no pertenece a  $S$ , siendo  $S$  un conjunto. Pero las palabras de una lengua natural no son generalmente del tipo «de-sí-o-no»: seguramente hay cosas de las que la descripción «árbol» es claramente verdadera y cosas de las que «árbol» es claramente falsa, pero hay multitud de casos dudosos. Peor aún: la línea divisoria entre los casos claros y los casos dudosos es ella misma difusa. Así, la idealización involucrada en la noción de *extensión* —la implicada al suponer que hay una cosa como el conjunto de las que es verdadero el término «árbol»— es realmente muy extrema.

Recientemente, algunos matemáticos han investigado la noción de *conjunto difuso* —es decir, de un objeto al que pertenecen o no otras cosas con una cierta probabilidad o hasta un grado de probabilidad dado, en vez de pertenecer a él al modo «sí-o-no». Si se deseara realmente formalizar la noción de extensión, en tanto que aplicada a los términos de una lengua natural, sería preciso emplear «conjuntos difusos» en vez de conjuntos en el sentido clásico.

El problema que supone que una palabra tenga más de un sentido se solventa comúnmente considerando cada uno de los sentidos como una palabra diferente (o mejor: considerando que la palabra lleva subíndices invisibles, como “gallina<sub>1</sub>” —animal de una cierta especie—; “gallina<sub>2</sub>” —cobarde—; y como si “gallina<sub>1</sub>” y “gallina<sub>2</sub>” o lo que fuere resultaran ser palabras completamente distintas)<sup>a</sup>. Esto conlleva de nuevo dos idealizaciones extremas: suponer que las palabras tienen muchos sentidos de una forma definida, y suponer que todo el repertorio de sentidos está fijado de una vez y para siempre.

<sup>a</sup> En este punto el texto inglés dice: «... (or rather, by treating the word as if carried invisible subscripts, thus “rabbit<sub>1</sub>” —animal of a certain kind; “rabbit<sub>2</sub>” —coward; and as if “rabbit<sub>1</sub>” and “rabbit<sub>2</sub>” or what ever were different words entirely).»

Paul Ziff ha investigado recientemente <sup>3</sup> en qué medida distorsionan estas dos suposiciones la situación real en la lengua natural; sin embargo, seguiremos ateniéndonos a estas idealizaciones.

Consideremos ahora los términos compuestos «criatura con corazón» y «criatura con riñones». Suponiendo que toda criatura con corazón posee un riñón, y viceversa, la extensión de estos dos términos será la misma. Pero obviamente difieren en cuanto a su significado. Asumiendo que hay un sentido de «significado» según el cual el significado = la extensión, debe haber otro sentido de «significado» según el cual el significado de un término no es su extensión, sino otra cosa: digamos que el «concepto» asociado con el término. Denominemos a esta «otra cosa» la *intensión* del término. El concepto de una criatura con corazón es un concepto claramente diferente del concepto de criatura con riñones. Así que estos dos términos tienen diferente *intensión*. Cuando decimos que tienen diferente «significado», el significado = la *intensión*.

#### INTENSIÓN Y EXTENSIÓN

Algo parecido a lo dicho en el párrafo precedente se da en las exposiciones usuales de las nociones de «*intensión*» y «*extensión*». Pero no es satisfactorio en absoluto. El por qué no lo es constituye, en un sentido, el tema principal de todo este ensayo. Sin embargo, ya desde el principio puede hacerse alguna observación: primera de todas, ¿qué pruebas hay de que «*extensión*» sea un sentido de la palabra «significado»? La explicación canónica de las nociones de «*intensión*» y «*extensión*» se parece mucho a esto: «en un sentido, “significado” significa *extensión* y en el otro sentido “significado” significa *significado*». El hecho es que aunque se precise la noción de «*extensión*», por referencia a la noción lógica fundamental de *verdad* (y sin olvidar las extremas idealizaciones subrayadas más arriba), la noción de «*intensión*» no se hace más precisa que la vaga (y, como hemos visto, engañosa) noción de «concepto». Es como si alguien explicara la noción de «probabilidad» diciendo: «en un sentido “probabilidad” significa frecuencia y en otro sentido significa *propensión*». «Probabilidad» *nunca* significa «frecuencia», y «*propensión*» es al menos tan poco claro como «probabilidad».

Por no ser clara, la doctrina tradicional de que la noción de «significado» posee la ambigüedad *extensión/intensión* tiene ciertas consecuencias típicas. Muchos filósofos tradicionales concibieron el concepto como algo *mental*. Así, la doctrina de que el significado de un término (es decir, el significado «en el sentido de *intensión*») es un

<sup>3</sup> Esto lo discute Ziff en *Understanding understanding* (especialmente en el capítulo VIII), Nueva York, 1972.

concepto que llevaba consigo la consecuencia de que los significados son entidades mentales. Frege, y más recientemente Carnap y sus seguidores, se han rebelado, no obstante, contra este «psicologismo», según lo bautizaron. Apercibiéndose de que los significados son propiedad pública —que el mismo significado puede ser «captado» por más de una persona y por personas situadas en tiempos diferentes— sostuvieron que los conceptos (y, por lo tanto, las «intensiones» o significados) eran entidades abstractas, en vez de entidades mentales. Sin embargo, «captar» estas entidades abstractas siguió siendo un acto psicológico individual. Ninguno de estos filósofos puso en duda que entender una palabra (conocer su intensión) no fuese precisamente cosa de estar en un cierto estado psicológico (de alguna forma, al modo en que saber cómo extraer mentalmente los factores de un número es exactamente cosa de estar en un cierto estado psicológico muy complejo).

En segundo lugar, el manido ejemplo de los términos «criatura con corazón» y «criatura con riñones» pone de manifiesto que dos términos pueden tener la misma extensión y diferir en cuanto a su intensión. Pero se consideró obvio que la inversa era imposible: dos términos no pueden diferir en extensión y tener la misma intensión. Es interesante darse cuenta de que nunca se ofreció argumento en pro de esa imposibilidad. Probablemente, eso refleja la tradición de los filósofos antiguos y medievales, que asumieron que el concepto que corresponde a un término no es sino una conjunción de predicados y, por lo tanto, que tal concepto proporciona *siempre* una condición necesaria y suficiente [que algo debe poseer] para caer en la extensión del término<sup>4</sup>. Para filósofos como Carnap, que aceptaban la teoría verificacionista del significado, el concepto que corresponde a un término proporciona (en el caso ideal, cuando el término tiene un

<sup>4</sup> Esta tradición arraigó porque el término cuyo análisis provocó toda la discusión dentro de la filosofía medieval fue el término «Dios» y porque se pensó que el término «Dios» se definía por medio de la conjunción de los términos «Bueno», «Omnipotente», «Omnisciente», etc. —las llamadas «Perfecciones»—. Había no obstante, un problema, porque se supuso que Dios era una Unidad, y la Unidad excluiría que Su esencia fuese de algún modo compleja —es decir, «Dios» habría de definirse a través de una conjunción de términos, pero Dios (sin comillas) no podía ser el producto lógico de dos o más propiedades distintas, pues se consideraba que incluso esta clase de «complejidad» altamente abstracta era incompatible con Su perfección de Unidad—. Es ésta una paradoja teológica con la que los teólogos judíos, árabes y cristianos lucharon denodadamente durante siglos (por ejemplo, en la doctrina de la Negación de Privación, de Maimónides y Santo Tomás). Tiene gracia que teorías que en la actualidad conservan su interés, como el conceptualismo y el nominalismo, fuesen propuestas inicialmente como soluciones al problema de la predicación para el caso de Dios. Tiene gracia también que el modelo de definición favorito de toda esta teología —el modelo de conjunción-de-propiedades— haya sobrevivido, al menos por sus consecuencias, en la filosofía del lenguaje hasta el presente día.

«significado completo») un *criterio* de pertenencia a la extensión [no sólo en el sentido de «condición necesaria y suficiente», sino en el más fuerte de *modo de determinar* (“*way of recognizing*”) si una cosa cae o no dentro de la extensión del término]. Por eso, estos filósofos positivistas se sentían perfectamente felices manteniendo al respecto el punto de vista tradicional. Con ello la teoría del significado viene a descansar en dos supuestos no cuestionados:

(I) Que conocer el significado de un término no es sino cosa de estar en un cierto estado psicológico (en el sentido de «estado psicológico» según el cual los estados de memoria y las disposiciones psicológicas son «estados psicológicos»; naturalmente, nadie pensó que conocer el significado de una palabra fuese un estado de conciencia continuo).

(II) Que el significado de un término (en el sentido de «intención») determina su extensión (en el sentido de que mismidad de intención implica mismidad de extensión).

Argüiré que no hay *ninguna* noción —no hablemos ya de la de significado— que satisfaga al mismo tiempo estos dos supuestos. El concepto tradicional de significado es un concepto que *se basa* en una teoría falsa.

#### «ESTADO PSICOLÓGICO» Y SOLIPSISMO METODOLÓGICO

Para mostrar esto, necesitamos primero aclarar la noción tradicional de estado psicológico. En un sentido, un estado es simplemente un predicado diádico cuyos argumentos son un individuo y un tiempo. En este sentido, *tener una altura de metro ochenta*, *tener un dolor*, *conocer el alfabeto* e incluso *hallarse a mil kilómetros de París* son todos ellos estados. (Adviértase que usualmente el *tiempo* se da por sabido tácita o «contextualmente»; la forma completa de una oración atómica elaborada a partir de estos predicados sería «*en el tiempo t, x tiene una altura de metro ochenta*», «*en el tiempo t, x tiene dolor*», etc.) No obstante, en ciencia se acostumbra a limitar el [uso del] término estado a propiedades que se definen a partir de parámetros del individuo que son fundamentales desde el punto de vista de la ciencia dada. Así, medir metro ochenta es un estado (desde el punto de vista de la física); tener un dolor es un estado (desde el punto de vista de la psicología mentalista); conocer el alfabeto podría ser un estado (desde el punto de vista de la psicología cognitiva), aunque es difícil asegurarlo; pero no sería natural pensar que hallarse a mil kilómetros de París sea un *estado*. En cierto sentido, un estado psicológico es, simplemente, un estado descrito por la psicología. En este sentido, puede ser trivialmente verdadero decir que *conocer el significado de la palabra «agua»*, por poner un ejemplo, sea un «esta-

do psicológico» (visto desde la perspectiva de la psicología cognitiva). Pero no es éste el sentido de estado psicológico que hace al caso a propósito del supuesto (I).

Cuando los filósofos tradicionales hablaban de estados psicológicos (o estados «mentales»), adoptaron un supuesto al que podemos dar el nombre de supuesto de solipsismo metodológico. Este supuesto es el de que ningún estado psicológico propiamente dicho presupone la existencia de otro individuo que no sea aquél al que se atribuye el estado. (De hecho, el supuesto era que ningún estado psicológico presupone ni siquiera la existencia del *cuerpo* de ese sujeto: si *P* es un estado psicológico propiamente dicho, debe de ser lógicamente posible que una «mente desencarnada» esté en *P*.) Este supuesto se halla muy explícito en Descartes, pero está implícito en la mayor parte de la psicología filosófica tradicional. Adoptar este supuesto es, naturalmente, adoptar un *programa restrictivo* —un programa que limita de forma deliberada el alcance y la naturaleza de la psicología para cuadrar con ciertas ideas mentalistas preconcebidas o, en algunos casos, con una reconstrucción idealista del conocimiento y del mundo. Lo que a menudo pasa desapercibido es *cuán* restrictivo resulta este programa. Por ejemplo, estados psicológicos tan comunes y espectaculares como el *estar celoso* han de reconstruirse, si es que se trata de conservar el supuesto del solipsismo metodológico. Pues, en su uso común, *x está celoso de y* implica que *y* existe; y *x está celoso por la atención que y le presta a z* implica que tanto *y* como *z* existen (aparte de *x*, por supuesto). Así, *estar celoso* y *estar celoso por la atención que alguien le presta a otro* no son estados psicológicos permitidos por el supuesto del solipsismo metodológico. (Los denominaremos «estados psicológicos en sentido amplio», y nos referiremos a los estados permitidos por el solipsismo metodológico como a los «estados psicológicos en sentido estricto»). Lo que el solipsismo metodológico exige es reconstruir *tener celos* de modo que yo pueda tener celos de mis propias alucinaciones o en las ficciones de mi imaginación, etc. Sólo si asumimos que los estados psicológicos en sentido estricto tienen un grado significativo de cierre causal (de modo que se facilite la enunciación de *leyes* psicológicas, al limitarnos a considerar estados psicológicos en sentido estricto) está justificado que nos embarquemos en esta reconstrucción o en adoptar el supuesto del solipsismo metodológico. Pero, en mi opinión, tres siglos de fracaso de la psicología mentalista es una tremenda evidencia en contra de esta forma de proceder.

Sea como fuere, podemos enunciar ahora de modo más preciso lo que sostuvimos ya al final de la sección precedente. Sean *A* y *B* dos términos que difieran en cuanto a su extensión. Por el supuesto (II), su significado (en el sentido de «intensión») debe ser diferente. Por el supuesto (I), *conocer el significado de A* y *conocer el significado de B*



son estados psicológicos en sentido estricto —porque así fue como entendimos el supuesto (I)—. *Pero estos estados psicológicos deben determinar la extensión de los términos A y B en la misma medida en que los significados (las «intensiones») lo hacen.*

Para ver esto, tratemos de aceptar lo contrario. Naturalmente, no puede haber dos términos A y B tales que *conocer el significado de A* sea el mismo estado que *conocer el significado de B*, incluso aunque A y B posean extensiones diferentes. Porque *conocer el significado de A* no es tan sólo «captar la intensión» de A, con independencia de qué pueda significar esto; es también saber que la «intensión» que uno ha «captado» es la intensión de A. (Por eso, quien conozca el significado de «rueda» capta presuntamente la intensión de su sinónimo alemán *Rad*; pero si no sabe que la «intensión» en cuestión es la intensión de *Rad*, no se dirá de él que «conoce el significado de *Rad*».) Si A y B son términos diferentes, *conocer el significado de A* es un estado diferente de *conocer el significado de B*, sean o no los mismos los significados de A y de B. Pero, por el mismo argumento, si  $I_1$  e  $I_2$  son diferentes *intensiones* y A es un término, *saber que  $I_1$  es el significado de A* es un estado psicológico diferente de *saber que  $I_2$  es el significado de A*. Así pues, no puede haber dos mundos lógicamente posibles  $L_1$  y  $L_2$ , que sean diferentes, tales que, por ejemplo, Óscar esté en el mismo estado psicológico (en sentido estricto) en  $L_1$  y en  $L_2$  (en todos los respectos), pero que en  $L_1$  Óscar entienda A con el significado  $I_1$  y en  $L_2$  Óscar entienda A con el significado  $I_2$ . Porque, si los hubiere, en  $L_1$  Óscar estaría en el estado psicológico *saber que  $I_1$  es el significado de A* y en  $L_2$  Óscar estaría en el estado psicológico *saber que  $I_2$  es el significado de A*, y éstos serían estados psicológicos (en sentido estricto) diferentes e incluso —suponiendo que A tenga tan sólo un significado en cada mundo para Óscar— incompatibles.

Dicho brevemente: si S es del tipo de estado psicológico que hemos estado discutiendo —un estado psicológico de la forma *saber que I es el significado de A*, donde I es una «intensión» y A es un término—, entonces en todo mundo lógicamente posible en el que el hablante esté en el estado psicológico S «vale» la misma condición necesaria y suficiente de pertenencia a la extensión de A. Porque el estado S *determina* la intensión I y, por el supuesto (II), la intensión equivale a una condición necesaria y suficiente de pertenencia a la *extensión*.

Si nuestra interpretación de la doctrina tradicional de la intensión y la extensión les hace justicia a Frege y a Carnap, en lo que respecta a la teoría del significado, todo el tema del psicologismo/platonismo parece una tempestad en un vaso de agua. (Naturalmente que este dualismo significa algo muy importante dentro de la filosofía de la matemática.) Porque incluso en el caso de que los significados sean entidades «platónicas», en vez de mentales, según el punto de vista

de Frege-Carnap, «captar» estas entidades es presuntamente un estado psicológico (en sentido estricto). Más aún: el estado psicológico determina unívocamente la entidad «platónica». De modo que tome uno la entidad platónica o el estado psicológico como lo que es el «significado», parece cosa de convención. Y optar por el estado psicológico apenas si tendría la consecuencia que Frege temía: la de que los significados dejen de ser públicos. Porque los estados psicológicos son «públicos» en el sentido de que diferentes personas (e incluso gentes de diversas épocas) pueden estar en el mismo estado psicológico. De hecho, el argumento de Frege en contra del psicologismo es únicamente un argumento contra la identificación de los conceptos con eventos mentales particulares, no con entidades mentales en general.

El carácter «público» de los estados psicológicos entraña, en particular, que si Óscar y Ernesto entienden una palabra *A* de modo distinto, deben estar en estados psicológicos *diferentes*. Pues el estado de *saber que la intensión de A es, por ejemplo, I* es el mismo estado al margen de si es Óscar o Ernesto quien esté en él. Por lo tanto, dos hablantes no pueden estar en el mismo estado psicológico en todos los respectos y comprender el término de modo diferente; el estado psicológico del hablante determina la «intensión» [y de aquí que, por el supuesto (II), la extensión también] de *A*.

Es esta última consecuencia conjunta de los supuestos (I) y (II) la que afirmamos que es falsa. Defendemos que es posible que dos hablantes se hallen exactamente en el mismo estado psicológico (en sentido estricto), a pesar de que la extensión de un término *A* en el idiolecto de uno difiera de la extensión de ese término en el idiolecto del otro. La extensión no queda determinada por el estado psicológico.

En secciones posteriores pondré esto de manifiesto con detalle. Si es correcto, hay dos vías para quien desee rescatar al menos uno de los supuestos tradicionales: renunciar a la idea de que el estado psicológico (en sentido estricto) determina la *intensión*, o renunciar a la idea de que la *intensión* determina la *extensión*. Consideraremos estas dos posibilidades después.

### ¿ESTÁN LOS SIGNIFICADOS EN LA CABEZA?

Con la ayuda de un ejemplo de ciencia-ficción menor explicaré ahora que los estados psicológicos no determinan las extensiones. Pensando en los ejemplos de ciencia-ficción que vienen a continuación, supondremos que en algún lugar de la galaxia hay un planeta al que llamaremos la Otra Tierra («*Twin Earth*»). La Otra Tierra es en gran medida como la Tierra; de hecho, allí la gente habla *español* incluso. En realidad, salvo las diferencias que especificaremos en nues-

tros ejemplos de ciencia-ficción, el lector puede suponer que la Otra Tierra es *exactamente* como la Tierra. Puede llegar a suponer, si así lo desea, que él mismo tiene un *Doppelgänger* —una copia suya idéntica— allí, aunque mi historia no dependerá de esto.

Aunque alguna de la gente de la Otra Tierra (digamos que los que se llaman a sí mismos «mejicanos», los que se llaman «argentinos», los que se llaman «españoles», etc.) hablen español, hay unas pocas pequeñas diferencias, como era de esperar, que describiremos a continuación entre los dialectos del español que se hablen en la Otra Tierra y el español de aquí. Estas diferencias dependen de las peculiaridades de ese planeta.

Una de las peculiaridades de la Otra Tierra es que el líquido llamado «agua» no es  $H_2O$ , sino uno diferente cuya fórmula química es muy larga y complicada. Abreviaré esta fórmula hablando del XYZ. Supondré que XYZ no se distingue del agua a temperaturas y presiones normales. En particular, sabe como el agua y como el agua apaga la sed. Imaginaré también que los océanos, lagos y mares de la Otra Tierra contienen XYZ, y no agua; que llueve XYZ, y no agua, etcétera.

Si una nave espacial de la Tierra llegara alguna vez a la Otra Tierra, entonces al principio se supondrá que «agua» tiene el mismo significado allí que en la Tierra. Cuando se descubriera que «agua» en la Otra Tierra es XYZ, esta suposición se rectificaría; y la nave espacial terráquea transmitiría entonces el siguiente informe:

«En la Otra Tierra, la palabra “agua” significa XYZ».

(A propósito, éste es el tipo de uso de la palabra «significado» que explica la doctrina de que la extensión sea un sentido del «significado». Pero tomemos nota de que, si bien «significa» sí que significa algo como *tiene por extensión a*, en este ejemplo, uno *no* diría

«En la Otra Tierra, el significado de la palabra “agua” es XYZ,

a menos que, como mera posibilidad, todo adulto que hable español en la Otra Tierra sepa que «agua es XYZ». Podemos dar cuenta de esto en términos de la teoría del significado que desarrollaremos más abajo; de momento, nos limitamos simplemente a observar que, aunque el verbo «significa» algunas veces significa «tiene por extensión a», la forma nominal «significado» *nunca* significa «extensión».)

Dejándonos llevar por un principio de simetría, si una nave espacial de la Otra Tierra visitara nuestro planeta alguna vez, la primera suposición sería la de que «agua» tiene el mismo significado en

ambos planetas. En cuanto se descubriera que en la Tierra «agua» es  $H_2O$ , la nave de la Otra Tierra enviaría este informe:

«En la Tierra <sup>5</sup>, la palabra “agua” significa  $H_2O$ ».

Démonos cuenta de que no hay problema alguno que tenga que ver con la extensión del término «agua». La palabra (como solemos decir) tiene dos significados diferentes: en el sentido en que se la usa en la Otra Tierra, el sentido de agua<sub>OT</sub>, lo que *nosotros* llamamos «agua» no es agua en absoluto; mientras que en el sentido en que se la usa en la Tierra, el sentido de agua<sub>T</sub>, lo que los habitantes de la Otra Tierra llaman «agua» pura y simplemente no es agua. La extensión de «agua» en el sentido de agua<sub>T</sub> es el conjunto de las totalidades formadas por moléculas de  $H_2O$ , o algo parecido; la extensión de «agua» en el sentido de agua<sub>OT</sub> es el conjunto de las totalidades formadas por moléculas de  $XYZ$ , o algo parecido a eso.

Retrocedamos ahora en el tiempo hasta 1750. Por aquel entonces ni en la Tierra ni tampoco en la Otra Tierra se había desarrollado la química. Los típicos terráqueos que hablaban español no sabían que el agua constaba de hidrógeno y oxígeno, y el típico otro-terráqueo que hablara español no sabía que el «agua» constase de  $XYZ$ . Sea Óscar<sub>1</sub> un típico terráqueo que hablase español; y sea Óscar<sub>2</sub> su contrapartida en la Otra Tierra. Podemos suponer que no hay ninguna creencia que Óscar<sub>1</sub> tuviera acerca del agua que no la tuviese Óscar<sub>2</sub> acerca del «agua». Si así se desea, se puede suponer incluso que Óscar<sub>1</sub> y Óscar<sub>2</sub> eran el uno el doble del otro en cuanto a su apariencia, sentimientos, pensamientos, monólogo interior, etc. Sin embargo, la extensión del término «agua» en la tierra era  $H_2O$  tanto en 1750 como en 1950; y la extensión del término «agua» en la Otra Tierra era  $XYZ$  tanto en 1750 como en 1950. Óscar<sub>1</sub> y Óscar<sub>2</sub> entendían el término «agua» de forma diferente en 1750, *pese a estar en el mismo estado psicológico* y pese a que, dado el estado de la ciencia en ese tiempo, a sus respectivas comunidades científicas les llevaría unos cincuenta años descubrir que entendían el término «agua» de modo diferente. Así pues, la extensión del término «agua» (y en realidad su «significado», en el uso preanalítico de este término) *no* es una función del mero estado psicológico del hablante.

Sin embargo, podría objetarse, ¿por qué habríamos de aceptar que el término «agua» tiene la misma extensión en 1750 que en 1950 (en ambos casos en la Tierra)? La lógica de los términos de clase natural, como «agua», es materia compleja, pero lo que sigue puede

<sup>5</sup> O mejor, dirán: «En la Otra Tierra (el nombre que darían a la Tierra en este otro planeta —H.P.—), la palabra “agua” significa  $H_2O$ ».

valer como esquema de una respuesta. Supongamos que señalo un vaso de agua y digo «este líquido se llama agua» (o bien: «a esto se le llama agua», suponiendo por el contexto que el marcador «líquido» se da como obvio). Mi «definición ostensiva» de agua tiene el siguiente presupuesto empírico: que la porción de líquido que estoy señalando se encuentra en una cierta relación de mismidad (por ejemplo, *x es el mismo líquido que y*, o *x es lo mismo que y*) con la mayor parte de la sustancia a la que yo y otros hablantes de mi comunidad lingüística hemos llamado «agua» en otras ocasiones. Si este presupuesto fuese falso, porque —por poner un caso— ignoro que estoy señalando un vaso de ginebra y no un vaso de agua, entonces no pretendería yo que mi definición ostensiva hubiese de aceptarse. Por ello, la definición ostensiva transmite lo que podríamos llamar una condición, necesaria y suficiente, falible («*a defeasible necessary and sufficient condition*»): la condición necesaria y suficiente que debe reunir algo para ser agua es la de hallarse en la relación mismo<sub>L</sub> con la sustancia del vaso; pero ésta es la condición necesaria y suficiente sólo si el presupuesto empírico se cumple. Si no es así, queda activada una de entre una serie de condiciones de «marcha atrás», por decirlo de este modo.

El punto clave es que la relación mismo<sub>L</sub> es una relación *teórica*: si algo es o no es el mismo líquido que *éste*, es cosa que, para ser establecida, puede llevar una cantidad indeterminada de investigación. Más aún: incluso habiéndose obtenido una respuesta «definida» por medio de la investigación científica o de la aplicación de alguna prueba de «sentido común», la respuesta es *falible*: incluso en el caso más «seguro» la investigación futura podría llegar a darle la vuelta. Así, el hecho de que un hispano-hablante podría haber llamado «agua» a XYZ en 1750, aunque él o los que siguiesen no habrían llamado agua al XYZ en 1800 o en 1850, no significa que el «significado» de «agua» cambiara en ese intervalo para el hablante medio. En 1750 o en 1850 o en 1950 uno podría haber apuntado con el dedo al líquido del lago Michigan en tanto que ejemplo de «agua». Lo que cambió fue que en 1750 habríamos pensado erróneamente que XYZ guardaba la relación mismo<sub>L</sub> con el líquido del lago Michigan, mientras que en 1800 o en 1850 habríamos sabido que ése no era el caso (ignoro, naturalmente, el hecho de que el líquido del lago Michigan era en 1950 un agua dudosa).

Modifiquemos ahora nuestra historia de ciencia-ficción. No sé si puede hacer uno ollas y sartenes con molibdeno; y si puede haberlas así, ignoro si se las podría distinguir fácilmente de las ollas y sartenes de aluminio. (No sé nada de todo esto, incluso a pesar de tener en mi léxico la palabra «molibdeno».) Así que supondré que las ollas y las sartenes de molibdeno *no pueden* distinguirse de las ollas y sartenes de aluminio, a no ser por un experto. (Para subrayar este punto, repi-

to que, hasta donde yo sé, esto podría ser verdadero; y *a fortiori* podría ser verdadero, en lo que se me alcanza, por «conocer el significado» de las palabras *aluminio* y *molibdeno*.) Supongamos ahora que el molibdeno es tan común en la Otra Tierra como lo es el aluminio en la Tierra, y que éste es tan raro allí como el molibdeno lo es aquí. En particular, supongamos que en la Otra Tierra las ollas y las sartenes de «aluminio» están hechas de molibdeno. Finalmente, supongamos que las palabras «aluminio» y «molibdeno» *tienen los papeles cambiados* en la Otra Tierra: «aluminio» es el nombre del *molibdeno* y «molibdeno» es el nombre del *aluminio*.

Este ejemplo comparte algunos rasgos comunes con los ejemplos anteriores. Si un vehículo espacial de la Tierra visitara la Otra Tierra, los visitantes no sospecharían que las ollas y las sartenes de «aluminio» que ahí encontrarán no estaban hechas de aluminio, especialmente cuando los habitantes de la Otra Tierra se lo dijeran. Sin embargo, hay una importante diferencia entre ambos casos. Un metalúrgico de la Tierra podría descubrir muy fácilmente que el «aluminio» era molibdeno, y un metalúrgico de la Otra Tierra podría hallar con igual facilidad que el aluminio era «molibdeno». (Las comillas de la oración precedente pretenden indicar los usos de los hispano-hablantes de este último planeta.) Mientras que en 1750 no había nadie, ni en la Tierra ni en la Otra Tierra, que pudiera haber distinguido el agua del «agua», la confusión del aluminio con el «aluminio» afecta tan sólo a una parte de las comunidades lingüísticas implicadas.

El ejemplo conduce a lo mismo que el precedente. Si Óscar<sub>1</sub> y Óscar<sub>2</sub> fuesen hablantes canónicos del español de la Tierra y del que se habla en la Otra Tierra, respectivamente, y ninguno posee una instrucción profunda de química o de metalurgia, puede que no haya diferencias en sus estados psicológicos cuando usan la palabra «aluminio»; sin embargo, hemos de decir que en el idiolecto de Óscar<sub>1</sub> «aluminio» tiene por extensión *aluminio* y en el idiolecto de Óscar<sub>2</sub> tiene por extensión *molibdeno*. (Hemos de decir también que Óscar<sub>1</sub> y Óscar<sub>2</sub> dicen cosas diferentes con «aluminio», que «aluminio» tiene en la Tierra un significado diferente del que tiene en la Otra Tierra, etc.) Vemos otra vez que el estado psicológico del hablante no determina la extensión (o el «significado», hablando en términos pre-analíticos) de la palabra.

Antes de seguir comentando todavía este ejemplo, permítaseme poner otro que *no* es de ciencia-ficción. Supongamos que usted es como yo y que no puede distinguir un olmo de un haya. Seguiremos diciendo aún que en mi idiolecto «olmo» tiene la misma extensión que en el de cualquier otra persona: a saber, el conjunto de todos los olmos, y que el conjunto de todas las hayas es la extensión de «haya» en todos nuestros idiolectos. Así pues, en mi idiolecto «olmo» tiene

una extensión diferente de la que en el suyo tiene «haya» (como debería ser). ¿Es realmente creíble que esta diferencia de extensión esté causada por alguna diferencia en nuestros *conceptos*? Mi concepto de olmo es exactamente el mismo que mi concepto de haya (he de confesarlo). (Esto muestra, de paso, que identificar significado «en el sentido de intensión» con *concepto* no puede ser correcto.) Si alguien intentara heroicamente mantener que la diferencia entre la extensión de «olmo» y la extensión de «haya», en *mi* idiolecto, se explica por una diferencia en mi estado psicológico, siempre podemos refutarle construyendo un ejemplo del estilo de los del «caso de la Otra Tierra» —basta con invertir los papeles de las palabras «olmo» y «haya» en tal planeta (al modo como lo hicimos con «aluminio» y «molibdeno» en el ejemplo anterior—). Más aún, supongamos que tenga yo un *Doppelgänger* en la Otra Tierra que es «idéntico» a mi molécula por molécula (en el sentido en que dos corbatas puedan ser «idénticas»). Si es usted un dualista, supongamos también que mi *Doppelgänger* tenga los mismos pensamientos verbalizados que yo, que tenga los mismos datos sensoriales, las mismas disposiciones, etc. Sería absurdo pensar que *su* estado psicológico difiera en algo del mío; sin embargo, él «dice» *haya* al proferir «olmo» y yo «diría» *olmo* al proferir «olmo». ¡Córtese el pastel como se quiera, pero los «significados» no pueden estar en la *cabeza*!

#### UNA HIPÓTESIS SOCIOLINGÜÍSTICA

Los dos últimos ejemplos dependen de un hecho lingüístico que, aunque sorprenda decirlo, parece no haber sido nunca señalado: que hay una *división del trabajo lingüístico*. Apenas podríamos emplear palabras como «olmo» y «aluminio» si no hubiera nadie que estuviera en posesión de una pauta para reconocer los olmos y el aluminio; pero no todos aquellos para quienes la distinción es importante son capaces de hacerla. Cambiemos de ejemplo: consideremos *oro*. El oro es importante por muchas razones: es un metal precioso, cumple una función monetaria, tiene valor simbólico (para mucha gente importa que el anillo de bodas que llevan sea *realmente* oro y que no sólo *parezca* ser oro), etc. Consideremos a nuestra comunidad una «factoría»: en esta «factoría» algunos hacen el «trabajo» de *llevar anillos de boda que sean de oro*, otros hacen el «trabajo» de *vender anillos de boda que sean de oro*, otros hacen el «trabajo» de *decir si algo es realmente o no de oro*. No es en absoluto necesario y tampoco es eficaz que todo el mundo lleve un anillo de oro (o unos gemelos de oro, etc.), o que el que discuta el «precio del oro», etc., se encargue de comprar y vender oro. Ni es en absoluto necesario ni eficaz que todo el que compre y venda oro sea capaz de decir si algo es o no

realmente oro, en una sociedad en la que la falta de honradez en este respecto sea poco común (vendiendo invitaciones) y en la que, en caso de duda, pueda uno consultar fácilmente a un experto. Y *ciertamente* no es ni necesario ni eficaz que todo el que tenga ocasión de comprar o de llevar oro encima sea capaz de distinguir con todo el grado de fiabilidad preciso si algo es o no realmente oro.

Los hechos precedentes son sólo ejemplos de división mundana del trabajo (en sentido amplio). Pero engendran una división del trabajo lingüístico: todo aquel para el cual el oro sea importante por cualquier razón tiene que aprender la palabra «oro»; pero no tiene que aprender el *método de determinación* de si algo es o no es oro. Puede fiarse de una subclase especial de los hablantes. Los rasgos de los que generalmente se piensa que se hallan presentes en el caso de un nombre general —condiciones necesarias y suficientes de pertenencia a la extensión, modos de determinar si algo está en la extensión («criterios»), etc.— se encuentran todos ellos presentes en la comunidad lingüística *considerada como un cuerpo colectivo*; pero ese cuerpo colectivo divide el «trabajo» de conocer y emplear estas varias partes del «significado» de «oro».

Esta división del trabajo lingüístico descansa en, y presupone, naturalmente, la división del trabajo no-lingüístico. Si sólo la gente que sabe cómo distinguir el oro de lo que no lo es pudiese explicar por qué tiene en su vocabulario la palabra «oro», entonces «oro» sería como «agua» en 1750 para esa subclase de hablantes, y los otros no la aprenderían en absoluto. Y algunas palabras no exhiben división alguna del trabajo lingüístico: «silla», por ejemplo. Pero con el aumento de la división del trabajo en la sociedad y con el alza de la ciencia, más y más palabras comienzan a exhibir este tipo de división del trabajo. «Agua», por ejemplo, no la mostraba en absoluto antes del surgimiento de la química. Hoy es necesario obviamente que todo hablante pueda reconocer el agua (fiablemente, en condiciones normales), y es probable que todo hablante adulto conozca incluso la condición necesaria y suficiente «el agua es H<sub>2</sub>O», pero sólo unos pocos de ellos podrían distinguir el agua de líquidos que superficialmente se le parecen. En caso de duda, los otros hablantes se fiarían del juicio de estos «expertos». Así, el modo de identificación de estos hablantes expertos lo posee también, a través suyo, el cuerpo lingüístico colectivo, aunque no esté en posesión de cada miembro individual de este cuerpo; y de esta forma el hecho más buscado que tenga que ver con el agua puede llegar a formar parte del significado *social* de la palabra, a la par que siga siendo desconocido para casi todos los hablantes que la aprendan. Me parece que será muy importante para los sociolingüistas investigar el fenómeno de la división del trabajo lingüístico.

Referida a él, quisiera proponer la siguiente hipótesis:



### *Hipótesis de la universalidad de la división del trabajo lingüístico:*

En toda comunidad lingüística se da clase de división del trabajo lingüístico acabada de describir: es decir, toda comunidad lingüística posee al menos algunos términos cuyos correspondientes «criterios» sólo los conoce un subconjunto de los hablantes que los aprendieron y cuyo uso por parte de los restantes depende de una cooperación estructurada entre éstos y los hablantes de los subconjuntos relevantes.

En particular, tendría interés descubrir si los pueblos extremadamente primitivos son excepciones ocasionales a esta hipótesis (lo cual indicaría que la división del trabajo lingüístico es un producto de la evolución social) o si incluso ellos la exhiben. En este segundo caso, se podría conjeturar que la división del trabajo, incluyendo el lingüístico, es un rasgo fundamental de nuestra especie.

Es fácil ver cómo este fenómeno explica algunos de los ejemplos dados más arriba del fracaso de los supuestos (I) y (II). Cuando quiera que un término se halle sujeto a la división del trabajo lingüístico, el hablante «medio» que lo aprenda no adquiere también nada que fije su extensión. En particular, *ciertamente* no es su estado psicológico individual lo que fija su extensión; es sólo el estado sociolingüístico del cuerpo lingüístico colectivo al que pertenece el hablante lo que determina la extensión.

Podemos resumir esta discusión señalando que hay dos clases de herramientas en el mundo: herramientas, como un martillo o un destornillador, que pueden ser usadas por una sola persona; y hay herramientas, como un barco a vapor, que, para ser usadas, requieren la actividad cooperativa de un cierto número de personas. De las palabras se ha pensado en demasía que respondían al primer modelo de herramienta.

### INDICABILIDAD Y RIGIDEZ <sup>6</sup>

El primero de nuestros ejemplos de ciencia-ficción —«agua» en la Tierra y en la Otra Tierra hacia 1750— no conllevaba división del trabajo lingüístico o, al menos, no la implicaba del mismo modo que los ejemplos de «aluminio» y «olmo». No había (por lo menos en nuestra historia) «expertos» en agua hacia 1750 en la Tierra, ni expertos en «agua» hacia 1750 en la Otra Tierra. El ejemplo *puede* presentarse, no obstante, de modo que suponga división del trabajo lin-

<sup>6</sup> Lo esencial de esta sección lo presenté en una serie de conferencias que di en la Universidad de Washington (en el Summer Institute in Philosophy) en 1968, así como en una conferencia en la Universidad de Minnesota.

gubático a través del tiempo. (Aquí no desarrollaré la forma de analizar el ejemplo así.) El ejemplo *sí* que entraña cosas —que ahora discutiremos— que son de fundamental importancia para la teoría de la referencia y para la teoría de la verdad necesaria.

Hay dos formas obvias de contarle a alguien lo que uno dice con una palabra de clase natural, como «agua» o «limón». Uno puede dar una de las llamadas definiciones ostensivas —«este *líquido* es agua»; «este *animal* es un tigre»; «esta *fruta* es un limón» (donde la cursiva trata de indicar que los «marcadores» *líquido*, *animal*, *fruta* pueden hallarse explícitos o implícitos). O bien uno puede proporcionar una *descripción*. En este caso, la descripción consta, como es típico, de uno o más marcadores, junto a un *estereotipo*<sup>7</sup> —una descripción normalizada de rasgos de la clase que son «típicos» o «normales» o, en cualquier caso, estereotipados—. Generalmente, los rasgos centrales del estereotipo son *criterios* —rasgos que en situaciones normales constituyen una forma de establecer si una cosa pertenece a la clase o, al menos, condiciones necesarias (o condiciones necesarias de carácter probabilístico) de pertenencia a la clase—. No todos los criterios empleados por la comunidad lingüística, en tanto que cuerpo colectivo, se incluyen en el estereotipo, y en algunos casos los estereotipos pueden ser bastante vagos. Por eso (a menos que yo sea un hablante poco representativo), el estereotipo de un olmo es precisamente el de un árbol de hoja caduca. De hecho, estos rasgos son condiciones necesarias de pertenencia a la clase (quiero decir, «necesarias» en sentido laxo; no creo que «los olmos son árboles de hoja caduca» sea *analítico*), pero están lejos de ser modos de identificación de los olmos. Por otro lado, el estereotipo de un tigre nos capacita para reconocer a los tigres (a menos que sean albinos o que presenten alguna otra circunstancia atípica), y el estereotipo de un limón nos permite generalmente reconocer los limones. En el caso extremo, el estereotipo puede ser *sólo* el marcador: el estereotipo de molibdeno podría ser *justamente* que el molibdeno es un metal. Consideremos estas dos formas de incrementar el vocabulario de una persona con un término así.

Supongamos que señale un vaso de agua y que diga «esto es agua», para enseñarle a alguien la palabra «agua». Hemos descrito algunos de los presupuestos de este acto y el modo en que esta clase de explicación semántica es falible. Trataremos ahora de aclarar más aún cómo se supone que ha de interpretarse.

A continuación, adoptaré como primitiva la noción de «mundo posible». Hago esto para creer que la noción es cabal en varios de sus sentidos y que, incluso en el caso de que fuera necesario hacerla más precisa, es científicamente importante. Asumiré, además, que al

<sup>7</sup> Véase «¿Es posible la semántica?», *Philosophical Papers*, vol. 2.

menos en algunos casos cabe hablar de un mismo individuo como si existiera en más de un mundo posible<sup>8</sup>. Nuestra discusión tiene en buena medida en cuenta la obra de Saul Kripke, aunque las conclusiones suyas y mías se obtuvieron de forma independiente.

Sean  $M_1$  y  $M_2$  dos mundos posibles en los cuales yo exista, en los que exista este vaso y en los cuales dé una explicación del significado [de una palabra] señalando este vaso y diciendo «esto es agua». (No supongo que ese líquido del vaso sea el mismo en ambos mundos.) Aceptemos que en  $M_1$  el vaso esté lleno de  $H_2O$  y que en  $M_2$  el vaso esté lleno de  $XYZ$ . Supondremos también que  $M_1$  es el mundo *real* y que  $XYZ$  es la sustancia a la que comúnmente se llama «agua» en el mundo  $M_2$  (de modo que la relación entre los hispanohablantes de  $M_1$  y los hispanohablantes de  $M_2$  sea exactamente la misma que la relación entre los hispanohablantes de la Tierra y los hispanohablantes de la Otra Tierra). Entonces, hay dos teorías que uno podría barajar acerca del significado de «agua».

(1) Uno podría decir que «agua» era *relativa-al-mundo*, pero de significado *constante* (es decir, la palabra tendría un *significado relativo constante*). Según esta teoría, «agua» *significa lo mismo* en  $M_1$  que en  $M_2$ ; sólo que en  $M_1$  el agua es  $H_2O$  y en  $M_2$  es  $XYZ$ .

(2) Uno podría decir que el agua es  $H_2O$  en todos los mundos (la sustancia llamada «agua» en  $M_2$  no es agua), pero que «agua» no tiene el mismo significado en  $M_1$  y en  $M_2$ .

Si lo que se dijo anteriormente sobre el caso de la Otra Tierra era correcto, entonces no hay duda de que la teoría buena es (2). Cuando digo «este líquido es agua», el «este» es, por decirlo así, un «este» *de re* en particular, el sentido de mi explicación es que «agua» es cualquier cosa que está en una cierta relación de equivalencia (la relación que llamamos «mismo<sub>L</sub>» más arriba) con la porción de líquido referida mediante «este» *en el mundo real*.

Podríamos simbolizar la diferencia entre las dos teorías del siguiente modo: según la teoría (1), es verdad que

(1) (Para todo mundo  $M$ ) (Para todo  $x$  de  $M$ ) ( $x$  es agua  $\equiv$   $x$  está en la relación mismo<sub>L</sub> con la entidad a la que «este» se refiere en  $M$ ),

mientras que según la teoría (2):

(2) (Para todo mundo  $M$ ) (Para todo  $x$  de  $M$ ) ( $x$  es agua  $\equiv$   $x$  está en la relación mismo<sub>L</sub> con la entidad a la que «este» se refiere en el mundo real  $M_1$ ).

<sup>8</sup> En realidad, este supuesto no hace falta para lo que sigue. Lo que sí se precisa es que la misma *clase natural* pueda existir en más de un mundo posible.

(Llamo a esto una diferencia en el «alcance» porque en (1') «la entidad a la que “este” refiere» se halla dentro del alcance del «Para todo el mundo  $M$ » —como explicita la frase cualificadora «en  $M$ »—, mientras que en (2') «la entidad a la que “este” refiere» significa «la entidad a la que “este” refiere en el mundo real» y tiene por ello referencia con independencia de la variable ligada « $M$ ».)

Kripke denomina «rígido» a un designador (en una oración dada) si en esa oración refiere al mismo individuo en todo mundo en el que el designador designe algo. Si hacemos extensiva la noción de rigidez a los nombres de sustancias, entonces podemos formular la teoría de Kripke y la mía diciendo que el término «agua» es *rígido*.

La rigidez del término «agua» se sigue del hecho de que cuando doy la definición ostensiva «este líquido es agua», me decanto por (2') y no por (1').

Podemos decir también, siguiendo a Kripke, que cuando doy la definición ostensiva «este líquido es agua», el demostrativo «este» es *rígido*.

Lo que Kripke fue el primero en observar es que esta teoría del significado (o del «uso», o de lo que fuere) de la palabra «agua» (y también de otros términos que designen clases naturales) tiene consecuencias sorprendentes para la teoría de la verdad necesaria.

A fin de explicar esto, permítaseme introducir la noción de *relación transmundana*. Una relación diádica recibirá el calificativo de *transmundana* cuando se la entienda de modo que su extensión sea un conjunto de pares ordenados de individuos *que no estén todos en el mismo mundo posible*. Por ejemplo, es fácil darse cuenta de que la relación *igual de alto que* es una relación transmundana: basta con percibirse de que si, por ejemplo,  $x$  es un individuo de un mundo  $M_1$  que mide metro setenta (en  $M_1$ ) e  $y$  un individuo de  $M_2$  que mide metro setenta (en  $M_2$ ), entonces el par ordenado  $\langle x, y \rangle$  pertenece a la extensión de *igual de alto que*. (Puesto que un individuo puede tener diferentes alturas en diferentes mundos posibles en los que ese mismo individuo exista, no es, propiamente hablando, el par  $\langle x, y \rangle$  un elemento constituyente de la extensión de *igual de alto que*, sino en realidad  $\langle x\text{-en-el-mundo-}M_1, y\text{-en-el-mundo-}M_2 \rangle$ .)

Análogamente, podemos entender la relación *mismo<sub>L</sub>* (mismo líquido que) como una relación transmundana, pensando que un líquido del mundo  $M_1$  que tenga las mismas propiedades físicas (en  $M_1$ ) que las que posea (en  $M_2$ ) un líquido de  $M_2$  está con éste en la relación *mismo<sub>L</sub>*.

Entonces, la teoría que hemos estado presentando puede resumirse diciendo que una entidad  $x$ , de un mundo posible arbitrario, es *agua* si, y sólo si, se encuentra en la relación *mismo<sub>L</sub>* (tomada como relación transmundana) con la *sustancia* que llamamos «agua» en el mundo real.

Supongamos ahora que no se hayan descubierto todavía las propiedades físicas importantes del agua (en el mundo real) —en particular, todavía ignoro que el agua es  $H_2O$ —. Puedo disponer, sin embargo, de medios satisfactorios para *reconocer* el agua (naturalmente, cabe que cometa un pequeño número de errores que no sea capaz de detectar hasta alcanzar un estadio posterior del desarrollo científico), pero no conozco la micro-estructura del agua. Sí estoy de acuerdo en que un líquido con las propiedades superficiales del agua, aunque con una micro-estructura distinta, *no es agua en realidad*, mis medios para identificar el agua (mi «definición operacional», por decirlo así) no pueden considerarse una especificación analítica de *qué* ha de tener algo para ser agua. Mejor aún: la definición operacional, al igual que la definición ostensiva, es simplemente una forma de indicar un patrón —de indicar esa sustancia del *mundo real* tal que para que  $x$  sea agua, en un mundo *cualquiera*,  $x$  ha de estar en la relación *mismo<sub>L</sub>* con los miembros *normales* de la clase de las entidades *locales* que cumplen con la definición operacional—. El «agua» de la Otra Tierra no es agua, incluso aunque sea como dice la definición operacional, porque no está en la relación *mismo<sub>L</sub>* con la sustancia *local* que cuadra con la definición operacional; y la sustancia local que cumple con la definición operacional, pero que posea una micro-estructura distinta de la del resto de las sustancias locales que se ajustan a la definición operacional, tampoco es agua, ya que no se halla en la relación *mismo<sub>L</sub>* con los ejemplos *normales* del «agua» local.

Supongamos ahora que descubra cuál es la micro-estructura del agua —que el agua es  $H_2O$ —. En este momento, seré capaz de decir que la sustancia que hay en la Otra Tierra, a la que antes *confundí* con el agua, no es realmente agua. Si usted me describe, no ya otro planeta de este universo nuestro, sino de otro universo posible en el que haya una sustancia con la fórmula química  $XYZ$ , que pase el «test operacional» del *agua*, habremos de decir que esa sustancia no es agua, sino tan sólo  $XYZ$ . Usted no habrá descrito un mundo posible en el que el agua es  $XYZ$ , sino pura y simplemente un mundo posible en el que habría lagos de  $XYZ$ , en el que la gente bebería  $XYZ$  (y no agua), etc. De hecho, una vez que hemos descrito la naturaleza del agua, no habrá ya mundo posible alguno en el que el agua no tenga esa naturaleza. Una vez que hemos descubierto que el agua es (en el mundo real)  $H_2O$ , *no habrá ya mundo posible alguno en el que el agua no sea  $H_2O$* . Concretamente, si un enunciado «lógicamente posible» es el que es verdadero en algún «mundo lógicamente posible», *no es lógicamente posible que el agua no sea  $H_2O$* .

Por otro lado, podemos imaginar perfectamente bien experiencias que nos convencerían de (y que harían que fuese algo racional creer) que el agua *no* es  $H_2O$ . En ese sentido, es concebible que el

agua no sea H<sub>2</sub>O. ¡Es concebible pero no es lógicamente posible! La posibilidad de concebir no es prueba de la posibilidad lógica.

Kripke denomina enunciados epistémicamente necesarios a los enunciados que son racionalmente intocables (suponiendo que los haya). A los enunciados que son verdaderos en todos los mundos posibles los califica, simplemente, de necesarios (o, a veces, de «metafísicamente necesarios»). De acuerdo con esta terminología, la idea que acabo de exponer puede reformularse así: un enunciado puede ser (metafísicamente) necesario y epistémicamente contingente. La intuición humana carece de acceso privilegiado a la necesidad metafísica.

Desde Kant ha habido una gran distancia entre los filósofos que pensaban que todas las verdades necesarias eran analíticas y los filósofos que pensaban que algunas verdades necesarias eran sintéticas *a priori*. Pero ninguno de estos filósofos cayó en la cuenta de que una verdad (metafísicamente) necesaria podía no ser *a priori*: la tradición kantiana fue tan culpable como la tradición empirista al decir que era lo mismo verdad metafísica que verdad epistémica. En este sentido, el reto de Kripke a la doctrina heredada va más allá de la oscilación usual empirismo/kantismo.

En este escrito, pese a todo, mi interés está puesto en la teoría del significado, y no en la teoría de la verdad necesaria. La noción de *indicabilidad*<sup>a</sup> ha servido para formular algunas ideas que tienen mucho que ver con las de Kripke. Durante mucho tiempo se ha reconocido que palabras como «ahora», «esto», «aquí» son *indicadoras* o [palabras] *de ejemplares reflexivos* («*tokenreflexive*»), es decir, que tienen una extensión que varía de contexto a contexto o de caso a caso. Nadie ha defendido, a propósito de estas palabras, la teoría tradicional de que «la intensión determina la extensión». Volviendo a nuestro ejemplo de Otra Tierra: si tengo un *Doppelgänger* en la Otra Tierra, cuando pienso «yo existo», él piensa «yo existo»<sup>b</sup>. Pero en su pensamiento verbalizado la extensión del término «yo» es él mismo (o su clase unitaria, para ser más exactos), mientras que la extensión de «yo» en mi pensamiento verbalizado soy yo (o mi clase unitaria, siendo más exactos). Así pues, la misma palabra tiene dos extensiones en dos diferentes idiolectos; pero no se sigue de eso que el concepto del yo sea de algún modo diferente del concepto que mi *Doppelgänger* tenga de sí mismo.

Ahora bien, hemos mantenido que la indicabilidad va más allá de

<sup>a</sup> Expuse estas ideas en las conferencias que di en la Universidad de Washington y en la Universidad de Minnesota.

<sup>b</sup> La oración que le sirve de ejemplo a Putnam es «I have a headache», pero he puesto en el texto otra en su lugar para que no se pierda el «yo» con la traducción de la original y para no hacer excesivamente prolija la discusión consiguiente.

los morfemas y palabras indicadoras *obvias* (también, por ejemplo, de los tiempos verbales). Nuestra teoría puede resumirse diciendo que palabras como «agua» tienen un elemento indicador oculto: el «agua» es una sustancia que guarda con el agua *de por aquí* una cierta relación de similaridad. En un tiempo o en un lugar distintos, o incluso en otro mundo posible, el agua, *si es que ha de ser agua*, ha de estar con *nuestra* «agua» en la relación mismo<sub>L</sub>. Así pues, la teoría de que (1) las palabras tienen «intensiones», que son algo parecido a los conceptos vinculados a las palabras de los hablantes; y que (2) la intensión determina la extensión, no puede ser verdadera en lo que toca a las palabras que designan clases naturales, como «agua», por la misma razón por la que no puede ser verdadera para el caso de palabras obviamente indicadoras, como «yo».

La teoría de que las palabras que, como «agua», designan clases naturales son indicadoras deja, sin embargo, abierta la cuestión de si en el dialecto español que se hable en la Otra Tierra «agua» tiene el mismo *significado* que «agua» en el dialecto de la Tierra, aunque sus extensiones no sean las mismas (que es lo que normalmente decimos de «yo» en los diferentes idiolectos, renunciando así a la doctrina de que «el significado (la intensión) determina la extensión»; o bien optar, como hemos preferido hacer, por [la tesis de] que la diferencia de extensión sea *ipso facto* una diferencia en el significado, renunciando con ello, en lo que respecta a las palabras que designan clases naturales, a la doctrina de que los significados son conceptos o, en realidad, entidades mentales de *cualquier* tipo que uno desee.

No debería haber dudas, sin embargo, en cuanto a que la doctrina de Kripke acerca de las palabras que designan clases naturales son designadores rígidos y la nuestra, según la cual son palabras indicadoras, sean dos formas de hacer la misma observación. De todo corazón aceptamos lo que dice Kripke al escribir:

«Supongamos que fijamos la referencia de un nombre por medio de una descripción. Incluso haciéndolo así, no convertimos en sinónimos el nombre y la descripción, sino que, en vez de esto, usamos el nombre rígidamente para referirnos al objeto así nombrado, incluso al hablar de situaciones contrafácticas en las que la cosa nombrada no satisfaría la descripción en cuestión. Ahora bien, es esto lo que creo que es verdad en tales casos de denominación ("*naming*") en los cuales la referencia se fija por medio de una descripción. Sin embargo, de hecho creo también que, a pesar de lo que muchos han dicho recientemente, la referencia de los nombres raramente o casi nunca se fija mediante una descripción. Y es por esto por lo que no estoy completamente de acuerdo con Searle cuando dice: "No es una descripción, sino más bien un haz de descripciones, una familia de propiedades, lo que fija la referencia." Afirmando que, en este sentido de propiedad, en absoluto se usan propiedades»<sup>10</sup>.

<sup>10</sup> S. Kripke, «Identity and Necessity», en M. Munitz (com.), *Identity and Individuation*, Nueva York, 1972, 135-164, p. 157; versión castellana en pp. 98-130 de este volumen.

## ¡¡¡ SOMOS REALISTAS

Deseo comparar ahora mis ideas con otras que, al menos entre los estudiantes, son populares (y que, al parecer, han surgido espontáneamente). Adoptemos para esta discusión la palabra *oro* como ejemplo de palabra que designa una clase natural. No haremos distinciones entre «oro» y su pareja en lengua griega, latina, etc., y nos limitaremos a «oro» en el sentido de oro en estado sólido. Dicho esto, afirmo: en dos mil años «oro» no ha cambiado de *extensión* (o no ha cambiado de modo significativo). Nuestros métodos de identificación del oro se han hecho increíblemente sofisticados. Pero la extensión de χρυσός en el dialecto del griego de Arquímedes es la misma que la extensión de *oro* en mi dialecto del español.

Es posible (y podemos suponer que así son las cosas) que hubiera fragmentos de metal de los que no pudo establecerse que *no* fueran de oro, con anterioridad a Arquímedes, que hubiese o haya también fragmentos de metal de los que no se pudiera determinar en tiempos de Arquímedes que *no* fueran de oro, pero que con las técnicas modernas podemos distinguir del oro con bastante facilidad. Sea *X* un fragmento de metal así. Obviamente, *X* no se encuentra en la extensión de χρυσός, en el griego del Ática, aunque un griego antiguo habría confundido a *X* con oro (o, mejor, con χρυσός).

La tesis alternativa es que «oro» *significa* lo que cumpla la «definición operacional» *contemporánea* de *oro*. Hace cien años, «oro» significó lo que se ajustaba a la «definición operacional» de *oro* en uso de hace cien años; «oro» significa ahora lo que atiene a la definición operacional de *oro* al uso de 1973; y χρυσός significaba lo que cuadraba con la definición operacional de χρυσός que se usó *entonces*.

Un motivo común para adoptar este punto de vista es un cierto escepticismo hacia la *verdad*. Según la tesis por la que abogo, cuando Arquímedes afirmó que algo era oro (χρυσός), no se limitaba a decir que tenía los rasgos superficiales del oro (en casos excepcionales, algo puede pertenecer a una clase natural y *no* tener, de hecho, las características aparentes de esa clase natural); estaba diciendo que tenía la misma *estructura oculta* (la misma «esencia» por así decirlo) de cualquier pedazo normal de oro. Arquímedes había dicho que nuestro hipotético fragmento de metal *X* era oro, pero habría estado *equivocado*. Sin embargo, ¿quién *ha de decir* que se habría equivocado?

La respuesta obvia es: *somos nosotros* (empleando para ello la mejor teoría que tengamos a mano). Hay mucha gente para quien la pregunta (*¿quién ha de decirlo?*) tiene fuerza y para quien nuestra respuesta carece de ella; o bien para quien nuestra respuesta tiene fuerza, pero para quien la pregunta no la tiene. ¿Por qué sucede esto?



La razón, creo yo, estriba en que la gente tiende a ser en sus intuiciones anti-realista o decididamente realista. Para alguien acusadamente anti-realista, carece de sentido decir que lo que esté en la extensión del término de Arquímedes χρυσός haya dedeterminarse utilizando *nuestra* teoría. Porque el anti-realista no considera nuestra teoría y la de Arquímedes dos descripciones aproximadamente correctas en algún dominio fijo de entidades independientes-de-toda-teoría, y tiende a ser escéptico hacia la idea de «convergencia» en la ciencia: no cree que nuestra teoría sea una descripción *mejor* de las *mismas* entidades que Arquímedes estaba describiendo. Pero si nuestra teoría es *únicamente* nuestra teoría, entonces es tan arbitrario utilizarla para determinar si *X* está o no en la extensión de χρυσός como valerse para ello de la teoría del hombre de Neanderthal. La única teoría que *no* resulta arbitrario utilizar es aquella a la que el hablante se adscribe.

La dificultad estriba en que para un decidido anti-realista la verdad no tiene sentido más que como una noción intra-teórica<sup>11</sup>. El anti-realista puede usar la verdad intra-teóricamente, en el sentido de una «teoría de la verdad como redundancia»; pero carece de las nociones de verdad y referencia disponibles extra-teóricamente. Pero la extensión está atada a la noción de verdad. La extensión de un término es justamente aquello de lo cual (el término) es verdadero. Mejor que tratar de conservar la noción de extensión vía un operacionalismo problemático, el anti-realista debería rechazar la noción de extensión como lo hace con la de verdad (en el sentido extra-teórico que haga al caso). Como Dewey, por ejemplo, él puede refugiarse en una noción de «afirmabilidad fundada» («warranted assertability»), en vez de en la de verdad (relativizada al método científico, si se piensa que hay un método científico *fijo*, o al mejor de los métodos disponibles en el momento, si está de acuerdo con Dewey en que el método científico mismo evoluciona). Entonces puede decir él que «*X* es oro (χρυσός)» era fundadamente afirmable en el tiempo de Arquímedes y que no lo es hoy (de hecho, esta afirmación es *mínima*, en el sentido de que representa aquello en lo que el realista y el antirrealista pueden ponerse de acuerdo); pero la afirmación de que *X* estaba en la extensión de χρυσός se rechazará como carente de significado, al igual que la de que “*X* es oro (χρυσός)” era verdadera.

Es bien sabido que el operacionalismo estrecho no puede explicar con éxito el uso real de los términos científicos ni el de los de sentido común. Versiones debilitadas del operacionalismo, como la versión

<sup>11</sup> Para una discusión de este punto, véase mi ensayo «Explanation and Reference», en G. Pearce y P. Maynard (comps.), *Conceptual Change*, Reidel, Dordrecht, 1973, 199-221. Reimpreso en *Philosophical Papers*, vol. 2, 196-214.

que hizo Carnap de la teoría de Ramsey, cuadran con, si no es que aplican, el uso científico real (¡principalmente porque las versiones debilitadas cuadran con todo uso posible!), pero al coste de hacer que la comunicabilidad de los resultados científicos sea un milagro. Está más allá de toda duda que los científicos usan los términos como si los criterios a ellos asociados no fueran *condiciones necesarias y suficientes*, sino caracterizaciones *aproximadamente* correctas de algún mundo de entidades independientes-de-la-teoría, y que hablan como si en una ciencia madura las teorías posteriores fueran, en general, descripciones *mejores* de las *mismas* entidades que aquéllas a las que las teorías previas se referían. En mi opinión, la hipótesis de que esto es *correcto* es la única hipótesis que puede explicar la comunicabilidad de los resultados científicos, el cierre de las teorías científicas aceptables por la lógica de primer orden, así como muchos otros rasgos del método científico<sup>12</sup>. Pero no es tarea mía defender esto aquí. Mi tesis es que si hemos de usar las nociones generales de verdad y de extensión de un modo extra-teórico (en particular, si hemos de considerar definidas estas nociones por enunciados expresados en lenguas de otras teorías que no sean la nuestra), debemos aceptar la perspectiva realista, a la cual pertenecen estas nociones. La duda de si podemos decir *nosotros* que *X* no está en la extensión de «oro», tal como Juan usó esta palabra, es la *misma* duda que la de si tiene sentido pensar en si el enunciado de Juan «*X* es oro» es *verdadero o falso* (y no solo afirmable fundamentalmente por Juan, pero no por nosotros). No hay progreso alguno en la idea de que debamos adherirnos a una teoría del significado insostenible, a fin de que la noción de verdad, que es esencialmente realista, cuadre con los propios prejuicios anti-realistas.

Un segundo motivo para adoptar una explicación operacionista extrema es una aversión hacia las hipótesis inverificables. A primera vista, puede parecer como si estuviésemos diciendo que «*X* es oro (χρυσός)» era falsa en los tiempos de Arquímedes, aunque Arquímedes no pudo haber sabido *en principio* que era falsa. Pero no es ésta exactamente la situación. El hecho es que hay muchísimas situaciones que *nosotros* podríamos descubrir (usando la misma teoría que nos dice que *X* no es oro) en las que *X* se habría comportado de un modo muy diferente del [modo en que se comportaría el] resto de las sustancias que Arquímedes consideraba oro. Quizá, al fundirse, *X* se habría convertido en dos metales diferentes, o su conductividad habría tenido otras propiedades, o se habría evaporado a una temperatura diferente, o lo que sea. Si hubiésemos llevado a cabo los expe-

<sup>12</sup> Para una iluminadora discusión de precisamente estos tres puntos, véase *Realism and Scientific Epistemology*, de R. Boyd (sin publicar, copia xerografiada distribuida por el autor, Departamento de Filosofía, de Cornell).

rimentos con Arquímedes de testigo, podría no haber conocido la teoría, pero habría podido comprobar la regularidad empírica «*X* se comporta de un modo diferente del propio de las sustancias que, en varios aspectos, catalogo como χρυσός». Eventualmente, habría llegado a concluir «*X* puede no ser oro».

La idea es que incluso en el caso de que algo cumpla los criterios usados en un momento de tiempo dado para identificar el oro (es decir, para establecer si algo es oro), puede comportarse, en una o más situaciones, de un modo diferente del que caracteriza al resto de las sustancias que cumplen los criterios. Puede que esto *no demuestre* que no sea oro, pero sugiere la hipótesis de que, a la larga, cabe que no sea oro, incluso en ausencia de toda teoría. Ahora bien, si hubiésemos procedido a informarle a Arquímedes de que el oro tenía tal y tal estructura molecular y de que *X* se comportaba de modo distinto por tener una estructura molecular diferente, ¿hay alguna duda de que estaría de acuerdo con nosotros en que *X* no es oro? Sea como sea, preocuparse porque cosas que pueden ser verdaderas (en un cierto tiempo) no puedan *verificarse* (entonces), me parece ridículo. En cualquier respecto razonable hay seguramente cosas que son verdaderas y que no pueden verificarse en *ningún* tiempo. Por ejemplo, supongamos que haya infinitas estrellas binarias. ¿Debemos verificar esto, ni siquiera *en principio*?<sup>13</sup>

Hasta aquí hemos tratado de razones *metafísicas* para rechazar nuestra explicación. Alguien, sin embargo, podría discrepar de nosotros en cuanto a los hechos empíricos que tienen que ver con las intenciones de los hablantes. Esto sucedería, por ejemplo, si alguien pensara que Arquímedes (en el *Gedankenexperiment* antes descrito) hubiera dicho: «No importa si *X* se comporta *verdaderamente* de un modo diferente del de los otros pedazos de oro; *X* es un pedazo de oro, porque *X* tiene tales-y-cuales propiedades, y esto es todo lo que conlleva ser oro». Aunque no podamos en realidad estar seguros de que todas las palabras del antiguo griego que designan clases naturales tenían las propiedades de las correspondientes palabras del español actual, no cabe ninguna duda en cuanto a las propiedades de estas segundas. Si dejamos a un lado los prejuicios filosóficos, entonces creo que sabemos perfectamente bien que ninguna definición operacional proporciona una condición necesaria y suficiente de aplicación de una palabra de esa especie. Podemos dar una «definición operacional», un haz de propiedades o lo que sea, pero la intención

<sup>13</sup> Véanse mis ensayos «Logical Positivism and the Philosophy of Mind», en *Philosophical Papers*, vol. 2, 441-451; «“Degree of confirmation” and inductive logic», en *Philosophical Papers*, vol. 1; *Mathematics, Matter and Method* (Cambridge University Press, 1975), 270-292; «Probability and Confirmation», *Philosophical Papers*, vol. 1, 293-304.

no es nunca la de «hacer al nombre *sinónimo* de la descripción». En lugar de esto, «usamos el nombre *rigidamente*» para referirnos a cualesquiera cosas que compartan la *naturaleza* que poseen las cosas que pertenecen a la descripción.

### OTROS SENTIDOS

Hasta el momento hemos analizado el sentido predominante de las palabras que designan clases naturales (o, mejor dicho, la *extensión* predominante). Pero es típico de estas palabras que posean varios sentidos. (Ziff ha sugerido incluso que poseen un *continuo* de sentidos.)

Parte de esto puede explicarse recurriendo a nuestra teoría. Para que algo sea agua, por ejemplo, ha de guardar la relación mismo<sub>L</sub> con ciertas cosas. Pero, ¿qué relación es mismo<sub>L</sub>?

$x$  está con  $y$  en la relación mismo<sub>L</sub> sólo en caso de que (1)  $x$  e  $y$  sean los dos líquidos, y (2)  $x$  e  $y$  compartan importantes propiedades físicas. El mismo término «líquido» es una palabra que designa una clase natural que no intentaré analizar aquí. El término «propiedad» es un término de amplio espectro que hemos analizado en otros escritos<sup>14</sup>. Quiero detenerme ahora en la noción de *importancia*. La importancia es una noción relativa a intereses. Normalmente, las propiedades «importantes» de un líquido, de un sólido, etc., son las propiedades que resultan importantes desde un punto de vista *estructural*: las que especifican de qué se halla últimamente hecho un líquido, un sólido, etc.: si de partículas elementales, de hidrógeno u oxígeno, de tierra, aire, fuego, agua o de lo que sea, y de cómo se dispone o combinan estos componentes para producir las características superficiales. Vistas así las cosas, lo que caracteriza a una típica gota de agua es consistir en H<sub>2</sub>O. Puede que sí o puede que no sea importante que haya impurezas; así, en un contexto «agua» puede significar *agua químicamente pura*, mientras que en otro significa: «sustancia del lago Michigan». Y un hablante puede referirse a veces a XYZ como si se tratase de agua, si uno *usa* la palabra así. También es corrientemente importante que el agua esté en estado líquido; pero en ocasiones no lo es, y uno puede llamar agua a una simple molécula de H<sub>2</sub>O, o agua al vapor de agua («el agua del aire»).

Incluso sentidos suficientemente lejanos como para considerarlos algo «marginales» pueden tener una relación definida con el sentido

<sup>14</sup> «On Properties», en N. Rescher *et al.* (comps.), *Essays in Honor of Carl G. Hempel*, D. Reidel, Dordrecht, Holland, 1970. Reimpreso en *Philosophical Papers*, vol. 1, 305-322. [En el texto, Putnam no se refiere a ninguno de sus escritos en particular, pero si hay una referencia obvia entre sus escritos, ésta es la mencionada en la nota (N del T.)]

medular. Por ejemplo, podría decir «¿viste el limón?» queriendo decir el limón *de plástico*. Un caso menos marginal es éste: descubrimos que hay «tigres» en Marte. Es decir, se parecen completamente a los tigres, pero la base de su química no es el carbono, sino la silicón. (¡Un caso notable de evolución paralela!) ¿Son tigres los «tigres marcianos»? Depende del contexto.

En el caso de esta teoría, como en el de cualquier otra ortogonal al modo en que la gente ha pensado previamente acerca de algo, es inevitable que surjan malas interpretaciones. Una que ha aparecido ya es la siguiente: un crítico ha sostenido que el sentido *predominante* de, digamos, «limón» es aquel según el cual todo lo que tenga (un número suficiente de las) propiedades de un limón es un limón. El mismo crítico ha sugerido que tener la estructura oculta —el código genético— de un limón es una condición necesaria para que algo sea un limón cuando «limón» se usa como término de la *ciencia*. Me parece que estos dos asertos se deben a una confusión o, quizá, a un par de confusiones complementarias.

El sentido en el que, literalmente, *cualquier* cosa con las características superficiales de un limón es necesariamente un limón, lejos de ser el dominante, es extremadamente marginal. En este sentido, algo sería un limón con tal de saber y de parecerse a un limón, incluso si su química tuviera como base la silicona, por ejemplo, o si un microscopio electrónico revelara que era una *máquina*. (Aun incluyendo el crecer «como un limón» entre sus características superficiales, no se excluirán los limones mecánicos; ¡puede que el limonero fuese también una máquina!)

Al mismo tiempo, el sentido en el que se tiene que tener el código genético de un limón para serlo *no* es el mismo sentido que el técnico (si lo hay, que lo dudo). El sentido técnico, tal como yo lo veo, sería uno en el que «limón» resultaría *sinónimo* de una descripción en la que se especificase el código genético. Pero cuando dijimos (por variar de ejemplo) que para ser agua se tiene que ser  $H_2O$ , no quisimos decir, como aclaramos, que el *hablante* haya de *saber* esto. Únicamente al confundir la necesidad *metafísica* con la epistémica cabe que se concluya que, si la condición veritativa (metafísicamente necesaria) que ha de reunirse para ser agua es la de ser  $H_2O$ , entonces «agua» ha de ser sinónima de  $H_2O$  —en cuyo caso se trata ciertamente de un término de la ciencia—. Y análogamente: incluso aunque el tener el código genético de un limón proporcione el sentido predominante de «limón» (como yo creo), no se sigue de esto que «limón» sea sinónimo de una descripción que especifique el código genético, explícitamente o de otro modo.

El error de pensar que hay un sentido importante de «limón» (quizá, el sentido predominante) según el cual tener las características superficiales de un limón es al menos una condición *suficiente* de

algo, si es que ha de tratarse de un limón, se hace más plausible cuando se incluye entre las características superficiales la de *ser fértil en fecundación cruzada con limones*. Pues la característica de la fertilidad en la fecundación cruzada con limones presupone la noción de limón. Así pues, aun obteniendo de *este* modo una condición suficiente, considerar que esto es inconsistente con la caracterización ofrecida aquí es pedir el principio. Más aún, la caracterización que procede en términos de «características superficiales» que presuponen saber qué es un *limón* (como la de ser fértil en la fecundación cruzada con *limones*) no proporciona condiciones veritativas que nos permitan decidir qué objetos de otros mundos posibles (o qué objetos a un millón de años-luz de aquí) son limones. (Además, no creo que esta caracterización, pidiendo el principio como lo hace, sea *correcta*, ni siquiera como condición suficiente. Creo que uno podría imaginar casos en los que algo que no fuese un limón fuera fértil al fecundarlo con otros limones, y que pareciera un limón, etc.)

También podría tratarse de excluir el caso del limón mecánico (¿máquina limonosa?) que «crece» en un árbol mecánico (¿máquina arbórea?) diciendo que su «crecimiento» no era realmente *crecimiento*. Esto es cierto: pero lo es porque *crecer* es un *verbo* que comparte la naturaleza de las palabras que designan clases naturales, y a él se aplica el género de análisis que hemos estado presentando.

Otra confusión que habría que evitar es la siguiente: interpretar el análisis que hemos desarrollado como si de él se dedujera que los miembros de la extensión de las palabras de clase natural *tienen* necesariamente una estructura oculta común. Podría haber resultado que las porciones de líquido que llamamos «agua» *no* poseyeran características físicas comunes de importancia, *a no ser* las superficiales. En ese caso, la condición necesaria y suficiente que ha de reunirse para ser «agua» sería la posesión de un número suficiente de características superficiales.

Incidentalmente, hay que decir que el último enunciado no implica que el agua podría haber tenido una estructura oculta (o que el agua podría haber sido cualquier otra cosa excepto H<sub>2</sub>O). Cuando afirmamos que podría haber *resultado* que el agua careciera de estructura oculta, lo que queremos decir es que un líquido sin estructura oculta (por ejemplo, muchas porciones de distintos líquidos, con algo en común *salvo* características superficiales) podría haberse parecido al agua, haber sabido como el agua, haber llenado los lagos, etc., que están realmente llenos de agua. Brevemente dicho: podríamos haber estado en la misma situación epistemológica, con respecto a un líquido sin estructura oculta, que aquélla en la cual estuvimos en un tiempo con respecto al agua <sup>15</sup>.

<sup>15</sup> Compárese esto con lo que dice Kripke sobre el «atril de hielo» en «Naming and

Hay, en realidad, prácticamente un continuo de casos. Ha resultado, por ejemplo, que algunas enfermedades no tenían estructura oculta (lo único que tenían en común los casos paradigmáticos era un haz de síntomas), mientras que se ha visto que otras poseían una estructura oculta común, en el sentido de una etiología (como la tuberculosis). En ocasiones, todavía no sabemos qué pasa; hay aún en la actualidad una controversia aguda a propósito de la esclerosis múltiple.

Un caso interesante es el del *jade*. Aunque los chinos no reconocen la diferencia, el término «jade» se aplica a dos minerales: a la jadeíta y a la nefrita. Químicamente, la diferencia es considerable. La jadeíta es una combinación de sodio y aluminio. La nefrita está hecha de calcio, magnesio y hierro. ¡Estas dos, bastantes diferentes, microestructuras producen las mismas cualidades de textura!

Volvamos por un momento al ejemplo de la Otra Tierra. Si tanto el H<sub>2</sub>O como el XYZ hubiesen abundado ambos en la Tierra, se habría presentado un caso similar al de la jadeíta/nefrita: habría sido correcto decir que había *dos clases de «agua»*. Y en lugar de decir que «resultó que la sustancia que hubiese en la Otra Tierra no era en realidad agua», tendríamos que decir «resultó que era *la clase XYZ de agua»*.

Resumiendo: si hay una estructura oculta, ésta determina generalmente qué sea ser un miembro de una clase natural, no sólo en el mundo real, sino en todos los mundos posibles. Dicho de otra forma determina qué podemos y qué no podemos suponer contrafácticamente acerca de la clase natural («¿podría haber sido vapor todo el agua?», sí / «¿podría haber sido el XYZ agua?», no). Pero el agua local, o lo que fuere, puede tener dos o más estructuras ocultas —o tantas que la «estructura oculta» se haga irrelevante y las características superficiales se conviertan en decisivas—.

## OTRAS PALABRAS

Hasta ahora hemos puesto como ejemplos palabras de clase natural; pero las afirmaciones que hemos hecho se aplican también a muchos otros tipos de palabras. Valen para la gran mayoría de los nombres, así como para otras partes del discurso.

Consideremos por un momento los nombres de artefactos —palabras como «lápiz», «silla», «botella», etc.—. La tesis tradicional es que estas palabras se definen mediante conjunciones o, posiblemente, haces de propiedades. Cualquier cosa que tenga las pro-

«Necessity», en G. Harman y D. Davidson (comps.). *The semantics of Natural Language*, Reidel, Dordrecht, 1972.

propiedades de la conjunción (o un número suficiente de propiedades del haz, según este modelo) es necesariamente un lápiz, una silla, una botella o lo que fuere. Además, algunas de las propiedades del haz (según el modelo del haz) usualmente se considera que son *necesarias* (según el modelo de la conjunción de propiedades, *todas* las propiedades de la conjunción son necesarias). Se supone que ser un *artefacto* es necesario, y lo mismo se supone de la pertenencia a una clase de fines canónicos: por ejemplo, «los lápices son artefactos» y «los lápices se utilizan normalmente para escribir». Finalmente, se afirma que este tipo de necesidad es *epistémica*: la analiticidad, de hecho.

Metámonos una vez más en ciencia ficción. Esta vez recurrimos a un ejemplo elaborado por Roger Albritton. Imaginemos que algún día se descubre que *los lápices son organismos*. Hacemos un corte en ellos y los examinamos con el microscopio electrónico, y vemos la trama casi invisible de nervios y otros órganos. Los espiamos y les vemos desovar y asistimos a su desarrollo hasta convertirse en lápices completos. Descubrimos que estos organismos no están imitando a los otros lápices (artefactos). Nunca hubo, ni hay todavía, más lápices que estos organismos. Seguro que es extraño, pero hay *inscripciones* en muchos de estos organismos —como: A. W. FABER CASTELL 9000 3H—<sup>c</sup>; quizá se trate de organismos inteligentes, y ésta es su forma de camuflarse. (También hemos de explicar por qué nadie intentó jamás manufacturar lápices, etc., pero éste es, en algún sentido, un mundo posible.)

Si esto fuera concebible, y estoy de acuerdo con Albritton en que lo es, sería epistémicamente posible que *los lápices fueran organismos*. Se sigue que *los lápices son artefactos* no es epistémicamente necesario en el sentido más fuerte y que no es, *a fortiori*, analítico.

Vayamos con cuidado, sin embargo. ¿Hemos mostrado que hay un mundo posible en el que los lápices son organismos? Creo que no. Lo que hemos mostrado es que hay un mundo posible en el que ciertos organismos serían las *contrapartidas epistémicas* de los lápices (según la frase de Kripke). Empleando de nuevo el recurso de la Otra Tierra, imaginemos esta vez que los lápices que hay en la Tierra sean lo que pensamos que son: artefactos manufacturados con los que escribir, mientras que en aquel otro planeta sean organismos a la Albritton. Imaginemos, además, que los habitantes de ese planeta lo ignoren por completo —que tengan sobre los lápices exactamente las opiniones que nosotros tenemos—. Cuando descubriésemos esto, no diríamos: «algunos lápices son organismos».

Supongamos ahora que la situación, tanto en la Tierra como en la Otra Tierra, sea como en el ejemplo de Albritton. Entonces diríamos

<sup>c</sup> En el original: «BONDED Grants DELUXE made in U.S.A. N.º 2».



que «los lápices son organismos». Por ello, el que esos «organismos-lápiz» de la Otra Tierra (o de otro mundo posible) sean o no realmente *lápices* está en función de si los lápices *del lugar* son o no son organismos. Si los lápices del lugar fueran lo que pensamos que son, un mundo posible en el que haya organismos-lápiz no es un mundo posible en el que *los lápices sean organismos*: tampoco ahora hay ningún mundo posible en el que los lápices sean organismos (como, naturalmente, no lo son en el mundo real). Es necesario, en el sentido de ser verdadero en todos los mundos posibles —metafísicamente necesario, por tanto—, que los lápices sean artefactos. Pero no se sigue de ahí que sea epistémicamente necesario.

Se infiere que «lápiz» no es sinónimo de ninguna descripción —ni siquiera aproximadamente sinónimo de una descripción *aproximada*—. Cuando usamos la palabra «lápiz», pretendemos referirnos a todo lo que tenga la misma *naturaleza* que los ejemplos normales de lápices cercanos que haya en el mundo real. «Lápiz» es tan *indicadora* como «agua» u «oro».

En un sentido, el caso de los lápices que resultaban ser organismos es complementario del caso que discutí hace algunos años acerca de los gatos que resultan ser robots<sup>16</sup>. Katz ha argüido<sup>17</sup> que describí erróneamente ese caso: que debería haberlo hecho de modo que *resultara que en ese mundo no hubiera gatos*. Katz admite que podríamos *decir* «ha resultado que los gatos no eran animales, sino robots»; pero aduce que esta oración presenta una desviación semántica, y que lo que quise decir se expresa mediante «las cosas a las que llamo “gatos” han resultado no ser animales, sino robots». Sin embargo, la teoría de Katz es un caso de mala lingüística. Primero de todo, la explicación de cómo es que podemos *decir* «los gatos son robots» es simplemente una explicación todo terreno que se aplica a *cualquier cosa* que podamos decir (*«is simply an all-purpose explanation of how can say anything»*). Más importante es que la teoría de Katz predice que «los gatos son robots» presenta una *desviación* [semántica], mientras que «No hay gatos en el mundo» no presenta tal desviación, sino que es canónica, en el caso descrito. Ahora bien, no niego que *haya* algún caso en el cual «No hay (ni nunca hubo) ningún gato en el mundo» sea canónica (hablando epistémicamente): podríamos descubrir que habíamos sufrido una alucinación colectiva. (Los «gatos» son como los elefantes rosados.) Pero en el caso descrito, «Ha resultado que los gatos eran movidos por control remoto desde Marte» no es con seguridad una oración que presente una desviación semántica, y sí que lo es, y mucho, «No hay gatos en el mundo».

<sup>16</sup> «It ain't necessarily so», *Philosophical Papers*, vol. 1, 237-249.

<sup>17</sup> «Logic and Language: A Defense of Intentionalism», en K. Gunderson (com.), *Minnesota Studies in the Philosophy of Science*, vol. VIII, s.f.

Incidentalmente, el análisis de Katz no es sólo un caso de mala lingüística; lo es también de mala reconstrucción. La razón por la que no usamos «gato» como sinónimo de una descripción es que sabemos lo suficiente de los gatos como para no ignorar que tienen una estructura oculta, y que es buena metodología científica usar del nombre para referirnos rígidamente a las cosas que posean esta estructura oculta, y no a lo que cuadre con alguna descripción. Como es natural, si *conociésemos* la estructura oculta podríamos elaborar una descripción a partir de *ésta*; pero no la conocemos de momento. En este sentido, el uso de las palabras de clase natural refleja un hecho de nuestra relación con el mundo: sabemos que hay clases de cosas con una estructura oculta común, pero no tenemos todavía el conocimiento [suficiente] para describir todas esas estructuras.

Sin embargo, en el caso de «lápiz» las ideas de Katz son más plausibles que en el caso de «gato». Creemos *conocer* la condición necesaria y suficiente que se precisa para ser un *lápiz*, aunque se trate de una condición vaga. Por ello, cabe que hagamos «lápiz» sinónima de una descripción imprecisa. *Podríamos* decir, en caso de que «los lápices resultaran ser organismos», *bien* que «Ha resultado que los lápices eran organismos» *bien* que «No hay lápices en el mundo» —es decir, podríamos usar «lápiz» como una palabra de clase natural o como una palabra «de un solo criterio»<sup>18</sup>—.

Por otro lado, podríamos dudar de que en la lengua natural *haya* palabras «de un solo criterio» —dejando a un lado los contextos en los que se estipule algo—. ¿No podría ser que los pediatras no sean médicos, sino espías marcianos? Responda: «sí», y habrá abandonado la sinonimia de «pediatra» y «médico especialista en el cuidado de los niños». Parece haber una fuerte tendencia por parte de las palabras que se introducen como si fuesen «de un solo criterio» a desarrollar un sentido «de clase natural», con toda su rigidez e indicabilidad concomitantes. En el caso de los nombres de artefactos, este último sentido parece ser el dominante.

(Hay un chiste sobre un paciente que está a punto de salir de un asilo para locos. Durante algún tiempo, los doctores le han estado interrogando, y ha dado respuestas perfectamente cuerdas. Deciden dejarle ir y, al final de la entrevista, uno de los doctores pregunta por casualidad: «¿Qué quiere ser cuando salga?» «Una tetera», responde aquél. El chiste sería ininteligible si literalmente no se pudiera concebir que una persona pudiese ser una tetera.)

Hay palabras, sin embargo, que conservan un sentido casi puro

<sup>18</sup> La idea de palabra «de un solo criterio» y de una teoría de la analítica basada sobre tal noción aparecen en «The Analytic and the Synthetic», *Philosophical Papers*, vol. 2, 33-69.

de expresión de un solo criterio. Se trata de palabras cuyo significado deriva de una transformación: *cazador* = *el que caza*.

El análisis ofrecido aquí no sólo se aplica a la mayoría de los nombres, sino también a otras partes del discurso. Verbos como «crecer», adjetivos como «rojo», etc., tienen todos ellos rasgos indicadores. Por otro lado, algunas palabras sincategoremáticas parecen poseer algo más que el sentido de expresión de un solo criterio. «Todo» («*whole*»), por ejemplo, puede explicarse así: *El ejército rodeó la ciudad* («*The army surrounded the town*») podría ser verdadera incluso en el caso de que la división *A* no tomase parte. *Todo el ejército rodeó la ciudad* («*The whole army surrounded the town*») significa que cada parte del ejército (del tipo relevante, como la división *A*, por ejemplo) tomó parte en la acción designada por el verbo <sup>19</sup>.

### SIGNIFICADO

Veamos ahora dónde estamos con respecto a la noción de significado. Hemos visto hasta ahora que la extensión de un término no se fija mediante un concepto que el hablante individual tenga en su cabeza, y esto es verdadero tanto porque la extensión está, en general, *socialmente* determinada —hay una división del trabajo lingüístico de igual modo que la hay del trabajo «real»— cuanto porque, en parte, la extensión se halla determinada *indicadoramente*. La extensión de nuestros términos depende de la naturaleza real de las cosas particulares que sirven de paradigmas <sup>20</sup>, y esta naturaleza real no es conocida por los hablantes en general. La teoría semántica tradicional omite solamente dos formas de contribuir a la extensión —¡la contribución de la sociedad y la contribución del mundo real!

Vimos al principio que el significado no puede identificarse con la extensión. Sin embargo, tampoco puede identificarse con la «intensión», si es que la intención es algo parecido a un *concepto* de un hablante individual. ¿Qué hemos de hacer?

Se puede tomar una de dos rutas plausibles. Una ruta sería la de seguir identificando significado y concepto, y pagar el precio de la renuncia de que el significado determina la extensión. Si siguiéramos esta ruta, podríamos decir que «agua» tiene el mismo *significado* en la Tierra y en la Otra Tierra, aunque diferentes *extensiones*. (No sólo una diferente extensión *local*, sino una diferente extensión *global*). El XYZ que haya en la Otra Tierra no estaría en la extensión de las diversas preferencias de «agua» que yo hiciera, pero sí en las que

<sup>19</sup> Este ejemplo procede de un análisis hecho por Anthony Kroch (en su tesis doctoral en el M.I.T., 1974, Departamento de Lingüística).

<sup>20</sup> No tengo en mente la noción flewiana de «paradigma», según la cual paradigma de un *K* es *necesariamente* un *K* (en la realidad).

haga mi *Doppelgänger* de ese mismo término; y esto no sólo porque ese planeta esté lejos de donde me hallo, pues las moléculas de H<sub>2</sub>O están en la extensión de las preferencias de «agua» que yo haga, no importa cuán lejos estén de mí en el espacio y el tiempo. Del mismo modo, lo que yo pueda suponer contrafácticamente acerca del agua es distinto de lo que pueda suponer contrafácticamente mi *Doppelgänger*.) Pese a que es ésta la ruta correcta que hay que seguir en el caso de una palabra *absolutamente* indicadora, como «yo», parece ser inadecuada en el caso de las palabras que hemos estado discutiendo. Consideremos «haya» y «olmo», por ejemplo. Si éstas «intercambiaran sus papeles» en la Otra Tierra, seguramente *no* diríamos que «olmo» tiene allí el mismo significado que en la Tierra, aunque el estereotipo que tenga mi *Doppelgänger* de una haya (o de un «olmo», como él diría) sea idéntico a mi estereotipo de un olmo. En lugar de esto, diríamos que en el idiolecto de mi *Doppelgänger* «olmo» significa *haya*. Por esta razón parece preferible seguir una ruta diferente y decir que el «significado» es un par ordenado (o que puede ser una *enéada* ordenada) de entidades, *una de las cuales sea la extensión*. (Los otros componentes del «vector del significado», por llamarlo así, se especificarán más tarde.) Optanto por esto, que *el significado determina la extensión* (es decir, que una diferencia de extensión es *ipso facto* una diferencia de significado) es trivialmente verdadero, aunque se abandone completamente la idea de que si hubiera una diferencia en cuanto al significado que yo y mi *Doppelgänger* le atribuyésemos a una palabra, *habría* de deberse a una diferencia en nuestros conceptos (o a nuestro estado psicológico). Siguiendo esta ruta podemos decir que al proferir la palabra «olmo» mi *Doppelgänger* y yo *decimos algo diferente*; pero este aserto no versará sobre nuestros estados psicológicos. Todo lo que significa es que las preferencias que haga él de la palabra tienen una extensión diferente de la que tengan las mías; pero esta diferencia de extensión no refleja diferencia alguna en nuestra competencia lingüística individual, considerada por separado.

Si esto es correcto, y yo creo que lo es, el problema tradicional del significado se divide en dos problemas. El primero es el de explicar la *determinación de la extensión*. Puesto que a menudo la extensión se determina social y no individualmente, debido a la división del trabajo lingüístico, creo que este problema es estrictamente un problema de sociolingüística. Resolverlo conllevaría analizar detalladamente cómo tiene lugar la división del trabajo lingüístico. La llamada «teoría causal de la referencia», introducida por Kripke para los nombres propios y extendida por nosotros a los términos de magnitud física, en otro escrito <sup>21</sup>, y a las palabras de clase natural, aquí, cae dentro

<sup>21</sup> En «Explanation and Reference», *loc. cit.*

de este ámbito. Pues el hecho de que en muchos contextos atribuyamos a las preferencias que yo haga de un nombre el mismo referente que le atribuyó, en sus preferencias del mismo nombre, la persona de quien yo lo aprendí (de modo que la referencia se transmitiera de hablante a hablante, comenzando por los que estaban presentes en la «ceremonia bautismal», a pesar de no transmitirse ninguna *descripción*) es simplemente un caso especial de cooperación en la determinación de la referencia.

El otro problema es el de describir la *competencia individual*. La extensión puede determinarse, en muchos casos, socialmente, pero *importa cómo* use Juan la palabra *W* a la hora de atribuir una extensión canónica a sus preferencias de *W*. A fin de que pueda desempeñar su papel en la división lingüística del trabajo, Juan ha de tener algunas ideas y destrezas referidas a *W*. Una vez que hayamos renunciado a la idea de que la competencia individual deba de ser tan grande que determine realmente la extensión, podremos comenzar a estudiarla con la mente bien despejada.

A este respecto, es instructivo observar que nombres como «tigre» o «agua» son muy diferentes de los nombres propios. Cabe que usemos correctamente el nombre propio «Napoleón» sin saber nada del referente, excepto que se llama «Napoleón» —e incluso no ser esto cierto: «Hace muchos, muchos años, vivía en un lejano país un rey llamado Midas»<sup>d</sup>—. Pero no puede usarse correctamente la palabra «tigre», a no ser *per accidens*, sin saber bastantes cosas de los tigres o sin haberse forjado un cierto concepto de los tigres. En este sentido, los conceptos tienen *mucho* ver con el significado.

De igual modo que el estudio del primer problema es propiamente un tema de la sociolingüística, el estudio del segundo es estrictamente un tema de la psicolingüística. A este segundo tema nos dedicaremos ahora.

## ESTEREOTIPOS Y COMUNICACIÓN

Supongamos que un hablante sepa que «tigre» tiene por extensión un conjunto de objetos físicos, y que no sepa más que esto. Si en otros respectos posee una competencia lingüística normal, podría usar «tigre» en *algunas* ocasiones: por ejemplo, «los tigres tienen

<sup>d</sup> El texto original da otro ejemplo: «(“Once upon a time, a very long time ago, about last Friday, Winnie-the-Pooh lived in a forest all by himself under the name of Sanders”).» A muchos lectores del español les será más familiar el cuento del rey Midas —al menos, lo he supuesto así—, pero esta libertad que me he tomado no interfiere con la intención del autor. De Winnie-the-Pooh se nos dice, en el inicio de su cuento, que vivía solo en un bosque, así como de Midas se dice en el otro que era un rey. Por otro lado, al margen de esta mínima información todo lo que se añade es cómo se llamaban.

masa», «los tigres ocupan espacios», «¡dame un tigre!», «¿es eso un tigre?», etc. Más aún, en la extensión *socialmente determinada* de «tigre», en esas oraciones, sería la canónica: a saber, el conjunto de los tigres. Sin embargo, no diríamos de ese hablante que «conoce el significado» de la palabra «tigre». ¿Por qué no [lo conoce]?

Antes de tratar de dar una respuesta a esta pregunta, reformulémosla una pizca. Diremos de una persona que ha *adquirido* la palabra «tigre» si es capaz de usarla de forma tal que (1) su uso sea cabal (es decir, que la gente no diga de él cosas como: «no sabe qué es un tigre», «no conoce el significado de la palabra “tigre”», etc.); (2) que su forma íntegra de hallarse situado en el mundo y en su comunidad lingüística sea tal que la extensión socialmente determinada de la palabra «tigre», en su idiolecto, resulte ser el conjunto de los tigres. La cláusula (1) significa, a grandes rasgos, que no se considera que hablantes como ese hipotético del párrafo precedente hayan aprendido la palabra «tigre» (u otra cualquiera). Podríamos decir de ellos, en algunos casos, que han *adquirido parcialmente* la palabra; pero dejemos esto por el momento. La cláusula (2) significa que no se considera que los hablantes de la Otra Tierra que tengan nuestros mismos hábitos lingüísticos hayan adquirido la palabra «tigre» sólo porque la extensión de «tigre» sea, en su idiolecto, el conjunto de los tigres. Lo esencial de las secciones precedentes de este escrito es que *no* se sigue que la extensión «tigre» en el dialecto (o en los idiolectos) de la Otra Tierra sea el conjunto de los tigres solamente porque sus hábitos lingüísticos sean los mismos que los nuestros: la naturaleza de los tigres que allí vivan es relevante también. (Si los organismos de la Otra Tierra tuvieran, digamos, una química basada en la silicona, entonces sus «tigres» no serían realmente tigres, incluso aunque se parecieran a los tigres, aunque los hábitos lingüísticos del lego de aquel lugar se correspondieran exactamente con los de los hablantes de la Tierra.) Así pues, la cláusula (2) significa que en este caso hemos decidido decir que los hablantes de la Otra Tierra no habrían adquirido nuestra palabra «tigre» (aunque hayan adquirido otra palabra que se deletrea y se pronuncia igual).

Nuestra razón para adherirnos a esta forma de expresarlo es que la pregunta «¿conoce el significado de la palabra “tigre”?» nos predispone a favor de la teoría según la cual aprender una palabra no es sino llegar a poseer una cosa denominada su «significado». Digamos que esta cosa es un concepto y estaremos de vuelta en la teoría de que una condición suficiente del aprendizaje de una palabra es la de su asociación con el justo concepto (o, más en general, la de estar en el estado psicológico adecuado con respecto a él) —la misma para cuya refutación hemos empleado todo este tiempo—. Así que, consistentemente, «adquiriremos» palabras, en vez de «aprenderemos su significado».

Ahora podemos reformular la pregunta con la que se inició esta sección. El uso del hablante que describimos no recibirá el visto bueno, aunque no nos lleve a atribuir una extensión no-canónica a la palabra «tigre» en su idiolecto. ¿Por qué no recibirá el visto bueno?

Supongamos que nuestro hipotético hablante señale una bola de nieve y pregunte: «¿es eso un tigre?». Obviamente no tiene mucho sentido que hablemos de tigres con él. La comunicación significativa exige que la gente sepa algo de aquello de lo cual se habla. Es seguro que oímos a la gente «comunicarse» entre sí sin saber nada de las cosas sobre las cuales hablan; pero el sentido en que no sabe nada el hombre que apunta con el dedo a una bola de nieve y pregunta: «¿es eso un tigre?», es mucho más radical que el que piense que Sudán ganará la Copa Davis<sup>e</sup> o que la guerra del Vietnam se libró para ayudar a los sudvietnamitas. El problema de la gente que piensa que Sudán ganará la Copa Davis o el del que cree que la guerra del Vietnam se libró para ayudar a los sudvietnamitas no puede resolverse con la adopción de convenciones lingüísticas; pero ignorar acerca de qué habla uno en este frustrante segundo sentido es cosa que se puede prevenir, y que se previene, casi del todo por medio de nuestras convenciones lingüísticas. Lo que yo pienso es que a los hablantes se les *pide* que sepan algo (estereotipado) sobre los tigres, si es que ha de considerarse que han adquirido la palabra «tigre»; algo concerniente a los olmos (o en cualquier caso, de su estereotipo), para que se juzgue que han adquirido la palabra «olmo», etc.

Esta idea no habría de sorprender en demasía. Después de todo, no permitimos que la gente conduzca por las autopistas sin pasar primero algunas pruebas que determinen si tienen un nivel *mínimo* de competencia; y no comemos con la gente que no haya aprendido a usar cuchillo y tenedor. También la comunidad lingüística tiene sus exigencias mínimas, tanto en lo que hace a la sintaxis como en lo que respecta a la «semántica».

La naturaleza del mínimo nivel de competencia exigido depende, no obstante, en gran parte tanto de la cultura como del tema. En nuestra cultura, a los hablantes se les pide que sepan cuál es el aspecto de los tigres (si es que han de adquirir la palabra «tigre», y esto es virtualmente obligatorio); no se les pide que conozcan los pequeños detalles (como la forma de las hojas) de la apariencia de un olmo. A los hablantes hispanos *su comunidad lingüística les pide* que puedan distinguir los tigres de los leopardos; no les pide que sepan distinguir las hojas de los olmos de las hojas de las hayas.

Las cosas podrían haber sido distintas con facilidad. Imaginemos

<sup>e</sup> En el texto original se hace referencia a «the man who thinks that Vancouver is going to win the Stanley Cup».

una tribu india, como la de los Cheroqui, que tenga palabras como *uhaba'* y *wa'arabi* para las hojas de los olmos y las hojas de las hayas, respectivamente. De un cheroqui que no supiese reconocer un olmo se diría que no sabe qué es una *uhaba'*, no que ignore el significado de la palabra *uhaba'* (quizá, que no sabe la palabra o que no tiene [en su léxico] la palabra); justo igual que en el caso de un hablante hispano que no tuviera idea de que la piel de los tigres es rayada, del cual se diría que no sabe qué es un tigre, y no que ignora el significado de la palabra «tigre» (naturalmente, si supiera al menos que los tigres son grandes felinos, podríamos decir que conoce parte del significado o que conoce parcialmente el significado), etc. Entonces, traducir *uhaba'* como «olmo» y *wa'arabi* como «haya» sería, según nuestro punto de vista, sólo *aproximadamente* correcto. En este sentido, hay una dificultad real en el caso de la traducción radical<sup>22</sup>, pero no es la dificultad abstracta de la que Quine nos habla<sup>23</sup>.

#### QUÉ SON LOS ESTEREOTIPOS

Introduje la noción de «estereotipo» en mis Conferencias de la Universidad de Washington y en el Minnesota Center for the Philosophy of Science, en 1968. El trabajo posteriormente publicado, «Is Semantics Possible?»<sup>24</sup>, sigue el argumento dado, y en el presente escrito deseo introducir la noción otra vez y responder a algunas preguntas que se me han hecho [al respecto].

En el habla ordinaria, un «estereotipo» es una idea convencional (frecuentemente maliciosa y que puede ser hartamente imprecisa) de cómo parece ser, de cómo es o de cómo se comporta un *X*. Obviamente, exploto algunos de los rasgos del habla común. No me ocupo de estereotipos maliciosos (salvo donde el lenguaje mismo lo sea); lo hago de ideas convencionales, que pueden ser imprecisas. Sugiero que ideas convencionales así se hallan asociadas a «tigre», a «oso», etc., y más aún: que esto es el solo elemento de verdad que hay en la teoría del «concepto».

De acuerdo con esta tesis, a quien sepa lo que significa «tigre» (o, como hemos decidido hacer en su lugar, quien haya adquirido la palabra «tigre») se le *pide* que sepa que los tigres *estereotípicos* ("stereotypical") tienen la piel rayada. Dicho en términos más preci-

<sup>22</sup> El término se debe a Quine (en *Palabra y objeto*): significa traducción en la que no se puede recurrir a indicios debidos ni a la cultura ni a otras cosas afines que se compartan.

<sup>23</sup> Para una discusión de la presunta imposibilidad de una única traducción correcta, véase «The refutation of conventionalism», en *Philosophical Papers*, vol. 2, 153-191.

<sup>24</sup> Véase «¿Es posible la semántica?», *Philosophical Papers*, vol. 2.



los: hay *un* estereotipo de los tigres (él puede tener otros) que la comunidad lingüística como tal exige: se le pide que tenga este estereotipo y que sepa (implícitamente) que es obligatorio. Este estereotipo debe incluir el rasgo de las rayas en la piel, para que su adquisición se juzgue conseguida.

El hecho de que un rasgo (por ejemplo, el de la piel rayada) se incluya en el estereotipo asociado a una palabra *X* no significa que sea una verdad analítica que todos los *X* tengan ese rasgo, ni que lo tengan muchos *X*, ni que todos los *X* normales lo posean, ni que algunos *X* lo posean <sup>25</sup>. Los tigres con tres patas y los tigres albinos no son entidades lógicamente contradictorias. Descubrir que nuestro estereotipo ha estado basado en miembros no-normales o no-representativos de una clase natural no es descubrir una contradicción lógica. Si los tigres perdieran sus rayas, no por eso dejarían de ser tigres; ni las mariposas dejarían necesariamente de ser mariposas, si perdieran las alas.

(Hablando propiamente, la situación es más complicada de lo que esto sugiere. Es posible darle a una palabra como «mariposa» un sentido según el cual las mariposas dejarían de ser mariposas de perder sus alas —por mutación, por ejemplo—. Por ello, uno puede hallar un sentido de «mariposa» de acuerdo con el cual «las mariposas tienen alas» sea analítico. Pero el sentido más importante del término es, creo, aquel en el que las mariposas sin alas siguen siendo mariposas.)

Llegados a este punto, el lector puede preguntarse cuál es el valor que tiene para la comunidad lingüística poseer estereotipos, si la «información» contenida en ellos no es necesariamente correcta. En realidad, no hay ningún misterio en esto. La mayoría de los estereotipos capturan verdaderamente rasgos poseídos por los miembros paradigmáticos de la clase en cuestión. Incluso cuando los estereotipos yerran, el modo en que esto ocurre ilumina la contribución que hacen normalmente los estereotipos a la comunicación. El estereotipo del oro, por ejemplo, contiene el rasgo *amarillo*, pese a que el oro químicamente puro sea casi blanco. Pero el oro que vemos en las joyas es típicamente amarillo (debido a que contiene cobre), por lo que la presencia de este rasgo en el estereotipo es útil incluso en contextos de lego. El estereotipo ligado a *bruja* es más acusadamente erróneo, al menos si se piensa en su significación existencial. Creer (con significación existencial) que las brujas pactan con Satán, que causan enfermedad y muerte, etc., facilita la comunicación sólo en el sentido de hacerlo para quien pueda creer en la verdad de la teoría de las brujas. No facilita la comunicación en ninguna situación en la que lo que se necesite sea un acuerdo con el mundo mayor que con la teoría de

<sup>25</sup> Véase «The Analytic and the Synthetic», *loc. cit.*

otros hablantes. (Estrictamente hablando, me refiero al estereotipo que existió en Nueva Inglaterra hace trescientos años; hoy, el que las brujas no sean reales es parte del estereotipo, y los funestos efectos de la teoría de las brujas están, por lo tanto, neutralizados.) Pero el hecho de que nuestro lenguaje tenga *algunos* estereotipos que impiden, en lugar de facilitar, nuestros tratos con el mundo y con los demás sólo indica que no somos seres infalibles. ¿Y qué hay de raro en esto? El hecho es que apenas podríamos comunicarnos bien, si la mayoría de nuestros estereotipos no fueran tan atinados como lo son.

#### EL SIGNIFICADO «OPERACIONAL» DE LOS ESTEREOTIPOS

Una cuestión más enrevesada es la siguiente: ¿en qué medida puede «definirse operacionalmente» la noción de estereotipo? Aquí es necesario ser extremadamente cuidadosos. Los intentos de las ciencias físicas por especificar *literalmente* definiciones operacionales de sus términos han fracasado estrepitosamente; y no hay razón por la que habiendo fallado en la física el intento deba tener éxito en la lingüística. A veces, los argumentos de Quine contra la posibilidad de una teoría del significado parecen limitarse a la petición de definiciones operacionales en lingüística; siempre que el caso sea éste, los argumentos habrán de ignorarse. Pero ocurre frecuentemente que los términos tienen definiciones operacionales, no en el mundo real, sino en circunstancias idealizadas. Como es frecuente con las idealizaciones, las «definiciones operacionales» tienen valor heurístico sólo cuando vemos en la definición operacional algo más que una idealización conveniente; si no, resultan dañinas. Así, podemos preguntar cuál es el «significado operacional» del enunciado de que una palabra tenga tal y cual estereotipo, sin suponer que la respuesta a esta pregunta valga como una explicación teórica de qué significa ser un estereotipo.

La explicación teórica de lo que significa ser un estereotipo procede en términos de la noción de *obligación lingüística*, una noción que creemos fundamental para la lingüística y que no intentaremos dilucidar aquí. Qué significa decir que tener la piel rayada sea parte del estereotipo (lingüístico) de «tigre» es que es *obligatorio* adquirir la información de que los tigres estereotípicos tienen la piel rayada cuando uno aprende «tigre», en el mismo sentido de «obligatorio» en que lo es aclarar si uno habla de varones o de hembras al emplear la palabra española «hombres»<sup>1</sup>. No es difícil describir una prueba ex-

<sup>1</sup> En el original se usa la palabra «lions» y la distinción singular/plural.

perimental de esta hipótesis. Introduzcamos una persona a la que podemos llamar el *cómplice* del lingüista. El cómplice será (o pretenderá ser) un adulto cuyo conocimiento del español sea excelente en términos generales, aunque por alguna razón (¿creció en una cultura distinta?, ¿se crió en un monasterio?) no haya adquirido en absoluto la palabra «tigre». El cómplice dirá la palabra «tigre» o, mejor todavía, tratará de decirla (como si no estuviese seguro de cómo se pronuncia) y preguntará «¿qué significa esta palabra?» o «¿qué es esto?» u otras cosas de este estilo. Ignorando todo aquello que no vaya bien en la práctica experimental, lo que nuestra hipótesis implica es que los informadores típicos deberían decirle que los tigres tienen, *inter alia*, la piel rayada.

En lugar de fiarnos de los cómplices, cabría esperar que el lingüista estudie a los niños que estén aprendiendo español. Pero a los niños que están aprendiendo su lengua materna u otra no se les enseña ni mucho menos, como suponen algunos filósofos; la aprenden, pero no se les enseña, como Chomsky ha subrayado. Sin embargo, los niños preguntan a veces: «¿qué es un tigre?», y nuestra hipótesis implica que también en estos casos los informadores habrían de decirles, *inter alia*, que los tigres tienen la piel rayada. Un problema, no obstante, es el de que los informadores sean probablemente los propios padres, y que habría que dejar cabida a variaciones en el humor, la atención y el tiempo que les prestan a sus hijos.

Sería fácil especificar un gran número de consecuencias «operacionales» de nuestra hipótesis, pero no tendría ningún valor hacer tal cosa. El hecho es que somos hablantes plenamente competentes del español, con un toque de buen sentido para nuestras obligaciones lingüísticas. Pretender que, con respecto al español, estemos en la situación de los marcianos no es el camino de la claridad metodológica; después de todo, sólo al dejarse de lado el enfoque operacional la lingüística transformacional floreció como ciencia.

Así pues, si alguien me preguntase el significado de «tigre», sé muy bien que le diría. Le diría que los tigres son felinos, le diría algo acerca de su tamaño, que son amarillos con rayas negras, que (a veces) viven en la jungla y que son fieros. Podría decir también otras cosas, según cual fuera el contexto y la razón por la que se me hiciera la pregunta; consideraría que la información anterior, salvo quizás por lo que toca a la jungla, es *obligatoria* de transmitir, y estoy seguro de que es esto aproximadamente lo que sería también obligatorio decir para otros hablantes. Naturalmente, habrá alguna diferencia de idiolecto a idiolecto; el rasgo de tener la piel rayada (dejando a un lado las relaciones figura-fondo: por ejemplo, ¿se trata de rayas negras sobre fondo amarillo, que es como lo veo yo, o de rayas amarillas sobre fondo negro?) se encontrará en todos los idiolectos normales, aunque algunos hablantes podrían considerar obligatoria la informa-

ción de que los tigres habitan (estereotípicamente) en las junglas, mientras que para otros esto no sería así. De modo alternativo, se podría juzgar obligatorio algún rasgo del estereotipo (gato-grande-aninas), y los otros *opcionales*, según el modelo de ciertos rasgos sintácticos. Sin embargo, no nos ocuparemos aquí de explorar esta posibilidad.

### EL TORNO A LOS «DOS DOGMAS» DE QUINE

En «Dos dogmas del empirismo», Quine lanzó un poderoso y sano ataque a la, por entonces de moda, distinción analítico-sintético. La distinción había crecido hasta convertirse en una verdadera devoradora filosófica de hombres: lo analítico se *identificaba con* lo necesario [y esto] se *identificaba con* lo que por principio no se halla sujeto a revisión [y esto] se *identificaba con* cualquier verdad que cualquier filósofo individual considerara obvia («*whatever truth the individual philosopher wished to explain away*»). Sin embargo, el mismo ataque de Quine fue demasiado lejos en ciertos aspectos; tenemos la impresión de que alguna clase reducida de oraciones puede quedar a salvo. Más importante que esto: con posterioridad, el propio Quine y otros entendieron el ataque como si el ocaso de la distinción analítico-sintético afectara a toda la noción de significado. Aunque hemos dejado claro que estamos de acuerdo en que la noción tradicional de significado se enfrenta a serios problemas, en este trabajo nuestro proyecto tiene carácter constructivo, no destructivo. Vamos a enmendar la noción de significado, no a enterrarla. De modo que será útil ver cómo quedan los argumentos de Quine tras nuestra revisión.

Los argumentos de Quine en contra de la noción de analiticidad pueden reducirse fundamentalmente a lo siguiente: que no puede uno atribuir a esta noción ningún significado conductual. Su argumento (simplificándolo de nuevo en alguna medida) era el de que solo había dos candidatos a ser índice conductual de la analiticidad y que ambos eran totalmente insatisfactorios, si bien por razones distintas. El primer índice conductual es la *centralidad*: muchos filósofos contemporáneos llaman analítica a una oración si, efectivamente, alguna comunidad lingüística sostiene que es inmune a toda revisión. Sin embargo, aduce Quine de un modo persuasivo, el máximo grado de inmunidad tal no es una prerrogativa exclusiva de las oraciones analíticas. Las oraciones que expresan leyes fundamentales de la física (como el principio de conservación de la energía) pueden bien gozar de una máxima inmunidad conductual a ser revisadas; aunque apenas si sería costumbre o si resultaría plausible que se las clasificase como analíticas. No obstante, no se deja llevar únicamente de que

no sería plausible clasificar como analíticos todos los enunciados a cuya renuncia fuésemos muy reacios; señala que «la inmunidad a ser revisados» es, en la historia real de la ciencia, una *cuestión de grado*. En la historia real de la ciencia no hay nada parecido a una *absoluta* inmunidad tal. Así pues, si identificásemos la analiticidad con la inmunidad a la revisión se alteraría aquella noción de dos formas fundamentales: la analiticidad se convertiría en algo gradual, y no habría ninguna oración analítica. Esto sería salirse de la noción de analiticidad de Carnap-Ayer-*et al.*, y Quine se da cuenta de que, si no es de ésta de la que queremos hablar, sería mucho menos equívoco introducir un término bien distinto: el de *centralidad*, pongamos por caso.

El segundo índice conductual es el de *llamarse «analítico»*. En efecto, algunos filósofos piensan que el sello de la analiticidad es el de que los informadores cualificados (digamos que los miembros de la Sociedad de Fellows, de Oxford) *llamen* analítica a una oración. Algunas variantes de este índice son: que la oración sea deducible de las oraciones de una lista en cuya cabecera alguien que haya guardado con Carnap el ancestro de la relación *alumno de* haya escrito las palabras «Postulado de significado»; que pueda obtenerse la oración a partir de un teorema de la lógica, sustituyendo sinónimos por sinónimos. Esta última variante parece prometedora, pero Quine lanza contra ella la pregunta: «¿Cuál es el criterio de sinonimia?». Un criterio posible sería el de que las palabras  $W_1$  y  $W_2$  son sinónimas si, y sólo si, el bicondicional « $\wedge x$  ( $x$  está en la extensión de  $W_1 \equiv x$  está en la extensión de  $W_2$ )» es *analítico*; pero esto nos lleva a un círculo vicioso. Otro podría ser el de que las palabras  $W_1$  y  $W_2$  son sinónimas si, y sólo si, informadores cualificados las *llaman* sinónimas; pero esto es una forma ligeramente distinta de nuestro segundo índice. Una línea prometedora es la [que dice] que las palabras  $W_1$  y  $W_2$  son sinónimas si, y sólo si,  $W_1$  y  $W_2$  son intercambiables (pueden ponerse la una donde la otra) *salva veritate* en todos los contextos de una clase adecuada. Pero Quine muestra convincentemente que también esta propuesta nos hace andar en círculo. Así pues, el segundo índice se reduce a esto: una oración es analítica si ella o alguna expresión, o bien una sucesión de pares ordenados de expresiones, o un conjunto de expresiones, relacionada(s) con aquélla de ciertos modos específicos, se encuentra en una clase a todos cuyos miembros los informadores cualificados aplican un *sonido* particular: o el *sonido* ANALÍTICO o el *sonido* POSTULADO DE SIGNIFICADO o el *sonido* SINÓNIMOS. En última instancia, esta propuesta hace de «analítico», etc., *sonidos sin elucidar*.

Pese a que Quine no discuta esto explícitamente, es manifiesto que no sería más satisfactorio adoptar [como índice conductual] la intersección de dos índices insatisfactorios; explicar la analiticidad de una oración como si consistiera en su centralidad y en que se la llame

ANALÍTICA es justamente decir que las oraciones analíticas son una subclase de las oraciones centrales, sin explicarnos en absoluto en qué consiste lo excepcional de esa subclase. En efecto, la conclusión de Quine es que la analiticidad o es centralidad mal entendida o no es nada.

A pesar del vigoroso argumento de Quine, muchos filósofos han seguido abusando de la noción de analiticidad, confundiéndola a menudo con la de máximo grado de centralidad. Enfrentados a las alternativas de Quine, han optado por identificar analiticidad y centralidad, y por pagar el precio —el de clasificar como «analíticas» oraciones con un aspecto de sintéticas tan obvio como el de «el espacio tiene tres dimensiones»— y el de decidirse a mantener la tesis de que, después de todo y pese a la impresionante evidencia en contra de ello, en la ciencia existe la irrevisabilidad absoluta. Pero esta línea puede ser dinamita emparejando el argumento de Quine con un importante argumento de Reichenbach<sup>26</sup>.

Reichenbach mostró que existe un *conjunto* de principios cada uno de los cuales Kant había considerado sintéticos *a priori*, pero cuya conjunción es incompatible con los principios de la relatividad especial y de la covarianza general. (Entre estos principios están los de inducción normal, continuidad del espacio y el del carácter euclídeo del espacio.) Un kantiano podría aferrarse sin contradecir a la geometría euclídea, sin importarle qué venga después; pero la experiencia puede obligarle a renunciar a la inducción normal o a la continuidad del espacio; pero entonces la experiencia le forzará a renunciar a la Geometría euclídea (esto sucede cuando el espacio físico no sea homomórfico con un espacio euclídeo cualquiera)<sup>27</sup>.

Aplicado en el presente contexto, lo que esto muestra es que hay principios que los filósofos aficionados a la marchita noción de analiticidad, y en particular los filósofos que identifican analiticidad e irrevisabilidad (máxima), catalogarían como analíticos, pero cuya conjunción tiene consecuencias empíricas comprobables. Así que o bien se ha de renunciar a identificar analiticidad y centralidad, de una vez y para siempre, o bien hay que renunciar a la idea de que la analiticidad se halle cerrada bajo la conjunción o bien uno debe digerir la molesta consecuencia de que una oración analítica tenga consecuencias empíricas comprobables (y de que pueda, por tanto, resultar que una oración *analítica* sea *empíricamente falsa*).

A propósito, no es ningún accidente que las oraciones que Kant habría clasificado como sintéticas *a priori* fueran consideradas analí-

<sup>26</sup> *The Theory of Relativity and A Priori Knowledge*, California, 1965.

<sup>27</sup> En la contribución de Reichenbach a la recopilación de P. Schilpp, *Albert Einstein Philosopher Scientist* (Nueva York, 1951), da en esencia el mismo argumento, aunque de una forma ligeramente distinta.

ticas por estos empiristas tardíos; al hinchar la noción de analiticidad, les guiaba precisamente el propósito de disolver el problema kantiano, identificando *aprioridad* y analiticidad, procediendo entonces a hacer lo mismo con la analiticidad y la verdad por convención. (También ha criticado Quine este último paso de una forma devastadora, pero nos alejaría de nuestro tema discutirlo.)

Otros filósofos han intentado responder a Quine distinguiendo entre *oraciones* y *enunciados*: están conformes con que todas las *oraciones* sean susceptibles de revisión, pero [afirman] que algunos *enunciados* no lo son. Revisar una oración no es cambiar de idea acerca del enunciado primeramente expresado por la oración en caso de que ésta (es decir, el objeto sintáctico y su significado) no signifique lo mismo antes y después de la revisión; es decir, cuando la revisión plantea un caso de cambio de significado, no de cambio de teoría. Pero (1) esto se convierte de golpe en la propuesta de explicar la analiticidad a partir de la sinonimia; y (2) si hay algo en lo que Quine ha hecho una contribución decisiva a la filosofía, es en haberse dado cuenta de que cambio de significado y cambio de teoría no pueden disociarse tajantemente. No coincidimos con Quine en que el cambio de significado no pueda definirse en absoluto, pero no se sigue de esto que la dicotomía «cambio de significado o cambio de teoría» sea defendible. Descubrir que vivimos en un mundo no-euclídeo *podría* alterar el significado de «línea recta» (esto sucedería en el caso, de algún modo improbable, de que algo como el postulado de las paralelas fuese parte del estereotipo de «línea recta»); pero no se trataría de un *simple* cambio de significado. En particular, no sería un cambio de *extensión*: por ello, no sería correcto decir que el postulado de las paralelas era «verdadero en el primero de los sentidos de esas palabras». Del hecho de que renunciar a una oración *O* lleve consigo un cambio de su significado no se sigue que *O* sea *verdadera*. Los significados pueden no ajustarse al mundo; y el cambio de significado puede venir impuesto por hallazgos empíricos.

Aunque no estemos tratando de analizar, en este trabajo, noción alguna de analiticidad, estamos tratando de hacer eso con una noción que podría parecer estrechamente ligada a ella: la de significado. Así, cabría pensar que los argumentos de Quine irían también en contra de nuestro intento. Comprobemos eso.

Desde nuestra perspectiva, hay un sentido perfectamente bueno en que tener la piel rayada es parte del significado de «tigre». Pero no se sigue de eso que, para nosotros, «los tigres tienen la piel rayada» sea analítico. Si se produjese una mutación, todos los tigres podrían ser albinos. La comunicación presupone que yo tenga un estereotipo de los tigres que incluya el tener la piel rayada, que usted tenga un estereotipo de los tigres que incluya el tener la piel rayada, que yo sepa que su estereotipo de usted contiene eso, que usted sepa que yo

... (à la Grice por siempre jamás). Pero no presupone que ningún estereotipo en particular sea *correcto*, o que la mayoría de nuestros estereotipos sigan por siempre siendo correctos. No se juzga que la obligatoriedad lingüística sea índice de la irrevisabilidad, y ni siquiera de la verdad; por ello podemos afirmar que «los tigres tienen la piel rayada» es parte del significado de «tigre», sin quedar atrapados por lo tanto en los problemas de la analiticidad.

Así pues, los argumentos de Quine contra la identificación de analiticidad y centralidad no son argumentos opuestos al pensamiento de que el que un rasgo «sea parte del significado» de *X* es lo mismo que el que esté obligatoriamente incluido en el estereotipo de *X*. ¿Qué decir del argumento quineano del «sonido»?

Naturalmente, las pruebas que se aduzcan acerca de lo que la gente *dice*, incluyendo las observaciones metalingüísticas explícitas, es importante tanto en «semántica» como en sintaxis. Así, si un hablante señala una *almeja* y pregunta «¿es eso un tigre?», probablemente la gente se reirá carcajadas. (Cuando dejen de reírse) puede que digan «no sabe el significado de “tigre”» o «no sabe qué son los tigres». Al lingüista pueden ayudarle comentarios así. Pero no estamos *definiendo* el estereotipo a partir de esos comentarios. Decir que ser «como-un-gato-grande» es parte del significado de «tigre» no es tan solo decir que predicar «tigre» de algo que no sea como-un-gato-grande (y tampoco un tigre) provocaría ciertos *sonidos*. Es decir que los hablantes adquieren la información de que «los tigres son (estereotípicamente) como-gatos-grandes», ya que adquieren la palabra «tigre» y se sienten obligados a garantizar que aquellos a quienes enseñen el uso de la palabra hagan lo mismo. La información relativa a las capacidades mínimas requeridas para entrar en la comunidad lingüística es información significativa; no existe aquí ninguna circularidad como la criticada por Quine.

#### TRADUCCIÓN RADICAL

En cualquier caso, lo que nuestra teoría no hace por sí sola es resolver el problema de la traducción radical de Quine (es decir, el de la traducción de un lenguaje/cultura ajenos del todo [a los nuestros]). No podemos traducir al español nuestro hipotético Cheroqui correccionando estereotipos, justamente porque el descubrir el estereotipo de, digamos, *wa'arabi* supone la traducción previa de oraciones del Cheroqui. Por otro lado, la restricción de que —en lo que concierne a los estereotipos— cada palabra del Cheroqui debería valer lo que su imagen en español bajo la función de traducción (o que hubiera de valer aproximadamente lo que su imagen, pues en muchos casos no es posible obtener una equivalencia exacta) *limita* severamente la función de traducción. Una vez que hayamos tenido éxito al



traducir el vocabulario básico del Cheroqui, podremos comenzar a disponer de estereotipos, y éstos servirán tanto para reducir el número de futuras traducciones como para comprobar la corrección interna de aquella parte de la función de traducción ya construida.

Incluso cuando podamos especificar los estereotipos (con respecto, digamos, a una traducción provisional del «vocabulario básico»), éstos no bastan, en general, para determinar una única traducción. Así, las palabras alemanas *Ulme* y *Buche* tienen el mismo estereotipo que olmo; pero *Ulme* significa «olmo», mientras que *Buche* significa «haya». En el caso del alemán, el hecho de que *Ulme* y «olmo» suenen parecidas podría indicar la traducción correcta (aunque esto diste de ser infalible —en general, no se puede decir que las palabras que suenan parecidas sean sinónimas—), pero en el caso del griego no tenemos ninguna pista acerca de cuál de las dos palabras ὄξυα y πτελέα significa *olmo* y cuál *haya*; tendríamos que encontrar un griego que pudiera distinguir los *olmos* de las *hayas* (las cosas que son *oxya* de las que son *ptelea*). Lo que esto ilustra es que puede que no sean las disposiciones de los hablantes *típicos* a asentir o a disentir lo que trate de descubrir el lingüista; a causa de la división del trabajo lingüístico, es necesario frecuentemente que el lingüista juzgue quiénes son los expertos en *oxya*, quienes en *wa'arabi*, en *gavagai* o en lo que sea, antes de que pueda barruntar la extensión socialmente determinada. Entonces esta extensión socialmente determinada y el estereotipo del hablante *típico*, por inexperto que sea, servirán *ambas cosas* de elementos constrictores de la función de traducción. El descubrimiento de que el estereotipo de *oxya* es muy distinto del de «olmo» descalificaría la traducción de *oxya* por «olmo» en todos los contextos, salvo en los más extensionales; pero el hallazgo de que la extensión de *oxya* no sea, ni siquiera aproximadamente, la clase de los olmos anularía la traducción en todos los contextos.

Se apreciará que hemos ampliado ya la totalidad de hechos que valen como prueba a favor de una función de traducción más allá de la ascética base que Quine permite en *Palabra y objeto*. Por ejemplo, el hecho de que los hablantes digan tal-y-tal cosa cuando el «cómplice» del lingüista señale la palabra *oxya* y pregunte «¿qué significa esto?», «¿qué es esto?» u otra cosa, es algo que Quine no admite (como cosa que el lingüista pueda «saber»), dado que este género de «conocimiento» presupone haber traducido ya la pregunta «¿qué significa esto?». Sin embargo, si Quine desea asumir que no puede conjeturar *de ningún modo* qué palabras del otro lenguaje significan *asentimiento* y *disentimiento*, no parece en absoluto nada razonable suponer que de alguna manera no se pueda hacer saber a un hablante nativo que uno no entiende una palabra. No es necesario que se descubra una expresión de la otra lengua que signifique literalmente «¿qué significa esta palabra?» (en tanto que cosa distinta de «no en-

tiendo esta palabra», «esta palabra no me es familiar» o «esta palabra me desconcierta», etc.). Quizás diciendo tan solo la palabra *oxya*, o lo que fuere, con tono de perplejidad, bastaría. ¿Por qué el *desconcierto* habría de serle al lingüista menos accesible que el *asentimiento*?

También sacamos partido del hecho de que segmentar [el discurso] en *palabras* ha resultado ser lingüísticamente universal (e incluso existen procedimientos de segmentación en palabras y morfemas que son independientes del significado). Obviamente, no hay ninguna razón especial para que accedamos a que el lingüista profiera oraciones completas y a que busque el asentimiento o el disentimiento y que nos neguemos a que profiera palabras y morfemas en un tono de desconcierto.

Repito: no estoy enunciando la idea de que incrementar la base evidencial de este modo resuelva el problema de la traducción radical. Lo que sí hago es añadir ulteriores restricciones a la clase de los candidatos admisibles a ser considerados una correcta traducción. Creo que incrementar la clase de las restricciones puede llegar a proporcionar una única traducción, o una que tenga tanto de única como seamos capaces de conseguir en la práctica. Pero, en mi opinión, habrá que usar restricciones que vayan más allá de la teoría lingüística en sentido estricto: habrá que poner restricciones a los géneros de creencias (y de conexiones entre creencias, así como entre éstas y el mundo) que razonablemente podamos imputar a la gente. Dejaremos para otro trabajo la discusión de estos temas.

#### UNA CRÍTICA DE LA TEORÍA SEMÁNTICA DE DAVIDSON

En una serie de publicaciones, Donald Davidson ha formulado la interesante sugerencia de que una teoría semántica de una lengua natural podría tomar como modelo lo que los lógicos matemáticos denominan una *definición de la verdad* para un lenguaje formalizado. Puestos a un lado los detalles técnicos, esta sugerencia viene a decir que uno podría tener un conjunto de reglas que especificasen (1) para cada palabra, bajo qué condición es esa palabra verdadera de algo (cuando se trate de palabras a propósito de las cuales tenga sentido el concepto de extensión; a las demás se les daría el tratamiento reservado a las sincategoremáticas); (2) para las oraciones de longitud mayor que la de la palabra, se da una regla que especifique las condiciones bajo las cuales la oración es verdadera en función del modo en que esa oración se construye a partir de otras oraciones más cortas (contando las palabras como si fuesen oraciones de una sola palabra; por ejemplo, «nieve» como «eso es nieve»). La elección de oraciones de una sola palabra como punto de partida es interpreta-

ción mía de lo que Davidson pretende; en cualquier caso, él nos dice que comencemos por una lista *finita* de oraciones *cortas*, cuyas condiciones veritativas tienen que darse *directamente*. La razón de ser de (2) no es la de que deba haber una regla por oración que no haya sido analizada en el apartado (1), puesto que eso supondría un número infinito de reglas, sino la de que tendría que haber una regla para cada *tipo* de oración. Por ejemplo, en un lenguaje formalizado una de las reglas de la especie (2) podría ser: si  $O$  es  $(O_1 \wedge O_2)$ , para oraciones  $O_1$  y  $O_2$ , entonces  $O$  es verdadera si, y sólo si,  $O_1$  y  $O_2$  son ambas verdaderas.

Se apreciará que, en el ejemplo de poner, la condición veritativa especificada para oraciones del tipo  $(O_1 \wedge O_2)$  lleva a cabo la tarea de dar el significado de « $\wedge$ ». De modo más preciso: da el significado de la estructura  $(— \wedge —)$ . En este sentido una definición de la verdad puede ser una teoría del significado. La idea de Davidson es que la *totalidad* de la teoría del significado de una lengua natural puede especificarse de este modo.

No hay ninguna duda de que reglas del tipo ilustrado puedan proporcionar el significado de algunas palabras y estructuras. La cuestión es: ¿qué razón hay para pensar que tan solo de este modo pueda darse el significado de la mayoría de las palabras?

La dificultad obvia es la siguiente: respecto de muchas palabras uno puede dar una definición de la verdad que sea extensionalmente correcta y que no sea, en ningún sentido, una teoría del significado de la palabra. Consideremos, por ejemplo, «*Agua* es verdadera de  $x$  si, y sólo si,  $x$  es  $H_2O$ ». Esta es una definición de la verdad extensionalmente correcta de la palabra «agua» (hablando con propiedad, no se trata de una definición de la verdad, sino de una definición de «verdad de» —es decir, *satisfacción-en-el-sentido-de-la-definición-de* Tarski, pero no nos ocuparemos aquí de estas sutilezas—). Lo es, al menos, si ignoramos el problema de que al agua con impurezas también se le llama «agua», etc. Ahora bien, supongamos que muchos hablantes no *sepan* que el agua es  $H_2O$ . Entonces, esta fórmula no nos dice nada sobre el *significado* de «agua». Podría tener interés para un químico, pero no vale como teoría del significado del término «agua». O bien: vale como teoría de la *extensión* del término «agua», aunque Davidson esté prometiéndonos más que eso.

Davidson es bien consciente de esta dificultad. Su respuesta (en conversación, desde luego) es que necesitamos desarrollar una teoría de la *traducción*. Al igual que para Quine, éste es el problema real. Por referencia a una teoría tal (por referencia a algo que reconocemos no tener), la teoría [de Davidson] viene a decir esto: deseamos un sistema de definiciones de verdad que sea simultáneamente un sistema de traducciones (o de traducciones aproximadas, si es que no pueden lograrse traducciones perfectas). Si tuviésemos una teoría

que especificara qué ha de ser una buena traducción, podríamos excluir la anterior definición de la verdad de «agua» por poco interesante, ya que *x es H<sub>2</sub>O* no es una traducción aceptable y ni siquiera una buena traducción de *x es agua* (en una comunidad pre-científica), incluso aunque suceda que es verdad que  $H_2O = \text{agua}$ .

Esto se aproxima peligrosamente a la afirmación de que una teoría del significado es una definición de la verdad más una teoría del significado. (Si tuviésemos jamón y huevos, tendríamos jamón y huevos —si tuviésemos jamón y tuviésemos huevos—.) Pero esta historia adolece, como veremos, de algo peor que el quedarse en una promesa.

Una segunda afirmación de Davidson es la de que la teoría de la traducción que todavía no tenemos es necesariamente una teoría cuyas unidades básicas son *oraciones* y no *palabras*, puesto que el fundamento de nuestras *pruebas* en lingüística consiste necesariamente en asentimientos y disentimientos de oraciones. Podemos apañárnoslas con las palabras, defiende Davidson, tratándolas como si fuesen oraciones («agua» como «eso es agua»).

¿Cómo le va, según la perspectiva aquí expuesta, a este ambicioso proyecto de construir una teoría del significado bajo la forma de una definición de la verdad limitada por una teoría de la traducción basada en «los únicos elementos de juicio que tenemos»: las disposiciones de las hablantes a usar oraciones?

Nuestra respuesta es que, en principio, la teoría no puede tener éxito. En casos especiales, como el de la palabra «y» en su sentido veritativo-funcional, una definición de la verdad (propriadamente hablando, una cláusula dentro de lo que los lógicos llaman una «definición de la verdad» —la suma total de todas las cláusulas es la definición inductiva de «verdad» para el lenguaje particular—) puede dar el significado de la palabra o estructura porque el estereotipo asociado a la primera (si es que quiere uno hablar de estereotipo en el caso de una palabra como «y») es tan fuerte que realmente constituye una condición necesaria y suficiente. Si todas las palabras fueran como «y» y como «soltero», el programa podría tener éxito. Y Davidson ciertamente hizo una contribución importante al señalar que la lingüística tiene que tratar de condiciones veritativas especificadas inductivamente. Pero para la gran mayoría de las palabras, las exigencias de una teoría de la verdad y las de una teoría del significado son mutuamente incompatibles, al menos en el caso del-español-al-español. Pero el caso del-español-al-español —en el cual pretendemos proporcionar una teoría efectiva del significado de las palabras españolas que esté ella misma formulada en español— es seguramente el caso básico.

El problema es que, en general, las únicas expresiones que son co-extensivas de *X* y que tienen el mismo estereotipo que *X* son ex-

presiones que contienen a la propia *X*. Si excluimos definiciones de la verdad (hablando con propiedad, cláusulas, aunque siga hablando de «definiciones de la verdad» por razones de simplicidad) como

«*X es agua*» es verdadera si, y sólo si, *X es agua*,

dado que nada dicen del significado de la palabra «agua», y excluimos definiciones de la verdad como

«*X es agua*» es verdadera si, y sólo si, *X es H<sub>2</sub>O*,

dado que lo que dicen no vale como descripción del significado de la palabra «agua», no nos quedará nada más.

El problema es que queremos que

*W es verdadera de x si, y sólo si, ———*

reúna las condiciones de que (1) la cláusula sea extensionalmente correcta (donde ——— se toma como una condición que contenga «*x*», como «*x es H<sub>2</sub>O*»); (2) que ——— sea una traducción de *W* — en nuestra teoría, esto significaría que el estereotipo asociado a *W* sería aproximadamente el mismo que el asociado a ———); (3) que ——— no contenga *W* ni variantes sintácticas de *W*. Si *W* fuese, por ejemplo, la palabra «olmo», no hay modo alguno de que las tres condiciones se cumplan simultáneamente. Cualquier oración de la forma mencionada que no contenga «olmo» y que sea extensionalmente correcta contendrá un ——— que como traducción de «olmo» será absolutamente terrible.

Incluso en el caso de que el lenguaje contenga dos sinónimos exactos, la situación no será mucho mejor. Así,

«*Brezo*» es verdadera de *x si, y sólo si, x es una aulaga*

es verdadera, al igual que

«*Aulaga*» es verdadera de *x si, y sólo si, x es un brezo*.

¿Es esto una teoría del significado de «brezo» y de «aulaga»?

Apercibámonos de que la condición (3) es precisamente lo que los lógicos *no* imponen a sus definiciones de la verdad.

«*La nieve es blanca*» es verdadera si, y sólo si, *la nieve es blanca*.

es el paradigma de definición de la verdad en sentido lógico. Pero los

lógicos tratan de dar la extensión de «verdadera» respecto de un lenguaje en particular, no el significado de «la nieve es blanca». Tarski habría ido lo suficientemente lejos como para sostener que estaba dando el *significado* (y no sólo la extensión) de «verdadero»; pero nunca habría afirmado que estaba diciendo *nada* sobre el significado de «la nieve es blanca».

Puede ser que lo que realmente piense Davidson es que la teoría del significado, en un sentido serio del término, sea imposible, y que todo lo que cabe construir es funciones de traducción. Si así fuera, Davidson podría muy bien pensar que la única «teoría del significado» posible para el español es la que diga «“olmo” es verdadera de  $x$  si, y sólo si,  $x$  es un olmo», «“agua” es verdadera de  $x$  si, y sólo si,  $x$  es agua», etc. y que sólo en raras ocasiones diga cosas como « $O_1 \wedge O_2$  es verdadera si, y sólo si, tanto  $O_1$  como  $O_2$  son verdaderas». Pero si la «teoría» de Davidson no es sino el escepticismo de Quine bajo el disfraz de una contribución positiva al estudio del significado, es una píldora amarga de tragar.

La pretensión de que los únicos elementos de juicio de que dispone el lingüista son las disposiciones de los hablantes para con oraciones en bloques es, todavía más, vacía en una interpretación y plenamente falsa en aquella en que no resulta vacía. Si las disposiciones a decir ciertas cosas *cuando se les pregunta* [a los hablantes] *acerca de palabras, morfemas o estructuras sintácticas* se incluyen en la noción de disposición a usar oraciones, no parece excluir nada en absoluto que nos limitemos [a considerar] disposiciones a usar oraciones. Según la interpretación no-vacía, lo que Davidson estaría diciendo es que el lingüista no puede tener acceso a datos como el de qué digan los informadores (incluyendo al propio lingüista) cuando se les pregunta el significado de una palabra, de un morfema o de una estructura sintáctica. No se ha dado nunca ninguna razón por la cual el lingüista no pueda tener acceso a tales datos, y es obvio que los lingüistas actuales ponen un notorio énfasis en el testimonio dado por los informadores de tales materias, en el caso de una lengua foránea, y sobre sus propias intuiciones, cuando se estudia su propia lengua. En particular, cuando tratamos de traducir toda una oración, no hay ninguna razón por la que no podamos guiarnos de nuestro conocimiento de las propiedades sintácticas y semánticas de los constituyentes de esa oración, incluida su estructura profunda. Como hemos visto, hay procedimientos para obtener información acerca de los constituyentes individuales. Es digno de tenerse en cuenta que el procedimiento que Quine y Davidson consideran que es el único *posible* —ir de las oraciones como un todo a sus palabras individuales— es el procedimiento *opuesto* al que ha servido de base a los éxitos logrados en el estudio del lenguaje natural.

presiones que contienen a la propia *X*. Si excluimos definiciones de la verdad (hablando con propiedad, cláusulas, aunque siga hablando de «definiciones de la verdad» por razones de simplicidad) como

*«X es agua» es verdadera si, y sólo si, X es agua,*

dado que nada dicen del significado de la palabra «agua», y excluimos definiciones de la verdad como

*«X es agua» es verdadera si, y sólo si, X es H<sub>2</sub>O,*

dado que lo que dicen no vale como descripción del *significado* de la palabra «agua», no nos quedará nada más.

El problema es que queremos que

*W es verdadera de x si, y sólo si, ———*

reúna las condiciones de que (1) la cláusula sea extensionalmente correcta (donde ——— se toma como una condición que contenga «*x*», como «*x* es H<sub>2</sub>O»); (2) que ——— sea una *traducción* de *W* — en nuestra teoría, esto significaría que el estereotipo asociado a *W* sería aproximadamente el mismo que el asociado a ———); (3) que ——— no contenga *W* ni variantes sintácticas de *W*. Si *W* fuese, por ejemplo, la palabra «olmo», no hay modo alguno de que las tres condiciones se cumplan simultáneamente. Cualquier oración de la forma mencionada que no contenga «olmo» y que sea extensionalmente correcta contendrá un ——— que como *traducción* de «olmo» será absolutamente terrible.

Incluso en el caso de que el lenguaje contenga dos sinónimos exactos, la situación no será mucho mejor. Así,

*«Brezos» es verdadera de x si, y sólo si, x es una aulaga*

es verdadera, al igual que

*«Aulaga» es verdadera de x si, y sólo si, x es un brezo.*

¿Es esto una teoría del *significado* de «brezo» y de «aulaga»?

Apercibámonos de que la condición (3) es precisamente lo que los lógicos *no* imponen a sus definiciones de la verdad.

*«La nieve es blanca» es verdadera si, y sólo si, la nieve es blanca.*

es el paradigma de definición de la verdad en sentido lógico. Pero los

lingüistas tratan de dar la extensión de «verdadera» respecto de un lenguaje en particular, no el significado de «la nieve es blanca». Tarski habría ido lo suficientemente lejos como para sostener que estaba dando el *significado* (y no sólo la extensión) de «verdadero»; pero nunca habría afirmado que estaba diciendo *nada* sobre el significado de «la nieve es blanca».

Puede ser que lo que realmente piense Davidson es que la teoría del significado, en un sentido serio del término, sea imposible, y que todo lo que cabe construir es funciones de traducción. Si así fuera, Davidson podría muy bien pensar que la única «teoría del significado» posible para el español es la que diga «“olmo” es verdadera de  $x$  si, y sólo si,  $x$  es un olmo», «“agua” es verdadera de  $x$  si, y sólo si,  $x$  es agua», etc. y que sólo en raras ocasiones diga cosas como « $O_1 \wedge O_2$  es verdadera si, y sólo si, tanto  $O_1$  como  $O_2$  son verdaderas». Pero si la «teoría» de Davidson no es sino el escepticismo de Quine bajo el disfraz de una contribución positiva al estudio del significado, es una píldora amarga de tragar.

La pretensión de que los únicos elementos de juicio de que dispone el lingüista son las disposiciones de los hablantes para con oraciones en bloques es, todavía más, vacía en una interpretación y plenamente falsa en aquella en que no resulta vacía. Si las disposiciones a decir ciertas cosas *cuando se les pregunta* [a los hablantes] *acerca de palabras, morfemas o estructuras sintácticas* se incluyen en la noción de disposición a usar oraciones, no parece excluir nada en absoluto que nos limitemos [a considerar] disposiciones a usar oraciones. Según la interpretación no-vacía, lo que Davidson estaría diciendo es que el lingüista no puede tener acceso a datos como el de qué digan los informadores (incluyendo al propio lingüista) cuando se les pregunta el significado de una palabra, de un morfema o de una estructura sintáctica. No se ha dado nunca ninguna razón por la cual el lingüista no pueda tener acceso a tales datos, y es obvio que los lingüistas actuales ponen un notorio énfasis en el testimonio dado por los informadores de tales materias, en el caso de una lengua foránea, y sobre sus propias intuiciones, cuando se estudia su propia lengua. En particular, cuando tratamos de traducir toda una oración, no hay ninguna razón por la que no podamos guiarnos de nuestro conocimiento de las propiedades sintácticas y semánticas de los constituyentes de esa oración, incluida su estructura profunda. Como hemos visto, hay procedimientos para obtener información acerca de los constituyentes individuales. Es digno de tenerse en cuenta que el procedimiento que Quine y Davidson consideran que es el único *posible* —ir de las oraciones como un todo a sus palabras individuales— es el procedimiento *opuesto* al que ha servido de base a los éxitos logrados en el estudio del lenguaje natural.



## CRÍTICA DE LA SEMÁNTICA DE CALIFORNIA

Deseo ahora tener en cuenta una aproximación a la teoría semántica de la que fue pionero el último Rudolf Carnap. Puesto que no quiero enredarme en cuestiones textuales, no atribuiré a ningún filósofo en particular la forma concreta [de análisis semántico] que voy a describir, sino que hablaré simplemente de la «semántica de California».

Damos por sabida la noción de *mundo posible*. Sea  $f$  una función definida sobre el «espacio» de todos los mundos posibles cuyo valor  $f(x)$  en cualquier mundo posible  $x$  es siempre un subconjunto del conjunto de las entidades que hay en  $x$ . A  $f$  se le llama entonces *intensión*. Un término  $T$  tiene significado para un hablante  $X$ , si  $X$  asocia  $T$  a una intención  $f_T$ . El término  $T$  es verdadero de una entidad  $e$  en un mundo posible  $x$  si, y sólo si,  $e$  pertenece al conjunto  $f(x)$ . En lugar de emplear el término «asociado a», el propio Carnap tendió a hablar de «captar» («*grasping*») las intensiones; pero, como es claro, lo que pretendía no era sólo que  $X$  «capta» la intención  $f$ , sino que capta *que*  $f$  es la intención *de*  $T$ —es decir, que  $X$  asocia  $f$  a  $T$  de algún modo—.

Obviamente, este cuadro de qué sea entender un término no está de acuerdo con la historia contada en este trabajo. La respuesta de un partidario de la semántica de California sería que ésta es una descripción de un lenguaje *ideal*; que el lenguaje real es *vago*. Con otras palabras, un término  $T$  de una lengua natural no tiene una intención precisa y simple; tiene un conjunto —puede que un conjunto difuso— de intensiones. No obstante, el primer paso en la descripción de una lengua natural es seguramente el de estudiar el caso ideal en que cada término  $T$  tiene exactamente una intención.

(En su libro *Meaning and Necessity*, Carnap emplea una formulación aparentemente distinta: una intención es simplemente una *propiedad*. Una entidad  $e$  pertenece a la extensión de un término  $T$  justo en el caso de que  $e$  tenga una propiedad que resulte ser la intención de  $T$ . La última formulación, en términos de funciones  $f$  como la descrita más arriba, evita que adoptemos como primitiva la noción de *propiedad*).

La primera dificultad contra la que choca esta concepción es el uso de la noción, totalmente inexplicada, de *captar* una intención (o bien, en la reformulación que hacemos de la concepción, de *asociar* una intención a un término). Decir que las intensiones son entidades conjuntistas,  $f$ , proporciona una ejemplificación «concreta» de la noción de intención en el estilo matemático usual (por referencia a las nociones de mundo posible y de conjunto), pero al precio de hacer muy difícil que se vea cómo pueda tener alguien una intención en su mente, o bien qué sea pensar, «captar» o «asociar» una a algo. No

valdrá decir que pensar una intensión es usar una palabra o un sustituto funcional de una palabra (por ejemplo, la contrapartida de una palabra en un «código cerebral», si, como parece probable, el cerebro «computa» en un «código» que tiene analogías con, y posiblemente préstamos de, el lenguaje; o una forma del pensamiento como una representación o un símbolo parecido, allí donde al pensar se empleen cosas así) que *refiera* a la intensión en cuestión, puesto que la *referencia* (es decir, ser la extensión de un término) acaba de definirse a partir de la *intensión*. Aunque la caracterización de qué sea pensar una entidad abstracta del tipo de una función o una propiedad sea ciertamente correcta, en el caso presente es patentemente circular. Pero nunca se ha ofrecido alguna caracterización de esta noción que no sea circular.

Esta dificultad está ligada a una general de la filosofía de la matemática que ha sido señalada por Paul Benacerraf<sup>28</sup>. Benacerraf ha observado que las filosofías de las matemáticas tienden a fracasar por no saber a qué carta quedarse: o explicar qué son los objetos matemáticos y en qué consiste la necesidad de la verdad matemática, dejando sin explicar el hecho de que la gente pueda *aprender* matemáticas, pueda *referirse* a objetos matemáticos, etc., o explicar estos últimos hechos y dejar de hacer lo mismo con las primeras cuestiones. La semántica de California explica qué *son* las intensiones, pero no proporciona ninguna explicación que no sea completamente circular de cómo es que las «captamos», de cómo asociamos con ellas términos, de cómo es que pensamos en ellas, nos *referimos* a ellas, etcétera.

A causa de su Verificacionismo puede que Carnap no haya reparado en esta dificultad. En sus primeros tiempos, concebía entender un término no como la posesión de la *capacidad de verificar* si una entidad cualquiera caía o no en la extensión del término. Recurriendo a las intensiones: «captar» una intensión vendría a ser, por lo tanto, poseer la capacidad de verificar si una entidad *e*, de un mundo posible *x*, pertenece o no a *f(x)*. El Carnap tardío modificó esa idea, reconociendo, como lo dice Quine, que las oraciones se presentan ante el tribunal de la experiencia colectiva y no individualmente. En general, no existe verificación alguna de que un término *T* sea verdadero de una entidad, independientemente del contexto de un conjunto particular de teorías, hipótesis auxiliares, etc. Quizá Carnap habría mantenido que algo parecido a la teoría inicial valía para una clase limitada de términos, los llamados «términos de observación». Nuestro propio punto de vista es que la teoría del significado como verificabilidad es falsa tanto en lo que hace a su idea central como en lo

<sup>28</sup> P. Benacerraf, «Mathematical Thought», *Journal of Philosophy*, vol. LXX (1973), 661-679.

que respecta a los términos de observación, pero no voy a discutir esto aquí. En cualquier caso, si uno *no* es un verificacionista, es difícil ver en la semántica de California teoría de ningún tipo, puesto que la noción de *captar* una intensión queda totalmente sin explicar.

En segundo lugar, si asumiésemos que «captar una intensión» (asociar una intensión a un término *T*) fuese estar en un *estado psicológico* (en sentido propio), entonces la semántica de California quedaría comprometida con los principios (I) y (II) que sometimos a crítica en la primera parte de este trabajo. Debe ser cierto que el estado psicológico del hablante determina la intensión de sus términos, lo cual determinará, a su vez, la extensión de sus términos. Se sigue de aquí que si dos seres humanos estuvieran en el mismo estado psicológico total, habrían de atribuir necesariamente la misma extensión a cada término que empleen. Como vimos, en lo que hace al lenguaje natural, esto probablemente es erróneo. La razón de que sea erróneo es, en parte, que la extensión se determina socialmente, y no mediante la competencia individual. Así pues, la semántica de California está forzada a hacer del lenguaje algo privado —a ignorar la división lingüística del trabajo—. Según esta escuela, la extensión de cada término está totalmente determinada por algo que hay en la cabeza del hablante individual. Una segunda razón por la que esto es un error estriba, como vimos, en que la mayoría de los términos son *rígidos*. Dentro de la semántica de California, a todo término se lo trata como si fuese una *descripción*. El componente *indicador* del significado —el hecho de que nuestros términos refieran a cosas que, en ciertos modos, se parecen a las que designamos *rígidamente*, a estas cosas, a la sustancia llamada «agua» o a la que sea— es ignorado.

Sin embargo, ¿qué decir de la defensa de que la semántica de California no se ocupa del lenguaje natural, sino de una idealización en la cual «se hace caso omiso de la vaguedad», y de que los términos del lenguaje natural pueden considerarse vinculados a un conjunto de intensiones, mejor que a una sola intensión bien definida?

La respuesta es que una palabra *indicadora* no puede representarse como si fuese una familia de palabras no-indicadoras. La palabra «yo», por tomar el caso extremo, es *indicadora*, pero no *vaga*. «Yo» no es sinónimo de ninguna *descripción*; ni de ningún conjunto difuso de descripciones. Análogamente, si estamos en lo cierto, «agua» no es sinónima ni de ninguna descripción ni de ningún conjunto difuso de descripciones (intensiones).

Análogamente, una palabra cuya extensión quede fijada socialmente, y no individualmente, no es lo mismo que una palabra cuya extensión quede fijada individualmente *de un modo vago*. La razón por la que la «aprehensión» individual que yo realizo de un olmo no fija la extensión de «olmo» no es que la palabra sea vaga —si el problema fuese el de la simple vaguedad, el hecho de que mis conceptos

no diferencien los olmos de las hayas implicaría que los olmos son hayas, en mi forma de usar el término, o en cualquier caso implicaría que los olmos son formas-límites de las hayas y que éstas son olmos o formas-límites de los olmos—. La razón es, mejor, que la extensión de «olmo» en mi idiolecto no queda fijada en absoluto por lo que el hablante medio «capte», o no «capte»; la fija la comunidad, incluyendo a los expertos, a través de un complejo proceso cooperativo. Un lenguaje en el que se dé la división del trabajo lingüístico no puede ser remedado satisfactoriamente con un lenguaje que tenga términos vagos y sin división lingüística del trabajo. Cooperación no es vaguedad.

Sin embargo, podría replicarse, ¿no podría uno sustituir nuestro actual lenguaje por uno en el que (1) los términos fuesen reemplazados por términos co-extensivos, pero *no* indicadores (por ejemplo, «agua» por «H<sub>2</sub>O», suponiendo que «H<sub>2</sub>O» no sea indicador), y (2) eliminásemos la división del trabajo lingüístico por el procedimiento de hacer de cada hablante un experto en todo?

Daremos una respuesta negativa a esta pregunta, pero supongámonos, por un momento, que la respuesta es: «sí». ¿Qué significación tendría esto? El lenguaje «ideal» no se parecería a nuestro lenguaje natural en ningún sentido: ni se debería la diferencia a «la vaguedad del lenguaje natural».

De hecho, uno no puede, sin embargo, llevar a cabo tal sustitución, por la muy buena razón de que *todas* las palabras de clase natural y de magnitud física son indicadoras del modo en que lo hemos descrito: «hidrógeno» y, por lo tanto, «H<sub>2</sub>O», lo mismo que «agua». Quizá los términos de «datos sensoriales», si es que los hay (dejando a un lado los términos con los que nos referimos al yo), no sean indicadores; pero «amarillo», en tanto que predicado de cosas, es indicador por la misma razón que «tigre»; incluso aunque algo *parezca* amarillo, puede no *ser* amarillo. Y no supone ninguna ayuda decir que las cosas que parecen amarillas en circunstancias normales (para quienes vean normalmente) son amarillas; aquí, «normal» tiene precisamente el rasgo de la indicabilidad. Simplemente, no existe ninguna razón para creer que, en principio, pueda llevarse a cabo el proyecto de reducir nuestro lenguaje a uno no-indicador.

«En principio», yo lo supongo, podríamos acabar con la división del trabajo lingüístico. Pero si, como sospecho, la división del trabajo lingüístico es un universal lingüístico, ¿qué interés tendría la posible existencia de un lenguaje al que le falta un rasgo constitutivo del lenguaje humano? Un mundo en el que todos fuésemos expertos en todo sería un mundo en el que las leyes sociales diferirían de las actuales hasta un punto difícil de imaginar. ¿Qué *motivo* hay para tomar tal mundo y tal lenguaje como modelos del análisis del lenguaje humano?

Incidentalmente, los filósofos que trabajan en la tradición de la semántica de California han comenzado recientemente a modificar su esquema para superar estos defectos precisamente. Así, se ha sugerido que una intensión podría ser una función cuyos argumentos no sólo sean mundos posibles sino, quizá, un mundo posible, un hablante y un contexto de proferencia no-lingüístico. Esto permitiría representar en el modelo algunas especies de indicabilidad y de división del trabajo lingüístico. Según desarrolla David Lewis estas ideas, «agua», por ejemplo, tendría la misma *intensión* (la misma función) en la Tierra que en la Otra Tierra, aunque una extensión distinta en cada caso. [En efecto, Lewis mantiene el supuesto (I) de la discusión de la primera parte de este escrito y renuncia a (II); nosotros optamos por renunciar a (I) y por conservar (II).] No hay ninguna razón por la que los modelos desarrollados por Carnap y por sus seguidores no puedan resultar valiosos, una vez modificados en esa dirección. Nuestro interés ha estado puesto aquí, no en la utilidad del formalismo matemático, sino en la filosofía del lenguaje que subyace a las primeras versiones del enfoque.

### MARCADORES SEMÁNTICOS

Si el enfoque sugerido aquí es correcto, entonces hay una gran cantidad de trabajo científico por hacer para (1) descubrir qué tipos de elementos pueden tener cabida en los estereotipos; y (2) para desarrollar un sistema conveniente de representaciones de estereotipos; etc. Este trabajo, sin embargo, no puede realizarse por medio de la discusión filosófica. Una idea que, creo, puede tener algún valor es la idea de *marcador semántico*. Esta idea procede de la obra de J. Katz y J. A. Fodor; aquí la modificamos en alguna medida.

Consideremos por un momento el estereotipo de «tigre». Contiene rasgos como el de ser un animal, el de ser-como-un-gato-grande, el de tener la piel a rayas sobre fondo amarillo (¿o rayas amarillas sobre fondo negro?), etc. Ahora bien, hay algo muy especial que tiene que ver con el rasgo *animal*. Si lo pensamos a partir de la noción quineana de *centralidad* o de *irreversibilidad*, es cualitativamente diferente de los otros elementos de la lista. No es imposible imaginar que los tigres puedan no ser animales (que fuesen robots). Pero, desarrollando esta posibilidad, deben haber sido robots desde siempre, no nos atrae la idea de contar un cuento en el que los tigres fuesen *suplantados* por robots, pues entonces los robots no serían tigres. O bien, si no fueron siempre robots, debieron haberse *convertido* en robots, lo cual es todavía más difícil de imaginar. Si los tigres son y fueron siempre robots, estos robots tampoco deben ser inteligentes o, si no, puede que no estemos en el caso de que los tigres no sean anima-

los — podríamos haber descrito, más bien, un caso en el que algunos robots fuesen animales—. Lo mejor es hacerlos robots «dirigidos por otros», es decir, que haya un operador que controle remotamente desde Marte cada movimiento suyo. Repito que cuesta desarrollar estas ideas y que es curioso lo difícil que resulta concebir el caso por el que comenzar, que es por lo que fácilmente comete uno el error de pensar que es «lógicamente imposible» que un tigre *no* sea un animal. Por otro lado, no hay dificultad alguna para imaginar un tigre que no tenga la piel rayada; podría ser albino. Ni es difícil imaginar un tigre que no parezca un gato grande: podría estar horriblemente deformado. Incluso podemos imaginar que la totalidad de los miembros de la especie haya perdido las rayas de su piel o que se haya deformado horriblemente. Pero, ¿que los tigres dejen de ser animales? ¡Una vez más: qué difícil es imaginarlo!

Démonos cuenta de que no estamos cometiendo el error, que Quine con tanta razón criticó, de atribuir a enunciados como «los tigres son animales», «los tigres no podrían dejar de ser animales y convertirse en otra cosa y seguir siendo tigres» una irrevisabilidad absoluta. De hecho, podemos describir casos inverosímiles en los que renunciaríamos a esos enunciados. Pero defendemos que es *cualitativamente* más difícil revisar «todos los tigres son animales» que «todos los tigres tienen la piel a rayas» —en realidad, el segundo enunciado no es ni siquiera verdadero—.

Rasgos como «animal», «ser vivo», «artefacto», «día de la semana» y «período de tiempo», no sólo ocupan un lugar muy central en las palabras «tigre», «almeja», «silla», «viernes» y «hora», respectivamente, sino que también forman parte de un *sistema de clasificación* importante y muy usado. La centralidad garantiza que los elementos clasificados bajo tales encabezamientos virtualmente no labrán de ser *reclasificados* nunca; así, estos encabezamientos son los que resultan natural utilizar, en plurales contextos, como indicadores de categoría. Me parece razonable que, así como en sintaxis usamos marcadores como «nombre», «adjetivo» y, de modo más propio, «nombre concreto» y «verbo cuyo sujeto es un nombre de persona y cuyo objeto es un nombre abstracto», etc., para clasificar palabras, así en semántica habrían de emplearse como marcadores estos indicadores de categoría.

Es interesante constatar que cuando Katz y Fodor introdujeron por primera vez la idea de marcador semántico, no propusieron que el significado —lo que llamamos nosotros el estereotipo— se agotara en una lista de marcadores así. En vez de esto, los marcadores se vieron reducidos tan sólo a los indicadores de categoría con un alto grado de centralidad, que es lo que nosotros proponemos. Al conjunto de los rasgos restantes se les puso el rótulo de «distinguidor». El esquema de Katz y Fodor no se puede comparar fácilmente con el

nuestro, porque ellos pretendían que los marcadores semánticos junto con el distinguidor dieran una condición necesaria y suficiente de pertenencia a la extensión del término de que se tratase. Puesto que se daba por supuesto que la cosa entera —marcadores y distinguidor— representaba lo que el hablante conoce implícitamente, estaban obligados a aceptar que todo hablante conoce de forma tácita una condición necesaria y suficiente de pertenencia a la extensión de «oro», «aluminio», «olmo» —lo cual, como hemos señalado ya, no es el caso—. Con posterioridad, Katz fue más allá y exigió que todos los rasgos constituyeran una condición *analíticamente* necesaria y suficiente de pertenencia a la extensión. Llegado a este punto, Katz dejó a un lado la distinción entre marcadores y distinguidores; si todos los rasgos tienen, por así decirlo, el grado infinito de centralidad, ¿por qué llamar a unos «marcadores» y a otros «distinguidores»? Desde nuestro punto de vista, su distinción original entre «marcadores» y «distinguidor» era correcta —siempre que uno abandone la idea de que el distinguidor proporciona (junto con los marcadores) una condición necesaria y suficiente, y la idea de que algo de esto sea una teoría de la *analiticidad*—. Sugerimos que la idea de marcador semántico es, cuando se la adopta como aquí se recomienda, una contribución importante.

#### EL SIGNIFICADO DE «SIGNIFICADO»

Podemos resumir ahora lo dicho bajo la forma de una propuesta acerca de cómo podría uno reconstruir la noción de «significado». Nuestra propuesta no es la única que cabría hacer considerando todas estas ideas, pero puede servir para encapsular algunos de los puntos más importantes. Además, tengo la impresión de que reivindica mucho tanto de la usanza común del habla cotidiano como de la teoría lingüística, como puede uno preservar convenientemente. Puesto que, según mi idea, algo muy parecido a los supuestos (I) y (II) de la primera parte de este trabajo, se halla profundamente enraizado en nuestra forma común de hablar del significado, y puesto que estos supuestos son inconsistentes, cuando se los toma simultáneamente, ninguna reconstrucción estará libre de consecuencias anti-intuitivas.

Brevemente expuesta, mi propuesta es la de definir «significado», no entresacando un objeto e identificándolo con el significado (aunque, si uno insiste, eso podría hacerse siguiendo el estilo conjuntista usual), sino especificando una forma normal (o mejor, un *tipo* de forma normal) de descripción del significado. Si supiésemos cual pudiera ser «una forma normal de descripción» del significado, entonces sabríamos, por lo que a mí respecta, qué *es* el significado en un sentido científicamente interesante.

Mi propuesta es que la forma normal de descripción del significado de una palabra debería ser una sucesión finita, o un «vector», entre cuyos componentes habrían de estar los siguientes (puede que resultara deseable tener también otros tipos de componentes): (1) los marcadores sintácticos que se aplican a la palabra, como «nombre»; (2) los marcadores semánticos que se aplican a la palabra, como «animal», «período de tiempo»; (3) una descripción de los rasgos adicionales del estereotipo, si los hubiere; (4) una descripción de la extensión.

Parte de esta propuesta consiste en la siguiente convención: los componentes del vector representan todos ellos, excepto la extensión, una hipótesis acerca de la competencia del hablante individual. Así pues, la descripción en forma normal de «agua» podría ser en parte:

MARCADORES SINTÁCTICOS	MARCADORES SEMÁNTICOS	ESTEREOTIPO	EXTENSIÓN
<i>nombre de masa; concreto;</i>	<i>clase natural; líquido</i>	<i>incoloro; transparente; insípido; que calma la sed; etc.</i>	<i>H<sub>2</sub>O (con o sin impurezas);</i>

esto *no* significa que ni al hablante individual ni a la sociedad se le impute el conocimiento del hecho de que el agua sea H<sub>2</sub>O—. Significa (decimos *nosotros*) que la extensión real término «agua» tal como *ellos* (los hablantes en cuestión) la usan es *de hecho* H<sub>2</sub>O. Antes hemos discutido la objeción de «quiénes somos *nosotros* para decir cual es en realidad la extensión de *su* término». Nótese que esta objeción afecta a la noción de *verdad*, y que la extensión es un pariente de la verdad y que hereda los problemas familiares.

Diremos que dos descripciones son *equivalentes* si son la misma excepto en la descripción de la extensión y si sus respectivas descripciones de la extensión son co-extensivas. Entonces, si el conjunto anotado por estas dos diferentes descripciones es, *en realidad*, la extensión de la palabra del caso, y si los otros componentes de la descripción son caracterizaciones correctas de los diversos aspectos de la competencia que ellos representan, *ambas* descripciones se consideran correctas. Las descripciones equivalentes son todas ellas correctas o todas ellas incorrectas. Esta es otra forma de decir que, aunque hayamos de usar una *descripción* de la extensión para *dar* la extensión, *ésta* (el *conjunto*) es nuestra concepción el componente, y no una descripción suya.

En particular, la representación de las palabras «agua» en dialecto terráqueo y «agua» en el dialecto hablado en la Otra Tierra sería la misma, excepto en que en la última columna de la descripción en



forma normal de esta segunda palabra tendríamos XYZ y no H<sub>2</sub>O. Esto significa, en vista de lo que acaba de decirse, que estamos atribuyendo la *misma* competencia lingüística a los hablantes típicos de estos dos planetas, aunque también una extensión diferente.

Esta propuesta significa que conservamos el supuesto (II) de nuestra primera discusión. El significado determina la extensión —digamos que por construcción—. Pero renunciamos a (I): el estado psicológico del hablante individual no determina «lo que él diga» («*what he means*»).

En muchos contextos, creo, esto cuadra con nuestra forma de hablar. Pero hay una paradoja: supongamos que Óscar sea un hablante bilingüe, por conocer tanto el español como el alemán. Según nuestro punto de vista, en su colección total de dialectos las palabras *Buche* y «haya» son *sinónimas exactas*. Las descripciones en forma normal de sus significados serían idénticas. ¡Sin embargo, podría ocurrir perfectamente que él no supiera que son sinónimas! ¡Un hablante puede tener en su vocabulario dos expresiones sinónimas y no saber que sean sinónimas!

Es instructivo darse cuenta de cómo el que, según nuestro análisis (I), sea falso está relacionado con la ceguera para con lo aparentemente obvio —que si  $S_1$  y  $S_2$  son sinónimas y Óscar entiende tanto  $S_1$  como  $S_2$ , entonces Óscar sabe que  $S_1$  y  $S_2$  son sinónimas—. Apercibámonos de que si hubiésemos optado por omitir la extensión como componente del «vector del significado», que es la propuesta de David Lewis, tal y como yo la entiendo, habríamos tenido la paradoja de que «olmo» y «haya» *significarían lo mismo*, a pesar de poseer extensiones diferentes.

Según cualquier teoría materialista, creer una proposición supone probablemente el procesamiento de alguna *representación* de esta proposición, bien sea una oración de algún lenguaje bien un trozo del «código cerebral», una forma de pensamiento o lo que sea. Los materialistas, aunque no sólo ellos, son reacios a pensar que uno pueda creer proposiciones *puras*. Pero incluso los materialistas tienden a pensar que, si uno cree una proposición, es inmaterial (perdón por el juego de palabras) *qué* representación utilice uno. Si  $R_1$  y  $R_2$  son las dos representaciones *disponibles*, entonces si creo la proposición expresada por  $R_1$ , según la representación  $R_1$ , debo creerla también bajo la representación  $R_2$  —debo hacerlo, al menos, si es que se me ha de atribuir algún atisbo de racionalidad—. Pero, como acabamos de ver, esto no es cierto. Óscar puede muy bien creer que *esto* sea una «haya» (dispone al respecto de una señal que diga «haya»), pero ni cree ni dejar de creer que esto sea un «*Buche*». No es únicamente que la creencia sea un proceso que involucre representaciones; él cree la proposición (si es que deseamos hablar de «proposiciones») según una representación y no la cree según otra.

Lo sorprendente de la teoría del significado es cuánto tiempo ha estado esta semántica bajo el peso de ideas filosóficas erróneas y cuán fuertes han sido estas ideas. Filósofo tras filósofo, se ha dicho que el significado era alguna condición necesaria y suficiente. También en la tradición empirista filósofo tras filósofo ha dicho que el significado era el método de verificación. Y tampoco ninguna de estas ideas ha tenido la virtud de hacerse con la exclusiva del caso. No pocos filósofos han afirmado que el significado = el método de verificación = una condición necesaria y suficiente.

Por otro lado, sorprende lo débil de la influencia ejercida por los hechos. Después de todo, lo que hemos señalado en este ensayo es poco más que verdades caseras relativas al modo en que usamos las palabras y a lo mucho (o mejor, a lo poco) que actualmente sabemos cuando las utilizamos. Mis propias reflexiones sobre estos temas comenzaron después de haber publicado yo un trabajo en el que decía, lleno de confianza, que el significado de una palabra era «una batería de reglas semánticas»<sup>29</sup>. Entonces empecé a preguntarme cómo podría explicarse de esta forma el significado de la palabra «oro». Y no es que los filósofos no se hayan detenido nunca en ejemplos como éste: ¡Locke, por citar un caso, usa esta palabra como ejemplo y no le preocupa la idea de que su significado sea una condición necesaria y suficiente!

Si existe una razón por la que tanto el erudito como el lego hayan ido tan lejos a propósito de una temática que tiene que ver, después de todo, con cuestiones que se encuentran en la experiencia de todo el mundo, cuestiones acerca de las cuales poseemos tantos datos que no sabemos qué hacer con ellos, cuestiones de las que poseemos intuiciones muy claras, si dejamos de lado ideas preconcebidas, esta razón tiene que ver con el hecho de que las ideas del lenguaje grotescamente erróneas que son y han sido siempre habituales reflejan dos tendencias filosóficas específicas y muy centrales: la tendencia a tratar la cognición como si fuera una cuestión puramente *individual* y la tendencia a *ignorar* el mundo en la medida en que consta de más que las «observaciones» de los individuos. Ignorar la división del trabajo lingüístico es ignorar la dimensión social de la cognición; ignorar lo que hemos denominado la *indicabilidad* de la mayor parte de las palabras es ignorar la contribución del medio. La filosofía tradicional del lenguaje, como mucha de la filosofía tradicional, deja fuera de consideración a los demás y al mundo; una mejor filosofía y una mejor ciencia del lenguaje deben abarcar ambos.

\* \* \*

<sup>29</sup> En «How not to talk about meaning», en *Philosophical Papers*, vol. 2, 117-131.

## NOTA DEL TRADUCTOR

Unas pocas palabras acerca de dos de las —supongo— muchas decisiones tomadas por mí que pueden ponerse en la picota. Primero, he traducido *stereotype* por «estereotipo». Ni la edición del Diccionario de la Academia que he consultado (la decimonovena) ni otras sabias obras reconocen esta palabra. La he preferido frente a otros candidatos, como «cliché», no sólo porque en el español de estos días se usa «estereotipo» con el significado que le da Putnam sino porque es mucho más flexible para formar palabras derivadas (aunque sean neologismos). Incluso en este caso encuentra uno palabras más familiares que hubieran podido preferirse. «Esquema» (segunda acepción de la Academia) es un ejemplo. Sin embargo, sobre todo esto ha pesado más lo conocido del término en lengua inglesa y lo que ha ayudado a formular y dar a conocer las ideas de Hilary Putnam. La convicción de que valía la pena buscarle a *stereotype* un compañero cuasihomofónico en español, útil para la expresión filosófica, es la última razón de mi preferencia.

En segundo lugar, he traducido *indexicability* como «indicabilidad». «Índice» es todo lo bueno que hace falta como contrapartida de *index* y el derivado surge por sí solo. En el mundo del procesamiento de datos, se habla de ficheros «indexados», pero «indexabilidad» me ha parecido excesivo, si tenemos en cuenta que «indicar» tiene mucho de las resonancias que precisa la teoría de Kripke-Putnam que se expone más arriba.

Finalmente, en algunos casos en los que me ha parecido que la versión dada podría discutirse en sus matices he dado también entre paréntesis el texto original. También he encerrado entre corchetes algunas palabras sin pareja en el texto inglés, aunque precisas al dar una versión no demasiado literal.

### III. SIGNIFICADO Y VERIFICACIÓN



La investigación sobre lo que sea el significado es, como se ha señalado anteriormente, una de las tareas centrales de la filosofía del lenguaje. En este apartado se incluye un grupo de artículos que tienen que ver con la teoría del significado del positivismo lógico, teoría dominante en el panorama filosófico, si bien durante breve tiempo, a partir de 1930.

Desde que en 1922 Moritz Schlick tomó posesión de la cátedra de Filosofía de las Ciencias Inductivas de la Universidad de Viena, un grupo de filósofos y científicos cuyo denominador común era su marcado talante antimetafísico, empezó a agruparse en torno suyo para dar lugar a lo que se denominaría «el Círculo de Viena». Ese talante tenía su fuente en la tesis, apasionadamente defendida por ellos, de que todas las proposiciones inteligibles se basaban en la experiencia. Había, sin embargo, un problema importante: ¿cómo reconciliar esa tesis con el hecho de que las proposiciones de la matemática y de la lógica son «ideales» y, con todo, inteligibles? (Ninguno de ellos, no es necesario recordarlo, se atrevería a abrazar la posición de Mill de que las proposiciones de la lógica y la matemática eran generalizaciones empíricas.) Fue, al parecer, el matemático Hans Hahn el primero en llamar la atención de los miembros del Círculo hacia la afirmación del *Tractatus* de Wittgenstein de que tales proposiciones eran *tautologías*. De este modo la tesis matriz sólo tuvo que sufrir un ligero cambio: toda proposición inteligible se basa en la experiencia *a menos que sea una tautología* que, al no decir nada sobre el mundo talgo que el metafísico no admitiría para sus proposiciones), no posee contenido empírico. Con esto se alcanzaba, como Schlick gustaba decir, un empirismo consistente y una formulación del principio de verificación (que ellos, erróneamente, suponían que se derivaba del *Tractatus*): el significado de una proposición reside en su método de verificación.

Casi simultáneamente a su formulación se produjeron las primeras disputas sobre el significado, el *status* y la plausibilidad del principio. En «Problemas y cambios del criterio empirista del significado», Carl Hempel, aunque receptivo a la concepción general del positivismo lógico, analiza y discute algunos de esos problemas mostrando que dicho principio es demasiado débil (incluye algunas cosas que se desearían excluir) y demasiado fuerte (excluye algunas cosas que se desearían incluir) y señalando que el significado cognitivo de un enunciado no reside sólo en las oraciones observacionales de un lenguaje, sino en las relaciones lógicas que mantiene con todos los demás enunciados de ese lenguaje.

«Dos dogmas del empirismo», quizá el artículo de filosofía más

importante de la segunda mitad de este siglo, pasa por ser la crítica más implacable del positivismo lógico y de su criterio de significatividad asociado: el principio de verificación. Ciertamente, en este artículo se lanza un despiadado ataque contra la distinción entre proposiciones analíticas y proposiciones sintéticas (empíricas) y se llega a la conclusión no tanto —como suele afirmarse— de que es imposible definir las nociones de «analítico» y «sintético» de manera no circular, como de que dichas nociones *no tienen aplicación alguna*: no existe la clase de las proposiciones analíticas como distinta de la de las sintéticas. Pero sería un error concluir a partir de esto que la posición de Quine no es verificacionista. Como Michael Dummett ha señalado («The Significance of Quine's Indeterminacy Thesis», *Synthese*, vol. 27, 1974), el modelo de lenguaje que Quine presenta en «Dos dogmas» es tan verificacionista como el de los empiristas lógicos: «nuestra comprensión del lenguaje reside, según este modelo, en nuestra capacidad de reconocer qué experiencias son recalcitrantes, esto es: qué experiencias nos obligan a hacer alguna revisión en nuestras asignaciones de valores de verdad a las oraciones del lenguaje, y qué revisiones constituyen una respuesta adecuada a una experiencia recalcitrante dada». Lo que Quine critica de los empiristas lógicos es, más bien, que su verificacionismo tomase un sesgo molecular en vez de holista, lo que los llevó a diferencias (erróneamente, según él) entre dos tipos de significados opuestos: el de las proposiciones de la lógica y la matemática y el de las proposiciones empíricas.

«Significado y traducción» —además de presentar un cuadro bastante completo de la famosa tesis de la indeterminación de la traducción— desarrolla y modifica partes importantes de «Dos dogmas». La admisión de oraciones observacionales con significado estimulativo determinado y de oraciones analíticas estimulativas —si bien no distinguibles de las oraciones empíricas generalmente aceptadas como verdaderas— y, por ende, de los conceptos de significado estimulativo y analiticidad estimulativa, obliga a plantearse qué se hizo del radicalismo inicial de «Dos dogmas».

## PROBLEMAS Y CAMBIOS EN EL CRITERIO EMPIRISTA DE SIGNIFICADO \*

CARL G. HEMPEL

### I INTRODUCCIÓN

La idea fundamental del empirismo moderno es la concepción de que todo conocimiento no analítico se basa en la experiencia. Llamemos a esta tesis el principio del empirismo <sup>1</sup>. El empirismo lógico contemporáneo ha agregado <sup>2</sup> la máxima de que una oración constituye una aserción cognoscitivamente significativa, y por lo tanto puede decirse que es verdadera o falsa, sólo si es o bien (1) analítica o contradictoria, o bien (2) capaz, por lo menos en principio, de ser comprobable empíricamente. De acuerdo con este *criterio empirista de significado cognoscitivo*, como se lo llama, muchas de las formulaciones de la metafísica y la gnoseología tradicionales carecen de significado cognoscitivo, por más que algunas de ellas sean ricas en connotaciones no cognoscitivas en virtud de su atractivo emocional o de la inspiración moral que ofrecen. Del mismo modo ciertas doctrinas que en algún momento se formularon dentro de las ciencias empíricas o en sus disciplinas colaterales están construidas de manera tal que resultan incapaces de ser comprobadas por ninguna prueba concebible; en consecuencia, se las califica de pseudohipótesis que no afirman nada y que por lo tanto están totalmente desprovistas de fuerza explicativa o predictiva. Este veredicto se aplica, por ejemplo, a las especulaciones neovitalistas acerca de entelequias o fuerzas vi-

\* Versión castellana de Mario Bunge, Emilio O. Colombo, Estela Arias y Lilia Fornasari.

<sup>1</sup> Este término es empleado por Benjamin (2) en un examen de los fundamentos del empirismo. Si se desea un tratamiento reciente de las ideas básicas del empirismo véase Russell (27), parte VI.

<sup>2</sup> Stace sostiene, en efecto, en su incitante artículo «Positivism» que el criterio de comprobabilidad del significado no se sigue lógicamente del principio del empirismo. (Véase (29), y especialmente la sección 11.) Esto es exacto; según el principio del empirismo una oración expresa conocimientos sólo si es analítica o si es corroborada por la prueba empírica; el criterio de comprobabilidad va más allá e identifica el dominio del discurso cognoscitivamente significativo con el del conocimiento potencial; es decir que otorga contenido cognoscitivo sólo a oraciones para las que es concebible una prueba empírica, salvo que sean analíticas o contradictorias.



tales, y a las «hipótesis telefinalistas» propuestas por Lecomte du Noüy<sup>3</sup>.

Sin embargo, las formulaciones precedentes del principio del empirismo y del criterio empirista del significado no ofrecen sino una caracterización general y muy vaga de un punto de vista básico, y por lo tanto necesitan ser elucidadas y ampliadas. Mientras que en las primeras fases de su desarrollo el empirismo lógico se ocupaba en gran parte de una crítica de las formulaciones filosóficas y científicas mediante aquellos principios fundamentales, en los últimos años ha experimentado una preocupación creciente por las tareas positivas de analizar en detalle la lógica y la metodología de la ciencia empírica y por aclarar y reformular las ideas básicas del empirismo a la luz de los resultados así obtenidos. En este trabajo me propongo discutir algunos de los problemas que han aparecido en dicha investigación y algunos de los resultados que parece haber establecido.

## 2. CAMBIOS EN EL CRITERIO DE COMPROBABILIDAD DEL SIGNIFICADO EMPÍRICO

Como lo muestra nuestra formulación, el criterio empirista del significado establece el requisito de comprobabilidad empírica para las oraciones cognoscitivamente significativas que no son ni analíticas ni contradictorias; denominémoslas oraciones de significado empírico. El concepto de comprobabilidad, que ha de precisar la vaga noción de estar fundado (o mejor de poder fundarse) en la experiencia, ha sufrido diversas modificaciones que revelan un análisis cada vez más fino de la estructura del conocimiento empírico. En esta sección examinaremos las etapas fundamentales de ese desarrollo.

Para facilitar la exposición introduciremos primero tres conceptos auxiliares, a saber, los conceptos de características observables, predicado observacional y oración observacional. Una propiedad o una relación de objetos físicos será una *característica observable* si, en condiciones apropiadas, su presencia o ausencia en una oración dada puede ser establecida por observación directa. Así, «verde», «blando», «líquido», «más largo que», designan características observables, mientras que no lo hace «bivalente», «radiactivo», «mejor conductor eléctrico» e «introvertido». Los términos que designan características observables se denominarán *predicados observacionales*. Finalmente, entenderemos por *oración observacional*, cualquier oración que afirme, correcta o incorrectamente, que uno o más objetos nombrados específicamente poseen o no poseen cierta caracteris-

<sup>3</sup> Cfr. (19), cap. XVI.

ta observable. Por ejemplo, las siguientes oraciones satisfacen esta condición: «La torre de Eiffel es más alta que los edificios vecinos», «La aguja de este instrumento no cubre el punto de la escala marcado con 3», y aun «El dinosaurio más grande que se exhibe en el Museo de Historia Natural de Nueva York tenía una lengua azul», ya que esta última oración asigna a un objeto específico una característica (la de tener una lengua azul) tal que en condiciones apropiadas (v. gr., en el caso de mi perro *chow*) puede establecerse su presencia o su ausencia por observación directa. Nuestro concepto de oración observacional pretende suministrar una interpretación precisa de la vaga idea de una oración que asevera algo que «en principio» puede establecerse por observación directa, aunque puede ocurrir que sea inobservable para mí, quizá también para mis contemporáneos, y aun para cualquier ser humano que jamás haya vivido o viva en el futuro. Cualquier prueba que pueda aducirse en la comprobación de una hipótesis empírica puede considerarse ahora expresada en oraciones observacionales de esta especie <sup>4</sup>.

Volvamos ahora nuestra atención sobre la concepción de la comprobabilidad, y por lo tanto del significado empírico. En los comienzos del Círculo de Viena se decía que una oración posee significado empírico si es capaz, por lo menos en principio, de ser completamente verificada por la prueba observacional; o sea si puede describirse una prueba observacional tal que de obtenerse realmente estableciera de manera concluyente la verdad de la oración <sup>5</sup>. Con la ayuda del

<sup>4</sup> Las oraciones observacionales de esta especie pertenecen a lo que Carnap ha llamado lenguaje-cosa (cfr. (7), pp. 52-3). En particular para el caso de procedimientos de comprobación intersubjetivos empleados tanto en la ciencia como en vastos sectores de investigación empírica al nivel del sentido común, es claro que se adecúan para formular datos que sirvan de base a los tests empíricos. En las discusiones gnoseológicas a menudo se supone que la prueba última de creencias acerca de cuestiones empíricas consiste en percepciones y sensaciones cuya descripción requiere un tipo de lenguaje fenomenista. No podemos discutir aquí los problemas específicos relacionados con la concepción fenomenista; pero debe señalarse que en todo caso todas las consideraciones críticas ofrecidas en este artículo respecto del criterio de comprobabilidad pueden aplicarse, *mutatis mutandis*, también al caso de un fundamento fenomenista.

<sup>5</sup> Originalmente, se consideraba que la prueba admisible se limita a lo que es observable por el interlocutor y sus semejantes durante sus respectivas vidas. Así interpretado, el criterio elimina a título de cognoscitivamente no significativos, a todos los enunciados acerca del pasado remoto y del futuro distante, como lo señalaron, entre otros, Ayer en (1), cap. I; Pap (21), cap. XIII, especialmente pp. 333 y ss.; y Russell (77), pp. 445-447. Sin embargo, esta dificultad desaparece si permitimos que la prueba consista en un conjunto finito de «datos observacionales lógicamente posibles», formulado cada uno de ellos en una oración observacional. Por ejemplo, la oración  $S_1$ : «La lengua del dinosaurio más grande que se exhibe en el Museo de Historia Natural de Nueva York era azul o negra» es completamente verificable en este sentido, puesto

concepto de oración observacional podemos reformular este requerimiento de la manera siguiente: Una oración  $S$  tiene significado empírico si y sólo si es posible indicar un conjunto finito de oraciones observacionales  $O_1, O_2, \dots, O_n$ , tal que si éstas son verdaderas, necesariamente  $S$  es también verdadera. Sin embargo, así formulada esta condición se satisface también si  $S$  es una oración analítica o si las oraciones observacionales dadas son lógicamente incompatibles entre sí. Mediante la siguiente formulación descartamos estos casos y a la vez formulamos más precisamente el criterio requerido:

### 2.1. REQUISITO DE COMPLETA VERIFICABILIDAD EN PRINCIPIO

Una oración tiene significado empírico si y solamente si no es analítica y se sigue lógicamente de alguna clase de oraciones observacionales finita y lógicamente coherente <sup>6</sup>.

que es una consecuencia lógica de la oración  $S_2$ : «La lengua del dinosaurio más grande que se exhibe en el Museo de Historia Natural de Nueva York era azul», y ésta es una oración observacional, como se demostró anteriormente.

Si a los conceptos de *verificabilidad en principio* y al más *general de confirmabilidad en principio* que consideraremos más adelante se los interpreta como referidos a la *prueba lógicamente posible*, tal como es expresada por las oraciones observacionales, entonces del mismo modo se sigue que la clase de los enunciados verificables, o por lo menos confirmables, incluyen en principio aserciones tales como la de que el planeta Neptuno y el continente Antártico existieron antes de ser descubiertos, y que la guerra atómica, si no se la detiene, puede conducir a la extinción de este planeta. Las objeciones que esgrime Russell [cfr. (27), pp. 445 y 447] contra el criterio de verificabilidad haciendo referencia a esos ejemplos, no se aplican, pues, si el criterio se entiende de la manera que hemos sugerido. Incidentalmente, los enunciados del tipo de los que menciona Russell, los cuales no son verificables por ningún ser humano, ya eran reconocidos como cognoscitivamente significativos por Schlick (en (28), parte V), quien sostenía que la imposibilidad de verificarlos es «meramente empírica». La caracterización de la verificabilidad con la ayuda del concepto de oración observacional tal como se ha sugerido aquí podría servir como una enunciación más explícita y rigurosa de esa concepción.

<sup>6</sup> Como se ha señalado a menudo en la literatura empirista, el término «verificabilidad» debe señalar, desde luego, la concebibilidad, o mejor, la posibilidad lógica de pruebas de tipo observacional, que de obtenerse, fuesen concluyentes con respecto a la oración dada; no se refiere a la posibilidad técnica de llevar a cabo los tests necesarios para obtener tales pruebas, y menos aun a la posibilidad de encontrar realmente fenómenos directamente observables que constituyan una prueba concluyente de dicha oración, lo que equivaldría a la existencia misma de tal prueba e implicaría la verdad de la oración. Observaciones análogas se aplican a los términos «refutabilidad» y «confirmabilidad». Esto ha sido omitido en algunas consideraciones críticas recientes del criterio de verificabilidad. Por ejemplo, Russell (cfr. (27), p. 448) interpreta la verificabilidad como la existencia real de un conjunto de hechos que constitu

Con todo, este criterio tiene varios defectos serios. El primero de los que hemos de mencionar ha sido señalado por diversos autores:

a) El requisito de verificabilidad elimina toda oración de forma universal, y por lo tanto todos los enunciados que pretendan expresar leyes generales, ya que éstas no pueden ser verificadas concluyentemente por ningún conjunto finito de datos observacionales. Y puesto que una parte importante de las teorías científicas está constituida por oraciones de este tipo, debe considerarse que el requisito de verificabilidad es excesivamente restrictivo a este respecto. Del mismo modo, el criterio mencionado descalifica todas las oraciones tales como ésta: «Para cualquier sustancia existe algún solvente», que contienen cuantificadores existenciales y universales (esto es, en las que aparecen los términos «todos» y «algún» o sus equivalentes), ya que ninguna de estas oraciones puede deducirse lógicamente de ningún conjunto finito de oraciones observacionales.

Hay otros dos defectos del requisito de verificabilidad que no parecen haber sido señalados generalmente:

b) Supongamos que *S* es una oración que satisface el criterio propuesto, mientras que *N* es una oración tal como «El absoluto es perfecto», a la cual el criterio no le atribuye significado empírico. Entonces la disyunción «*S* o *N*» (o sea, la expresión que se obtiene conectando ambas oraciones por medio de la palabra «o») satisface el criterio, porque si *S* es una consecuencia de alguna clase finita de ora-

---

una verificación concluyente. Esta concepción, que jamás ha sido sostenida por ningún empirista lógico, naturalmente debe resultar inadecuada, ya que de acuerdo con ella no puede establecerse la significación empírica de una oración sin obtener pruebas empíricas, y además en cantidad suficiente para constituir una prueba concluyente de la oración en cuestión (*sic.*). No es sorprendente, pues, que su extraordinaria interpretación de la verificabilidad lo lleve a la siguiente conclusión: «En realidad, no es verificable que una proposición sea verificable» (*loc. cit.*). Según la interpretación empirista de la completa verificabilidad, cualquier enunciado que asevere la verificabilidad de una oración *S* cuyo texto se cita es analítico o contradictorio; puesto que la cuestión acerca de si existe una clase de oraciones observacionales, sean éstas verdaderas o falsas, es asunto de lógica pura y no requiere información fáctica alguna.

Un malentendido similar se muestra en el siguiente pasaje en el cual W. H. Werkmeister pretende caracterizar una concepción sostenida por los positivistas lógicos. «Se dice que una proposición es "verdadera" cuando es "verificable en principio", es decir, cuando conocemos las condiciones que de darse haría posible la "verificación" (Ayer).» (Cfr. (31), p. 145.) La tesis citada, que tampoco ha sido sostenida por ningún positivista lógico, incluyendo el propio Ayer, es lógicamente absurda. Fácilmente podemos describir condiciones que de darse verificarían la oración «El edificio Chrysler está pintado por fuera de un amarillo brillante», pero del mismo modo podríamos describir condiciones para su negación; de aquí que, según el principio citado, deberían considerarse verdaderas tanto la primera oración como su negación. En realidad, el pasaje en discusión no concuerda con la correcta observación de Werkmeister (*loc. cit.*, p. 40) de que la verificabilidad caracteriza el significado de una oración, lo que muestra que es un criterio de significado cognoscitivo y no un criterio de verdad.

ciones observacionales es trivial que « $S$  y  $N$ » sea una consecuencia de esa clase. Pero es claro que el criterio empirista de significado no pretende justificar oraciones de este tipo. En consecuencia, en este respecto el requisito de verificabilidad es demasiado amplio.

c) Sea « $P$ » un predicado observacional. Entonces la oración puramente existencial « $(Ex) Px$ » («Existe por lo menos una cosa que tiene la propiedad  $P$ ») es completamente verificable, ya que se sigue de cualquier oración observacional que afirme que un determinado objeto tiene la propiedad  $P$ . Pero su negación, al ser equivalente a la oración universal « $(x)-Px$ » («Nada tiene la propiedad  $P$ ») evidentemente no es completamente verificable, como se desprende de lo dicho en el punto (a). De aquí que, según el criterio (2.1) las negaciones de algunas oraciones empíricamente (y por lo tanto cognoscitivamente) significativas son empíricamente no significativas, y puesto que no son ni analíticas ni contradictorias, carecen de significado cognoscitivo. Pero cualquiera que sea la manera en que delimitemos el dominio del discurso significativo, tendremos que insistir en que si una oración cae dentro de ese dominio, lo mismo debe ocurrir con su negación. Para decirlo más explícitamente: las oraciones que han de calificarse como cognoscitivamente significativas son precisamente aquellas de las cuales puede decirse significativamente que son o verdaderas o falsas. Pero en ese caso, la adhesión a (2.1) engendrará un serio dilema, como lo muestra la consecuencia que acaba de mencionarse: o bien tendríamos que desechar el principio lógico fundamental de que una oración es o verdadera o falsa, y por lo tanto su negación es falsa o verdadera respectivamente (siendo entonces cognoscitivamente significativa); o si no debemos negar, de una manera similar a la concepción intuicionista de la lógica y la matemática, que « $(x)-Px$ » sea lógicamente equivalente a la negación de « $(Ex) Px$ ». Evidentemente, el criterio expuesto en (2.1), que se ha mostrado inadecuado en varios otros aspectos, no justifica tan drásticas medidas para seguir sosteniéndolo; por lo tanto debe abandonarse <sup>7</sup>.

<sup>7</sup> Los argumentos que se sostienen aquí contra el criterio de verificabilidad también prueban la inadecuación de una concepción estrechamente relacionada con él, a saber, que dos oraciones tienen el mismo significado cognoscitivo si cualquier conjunto de oraciones observacionales que verificase a una de ellas también verificaría a la otra, y viceversa. Así, de acuerdo con este criterio, habría que asignar el mismo significado cognoscitivo a dos leyes generales cualesquiera, puesto que ninguna ley general es verificada por ningún conjunto de oraciones observacionales. Esta concepción debe distinguirse claramente de una posición que Russell examina en su exposición crítica del criterio positivista del significado. Se trata de «la teoría de que dos proposiciones cuyas consecuencias verificadas son idénticas, tienen el mismo significado» [(27), p. 448]. En verdad, esta concepción es insostenible, puesto que la cuestión acerca de cuáles son las consecuencias de un enunciado que han sido verificadas en un momento dado, constituye un accidente histórico que no puede servir para establecer la identi-

Exactamente las mismas consideraciones se aplican a un criterio opuesto, que hace de la completa refutabilidad en principio la característica definitoria del significado empírico. Formulémoslo así: Una oración tiene significado empírico si y sólo si en principio es capaz de ser completamente refutada por un número finito de datos observacionales, o más precisamente:

### 2.2. REQUISITO DE COMPLETA REFUTABILIDAD EN PRINCIPIO

Una oración tiene «significado empírico» si y sólo si su negación no es analítica y se sigue lógicamente de una clase finita y lógicamente coherente de oraciones observacionales<sup>8</sup>.

Este criterio califica a una oración como empíricamente significativa si su negación satisface el requisito de completa verificabilidad; como era de esperar, resulta inadecuado por los mismos motivos que el anterior. Éstos son:

a) Descarta hipótesis puramente existenciales, tales como "Existe por lo menos un unicornio", y todas aquellas oraciones cuya formulación exige una cuantificación mixta (esto es, universal y existencial), puesto que es imposible refutarlas concluyentemente por un número finito de oraciones observacionales.

b) Si una oración *S* es completamente refutable, mientras que la oración *N* no lo es, entonces su conjunción *S.N* (o sea la expresión que se obtiene conectando ambas oraciones por medio de la palabra «y») es completamente refutable; porque si la negación de *S* se sigue de alguna clase de oraciones observacionales, entonces *a fortiori* la negación de *S.N* se sigue de la misma clase. En consecuencia, el criterio otorga significado empírico a muchas oraciones que un criterio empirista adecuado tendría que rechazar, como, por ejemplo, «Todos los cisnes son blancos y el absoluto es perfecto».

c) Si «*P*» es un predicado observacional, entonces la afirmación de que todas las cosas tienen la propiedad *P* es calificada como significativa, pero su negación, al ser equivalente a una hipótesis puramente existencial, es descalificada (cfr. a). Por lo tanto el criterio (2.2) origina el mismo dilema que (2.1).

dad de significados cognoscitivos. Pero no creo que ningún positivista lógico se haya adherido a esa «teoría».

<sup>8</sup> La idea de emplear la refutabilidad teórica por pruebas observacionales como «criterio de demarcación», separando la ciencia empírica de la matemática y la lógica por un lado y de la metafísica por el otro, se debe a K. Popper (cfr. (22), sección 1-7 y 19-24; véase también (23), vol. II, pp. 282-5). No sé si Popper aceptaría la reformulación del criterio de refutabilidad propuesta.

En suma, pues, las interpretaciones del criterio de comprobabilidad en términos de completa verificabilidad o completa refutabilidad son inadecuadas por ser excesivamente restrictivas en una dirección y excesivamente amplias en la otra, y porque ambas requieren cambios importantes en los principios fundamentales de la lógica.

Se han llevado a cabo varios intentos de salvar estas dificultades interpretando el criterio de comprobabilidad como un requisito de confirmabilidad parcial y posiblemente indirecta de las hipótesis empíricas por medio de la prueba observacional.

(2.3) Una formulación sugerida por Ayer<sup>9</sup> es característica de estas tentativas para establecer un criterio de confirmabilidad claro y suficientemente amplio. Establece, en efecto, que una oración *S* tiene significado empírico si es posible deducir de *S*, en conjunción con hipótesis auxiliares apropiadas, oraciones observacionales que no son deducibles de las hipótesis auxiliares únicamente.

Esta condición es sugerida luego de una consideración más acabada de la estructura lógica de la comprobación científica, pero así formulada es demasiado liberal. En verdad, el mismo Ayer ha señalado en la segunda edición de su libro *Language, Truth and Logic*<sup>10</sup>, que su criterio otorga significado cognoscitivo a cualquier oración. Así, por ejemplo, si *S* es la oración «El absoluto es perfecto», basta con elegir como hipótesis auxiliar a la oración «Si el absoluto es perfecto, entonces esta manzana es roja» para que sea posible deducir la oración observacional «Esta manzana es roja», la cual claramente no se deduce de la hipótesis auxiliar aislada<sup>11</sup>.

<sup>9</sup> (1), cap. I. También Pap presentó de una manera muy clara la posición contraria a los requisitos de verificabilidad y refutabilidad y a favor de un requisito de confirmabilidad y disconfirmabilidad parciales en (21), cap. XIII.

<sup>10</sup> (1), 2.ª ed., pp. 11-2.

<sup>11</sup> Según Stace (cfr. (29), p. 218), el criterio de comprobabilidad parcial o indirecta, al cual denomina principio positivista, presupone (y por lo tanto implica lógicamente) otro principio, y llama a éste *Principio de especies observables*; reza así: «Para que una oración sea significativa debe aseverar o negar hechos de una especie o clase tal que resulta lógicamente posible observar directamente algunos hechos que sean casos de dicha especie o clase. Y si una oración asevera o niega hechos de una especie o clase tal que resultaría lógicamente imposible observar directamente ningún caso de esta clase o especie, entonces la oración es no significativa.» Creo que el argumento que ofrece Stace para probar que su principio se sigue del requisito de comprobabilidad no es concluyente, fundamentalmente, en virtud de la incorrecta y tácita suposición de que «según la concepción de la deducción como transformación» las premisas de un razonamiento deductivo válido deben ser condiciones necesarias de la conclusión (*loc. cit.*, p. 225). Sin ir más adelante, quiero agregar aquí una observación sobre el propio principio de especies observables. El profesor Stace no dice cómo habremos de determinar cuáles son los «hechos» que asevera o niega una oración dada, ni aun si asevera o niega «hecho» alguno. En consecuencia queda sin aclarar el contenido exacto del principio. Sin embargo, cualquiera que sea el criterio que elijamos respecto de la referen-

(2.4) Para hacer frente a esta objeción Ayer ha propuesto recientemente una versión modificada de su criterio de comprobabilidad. La modificación, en efecto, limita las hipótesis auxiliares mencionadas en (2.3) a oraciones que sean o bien analíticas, o bien que pueda demostrarse independientemente que son comparables en el sentido de la versión modificada del criterio <sup>12</sup>.

Pero en seguida puede demostrarse que este nuevo criterio, tal como el requisito de completa refutabilidad, concede significado empírico a cualquier conjunción  $S.N$  donde  $S$  satisfaga el criterio de Ayer, mientras que  $N$  puede ser una oración tal como «El absoluto es perfecto», la cual debe rechazarse según dicho criterio. En realidad, cualesquiera que sean las consecuencias que se deduzcan de  $S$  con la ayuda de hipótesis auxiliares admisibles, también pueden deducirse de  $S.N$  por medio de las mismas hipótesis subsidiarias, y puesto que el nuevo criterio de Ayer se formula esencialmente en términos de la deducibilidad de cierto tipo de consecuencia de una oración dada, admite a  $S.N$  lo mismo que a  $S$ . El profesor A. Church <sup>13</sup> ha señalado que si se dan tres oraciones observacionales cualesquiera y ninguna de ellas es deducible de las demás, entonces se sigue que cualquier oración  $S$  o su negación tiene significado empírico según el criterio revisado de Ayer.

### 1. TRADUCIBILIDAD A UN LENGUAJE EMPIRISTA COMO UN NUEVO CRITERIO DE SIGNIFICADO EMPÍRICO

Creo que sería inútil continuar la búsqueda de un criterio adecuado de comprobabilidad en términos de relaciones de deducibilidad con oraciones observacionales. El desarrollo de esta investigación, que hemos considerado en sus principales etapas, parece garantizar la presunción de que en tanto procuremos establecer un criterio de comprobabilidad de oraciones individuales en términos de sus rela-

---

ción fáctica de las oraciones, parece indudable que si una oración expresa hecho alguno, digamos  $f$ , entonces satisface el criterio establecido en la primera oración del principio, porque siempre podemos formar una clase que contenga a  $f$  junto con el hecho expresado por medio de alguna oración observacional a elección, lo que convierte a  $f$  en miembro de una clase de hechos, por lo menos de una clase directamente observable en principio. Por lo tanto, la primera parte del principio de especies observables es demasiado amplia, similar en esto a la primera formulación de Ayer del criterio empirista del significado.

<sup>12</sup> Esta restricción se expresa en forma recurrente y no involucra ningún círculo vicioso. Si se desea el enunciado completo del criterio de Ayer, véase (1), 2.ª ed., p. 13.

<sup>13</sup> Church (11).



ciones lógicas con oraciones observacionales, el resultado será o bien demasiado restrictivo, o bien demasiado amplio, o bien ambas cosas a la vez. En particular parece probable que tales criterios otorguen significado empírico, al modo de (2.1) *b* o (2.2) *b* ya sea a cualquier disyunción o a cualquier conjunción de dos oraciones de las cuales por lo menos una se califique como empíricamente significativa. Esta característica tiene consecuencias indeseables porque las liberales reglas gramaticales del castellano, así como las de cualquier otro lenguaje natural, acuerdan el carácter de oraciones a expresiones («El absoluto es perfecto» era nuestro ejemplo) que aun para los patrones empiristas más liberales no constituyen de ninguna manera aserciones, y tendrían que ser admitidas como componentes de enunciados empíricamente significativos.

Esta dificultad no surgiría, por supuesto, en un lenguaje artificial cuyo vocabulario y cuya gramática se eligieran de tal manera que impidiesen totalmente la posibilidad de formar cualquier oración que sea rechazada por el criterio empirista de significado. Llamemos *lenguaje empirista* a cualquier lenguaje de esas características. Esta reflexión sugiere un encaramiento del problema totalmente distinto: hacer una caracterización general del tipo de lenguaje que se consideraría empirista, y establecer luego el siguiente:

### 3.1. CRITERIO DE TRADUCIBILIDAD PARA EL SIGNIFICADO COGNOSCITIVO

Una oración tiene significado cognoscitivo si y sólo si es traducible a un lenguaje empirista.

Esta concepción del significado cognoscitivo parece estar en la base de muchos de los más recientes trabajos de los empiristas, aunque quizá no se la formule explícitamente. Hasta donde alcanzo a ver tiene su origen en el trabajo de Carnap «Testability and Meaning» (especialmente la parte IV).

Como cualquier otro lenguaje, el lenguaje empirista también puede caracterizarse indicando su vocabulario y las reglas que determinan su lógica; esta última incluye las reglas sintácticas de acuerdo con las cuales pueden formarse oraciones por medio del vocabulario dado. En efecto, el criterio de traducibilidad propone caracterizar las oraciones cognoscitivamente significativas por el vocabulario con el que puede construirse y por los principios sintácticos que gobiernan dicha construcción. Qué oraciones se considerarán cognoscitivamente significativas dependerá, entonces, de la elección del vocabulario y las reglas de construcción. Consideremos una de las posibilidades:

(3.2) Podríamos considerar empirista a un lenguaje *L* si satisface las siguientes condiciones:

(a) *El vocabulario de L* contiene:

(1) Las locuciones comunes de la lógica que se emplean en la formulación de oraciones, incluyendo especialmente las expresiones «no», «y», «o», «si... entonces», «todos», «algunos», «la clase de las cosas tales que...», «... es un elemento de la clase...»;

(2) Ciertos *predicados observacionales*. Éstos constituirán el vocabulario empírico básico de *L*;

(3) Toda expresión definible mediante las indicadas en (1) y (2).

(b) *Las reglas para la formación de oraciones en L* son las que se establecen en algún sistema lógico contemporáneo, v. gr., *Principia Mathematica*.

Puesto que pueden eliminarse todos los términos definidos a favor de los primitivos, estas reglas estipulan, en efecto, que un lenguaje *L* es empirista si todas sus oraciones son expresables, con la ayuda de las locuciones lógicas usuales, en términos de características observables de objetos físicos. Denominemos a cualquier lenguaje de esta especie lenguaje-cosa en sentido restringido. Inversamente, puede interpretarse el vocabulario empírico básico de un lenguaje empirista como consistiendo en términos fenomenistas, cada uno de los cuales se refiere a un aspecto del fenómeno perceptivo o sensitivo. Sin embargo, la construcción de lenguajes fenomenistas adecuados presenta dificultades considerables<sup>14</sup>, y en el empirismo más reciente se ha dirigido la atención especialmente a las potencialidades de los lenguajes cuyo vocabulario empírico básico consiste en predicados observacionales, pues éstos se prestan más directamente a la descripción del tipo de la prueba intersubjetiva que se requiere en la comprobación de las hipótesis científicas.

Si interpretamos los lenguajes empiristas en el sentido de (3.2), entonces el criterio de traducibilidad (3.1) elimina todos los inconvenientes señalados en nuestra discusión de las primeras formas del criterio de comprobabilidad.

(a) Nuestra caracterización de los lenguajes empiristas toma en cuenta explícitamente la cuantificación universal y existencial, o sea, el uso de los términos «todos» y «algunos»; de aquí que en general no se excluye ningún tipo de enunciado cuantificado del ámbito del discurso cognoscitivamente significativo.

(b) Oraciones tales como «El absoluto es perfecto» no pueden formularse en un lenguaje empirista (cfr. *d* más adelante); por lo tanto no existe el peligro de que se considere cognoscitivamente significativa una conjunción o disyunción que contenga una oración de ese tipo;

<sup>14</sup> Carnap (5) y Goodman (15) han hecho contribuciones importantes al problema.

(c) En un lenguaje  $L$  con reglas sintácticas conforme a *Principia Mathematica* la negación de una oración siempre es también una oración de  $L$ . Por lo tanto el criterio de traducibilidad no conduce a la consecuencia de que las negaciones de ciertas oraciones significativas son no significativas, como se sigue, en cambio, tanto de (2.1) como de (2.2);

(d) A pesar de su amplitud, el nuevo criterio no atribuye significado cognoscitivo a toda oración; por ejemplo, «El absoluto es perfecto» y «La nada anonada» no pueden traducirse a un lenguaje empirista porque sus términos claves no son definibles por medio de expresiones puramente lógicas y términos observacionales.

#### 4. EL PROBLEMA DE LOS TÉRMINOS DISPOSICIONALES Y LAS CONSTRUCCIONES LÓGICAS

Empero, este nuevo criterio es todavía demasiado restrictivo, como lo son sus predecesores, en un importante respecto que reclama nuestra atención. Si se definen los lenguajes empiristas de acuerdo con (3.2), entonces, como se señaló anteriormente, el criterio de traducibilidad (3.1) otorga significado cognoscitivo a una oración sólo si los términos empíricos que la constituyen pueden definirse explícitamente por medio de predicados observacionales. Pero como veremos enseguida, muchos términos aun de las ciencias físicas no pueden ser definidos de este modo; por lo tanto el criterio nos obligaría a rechazar por carentes de significado cognoscitivo a todas las hipótesis científicas que contengan esos términos, y ésta es una consecuencia totalmente inadmisibles.

Tal es el caso, por ejemplo, del concepto de temperatura. A primera vista parecería que la frase «El objeto  $x$  tiene una temperatura de  $c$  grados centígrados», o más brevemente « $T(x) = c$ » puede definirse mediante la siguiente oración ( $D$ ): « $T(x) = c$  si y sólo si satisface la condición siguiente: Si se pone en contacto un termómetro con  $x$  entonces registrará  $c$  grados sobre su escala.»

Sutilezas a un lado, puede afirmarse que el *definiens* dado aquí está formulado totalmente con términos observacionales. Sin embargo, tiene un aspecto muy cuestionable: en *Principia Mathematica* y otros sistemas similares se considera la frase «si  $p$  entonces  $q$ » sinónimo de «no  $p$  o  $q$ »; y en esta interpretación llamada material del condicional, un enunciado de la forma «si  $p$  entonces  $q$ » es obviamente verdadero si (aunque no sólo si) la oración que está en lugar de « $p$ » es falsa. Por lo tanto, si el significado de «si... entonces...» en el *definiens* de ( $D$ ) se entiende en el sentido material, dicho *definiens* es verdadero si (aunque no sólo si)  $x$  es un objeto que no está en contacto con un termómetro, cualquiera que sea el valor numérico que

otorguemos a *c*. Y puesto que el *definiendum* será verdadero en las mismas circunstancias, la definición (*D*) considerará verdadera la asignación de cualquier temperatura a todo objeto que no esté en contacto con un termómetro (*sic*). Consideraciones análogas se aplican a términos tales como «eléctricamente cargado», «magnético», «inteligente», «resistencia eléctrica», etc., en pocas palabras, a todos los términos disposicionales, o sea, términos que expresan la disposición de uno o más objetos a reaccionar de determinada manera en condiciones especificadas. No puede darse una definición de estos términos por medio de predicados observacionales a la manera de (*I*), por muy natural y obvio que parezca en primera instancia este modo de definición <sup>15</sup>.

Se puede tratar de resolver esta dificultad en dos direcciones principales. Por un lado, puede sostenerse que la definición de los términos disposicionales a la manera de (*D*) es perfectamente adecuada siempre que la frase «si... entonces...» que figura en el *definiens* se interprete en el sentido que obviamente se supone que posee, a saber, implicando en el caso (*D*) que aunque *x* no esté realmente en contacto con un termómetro, si ése *fuera* el caso, entonces el termómetro registraría *c* grados. En oraciones como ésta se dice que la frase «si... entonces...» se emplea contrafácticamente; y en este sentido «fuerte», que implica un condicional contrafáctico, tendría que interpretarse el *definiens* de (*D*). Esta sugestión ofrecería una respuesta al problema de definir términos disposicionales si no fuese que hasta el momento no se ha podido dar cuenta satisfactoriamente del significado exacto de los condicionales contrafácticos. Por lo tanto, el primer camino para salvar la dificultad posee el carácter de programa más bien que el de solución. La falta de una teoría adecuada de los condicionales contrafácticos es aún más deplorable por cuanto se la necesita también para el análisis del concepto de ley general en la ciencia empírica y otras ideas conexas. Un esclarecimiento de este conjunto de problemas constituye actualmente uno de los urgentes desiderata en la lógica y la metodología de la ciencia <sup>16</sup>.

Carnap sugirió otra manera de encarar el problema de la definición originado por los términos disposicionales, y la desarrolló en de-

<sup>15</sup> Esta dificultad en la definición de los términos disposicionales fue señalada y analizada por primera vez por Carnap (en (6); ver especialmente la sección 7).

<sup>16</sup> El concepto de implicación estricta tal como lo entiende C. I. Lewis no nos servirá para interpretar el «si... entonces...» fuerte tal como lo entendemos aquí, puesto que se refiere a una relación de implicación puramente lógica, mientras que el concepto en cuestión en general representará una relación nomológica, esto es, basada en leyes empíricas. Para discusiones recientes de los problemas de los condicionales contrafácticos y las leyes, véase Langford (18); Lewis (20); pp. 210-30; Chisholm (10); Goodman (14); Reichenbach (26), cap. VIII; Hempel y Oppenheim (16), parte III y Popper (24).

talle. Consiste en permitir la introducción de términos nuevos en un lenguaje empirista mediante las llamadas oraciones reductivas [*reduction sentences*], que tienen el carácter de definiciones parciales o condicionales<sup>17</sup>. Así, en nuestro último ejemplo puede introducirse el concepto de temperatura por medio de la siguiente oración reductiva (R): Si un termómetro está en contacto con un objeto  $x$ , entonces  $T(x) = c$  si y sólo si el termómetro registra  $c$  grados.

Esta regla, con la cual el condicional puede interpretarse en el sentido material, especifica el significado de «temperatura» [esto es, de los enunciados de la forma « $T(x) = c$ »] sólo parcialmente, a saber, con respecto a los objetos que estén en contacto con un termómetro; para todos los objetos restantes no determina el significado de « $T(x) = c$ ». La especificación del significado de «temperatura» entonces puede extenderse sucesivamente a casos no comprendidos en (R) formulando nuevas oraciones reductivas que indiquen la medición de la temperatura por medios distintos del empleo de termómetros.

Las oraciones reductivas no permiten, pues, formular con precisión las definiciones que suelen llamarse «operacionales»<sup>18</sup>. Al mismo tiempo muestran que estas últimas no son definiciones en el estricto sentido de la palabra, sino más bien especificaciones parciales de significado.

Estas consideraciones sugieren que en nuestra caracterización (3.2) de los lenguajes empiristas ampliamos la condición en  $a$ -(3) introduciendo en el vocabulario de  $L$  todos aquellos términos cuyo significado pueda especificarse en términos del vocabulario empírico básico por medio de definiciones u oraciones reductivas. Los lenguajes que satisfagan este criterio más amplio serán llamados lenguajes cosa en sentido amplio.

Si se amplía de esta manera el concepto de lenguaje empirista, el criterio de traducibilidad alcanza también, como tendría que hacerlo, a todos aquellos enunciados cuyos términos empíricos constitutivos incluyan «construcciones conceptuales empíricas», o sea, términos que no designan observables, pero que pueden introducirse mediante oraciones reductivas sobre la base de predicados observacionales.

Sin embargo, aun en esta formulación generalizada nuestro criterio de significado cognoscitivo puede no hacer justicia a las teorías

<sup>17</sup> Cfr. Carnap (6); puede encontrarse una exposición elemental y breve de la idea central en Carnap (7), parte III. La definición parcial (R) formulada anteriormente para la expresión « $T(x) = c$ » sólo ilustra el tipo más simple de oración reductiva, llamada oración reductiva bilateral.

<sup>18</sup> Acerca del concepto de definición operacional, que fue desarrollado por Bridgman, véase, por ejemplo, Bridgman (3, 4) y Feigl (12).

científicas avanzadas, las cuales se formulan en términos de «construcciones teóricas», como, por ejemplo, los términos «temperatura absoluta», «potencial gravitatorio», «campo eléctrico», «función  $\psi$ », etc. Hay razones para pensar que ni las definiciones ni las oraciones reductivas son adecuadas para introducir estos términos sobre la base de predicados observacionales. Por ejemplo, si se dispusiera de un sistema de oraciones reductivas para el concepto de campo eléctrico, entonces (simplificando algo la cosa) sería posible describir en términos de características observables condiciones necesarias y suficientes para la presencia de un campo eléctrico en una región dada, cualquiera que sea la complejidad de su descripción matemática. Sin embargo, en el mejor de los casos sólo pueden formularse tales criterios para algunas especies de campos suficientemente simples.

Ahora bien, las teorías de tipo avanzado mencionadas aquí pueden considerarse como sistemas hipotético-deductivos en los cuales todos los enunciados son consecuencias lógicas de un conjunto de suposiciones fundamentales. En un sistema tal, tanto los enunciados fundamentales como los derivados se formulan o bien en términos de ciertas construcciones conceptuales teóricas que no se definen en el sistema y que por lo tanto representan el papel de primitivos, o bien en términos de expresiones definidas en términos de estos últimos. Por lo tanto estos sistemas se asemejan por su estructura lógica a los sistemas axiomatizados no interpretados que se estudian en lógica y matemática. Se tornan aplicables a las cuestiones empíricas, adquiriendo así el *status* de teorías de la ciencia empírica, en virtud de una interpretación empírica. Ésta se lleva a cabo traduciendo algunas de las oraciones de la teoría, más a menudo derivadas que fundamentales, a un lenguaje empirista, que puede contener tanto predicados observacionales como construcciones empíricas. Y puesto que las oraciones a las que de este modo se les confiere significado empírico son consecuencias lógicas de las hipótesis fundamentales de la teoría, dicha traducción produce indirectamente una interpretación parcial de dichas hipótesis y de las construcciones conceptuales en términos de las cuales se las formula <sup>19</sup>.

<sup>19</sup> La distinción entre un sistema deductivo formal y la teoría empírica que resulta de ella en virtud de una interpretación ha sido elaborada en detalle por Reichenbach en sus profundos estudios acerca de las relaciones entre la geometría pura y la geometría física; cfr. Reichenbach (25). El método mediante el cual se confiere a un sistema formal un contenido empírico es caracterizado por Reichenbach como «definición coordinadora» de los primitivos de la teoría por medio de conceptos empíricos específicos. Sin embargo, como lo sugiere nuestra discusión de la reducción y de la interpretación de las construcciones conceptuales teóricas, el proceso en cuestión debe considerarse como una interpretación parcial de los términos no lógicos del sistema y no como una definición completa de estos últimos en términos de los conceptos de un lenguaje-cosa.

A los efectos de hacer de la traducibilidad a un lenguaje empirista un criterio adecuado de contenido cognoscitivo, extendemos así el concepto de lenguaje empirista de modo tal que incluya lenguajes en cosa en sentido amplio y en sentido restringido, como así también todos los sistemas teóricos interpretados del tipo de los que hemos mencionado <sup>20</sup>. Entendido de esta manera (3.1), puede servir finalmente como criterio general de significado cognoscitivo.

### 5. SOBRE «EL SIGNIFICADO» DE UN ENUNCIADO EMPÍRICO

En efecto, el criterio que acabamos de formular califica a una oración de cognoscitivamente significativa si sus componentes no lógicos se refieren a observables directamente o de ciertas y determinadas maneras indirectas. Pero no se pronuncia acerca de cuál es «el significado» de una oración cognoscitivamente significativa, y en particular no dice ni implica que el significado pueda caracterizarse exhaustivamente por lo que revelarían todos los tests posibles en términos de fenómenos observables. En realidad, *generalmente el contenido de un enunciado de contenido empírico no puede expresarse exhaustivamente por medio de ninguna clase de oraciones observacionales.*

Consideremos primero, entre los enunciados que admite nuestro criterio, cualquier hipótesis puramente existencial o cualquier enunciado en el que se dé una cuantificación mixta. Como señalamos anteriormente, en (2.2) a, los enunciados de esta especie no implican ninguna oración observacional; por lo tanto su contenido no puede expresarse por medio de una clase de oraciones observacionales.

En segundo lugar, aun la mayor parte de los enunciados de forma puramente universal (tales como «Todos los flamencos son rosados») implican oraciones observacionales (tales como «Esa cosa es rosa») sólo cuando se unen a otras oraciones observacionales adecuadas (tales como «Esa cosa es un flamenco»).

Puede generalizarse esta última observación; el empleo de hipótesis empíricas para predecir fenómenos observables requiere en casi

<sup>20</sup> No hemos caracterizado aquí a estos sistemas todo lo acabada y precisamente que sería deseable. En verdad, el carácter exacto de la interpretación empírica de las construcciones conceptuales teóricas y de las teorías en las cuales funcionan requiere aún ser investigado. Algunos de los problemas que surgen a este respecto, como por ejemplo, si puede decirse, y en qué sentido, que las construcciones conceptuales teóricas denotan algo, sin duda son de considerable interés gnoseológico. En Carnap (8), sección 24, y en Kaplan (17) pueden encontrarse algunas sugerencias referentes a la interpretación de los esquemas conceptuales teóricos. Para una excelente discusión de los aspectos gnoseológicos del problema véase Feigl (13).

todos los casos el uso de hipótesis empíricas auxiliares <sup>21</sup>. Por ejemplo, la hipótesis de que el agente de la tuberculosis tiene forma de bastón por sí sola no implica que si se observan a través de un microscopio esputos de un tuberculoso se hallarán objetos con forma de bastón; tendrán que emplearse como premisas adicionales de un gran número de hipótesis auxiliares, incluyendo la teoría del microscopio, para deducir esa predicción.

En consecuencia, lo que se denomina «el significado (cognoscitivo)» de una hipótesis científica dada no puede ser adecuadamente caracterizado en términos de pruebas potencialmente observacionales solamente, ni puede tampoco especificarse para la hipótesis tomada aisladamente; para entender «el significado» de una hipótesis formulada en un lenguaje empirista, no sólo tenemos que saber qué oraciones observacionales implica, sola o en conjunción con hipótesis auxiliares, sino también qué otras oraciones empíricas no observacionales implica, qué oraciones del lenguaje empirista dado la confirmarían o refutarían, y qué otras hipótesis ella misma confirmaría o refutaría. En otras palabras, el significado cognoscitivo de un enunciado en un lenguaje empirista surge de la totalidad de sus relaciones con todos los demás enunciados del lenguaje y no con las oraciones observacionales solamente. En este sentido, los enunciados de la ciencia empírica tienen un significado que excede a lo que puede ser expresado en términos de oraciones observacionales relevantes <sup>22</sup>.

## 6 EL STATUS LÓGICO DEL CRITERIO EMPIRISTA DE SIGNIFICADO

Se inquiere a menudo: ¿qué tipo de oración es el propio criterio empirista de significado? Ciertamente no es una hipótesis empírica; pero tampoco es analítica ni contradictoria; por lo tanto, si se la juzga con sus mismo términos, ¿no carece de significado cognoscitivo? Y en tal caso, ¿qué tipo de verdad o validez puede otorgársele?

Se lo podría interpretar como una definición que indique lo que los empiristas entienden por una oración cognoscitivamente significativa, así entendido no tendría el carácter de una aserción y no sería verdadero ni falso. Pero esta concepción asignaría al criterio un grado de arbitrariedad inconciliable con las ardientes controversias que ha originado, y aun menos con el hecho repetidamente señalado en este ar-

<sup>21</sup> En los criterios del significado cognoscitivo de Ayer, los cuales fueron considerados en la sección 2, se toma en cuenta claramente este punto.

<sup>22</sup> Para una discusión más completa de los problemas involucrados aquí, cfr. Feigl (13) y los comentarios acerca de la posición de Feigl que serán publicados junto con este artículo.



título, de que los cambios en su contenido específico siempre han sido determinados por la pretensión de hacerlo un índice más adecuado de significado cognoscitivo. Y este objetivo aclara el carácter del criterio empirista de significado: pretende ofrecer un esclarecimiento y *elucidación* de la idea de una oración que constituya una aserción inteligible<sup>23</sup>. Se admite que dicha idea sea vaga, y es la tarea de la elucidación filosófica reemplazarla por un concepto más preciso. En vista de esta diferencia de precisión no podemos exigir, por supuesto, que este «nuevo» concepto, el *explicatum*, sea idéntico al viejo, el *explicandum*<sup>24</sup>. ¿Cómo hemos de juzgar entonces la adecuación de la elucidación propuesta tal como se expresa en determinado criterio de significado cognoscitivo?

Ante todo, existe una clase numerosa de oraciones que generalmente se reconoce que constituyen aserciones inteligibles, y otra clase, igualmente numerosa, de las que se suele negar que lo sean. Necesitamos una elucidación adecuada que tome en cuenta estas esferas de uso común; por lo tanto tendrá que rechazarse toda elucidación que niegue significado cognoscitivo a las descripciones de acontecimientos pasados o generalizaciones expresadas en términos de observables. Como hemos visto, este primer requisito de adecuación ha representado un papel importante en el desarrollo del criterio empirista de significado.

Pero una elucidación adecuada del concepto de enunciado cognoscitivamente significativo debe satisfacer también otro requisito, aún más importante: además de la elucidación de algunos otros conceptos, tales como los de confirmación y probabilidad, debe ofrecer el esquema de una descripción teórica general de la estructura y los fundamentos del conocimiento científico. La elucidación, como aquí se la entiende, no es una mera descripción de los usos aceptados de los términos en cuestión; tiene que ir más allá de las limitaciones, ambigüedades e incoherencias del uso común, mostrando cómo tendríamos que interpretar el significado de dichos términos si hemos de lograr una teoría del conocimiento coherente y amplia.

<sup>23</sup> En el prefacio de la segunda edición de su libro, Ayer adopta una posición muy similar; sostiene que el criterio de comprobabilidad es una definición que con todo no es enteramente arbitraria, porque una oración que no satisfaga dicho criterio «no podría ser comprendida en el sentido en el que habitualmente se entienden tanto las hipótesis científicas como los enunciados de sentido común» (1), p. 16).

<sup>24</sup> Cfr. la caracterización que hace Carnap de la elucidación en su artículo (9), donde examina en sus líneas generales la elucidación del concepto de probabilidad. La definición de Frege-Russell de los enteros como clases de clases equivalentes, y la definición semántica de la verdad (cfr. Tarski (30)), son notorios ejemplos de elucidación. Para una lúcida discusión de varios aspectos del análisis lógico véase Pap (21), cap. 17.

Este tipo de consideraciones, que han sido muy influidas por el estudio de la estructura de teorías científicas, han impulsado los desarrollos más recientes del criterio empirista de significado. Estas ampliaciones han sido diseñadas para incluir en el ámbito de lo cognoscitivamente significativo varios tipos de oraciones que podrían aparecer en teorías científicas avanzadas, o que tendrían que admitirse a favor de una simplicidad sistemática y uniformidad<sup>25</sup>, pero acerca de cuyo significado o no significado cognoscitivo el estudio del significado del término «aserción inteligible» en el habla cotidiana no arrojaría luz alguna.

Como consecuencia de ello, el criterio empirista de significado, así como el resultado de cualquier otra elucidación, representa una propuesta lingüística que en sí misma no es ni verdadera ni falsa, pero a la cual se exige que sea adecuada en dos aspectos: primero, en el sentido de que suministre un minucioso *análisis* del significado del *explicandum* usualmente aceptado, y este requisito implica una aserción empírica; y en segundo lugar, en el sentido de que la elucidación logre una «reconstrucción racional» del *explicandum*, o sea, que suministre, quizá junto con otras elucidaciones, un esquema conceptual general que permita una reformulación coherente y precisa y una sistematización teórica de los contextos en los que se usa el *explicandum*, y este requisito implica por lo menos una aserción de carácter lógico.

Aunque por su forma es una propuesta, el criterio empirista de significado está lejos de ser una definición arbitraria; está sujeto a revisión en caso de descubrirse una violación de los requisitos de adecuación, o aun si se hallan maneras de satisfacerlos más completamente. En realidad cabe esperar que a corto plazo se esclarezcan algunos de los problemas abiertos surgidos en el análisis del significado cognoscitivo, y que entonces nuestra última versión del criterio empirista de significado sea sustituida por otra más adecuada.

<sup>25</sup> Por ejemplo, nuestro criterio califica de significativos a ciertos enunciados que contienen, digamos, miles de cuantificadores existenciales o universales, aun cuando tales oraciones pueden no aparecer en el lenguaje cotidiano, ni aun en el discurso científico. En realidad, desde un punto de vista sistemático, sería arbitrario e injustificable limitar la clase de los enunciados significativos a aquellos que contengan no más de un número determinado de cuantificadores. Para una discusión más completa de este punto cfr. Carnap (6), secciones 17, 24 y 25.

## REFERENCIAS BIBLIOGRÁFICAS

- (1) AYER, A. J.: *Language, Truth and Logic*, Oxford University Press, 1936, 2.ª ed., Londres, Gollancz, 1946. [Versión castellana: Martínez Roca, Barcelona, 1971.]
- (2) BENJAMIN, A. C.: «Is Empiricism Self-Refuting?», en *Journal of Philosophy*, 38, 1941.
- (3) BRIDGMAN, P. W.: *The Logic of Modern Physics*, Macmillan, Nueva York, 1927.
- (4) BRIDGMAN, P. W.: «Operational Analysis», en *Philosophy of Science*, 5, 1938.
- (5) CARNAP, R.: *Der logische Aufbau der Welt*, Berlín, 1928.
- (6) CARNAP, R.: «Testability and Meaning», en *Philosophy of Science*, 3, 1936, y 4, 1937.
- (7) CARNAP, R.: «Logical Foundations of the Unity of Science», en *International Encyclopedia of Unified Science*, I, 1, University of Chicago Press, 1939.
- (8) CARNAP, R.: «Foundations of Logic and Mathematics», en *International Encyclopedia of Unified Science*, I, 3, University of Chicago Press, 1939.
- (9) CARNAP, R.: «The Two Concepts of Probability», en *Philosophy and Phenomenological Research*, 5, 1945.
- (10) CHISHOLM, R. M.: «The Contrary-To-Fact Conditional», en *Mind*, 55, 1946.
- (11) CHURCH, A.: Reseña de (1), 2.ª ed., en *The Journal of Symbolic Logic*, 14, 1949, pp. 52-3.
- (12) FEIGL, H.: «Operationism and Scientific Method», en *Psychological Review*, 52, 1945. Reproducido en Feigl y Sellars (comps.), *Readings in Philosophical Analysis*, Nueva York, 1949.
- (13) FEIGL, H.: «Existential Hypothesis: Realistic vs. Phenomenalistic Interpretations», en *Philosophy of Science*, 17, 1950.
- (14) GOODMAN, N.: «The Problem of Counterfactual Conditionals», en *Journal of Philosophy*, 44, 1947.
- (15) GOODMAN, N.: *The Structure of Appearance*, Harvard University Press, Cambridge, 1951.
- (16) HEMPEL, C. G., y OPPENHEIM, P.: «Studies in the Logic of Explanation», en *Philosophy of Science*, 15, 1948.
- (17) KAPLAN, A.: «Definition and Specification of Meaning», en *Journal of Philosophy*, 43, 1946.
- (18) LANGFORD, C. H.: Reseña en *The Journal of Symbolic Logic*, 6, 67, 1941.
- (19) LECOMTE DU NOÛY: *Human Destiny*, Nueva York/Londres/Toronto, 1947.
- (20) LEWIS, C. I.: *An Analysis of Knowledge and Valuation*, Open Court Publ. La Salle, Ill., 1946.
- (21) PAP, A.: *Elements of Analytic Philosophy*, Macmillan, Nueva York, 1949.
- (22) POPPER, K. R.: *Logik der Forschung*, Springer, Viena, 1935. [Traducción inglesa: *The Logic of Scientific Discovery*, Hutchinson, Londres, 1959. Versión castellana: Tecnos, Madrid, 1972.]
- (23) POPPER, K. R.: *The Open Society and its Enemies*, 2 vols., Routledge, Londres, 1945. [Traducción castellana: *La sociedad abierta y sus enemigos*, Paidós, Buenos Aires, 1957.]
- (24) POPPER, K. R.: «A Note on Natural Laws and So-Called "Contrary-To-Fact Conditionals», en *Mind*, 58, 1949.
- (25) REICHENBACH, H.: *Philosophie der Raum-Zeit-Lehre*, Berlín, 1928. [Traducción inglesa: *The Philosophy of Space and Time*, Dover, Nueva York, 1958.]
- (26) REICHENBACH, H.: *Elements of Symbolic Logic*, Macmillan, Nueva York, 1947.
- (27) RUSSELL, B.: *Human Knowledge*, Simon and Schuster, Nueva York, 1948. [Versión castellana en Taurus, Madrid, 1977.]

- (28) SCHLICK, M.: «Meaning and Verification», en *Philosophical Review*, 45, 1946. Reproducido en Feigl y Sellars (comps.), *Readings in Philosophical Analysis*, Nueva York, 1949.
- (29) STACE, W. T.: «Positivism», en *Mind*, 53, 1944.
- (30) TARSKI, A.: «The Semantic Conception of Truth and the Foundations of Semantics», en *Philosophy and Phenomenological Research*, 4, 1944 [pp. 275-313 de este volumen].
- (31) WERKMEISTER, W. H.: *The Basis and Structure of Knowledge*, Harper, Nueva York/Londres, 1948.
- (32) WHITEHEAD, A. N. y RUSSELL, B.: *Principia Mathematica*, 3 vols., 2.ª ed., Cambridge, 1925-1927.

## DOS DOGMAS DEL EMPIRISMO \*

WILLARD V. O. QUINE

El empirismo moderno ha sido en gran parte condicionado por dos dogmas. Uno de ellos es la creencia en cierta distinción fundamental entre verdades que son *analíticas*, basadas en significaciones, con independencia de consideraciones fácticas, y verdades que son  *sintéticas*, basadas en los hechos. El otro dogma es el *reductivismo*, la creencia en que todo enunciado que tenga sentido es equivalente a alguna construcción lógica basada en términos que refieren a la experiencia inmediata. Voy a sostener que ambos dogmas están mal fundados. Una consecuencia de su abandono es, como veremos, que se desdibuja la frontera que se supone trazada entre la metafísica especulativa y la ciencia natural. Otra consecuencia es una orientación hacia el pragmatismo.

### 1. EL TRASFONDO DE LA ANALITICIDAD

La distinción kantiana entre verdades analíticas y verdades sintéticas fue anticipada por la distinción de Hume entre relaciones de ideas y cuestiones de hecho, y por la distinción leibniziana entre verdades de razón y verdades de hecho. Leibniz decía de las verdades de razón que son verdaderas en todos los mundos posibles. Dejando aparte ese pintoresquismo, lo que quería decir es que las verdades de razón son aquellas que no pueden ser falsas. En el mismo sentido vemos definir los enunciados analíticos como aquellos enunciados cuyas negaciones son autocontradictorias. Pero esta definición tiene escaso valor explicativo, pues la noción de autocontradictoriedad, en el muy amplio sentido requerido por esta definición de la analiticidad, necesita tanta clarificación como la misma noción de analiticidad. Las dos nociones son la cara y la cruz de una misma problemática moneda.

Kant concebía un enunciado analítico como aquel que no atribuye a su sujeto más de lo que ya está conceptualmente contenido en dicho sujeto. Esta formulación tiene dos insuficiencias: se limita a

---

\* Versión castellana de Manuel Sacristán.

enunciados de la forma sujeto-predicado, y apela a la noción de contenido, dejándola, al mismo tiempo, al nivel de una metáfora. Pero la intención de Kant, que se manifiesta en el uso que hace de la noción de analiticidad más que en su definición de ella, puede precisarse del modo siguiente: un enunciado es analítico cuando es verdadero por virtud de significaciones e independientemente de los hechos. Continuemos siguiendo esa línea el concepto de *significación* que queda presupuesto.

Recordemos que significar y nombrar no pueden identificarse <sup>1</sup>. El ejemplo de Frege de «el lucero de la tarde» y «el lucero del alba» y el ejemplo russelliano de «Scott» y «el autor de *Waverley*» ilustran el hecho de que diversos términos pueden nombrar o denotar la misma cosa y diferir por su significación o sentido. No menos importante es la distinción entre significar y nombrar al nivel de los términos abstractos. Los términos «9» y «el número de los planetas» nombran una sola y misma cosa, pero seguramente deben considerarse diversos en cuanto al sentido; pues para determinar la identidad de la entidad en cuestión hizo falta practicar observaciones astronómicas y no bastó la mera reflexión sobre significaciones.

Los anteriores ejemplos constan de términos singulares, concretos o abstractos. Con términos generales, o predicados, la situación es algo diversa, pero paralela. Mientras que un término singular pretende nombrar una entidad, abstracta o concreta, un término general o universal no tiene ese alcance, sino que es *verdadero* de una entidad, o de cada una de muchas, o de ninguna de ellas <sup>2</sup>. La clase de todas las entidades de las que es verdadero un término general se llama *extensión* del mismo. En paralelismo con el contraste que se da entre la significación o el sentido de un término singular y la entidad denotada por él tenemos que distinguir ahora análogamente entre el sentido de un término general y su extensión. Los términos generales «criatura con corazón» y «criatura con riñones», por ejemplo, son quizás iguales en extensión, pero desiguales en significación.

La confusión de la significación con la extensión es menos común en el caso de los términos generales que la confusión de significación con denotación en el caso de los términos singulares. Es, en efecto, un tópico filosófico la oposición entre intensión \* (o significación, o sentido) y extensión, o bien, en un léxico diverso, entre connotación y denotación.

La noción aristotélica de esencia fue sin duda la precursora de la noción moderna de intensión, significación y sentido. Para Aristóte-

<sup>1</sup> Cfr. ensayo anterior, p. 35 [se refiere a «Acerca de lo que hay» en W. V. O. Quine, *Desde un punto de vista lógico*, Ariel, Barcelona, 1962].

<sup>2</sup> Cfr. *supra*, p. 36 e *infra*, pp. 159-171 [se refiere a la obra citada en nota 1].

\* En la terminología tradicional: comprensión o comprensión. (*N. del T.*)

les, era esencial al hombre el ser racional, y accidental el ser bípedo. Pero hay una diferencia importante entre esa actitud y la teoría de la significación. Desde el punto de vista de la última puede en efecto concederse (pero sólo por necesidades de la discusión) que la racionalidad esté incluida en la significación de la palabra «hombres» mientras que el tener dos piernas no lo esté; pero el tener dos piernas puede al mismo tiempo considerarse incluido en la significación de «bípedo», mientras que la racionalidad no lo está. Así que, desde el punto de vista de la teoría de la significación, no tiene sentido decir de un individuo concreto, que sea a la vez hombre y bípedo, que su racionalidad le es esencial y que su tener dos piernas le es accidental o viceversa. Las cosas, según Aristóteles, tienen esencia, pero sólo las formas lingüísticas tienen significación. Significación es aquello en que se convierte la esencia cuando se separa de su objeto de referencia y se adscribe a la palabra.

Una cuestión capital para la teoría de la significación es la de la naturaleza de su objeto: ¿qué clase de cosas son las significaciones? La necesidad tradicionalmente sentida de recurrir a entidades mentadas puede deberse a la antigua ceguera para apreciar el hecho de que significación y referencia son dos cosas diversas. Una vez tajantemente separadas la teoría de la referencia y la de la significación, basta dar un breve paso para reconocer que el objeto primario de la teoría de la significación es, simplemente, la sinonimia de las formas lingüísticas y la analiticidad de los enunciados; las significaciones mismas, en tanto que oscuras entidades intermediarias, pueden abandonarse tranquilamente<sup>3</sup>.

Así nos encontramos, pues, de nuevo con el problema de la analiticidad. No hay que buscar mucho para dar con enunciados que sean analíticos por filosófica aclamación. Esos enunciados se distribuyen en dos clases. Los de la primera clase, que pueden llamarse *lógicamente verdaderos*, pueden tipificarse mediante el enunciado siguiente:

(1) Ningún hombre no casado es casado.

El rasgo relevante de ese ejemplo consiste en que no sólo es verdadero tal como queda enunciado, sino que sigue siéndolo para toda nueva interpretación de «hombre» y «casado». Si suponemos un inventario previo de partículas *lógicas*, con «no» y otras formas de negación, «si», «entonces» (en sentido ilativo, no temporal), «y», etc., puede decirse en general que una verdad lógica es un enunciado que es verdadero y sigue siéndolo para cualquier interpretación de sus componentes que no sean partículas lógicas.

<sup>3</sup> Cfr. *supra*, pp. 37 s., e *infra*, pp. 84 s. [se refiere a la obra citada en nota 1]

Pero hay además una segunda clase de enunciados analíticos, tipificable por:

(2) Ningún soltero es casado.

Lo característico de un enunciado como éste es que puede convertirse en una verdad lógica sustituyendo sinónimos por sinónimos; así (2) puede convertirse en (1) poniendo «hombre no casado» por su sinónimo «soltero». Pero seguimos careciendo de una caracterización adecuada de esta segunda clase de enunciado analítico y, por tanto, de la analiticidad en general, pues en la anterior descripción nos hemos basado en una noción de «sinonimia» que no necesita menos aclaración que la de analiticidad.

En años recientes Carnap ha tendido a explicar la analiticidad apelando a lo que llama descripciones de estado <sup>4</sup>. Una descripción de estado es cualquier asignación exhaustiva de valores veritativos a los enunciados atómicos, no compuestos, del lenguaje. Carnap admite que todos los demás enunciados del lenguaje se construyen a partir de sus cláusulas componentes por medio de los expedientes lógicos habituales, de tal modo que el valor veritativo de cualquier enunciado complejo queda fijado para cada descripción de estado por leyes lógicas especificables. Un enunciado se explica entonces como analítico cuando resulta verdadero para cualquier descripción de estado. Esta explicación es una adaptación de la idea leibniziana de «verdad en todos los mundos posibles». Pero nótese que esta versión de la analiticidad consigue su propósito sólo en el caso de que los enunciados atómicos del lenguaje sean recíprocamente independientes; a diferencia de lo que ocurre con «Juan es soltero» y «Juan es casado». Si no hay tal independencia, habrá una descripción de estado que asigne el valor verdad a «Juan es soltero» y a «Juan es casado», con lo que «Ningún soltero es casado» resultaría, bajo el criterio ofrecido, sintético en vez de analítico. Así pues, el criterio de analiticidad en términos de descripciones de estado no sirve más que para lenguajes que carezcan de pares sinónimos del tipo que precisamente da origen a la «segunda clase» de enunciados analíticos. Este criterio es pues, en el mejor de los casos, una reconstrucción de la verdad lógica, y no de la analiticidad.

No quiero decir con ello que Carnap se haga ilusiones en este punto. Su simplificado modelo lingüístico, con sus descripciones de estado, no está primariamente orientado hacia la solución del problema general de la analiticidad, sino hacia otro objetivo, a saber, la aclaración de los problemas de la probabilidad y la inducción. Nues-

<sup>4</sup> Carnap, pp. 9; [2], pp. 70 ss.



tro problema es en cambio la analiticidad; y en este campo la dificultad no se encuentra en la primera clase de enunciados analíticos, las verdades lógicas, sino más bien en la segunda clase, que depende de la noción de sinonimia.

## 2. DEFINICIÓN

Hay quien considera resolutoria la salida consistente en decir que los enunciados de la segunda clase se reducen a los de la primera, a las verdades lógicas, por *definición*; «soltero», por ejemplo, se *define* como «hombre no casado». Pero, ¿cómo descubrimos que «soltero» se define por «hombre no casado»? ¿Quién lo ha definido así, y cuando? ¿Es que basta con apelar al diccionario más a mano y con aceptar como una ley la formulación del lexicógrafo? Esto equivaldría a poner la carreta delante de los bueyes. El lexicógrafo es un científico empírico, cuya tarea consiste en recopilar hechos antecedentes; y si glosa la palabra «soltero» mediante «hombre no casado» es porque cree que se da una relación de sinonimia entre esas formas, relación implícita en el uso general o preponderante anterior a su propia obra. La misma noción de sinonimia, presupuesta por el lexicógrafo, tiene que ser aclarada, presumiblemente en términos referentes al comportamiento lingüístico. Está claro que la «definición», que no es más que el informe del lexicógrafo acerca de una sinonimia observada, no puede tomarse como fundamento de la sinonimia.

Pero la definición no es exclusivamente una actividad de filólogos. Filósofos y científicos tienen frecuentemente ocasión de «definir» un término abstruso parafraseándolo en términos de un vocabulario más familiar. No obstante, ordinariamente una tal definición, igual que la del filólogo, es mera cuestión de lexicografía, y afirma simplemente una relación de sinonimia anterior a la exposición en curso.

Lo que no está aclarado, ni mucho menos, es lo que significa el afirmar una sinonimia, qué son las interconexiones que resultan necesarias y suficientes para que dos formas lingüísticas puedan describirse correctamente como sinónimas; pero, cualesquiera que sean, esas interconexiones están ordinariamente basadas en el uso. Las definiciones que aportan casos seleccionados de sinonimia son, pues, informaciones acerca del uso.

Hay empero un tipo de actividad definatoria que no se limita a informar acerca de sinonimias preexistentes. Pienso al decir esto en la que Carnap llama *explicación*, actividad a la que son aficionados los filósofos y también los científicos en sus momentos más filosóficos. En la explicación, la intención no es meramente parafrasear el *definiendum* mediante un sinónimo palmario, sino perfeccionar real

mente el *definiendum*, afinando o completando su significación. Pero incluso la explicación, a pesar de no consistir meramente en recoger una sinonimia preexistente entre el *definiendum* y el *definiens*, des-cansa de todos modos en otras sinonimias preexistentes. Esta cuestión puede considerarse del modo siguiente. Toda palabra digna de explicación tiene algunos contextos que, en conjunto, son lo suficientemente claros y precisos como para resultar útiles; el objeto de la explicación es preservar el uso de esos contextos privilegiados y afinar el uso de otros contextos. Para que una determinada definición sea adecuada a fines de explicación, lo que se requiere no es, por tanto, que en el uso anterior el *definiendum* fuera sinónimo del *definiens*, sino sólo que todos y cada uno de los contextos privilegiados del *definiendum*, tomados como un todo en su uso anterior, sean sinónimos del contexto correspondiente del *definiens*.

Dos *definientia* alternativos pueden ser igualmente apropiados para los fines de una misma tarea de explicación, aun sin ser sinónimos entre sí; pues pueden ser ambos igualmente apropiados en los contextos privilegiados, y diferir en cambio en otros. Al escoger uno de esos *definientia* en vez de otro, una definición de tipo explicativo engendra, por un *fiat*, una relación de sinonimia entre *definiendum* y *definiens* que no existía antes. Pero, como se ha visto, una tal definición debe su propia función explicativa a sinonimias anteriores.

Queda, de todos modos, un tipo extremo de definición que no recurre en absoluto a sinonimias anteriores, a saber, la introducción, explícitamente convencional, de nuevas notaciones con fines de mera abreviación. Aquí el *definiendum* se hace sinónimo del *definiens* simplemente porque ha sido precisamente creado para ser sinónimo del *definiens*. Éste es un patente caso de sinonimia creada por definición; si esto ocurriera en todos los casos, todas las especies de sinonimia serían inteligibles sin más. Pero, en general, la definición descansa en la sinonimia más que explicarla.

La palabra «definición» ha llegado a cobrar un sonido peligroso por la tranquilidad que produce, seguramente a causa de la frecuencia con que aparece en los escritos lógicos y matemáticos. Será conveniente ahora hacer una breve digresión para apreciar el papel de la definición en el trabajo formal.

En los sistemas lógicos y matemáticos puede preferirse uno de dos tipos antagónicos de economía expresiva, cada uno de los cuales tiene su peculiar utilidad práctica. Por un lado, puede buscarse la economía de la expresión práctica, facilidad y brevedad en el enunciado de relaciones complejas. Este tipo de economía exige generalmente notaciones concisas y bien distintas para una gran cantidad de conceptos. Por otro lado, en cambio, puede buscarse una economía en la gramática y el vocabulario; podemos intentar hallar un mínimo de conceptos básicos tales que, una vez adjudicada una notación pre-

cisa a cada uno de ellos, sea posible expresar cualquier otro concepto ulterior que se desee mediante la mera combinación e iteración de nuestras notaciones básicas. Este segundo tipo de economía es poco práctico en un sentido, puesto que la pobreza en elementos idiomáticos básicos tiende necesariamente a ocasionar la dilatación del discurso. Pero es práctico en otro sentido: simplifica grandemente el discurso teórico *acerca del lenguaje*, puesto que minimiza el número de términos y de formas constructivas en que consiste el lenguaje.

Ambos tipos de economía, aunque incompatibles *prima facie*, son valiosos cada uno a su manera. Por eso se ha desarrollado la costumbre de combinar ambos, fijando en la práctica dos lenguajes tales que uno de ellos sea parte del otro. El lenguaje más amplio, aunque redundante en su gramática y en su vocabulario, es económico en cuanto a la longitud de las comunicaciones, mientras que el lenguaje parte, llamado notación primitiva, es económico en su gramática y en su vocabulario. El todo y la parte están relacionados por reglas de traducción gracias a las cuales cada elemento idiomático que no pertenezca a la notación primitiva se pone en ecuación con alguna construcción compleja de dicha notación primitiva. Esas reglas de traducción son las llamadas *definiciones* que aparecen en los sistemas formalizados. Lo mejor es considerarlas no como agregadas a un lenguaje, como apéndices de él, sino como correlaciones entre dos lenguajes, uno de los cuales es parte del otro.

Pero esas correlaciones no son arbitrarias. Se supone que muestran cómo las notaciones primitivas son capaces de cumplir todos los objetivos del lenguaje redundante, excepto su brevedad y su conveniencia. Por eso puede esperarse que, en cada caso, el *definiendum* y su *definiens* estén relacionados entre sí de uno de los tres modos antes indicados. El *definiens* puede ser una fiel paráfrasis del *definiendum* en la notación más reducida (primitiva), recogiendo una sinonimia directa<sup>5</sup> como las de usos preexistentes; o bien el *definiens* puede perfeccionar, en el sentido de la explicación, el anterior uso del *definiendum*; o bien, por último, el *definiendum* puede ser una notación creada *ad hoc* y a la que se asigna significación en ese momento y en ese contexto.

Así pues, tanto en el trabajo formal cuanto en el que no lo es, comprobamos que la definición —excepto en el caso extremo de la introducción explícitamente convencional de nuevas notaciones— se basa en relaciones de sinonimia anteriores. Tras reconocer, por tanto, que la noción de definición no contiene la clave de la sinonimia

<sup>5</sup> Según otro sentido importante de «definición», la relación recogida puede ser la relación, más débil, de mera concordancia en la referencia; cfr. *infra*, p. 191 [obra citada en nota 1]. Pero en el presente contexto será mejor olvidar ese sentido de «definición», que es irrelevante para la cuestión de la sinonimia.

y la analiticidad, volvamos a prestar atención a la sinonimia y dejemos ya la definición.

## 1. INTERCAMBIABILIDAD

Una ocurrencia muy natural y que merece atento examen es la de que la sinonimia de las formas lingüísticas consiste simplemente en su intercambiabilidad en todos los contextos sin que cambie el valor veritativo; intercambiabilidad *salva veritate*, según expresión de Leibniz<sup>6</sup>. Nótese que la sinonimia así concebida no se libera necesariamente de vaguedad, al menos en la medida en que es posible hacer compatibles vaguedades.

Pero no es completamente verdad que los sinónimos «soltero» y «hombre no casado» sean intercambiables en todo caso *salva veritate*. Es fácil construir verdades que resultan falsedades al sustituir «soltero» por «hombre no casado»; por ejemplo, con ayuda de comillas:

«soltero» tiene menos de diez letras.

Pero tales contraejemplos pueden probablemente darse de lado tratando el entrecomillado «soltero» como una palabra simple e indivisible (comillas incluidas), y estipulando que la intercambiabilidad *salva veritate* que debe ser piedra de toque de la sinonimia no se presume aplicable a instancias fragmentarias en el interior de una palabra. Esta explicación de la sinonimia, aún admitiendo que sea aceptable en todo lo demás, tiene el inconveniente de apelar a una previa concepción de «palabra» que puede a su vez, con toda probabilidad, presentar dificultades de formulación. No obstante, puede argüirse que se ha hecho algún progreso al reducir el problema de la sinonimia al problema de la naturaleza de las palabras. Sigamos pues un poco esta línea, considerando resuelto el problema «palabra».

Sigue en pie la cuestión de si la intercambiabilidad *salva veritate* (aparte de instancias en el interior de palabras) es una condición suficiente de sinonimia o si, por el contrario, hay expresiones heterónimas que pueden ser intercambiables del mismo modo. Tengamos bien claro que lo que nos preocupa aquí no es la sinonimia en el sentido de completa identidad de las asociaciones psicológicas o de la cualidad poética; en este sentido no hay dos expresiones sinónimas. Lo único que nos ocupa es lo que puede llamarse sinonimia *cognitiva*. No puede decirse, naturalmente, qué es esta sinonimia sino una vez

<sup>6</sup> Cfr. Lewis [1], p. 373.

rematado con éxito el presente estudio; pero sabemos algo de ella a causa de la necesidad que se presentó de ella en conexión con la analiticidad en el § 1. El tipo de sinonimia que allí se necesitó consistió meramente en que todo enunciado analítico pudiera convertirse en una verdad lógica sustituyendo sinónimos por sinónimos. Empezando ahora por el final y suponiendo explicada la analiticidad, podríamos explicar la sinonimia cognitiva en los términos siguientes (tomando los del ejemplo ya conocido): decir que «soltero» y «hombre no casado» son cognitivamente sinónimos no es ni más ni menos que decir que el enunciado

(3) Todos y sólo los solteros son hombres no casados

es analítico <sup>7</sup>.

Lo que necesitamos es una explicación de la sinonimia cognitiva que no presuponga la analiticidad, si es que queremos explicar, a la inversa, la analiticidad con ayuda de la sinonimia cognitiva, tal como se emprendió en el § 1. A nuestra consideración se ofrece ahora, precisamente, una tal independiente explicación de la sinonimia cognitiva: la intercambiabilidad *salva veritate* en todas partes excepto en el interior de palabras. La cuestión que se nos plantea —cojamos el cabo del hilo— es la de si esa intercambiabilidad es una condición suficiente de la sinonimia cognitiva. Podemos convencernos pronto de que lo es, mediante ejemplos del tipo siguiente. El enunciado

(4) Necesariamente todos y sólo los solteros son solteros

es evidentemente verdadero, incluso suponiendo que «necesariamente» se construye tan restrictivamente que no sea correctamente aplicable más que a enunciados analíticos. Si «soltero» y «hombre no casado» son intercambiables *salva veritate*, el resultado de poner «hombre no casado» por una de las instancias de «soltero» en (4), a saber,

(5) Necesariamente todos y sólo los solteros son hombres no casados

tiene que ser verdadero como (4). Pero decir que (5) es verdadero es decir que (3) es analítico y, por tanto, que «soltero» y «hombre no casado» son cognitivamente sinónimos.

<sup>7</sup> Esta es sinonimia cognitiva en un sentido primario y amplio. Carnap ([3], pp. 56 ss.) y Lewis ([2], pp. 83 ss.) han indicado como puede obtenerse, una vez que se tiene esta noción, un sentido más estricto de sinonimia cognitiva que es preferible para algunas finalidades. Pero esta especial ramificación en la construcción de conceptos cae fuera de nuestro presente objetivo y no debe confundirse con el tipo amplio de sinonimia cognitiva que aquí nos ocupa.

Veamos qué hay en esa argumentación que le da su aspecto de arte de birlibirloque. La condición de intercambiabilidad *salva veritate* tiene mayor o menor fuerza según la riqueza del lenguaje de que se trate. La anterior argumentación supone que estamos trabajando con un lenguaje lo suficientemente rico como para contener el adverbio «necesariamente» construido de tal modo que da el valor verdad siempre y sólo si se aplica a un enunciado analítico. Pero ¿podemos admitir un lenguaje que contenga ese adverbio? ¿Tiene realmente sentido ese adverbio? Suponer que lo tiene es suponer que hemos conseguido ya un sentido satisfactorio de «analítico». Y entonces, ¿para qué seguimos trabajando tan celosamente?

Nuestra argumentación no era un flagrante círculo vicioso, pero sí algo parecido. Por decirlo metafóricamente, tiene la forma de una curva cerrada en el espacio.

La intercambiabilidad *salva veritate* carece de sentido a menos que se relativice a un lenguaje cuya amplitud esté especificada en algunos importantes aspectos. Supongamos que consideramos un lenguaje que contiene precisamente los siguientes elementos. Hay una reserva indefinidamente grande de predicados monádicos (por ejemplo, « $F$ »; « $Fx$ » significa que  $x$  es un hombre) y poliádicos (por ejemplo, « $G$ »; « $Gxy$ » significa que  $x$  ama a  $y$ ), la mayoría de los cuales se refieren a materias extralógicas. El resto del lenguaje es lógico. Los enunciados atómicos consisten cada uno de ellos en un predicado seguido por una o más variables « $x$ », « $y$ », etc.; y los enunciados compuestos se construyen partiendo de los atómicos mediante funciones veritativas («no», «y», «o», etc.), y la cuantificación<sup>8</sup>. Un tal lenguaje goza de los beneficios de la descripción y, por tanto, de los términos singulares en general, los cuales pueden ser contextualmente definidos del modo visto<sup>9</sup>. También los términos singulares abstractos que denotan clases, clases de clases, etc., son contextualmente definibles con tal de que la reserva de predicados incluya el predicado diádico de pertenencia de individuo a clase<sup>10</sup>. Ese lenguaje puede ser adecuado para la matemática clásica y para el discurso científico en general, excepto en la medida en que este último incluye expedientes discutibles como los condicionales contrafactuales o adverbios modales como «necesariamente»<sup>11</sup>. Un lenguaje de este tipo es extensional en el siguiente sentido: siempre que dos predicados coin-

<sup>8</sup> En pp. 125 ss., *infra* [obra citada en nota 1], se encontrará una descripción de un lenguaje así, con la particularidad de que no contiene más que un predicado, el predicado diádico « $\epsilon$ » [pertenencia de miembro a clase]. (*N. del T.*)

<sup>9</sup> Cfr. *supra*, pp. 31-34; *infra*, pp. 130 ss., 237 s. [obra citada en nota 1].

<sup>10</sup> Cfr. *infra*, p. 134. [obra citada en nota 1].

<sup>11</sup> Sobre tales expedientes cfr. también el ensayo VIII. [obra citada en nota 1].

ciden extensionalmente (esto es, son verdaderos de los mismos objetos) son intercambiables *salva veritate*<sup>12</sup>.

En un lenguaje extensional, por tanto, la intercambiabilidad *salva veritate* no garantiza una sinonimia cognitiva del tipo deseado. Que «soltero» y «hombre no casado» son intercambiables en un lenguaje extensional *salva veritate* no nos garantiza absolutamente nada más que la verdad de (3). No hay ninguna seguridad de que la coincidencia extensional de «soltero» y «hombre no casado» descansa en la significación y no en circunstancias fácticas accidentales, como ocurre con la coincidencia extensional de «criatura con corazón» y «criatura con riñones».

Para muchos propósitos la coincidencia extensional es la mejor aproximación a la sinonimia que podemos conseguir. Pero sigue en pie el hecho de que la coincidencia extensional queda lejos de la sinonimia cognitiva del tipo requerido para explicar la analiticidad del modo emprendido en el § 1. El tipo de sinonimia cognitiva que se necesita tiene que ser tal que permita sentar la equivalencia de la sinonimia de «soltero» y «hombre no casado» con la analiticidad de (3) y no simplemente con la verdad de (3).

Tenemos pues que reconocer que la intercambiabilidad *salva veritate* construida en relación con un lenguaje extensional no es condición suficiente de la sinonimia cognitiva en el sentido requerido para derivar de ella la analiticidad a la manera del § 1. Si el lenguaje contiene un adverbio intensional, el adverbio «necesariamente», en el sentido antes indicado, u otras partículas que tengan el mismo efecto, la intercambiabilidad *salva veritate* será en ese lenguaje una condición suficiente de la sinonimia cognitiva; pero ocurre que un tal lenguaje no es inteligible más que si la noción de analiticidad se entiende ya por anticipado.

Es posible que el esfuerzo dirigido a explicar primero la sinonimia cognitiva para derivar luego de ella la analiticidad, como se apuntó en el § 1, yerre su dirección. En lugar de esforzarnos según esa línea podríamos intentar explicar la analiticidad de algún modo que no apele a la sinonimia cognitiva. Luego podríamos sin duda derivar la sinonimia cognitiva de la analiticidad de un modo plenamente satisfactorio. Hemos visto que la sinonimia cognitiva de «soltero» y «hombre no casado» puede explicarse como analiticidad de (3). La misma explicación sirve para todo par de predicados monádicos, como es natural, y puede generalizarse de modo obvio a los predicados poliádicos. También pueden incluirse en la explicación, de un modo paralelo, otras categorías sintácticas. Por lo que hace a los términos singulares puede decirse que son cognitivamente sinónimos

<sup>12</sup> Ésa es la sustancia de Quine, \* 121.

cuando el enunciado de identidad formado escribiendo «=» entre aquellos términos singulares es analítico. Por lo que hace a los enunciados, puede decirse simplemente que son cognitivamente sinónimos cuando su bicondicional (el resultado de unirlos mediante la conectiva «si y sólo si») es un enunciado analítico<sup>13</sup>. Si queremos reunir todas esas categorías sintácticas en una sola formulación, podemos hacerlo —al precio de volver a cargar con la noción de «palabra», a la que ya antes se apeló en esta sección— describiendo como cognitivamente sinónimo cualquier par de formas lingüísticas que sean intercambiables (aparte de instancias en el interior de palabras) *salva analyticitate* (y no ya *veritate* sólo). Surgen entonces ciertos problemas técnicos sobre casos de ambigüedad o de homonimia; pero no nos detendremos ahora en ellos, ya que aún nos encontramos en nuestra larga digresión. Abandonemos más bien el problema de la sinonimia y volvamos de nuevo al de la analiticidad.

#### 4. REGLAS SEMÁNTICAS

Pareció al principio que la manera más natural de definir la analiticidad consistía en apelar a un reino de significaciones. Afinando esa solución, la apelación a significaciones dio lugar a la apelación a la sinonimia o a la definición. Pero la definición mostró ser un fuego fatuo, y en cuanto a la sinonimia, resultó que ésta no puede entenderse correctamente sino mediante una previa apelación a la analiticidad misma. Y así volvemos al problema de la analiticidad.

No sé si el enunciado «Toda cosa verde es extensa» es analítico. Traiciona mi indecisión ante ese ejemplo una comprensión incompleta, una incompleta captación de las significaciones de «verde» y «extensa»? Yo creo que no. La dificultad no está en «verde» ni en «extensa», sino en «analítico».

Se dice a menudo que la dificultad de distinguir entre enunciados analíticos y enunciados sintéticos en el lenguaje ordinario se debe a la vaguedad de éste, y que la distinción es clara cuando se trata de un preciso lenguaje artificial con «reglas semánticas» precisas. Voy a intentar mostrar que eso es una confusión.

La noción de analiticidad en torno de la cual nos movemos es una relación entre enunciados y lenguajes: de un enunciado *E* se dice que es *analítico para* un lenguaje (o en un lenguaje) *L*, y el problema consiste en conseguir un sentido general de esa relación, es decir, para «*E*» y «*L*» como variables. La gravedad del problema no es menos

<sup>13</sup> Entendiendo «si y sólo si» en el sentido veritativo-funcional. Cfr. Carnap [1], p. 14.



perceptible en lenguajes artificiales que en lenguajes naturales. El problema de dar sentido a la frase «*E* es analítico para *L*», con «*E*» y «*L*» variables, sigue siendo correoso aunque limitemos el campo de la variable «*L*» a lenguajes artificiales. Intentaré ahora poner esto de manifiesto.

En materia de lenguajes artificiales y de reglas semánticas es natural dirigirse a los escritos de Carnap. Sus reglas semánticas toman varias formas, y para precisar mi tarea tendré que distinguir algunas de esas formas. Supongamos, para empezar, un lenguaje artificial  $L_0$  cuyas reglas semánticas tengan explícitamente la forma de una especificación —recursiva o de otro tipo— de todos los enunciados analíticos de  $L_0$ . Las reglas nos dicen que tales y cuales enunciados, y sólo ellos, son los enunciados analíticos de  $L_0$ . La única dificultad en este caso es que las reglas contienen la palabra «analítico»... que es la palabra que no comprendemos. Comprendemos cuáles son las expresiones a las que las reglas atribuyen analiticidad, pero no comprendemos qué es en realidad lo que las reglas les atribuyen. Dicho brevemente: para que podamos entender una regla que empieza diciendo «Un enunciado *E* es analítico para el lenguaje  $L_0$  si y sólo si...», tenemos que entender antes el término general relativo «analítico para»; tenemos que entender «*E* es analítico para *L*» siendo «*E*» y «*L*» variables.

Podemos naturalmente también considerar la llamada regla como una definición convencional de un nuevo símbolo simple, el símbolo «analítico para  $L_0$ », que valdrá más escribir, sin tendencia psicológica, «*K*» por ejemplo, para que no parezca indebidamente que arroja luz sobre la palabra que nos interesa, «analítico». Cualquiera número de clases, *K*, *M*, *N*, de enunciados de  $L_0$  puede especificarse en atención a diversas necesidades o sin ninguna finalidad. ¿qué significa entonces decir que *K*, a diferencia de *M*, *N*, etc., es la clase de los enunciados «analíticos» de  $L_0$ ?

Enumerando los enunciados que son analíticos para  $L_0$  explicamos «analítico para  $L_0$ », pero no «analítico» ni «analítico para». No explicamos la frase «*E* es analítico para *L*» con «*E*» y «*L*» variables, ni siquiera limitando el campo de «*L*» a los lenguajes artificiales.

En realidad, conocemos lo suficiente de la significación buscada de «analítico» como para saber que los enunciados analíticos se suponen verdaderos. Atendamos por ello ahora a otra forma de regla semántica, la cual no dice que tales o cuales enunciados son analíticos, sino, simplemente, que tales o cuales enunciados se incluyen entre los verdaderos. Una regla de este tipo no está sujeta a la crítica por contener la palabra «analítico», cuya comprensión se busca; por motivo de la argumentación podemos suponer que no hay dificultades a propósito del término, más amplio, «verdadero». No se supone que una regla semántica de este segundo tipo, una regla veritativa, espe-

«Dique todas las verdades de su lenguaje; sólo precisa —recursivamente o de otro modo— un cierto número de enunciados que deben considerarse verdaderos junto con otros que no especifica. Puede concederse que una tal regla es suficientemente clara. Sobre ella puede luego precisarse derivativamente la analiticidad del modo siguiente: un enunciado es analítico si es verdadero por la regla semántica (no simplemente verdadero).

Pero con ello no se ha conseguido ningún progreso real. En vez de apelar a la inexplicada palabra «analítico» estamos apelando ahora a la inexplicada frase «regla semántica». No todo enunciado verdadero que dice que los enunciados de una clase determinada son verdaderos puede tomarse como una regla semántica, pues entonces todas las verdades serían «analíticas» en el sentido de ser verdaderas por virtud de reglas semánticas. Todo parece indicar que la única característica de las reglas semánticas consiste en figurar en una página encabezada por el rótulo «Reglas Semánticas», y este rótulo carece por su parte de significación.

Podemos pues decir que un enunciado es *analítico para*  $L_0$  si y sólo si es verdadero según tales y cuales «reglas semánticas» precisamente especificadas, pero con ello volvemos a encontrarnos esencialmente en el mismo caso inicialmente discutido: « $E$  es analítico para  $L_n$  si y sólo si...». Y puesto que lo que queremos explicar es « $E$  es analítico para  $L$ » en términos generales para la variable « $L$ » (aunque admitiendo su limitación a los lenguajes artificiales), la explicación «verdadero según las reglas semánticas de  $L$ » es estéril, pues el término relativo «regla semántica de» necesita por lo menos tanta aclaración como «analítico para».

Puede ser instructivo comparar la noción de regla semántica con la de postulado. Dado un conjunto de postulados, es muy fácil decir qué es un postulado: es un miembro del conjunto dado. Y dado un conjunto de reglas semánticas, es también muy fácil decir qué es una regla semántica. Pero dada simplemente una notación matemática o de otro tipo, entendida como se quiera en cuanto a la traducción o a las condiciones veritativas de sus enunciados, ¿quién puede decir cuáles de sus enunciados verdaderos tienen el rango de postulados? Es obvio que la cuestión carece de sentido; tanto como la pregunta que inquiriera qué lugares de Ohio son puntos de partida. Toda selección finita (o infinita, pero efectivamente especificable) de enunciados (quizá preferiblemente verdaderos) es un conjunto de postulados con el mismo derecho que cualquier otra selección. La palabra «postulado» es significante sólo si es relativa a un concreto acto de investigación; aplicamos la palabra a un conjunto de enunciados en la medida en que al mismo tiempo pensamos en esos enunciados en relación con otros que pueden obtenerse de ellos mediante un conjunto de transformaciones al que hemos tenido que prestar nues-

tra atención. La noción de regla semántica es tan concreta y significativa como la de postulado si se concibe con el mismo espíritu relativo—relativo, en este caso, a la tarea de informar a alguna persona acerca de las condiciones suficientes de la verdad de enunciados en un determinado lenguaje, natural o artificial, *L*—. Pero desde este punto de vista ninguna indicación de una subclase de verdades de *L* es por derecho propio más regla semántica que otra, y si «analítico» significa «verdadero por reglas semánticas», ninguna verdad de *L* es más analítica que otra <sup>14</sup>.

Podría pensarse en argüir que un lenguaje artificial *L* (a diferencia de un lenguaje natural) es un lenguaje en el sentido ordinario de esa palabra *más* un conjunto de reglas semánticas explícitas—constituyendo el conjunto, digamos, un par ordenado—; y que las reglas semánticas de *L* son entonces simplemente especificables como el segundo elemento del par *L*. Pero, con el mismo resultado y menos dificultad, podemos construir un lenguaje artificial *L* como un par ordenado cuyo segundo elemento es la clase de sus enunciados analíticos; en este caso los enunciados analíticos de *L* son especificables sencillamente como los enunciados que componen el segundo elemento de *L*. O, mejor aún, podemos dejar de una vez de intentar levantarnos tirándonos de nuestra propias orejas.

No todas las explicaciones de la analiticidad conocidas por Carnap y sus lectores han sido explícitamente cubiertas por las anteriores consideraciones; pero no es difícil ver cómo pueden éstas ampliarse a las demás formas. Sólo habría que mencionar aún un factor adicional que interviene algunas veces: en ocasiones las reglas semánticas son en realidad reglas de traducción al lenguaje ordinario, caso en el cual los enunciados analíticos del lenguaje artificial se reconocen efectivamente por la analiticidad de sus especificadas traducciones al lenguaje ordinario. Realmente, en este caso no podrá decirse que el problema de la analiticidad quede eliminado por el lenguaje artificial.

Desde el punto de vista del problema de la analiticidad, la noción de lenguaje artificial con reglas semánticas es un *feu follet par excellence*. Las reglas semánticas como determinantes de los enunciados analíticos de un lenguaje artificial no tienen interés más que si hemos entendido ya la noción de analiticidad; pero no prestan ninguna ayuda en la consecución de esa comprensión.

La apelación a lenguajes hipotéticos de un tipo artificialmente sencillo podría probablemente ser útil para la aclaración de la analiticidad, siempre que el modelo simplificado incluyera algún esquema de los factores mentales, comportamentísticos o culturales relevantes.

<sup>14</sup> El párrafo anterior no figuraba en la primera edición de este ensayo. Ha sido inspirado por Martin (v. bibliografía), igual que el final del ensayo VII. [obra citada en nota 1].

tos para la analiticidad, cualesquiera que ellos sean. Pero es poco verosímil que un modelo que toma la analiticidad como un carácter irreductible pueda arrojar luz a la hora de intentar explicar la analiticidad.

Es obvio que la verdad en sentido general depende a la vez del lenguaje y del hecho extralingüístico. El enunciado «Bruto mató a César» sería falso si el mundo hubiera sido diverso en algunos aspectos de lo que ha sido, y también lo sería si resultara que la palabra «mató» tuviera el sentido de «procreó». Por eso se presenta la tentación de suponer que la verdad de un enunciado es algo analizable en una componente lingüística y una componente fáctica. Dada esa suposición, parece a continuación razonable que en algunos enunciados la componente fáctica se considere nula; y estos son los enunciados analíticos. Pero por razonable que sea todo eso *a priori*, sigue sin trazarse una línea separatoria entre enunciados analíticos y enunciados sintéticos. La convicción de que esa línea debe ser trazada es un dogma nada empírico de los empiristas, un metafísico artículo de fe.

### 3 LA TEORÍA DE LA VERIFICACIÓN Y EL REDUCTIVISMO

En el curso de estas sombrías reflexiones hemos conseguido una visión bastante oscura de la noción de significación primero, luego de la sinonimia cognitiva y, finalmente, de la de analiticidad. ¿Y la teoría de la verificación, qué es una teoría de la significación?, se preguntará. Esa frase —teoría de la verificación— se ha establecido tan firmemente como marca de fábrica del empirismo que habría sido muy poco científico no buscar antes por otros lados una posible clave del problema de la significación y demás problemas asociados con él.

La teoría de la verificación, tan destacada en la literatura a partir de Peirce, sostiene que el sentido o significación de un enunciado es el método de confirmación o confutación empírica del mismo. Un enunciado analítico es aquel caso límite que queda confirmado en cualquier supuesto.

Como se dijo en el § 1, podemos perfectamente obviar la cuestión de las significaciones como entidades y dirigirnos directamente hacia la de la identidad de significación, o sinonimia. Pues lo que la teoría de la verificación dice es que unos enunciados son sinónimos si y sólo si coinciden en cuanto al método de confirmación o invalidación empírica.

Es ésta una explicación de la sinonimia cognitiva de enunciados, y no de formas lingüísticas en general <sup>15</sup>. No obstante, partiendo del

<sup>15</sup> Pero la doctrina puede formularse con términos —en vez de enunciados— como

concepto de sinonimia de enunciados podemos derivar el concepto para otras formas lingüísticas mediante consideraciones bastante parecidas a las hechas al final del § 3. Presuponiendo la noción de «palabra», podemos en efecto explicar la sinonimia de dos formas cualquiera por el hecho de que la sustitución de una instancia de una forma en cualquier enunciado (aparte de instancias en el interior de una «palabra») por la otra forma produce un enunciado sinónimo. Por último, dado así el concepto de sinonimia para formas lingüísticas en general, podemos definir la analiticidad en términos de sinonimia y verdad lógica como en el § 1. En realidad, podemos definir la analiticidad más simplemente en términos de mera sinonimia de enunciados más verdad lógica; no es necesario apelar a la sinonimia de formas lingüísticas diversas de los enunciados. Pues un enunciado puede describirse como analítico con tal de que sea sinónimo de un enunciado lógicamente verdadero.

Así pues, si la teoría de la verificación puede aceptarse como explicación adecuada de la sinonimia de enunciados, la noción de analiticidad se salva en última instancia. Pensemos, de todos modos. La teoría dice que la sinonimia de enunciados es la igualdad de método de confirmación o invalidación empírica. Pero, ¿qué son esos métodos que hay que comparar para establecer su igualdad? Dicho de otro modo: ¿cuál es la naturaleza de la relación entre un enunciado y las experiencias que contribuyen a su confirmación o la impiden?

La concepción más ingenua de esta relación consiste en suponer que se trata de una referencialidad directa. Tal es el reductivismo radical, que sostiene que todo enunciado con sentido es traducible a un enunciado (verdadero o falso) acerca de experiencia inmediata. En una forma u otra, el reductivismo radical precede a la teoría de la verificación propiamente dicha. Así, por ejemplo, Locke y Hume sostenían que toda noción se origina directamente en la experiencia sensible, o bien es un compuesto de nociones así originadas. Recogiendo una indicación de Tooke, podemos reformular esta doctrina en la jerga técnica semántica diciendo que para ser significante un término tiene que ser el nombre de un dato sensible, o bien un compuesto de tales nombres o una abreviatura de un compuesto de esa naturaleza. Así formulada, la doctrina sigue siendo ambigua porque se refiere a la vez a datos sensibles como acaecimientos sensoriales y datos sensibles como cualidades sensibles; y es además vaga en cuanto a los

unidades. Así, Lewis define la significación de un término como «un criterio mental por referencia al cual somos capaces de aplicar, o negarnos a aplicar, la expresión en cuestión en el caso de cosas o situaciones presentes o imaginadas» ([2], p. 133).

—Para una instructiva exposición de las vicisitudes de la teoría de la verificación, centrada en la noción de significatividad y no en las de sinonimia y analiticidad, véase Hempel; pp. 199-219 de este volumen.

...admisibles de composición (de nombres de datos sensibles).  
Aun más: la doctrina es innecesaria e inadmisiblemente restrictiva por la casuística crítica de términos que impone. Más razonablemente — aunque sin rebasar los límites de lo que he llamado reduccionismo radical — podemos tomar como unidades significantes enunciados completos, y exigir que nuestros enunciados sean traducibles como totalidades al lenguaje de los datos sensibles, y no que lo sean término por término.

Esta corrección habría sido sin duda bien recibida por Locke, Mure y Tooke, pero históricamente no se produjo hasta el momento de la importante reorientación de la semántica por la cual se pasó a ver el vehículo primario de la significación en el enunciado y no en el término. Esta reorientación, ya explícita en Frege ([1], § 60), está en la base del concepto russelliano de símbolo incompleto definido por el uso<sup>16</sup>; también está implícita en la teoría de la significación que consideramos, la teoría de la verificación, puesto que los objetos de la verificación son enunciados.

El reduccionismo radical, concebido con los enunciados como unidades, se pone la tarea de especificar un lenguaje de los datos sensibles y de mostrar la forma de traducir a él, enunciado por enunciado, el resto del discurso significativo. En esta empresa se embarcó Carnap en *Der logische Aufbau der Welt*.

El lenguaje que Carnap adoptó como punto de partida no era un lenguaje de datos sensibles, en el sentido más estricto imaginable, pues incluía también notaciones lógicas hasta el nivel de la teoría de conjuntos superior. Incluía, en efecto, todo el lenguaje de la matemática pura. La ontología implícita en ese lenguaje — es decir, el campo de valores de sus variables — abrazaba no sólo acaecimientos sensoriales, sino también clases de clases, etc. Hay empiristas que se asustarían ante tal prodigalidad. En cambio, el punto de partida de Carnap es muy económico en su parte extralógica o sensorial. En una serie de construcciones en las que aprovecha con mucho ingenio los recursos de la lógica moderna, Carnap consigue definir una amplia colección de importantes conceptos adicionales de tipo sensorial que, a no ser por sus construcciones, nadie habría imaginado definibles sobre tan estrecha base. Carnap fue el primer empirista que, no contento con afirmar la reducibilidad de la ciencia a términos de experiencia inmediata, dio serios pasos hacia la realización de esa reducción.

Si el punto de partida de Carnap es satisfactorio, sus construcciones no eran en cambio, como él mismo subrayaba, más que un fragmento del programa entero. Incluso la construcción de los enuncia-

<sup>16</sup> Cfr. *supra*, p. 31 [obra citada en nota 1].

dos más sencillos acerca del mundo físico quedaba en un estado esquemático o de esbozo. A pesar de su carácter esquemático, las sugerencias de Carnap en este terreno eran realmente sugestivas—sugestivas—. Explicaba los puntos-instantes espacio-temporales como conjuntos de cuatro números reales, y estudiaba la asignación de cualidades sensibles a los puntos-instantes según ciertos cánones. Sumariamente resumido, el plan consistía en asignar cualidades a los puntos-instantes de tal modo que se consiguiera el mundo más perezo compatible con nuestra experiencia. El principio de acción mínima debía ser nuestra guía en la construcción de un mundo a partir de la experiencia.

Pero Carnap no parece haber visto que su tratamiento de los objetos físicos no alcanzaba la reducción no sólo por su carácter esquemático, sino por principio. Según sus cánones, había que atribuir valores veritativos a enunciados de la forma «La cualidad  $c$  se encuentra en el punto-instante  $x; y; z; t'$ », maximizando y minimizando ciertos rasgos generales, y con el enriquecimiento de la experiencia había que revisar progresivamente los valores veritativos dentro de esa misma línea. Creo que esto es una buena esquematización (sin duda deliberadamente simplificada) de lo que realmente hace la ciencia; pero no da la menor indicación, ni siquiera la más esquemática, sobre cómo podría traducirse al inicial lenguaje de datos sensibles y lógica un enunciado de la forma «La cualidad  $c$  se encuentra en  $x; y; z; t'$ ». La conectiva «se encuentra en» es una conectiva añadida y no definida; los cánones nos guían en su uso, pero no en su eliminación.

Carnap parece haber apreciado este problema más tarde, pues en sus posteriores escritos ha abandonado la noción de traducibilidad de los enunciados sobre el mundo físico a enunciados acerca de la experiencia inmediata. El reductivismo en su forma radical ha dejado de figurar en la filosofía de Carnap hace ya mucho tiempo.

Pero el dogma reductivista ha seguido influyendo en el pensamiento de los empiristas en una forma sutil y más tenue. Persiste la opinión de que con cada enunciado, o con todo enunciado sintético, está asociado un único campo posible de acaecimientos sensoriales, de tal modo que la ocurrencia de uno de ellos añade probabilidad a la verdad del enunciado, y también otro campo único de posibles acaecimientos sensoriales cuya ocurrencia eliminaría aquella probabilidad. Esta noción está sin duda implícita en la teoría de la verificación.

El dogma reductivista sobrevive en la suposición de que todo enunciado, aislado de sus compañeros, puede tener confirmación o invalidación. Frente a esta opinión, la mía, que procede esencialmente de la doctrina carnapiana del mundo físico en el *Aufbau*, es que nuestros enunciados acerca del mundo externo se someten como

cuerpo total al tribunal de la experiencia sensible, y no individualmente.<sup>17</sup>

Incluso en su forma atenuada, el dogma reductivista está en íntima conexión con el otro dogma, a saber, que hay una distinción entre lo analítico y lo sintético. Nosotros mismos nos hemos visto llevados de un problema a otro a través de la doctrina de la significación ofrecida por la teoría de la verificación. Aun más directamente, el primer dogma sostiene al segundo del modo siguiente: mientras se considere significativo en general hablar de la confirmación o la invalidación de un enunciado, parece también significativo hablar de un tipo límite de enunciados que resultan confirmados vacuamente, *ipso facto*, ocurra lo que ocurra; esos enunciados son analíticos.

Los dos dogmas son en efecto idénticos en sus raíces. Antes dijimos que en general la verdad de los enunciados depende obviamente del lenguaje y del hecho extralingüístico; y ya entonces notamos que esa circunstancia obvia lleva consigo, no por inferencia lógica, pero sí muy naturalmente, la sensación de que la verdad de un enunciado es algo analizable en una componente lingüística y otra factual. Desde un punto de vista empirista, la componente factual debe reducirse a un campo de experiencias confirmativas. En el caso extremo de que lo único que importe sea la componente lingüística, el enunciado es analítico. Pero creo que ahora estaremos bastante impresionados por la tenacidad con que la distinción entre analítico y sintético ha resistido a toda precisión. Personalmente me impresiona también lo confuso que ha sido siempre el problema de llegar a cualquier teoría explícita de la confirmación empírica de un enunciado sintético —dejando aparte los prefabricados ejemplos de las bolas blancas y negras en la urna—. Quiero sugerir en este momento que hablar de una componente lingüística y una componente factual en la verdad de cualquier enunciado particular es un sinsentido que da lugar a muchos otros sinsentidos. Tomada en su conjunto, la ciencia presenta esa doble dependencia respecto del lenguaje y respecto de los hechos; pero esta dualidad no puede perseguirse significativamente hasta los enunciados de la ciencia tomados uno por uno.

Como ya hemos observado, la idea de definir un símbolo por el uso fue un progreso respecto del imposible empirismo de los términos individuales propios de Locke y Hume. Con Frege, el enunciado llegó a ser reconocido, en vez del término, como la unidad relevante para una crítica empirista. Lo que ahora afirmo es que nuestra red sigue siendo de mallas demasiado estrechas incluso cuando tomamos el enunciado entero como unidad. La unidad de significación empírica es el todo de la ciencia.

<sup>17</sup> Esta doctrina fue bien argüida por Duhem, pp. 303-328. —Ver también Lowin-son, pp. 132-140.



## 6. EMPIRISMO SIN DOGMAS

La totalidad de lo que llamamos nuestro conocimiento, o creencias, desde las más casuales cuestiones de la geografía y la historia hasta las más profundas leyes de la física atómica o incluso de la matemática o de la lógica puras, es una fábrica construida por el hombre y que no está en contacto con la experiencia más que a lo largo de sus lados. O, con otro símil, el todo de la ciencia es como un campo de fuerzas cuyas condiciones-límite da la experiencia. Un conflicto con la experiencia en la periferia da lugar a reajustes en el interior del campo: hay que redistribuir los valores veritativos entre algunos de nuestros enunciados. La nueva atribución de valores a algunos enunciados implica la re-valoración de otros en razón de sus interconexiones lógicas —y las leyes lógicas son simplemente unos determinados enunciados del sistema, determinados elementos del campo. Una vez redistribuidos valores entre algunos enunciados, hay que redistribuir también los de otros que pueden ser enunciados lógicamente conectados con los primeros o incluso enunciados de conexiones lógicas. Pues el campo total está tan escasamente determinado por sus condiciones-límite —por la experiencia— que hay mucho margen de elección en cuanto a los enunciados que deben recibir valores nuevos a la luz de cada experiencia contraria al anterior estado del sistema. Ninguna experiencia concreta y particular está ligada directamente con un enunciado concreto y particular en el interior del campo, sino que esos ligámenes son indirectos, se establecen a través de consideraciones de equilibrio que afectan al campo como un todo.

Si esta visión es correcta, será entonces erróneo hablar del contenido empírico de un determinado enunciado —especialmente si se trata de un enunciado situado lejos de la periferia del campo—. Además, resulta entonces absurdo buscar una divisoria entre enunciados sintéticos, que valen contingentemente y por experiencia, y enunciados analíticos que valen en cualquier caso. Todo enunciado puede concebirse como valedero en cualquier caso siempre que hagamos reajustes suficientemente drásticos en otras zonas del sistema. Incluso un enunciado situado muy cerca de la periferia puede sostenerse contra una experiencia recalcitrante apelando a la posibilidad de estar sufriendo alucinaciones, o reajustando enunciados de las llamadas leyes lógicas. A la inversa, y por la misma razón, no hay enunciado alguno inmune a la revisión. Hasta una revisión de la ley lógica de tercio excluso se ha propuesto como un expediente para simplificar la mecánica cuántica; ¿y qué diferencia hay en principio entre un cambio así y el cambio por el que Kepler sustituyó a Ptolomeo, o Einstein a Newton, o Darwin a Aristóteles?

Por motivos de plasticidad he estado hablando de distancias respecto de una periferia sensible. Aclaremos ahora esta noción sin me-

otras. Algunos enunciados, aunque *se refieren* a objetos físicos y no a experiencia sensible, parecen hermanarse característicamente con la experiencia sensible —y, además, de un modo selectivo: esto es, tales enunciados con tales experiencias, tales otros con tales otras, etc.—. En nuestra metáfora, los enunciados que están especialmente hermanados con experiencias determinadas se describen como próximos a la periferia. Pero en esa relación de «hermandad» no veo nada que una laxa asociación que refleja la relativa probabilidad de que en la práctica escojamos un enunciado en vez de otro para someterlo a revisión caso de presentarse una experiencia negativa. Podemos, por ejemplo, imaginar experiencias negativas para acomodar a las cuales nuestro sistema nos inclinaríamos sin duda a cambiar los valores anteriormente atribuidos a un enunciado como el de que hay casas de adobe en el Paseo de Gracia \*, junto con otros asociados y relativos a ese mismo tema. Podemos imaginar otras experiencias negativas para acomodar a las cuales nuestro sistema nos inclinaríamos a dar un nuevo valor al enunciado de que no hay centauros y a otros emparentados con él. Según he dicho, una experiencia imprevista puede acomodarse en el sistema mediante una de varias nuevas valoraciones posibles en otros tantos sectores del sistema; pero en los casos que hemos imaginado, nuestra natural tendencia a perturbar lo menos posible el sistema en su conjunto nos lleva a centrar la revisión en esos específicos enunciados relativos a casas de adobe o a centauros. Por eso se tiene la sensación de que esos enunciados tienen una referencia empírica más precisa que los muy teóricos enunciados de la física, de la lógica o de la ontología. Puede considerarse que éstos están situados en una zona relativamente central de la red, lo que significa meramente que presentan poca conexión preferencial con algún dato sensible determinado.

Como empirista, sigo concibiendo el esquema conceptual de la ciencia como un instrumento destinado en última instancia a predecir experiencia futura a la luz de la experiencia pasada. Introducimos con razón conceptualmente los objetos físicos en esta situación porque son intermediarios convenientes, no por definición en términos de experiencia, sino irreductiblemente puestos con un estatuto epistemológico comparable al de los dioses de Homero<sup>18</sup>. Yo por mi parte, como físico lego que soy, creo en los objetos físicos y no creo en los dioses de Homero, y considero un error científico orientar su creencia de otro modo. Pero en cuanto a fundamento epistemológico, los objetos físicos y los dioses difieren sólo en grado, no en esencia. Ambas suertes de entidades integran nuestras concepciones sólo como elementos de cultura. El mito de los objetos físicos es episte-

\* Texto original: «...that there are brick houses on Elm Street...» (*N. del T.*)

<sup>18</sup> Cfr. *supra*, p. 44 [obra citada en nota 1].

mológicamente superior a muchos otros mitos porque ha probado ser más eficaz que ellos como procedimiento para elaborar una estructura manejable en el flujo de la experiencia.

Esa actitud que pone objetos físicos no se reduce al nivel macroscópico. También al nivel atómico se ponen objetos para que las leyes de los objetos macroscópicos —y, en última instancia, las leyes de la experiencia— sean más simples y manejables; y no debemos esperar ni pedir una plena definición de las entidades atómicas y subatómicas en términos de entidades macroscópicas, ni tampoco una definición de las cosas macroscópicas en términos de datos sensibles. La ciencia es una prolongación del sentido común que consiste en hinchar la ontología para simplificar la teoría.

Los objetos físicos, los grandes y los pequeños, no son las únicas entidades puestas. Otro ejemplo son las fuerzas; y efectivamente hoy nos dicen que la separación entre materia y energía está anticuada. Las entidades abstractas que son la sustancia de las matemáticas —en última instancia clases y clases de clases y así sucesivamente— son también entidades puestas en el mismo sentido. Epistemológicamente, todos esos son mitos con la misma base que los objetos físicos y los dioses, y por lo único que unos son mejores que otros es por el grado en que favorecen nuestro manejo de la experiencia sensible.

La extensa álgebra de los números racionales e irracionales está subdeterminada por el álgebra de los números racionales, pero es más cómoda y conveniente que ella, y la incluye como parte coja o manca<sup>19</sup>. La ciencia total —matemática, natural y humana— está análogamente subdeterminada por la experiencia, de un modo aun más extremado. El contorno del sistema tiene que cuadrar con la experiencia; el resto, con todos sus elaborados mitos y sus ficciones tiene como objetivo la simplicidad de las leyes.

Desde este punto de vista, las cuestiones ontológicas van a la par con las científico-naturales<sup>20</sup>. Considérese la cuestión de si deben admitirse las clases como entidades. Se trata, como he indicado en otros lugares<sup>21</sup>, de la cuestión de si deben cuantificarse variables que toman clases como valores. Carnap [6] ha sostenido que ésta no es una cuestión factual, sino de elección de la forma lingüística conveniente, del esquema o estructura conceptual conveniente para la ciencia. Puedo estar de acuerdo con esa opinión, siempre que se conceda lo mismo respecto de todas las hipótesis científicas en general. Carnap ([6], p. 32 n.) ha reconocido que sólo puede sostener una diversidad de criterios para las cuestiones ontológicas por un lado y

<sup>19</sup> Cfr. *supra*, p. 43 [obra citada en nota 1].

<sup>20</sup> «L'ontologie fait corps avec la science elle-même et ne peut en être séparée» Meyerson, p. 439.

<sup>21</sup> Cfr. *supra*, pp. 39 s.; *infra*, pp. 153 ss. [obra citada en nota 1].

para las hipótesis científicas por otro asumiendo una distinción absoluta entre lo analítico y lo sintético; y no es necesario repetir que ésta es una distinción que ya he rechazado <sup>22</sup>.

La cuestión de si hay o no hay clases parece más bien una cuestión relativa al esquema conceptual conveniente. Y la cuestión de si hay clases de adobe en el Paseo de Gracia o la de si hay centauros parecen más bien cuestiones de hecho. Pero he indicado que esta diferencia es sólo de grado y se basa en nuestra vaga inclinación pragmática a reajustar un determinado ramal de la red de la ciencia, en vez de otros u otros, cuando intentamos acomodar en ella alguna experiencia negativa inesperada. En esas decisiones desempeñan algún papel el conservadurismo y la búsqueda de la simplicidad.

Carnap, Lewis y otros adoptan una actitud pragmática en la elección entre formas lingüísticas o estructuras científicas; pero su pragmatismo se detiene ante la imaginaria frontera entre lo analítico y lo sintético. Al repudiar esa frontera expongo un pragmatismo más completo: Todo hombre recibe una herencia científica más un continuo y graneado fuego de estímulos sensoriales; y las consideraciones que le mueven a moldear su herencia científica para que recoja sus continuos estímulos sensoriales son, si racionales, pragmáticas.

#### BIBLIOGRAFÍA

- CARNAP, R. [1]: *Meaning and Necessity*, University of Chicago Press, Chicago, 1947.  
 CARNAP, R. [2]: *Logical Foundations of Probability*, Chicago University Press, Chicago, 1950.  
 DUHEM, P.: *La Théorie Physique: son object et sa structure*, París, 1906.  
 LEWIS, C. I. [1]: *A Survey of Symbolic Logic*, Berkeley, 1918.  
 LEWIS, C. I. [2]: *An Analysis of Knowledge and Valuation*, Open Court, Lasalle, Ill., 1946.  
 LOWINGER, A.: *The Methodology of Pierre Duhem*, Columbia University Press, Nueva York, 1941.  
 MARTIN, R. M.: «On "analytic"», *Philosophical Studies*, 3, 1952, pp. 42-47.  
 MEYERSON, E.: *Identité et réalité*, París, 1908, 4.ª ed., 1932.  
 QUINE, W. V. O.: *Mathematical Logic*, Norton, Nueva York, 1940; ed. revisada: Harvard University Press, Cambridge, 1951. [Version castellana en *Revista de Occidente*, Madrid, 1972 ]  
 WHITE, M.: «The Analytic and the Synthetic: an Untenable Dualism», en Sidney Hook (ed.), *John Dewey, Philosopher of Science and Freedom*, Dial Press, Nueva York, 1950, pp. 316-330.

<sup>22</sup> Se hallará una eficaz expresión de otros motivos para dudar de esta distinción en White.

## SIGNIFICADO Y TRADUCCIÓN \*

WILLARD V. O. QUINE

### I. SIGNIFICADO ESTIMULATIVO

Dado un discurso, y todas sus circunstancias estimulativas, su significado empírico es lo que queda después de despojarlo de toda verborrea; es lo que las oraciones de una lengua tienen en común con sus traducciones bien fundamentadas a una lengua completamente diferente. Por ello, si quisiéramos aislar el significado empírico, una perspectiva apropiada para hacernos una idea de en qué habría de consistir tal cosa sería la del lingüista dispuesto a comprender y traducir una lengua desconocida hasta el momento. Se dispondría de las preferencias inanalizadas de los indígenas y de las circunstancias observables que las rodean. Se necesitarían los significados; o las traducciones castellanas, pues una buena forma de informar sobre un significado es ofrecer una expresión de la lengua propia que tenga ese significado.

La traducción entre lenguas tan próximas como el frisón y el inglés se ve facilitada por la semejanza formal entre palabras afines. La traducción entre lenguas no emparentadas, como, por ejemplo, el húngaro y el inglés, puede estar facilitada por las tradicionales ecuaciones que se han ido estableciendo paralelamente al desarrollo de una cultura compartida. Para iluminar la naturaleza del significado, debemos pensar más bien en la traducción *radical*, es decir, la traducción de la lengua de un pueblo que ha permanecido aislado hasta ahora. Aquí es donde, en el caso de que ello sea posible, el significado estrictamente empírico se separa de las palabras que lo poseen.

En esta situación, las preferencias traducidas en primer lugar y con el menor riesgo de error serán por fuerza las que informan sobre observaciones manifiestamente compartidas por el lingüista y su informante. Un conejo pasa corriendo, el indígena dice «Gavagai» y nuestro lingüista de la jungla anota la oración «Conejo» (o «He aquí un conejo») como traducción de tanteo. Así, se abstendrá al principio de poner ninguna palabra en boca de su informante, bien que sólo sea porque no tiene palabras que poner. Cuando pueda, sin em

\* Versión castellana de Aurelio Pérez Fustegueras.

Por lo tanto, el lingüista tendrá que someter oraciones indígenas a la aprobación del informante, aun a riesgo de sesgar los datos por sugestión. Por lo demás, poco puede hacer a base de términos indígenas que tengan referencias en común. Supongamos, en efecto, que el lenguaje de la jungla cuenta con las oraciones  $S_1$ ,  $S_2$  y  $S_3$  que son, de hecho, traducibles por «Animal», «Blanco» y «Conejo», respectivamente. Las situaciones estimulativas son siempre diferentes, de forma relevante o no; y, dado que las respuestas aportadas tienen lugar una a una, las clases de las situaciones en las que el nativo afirma  $S_1$ ,  $S_2$  y  $S_3$  son, desde luego, mutuamente excluyentes, a pesar de los significados reales, aunque ocultos, de las palabras. Entonces, ¿de qué manera podrá percibir el lingüista que el nativo habría estado dispuesto a asentir a  $S_1$  en todas las situaciones en las que ha afirmado  $S_3$  y en algunas, aunque quizás no todas, en las que ha afirmado  $S_2$ ? Únicamente tomando la iniciativa e indagando sobre distintas combinaciones de oraciones indígenas y situaciones estimulativas, con objeto de ir reduciendo el número de sus hipótesis hasta quedarse, eventualmente, con la más satisfactoria.

Imaginemos, pues, al lingüista preguntando «¿Gavagai?» en situaciones estimulativas diversas y anotando cada vez si el indígena asiente, disiente o se abstiene. Aquí están implícitas algunas suposiciones sobre la capacidad de intuición del lingüista. En primer lugar, debe ser capaz de reconocer el asentimiento y el disentimiento en cualquier lengua. Por otra parte, debe ser capaz de adivinar la estimulación que su informante tiene en cuenta en cada momento —no desde un punto de vista neurológico sino en términos de referencia, aunque sea aproximada, al entorno—. Por último, debe poder conjeturar si esa estimulación impulsa realmente el asentimiento, o el disentimiento, del indígena a la pregunta concurrente; en este sentido, ha de estar en condiciones de eliminar los casos en los que el asentimiento, o el disentimiento, tiene su origen en una valoración de la oración en sí misma y no en la consideración del conejo que, ostensiblemente, pasa corriendo.

En un número suficiente de casos el lingüista tiene, ciertamente, éxito, y de igual modo podríamos tenerlo cualquiera de nosotros, aunque no fuéramos conscientes de nuestras pautas ni de nuestro método. Los gestos de asentimiento y disentimiento de los turcos son casi una inversión de los nuestros pero la expresión del rostro es reveladora y pronto nos pone en el buen camino. Por otra parte, lo que alguien observa en un momento dado puede inferirse, generalmente, de su orientación en el espacio, junto con nuestro conocimiento de los intereses humanos. El tercer y último punto a discernir es más difícil, a pesar de que con facilidad nos imaginamos cumpliéndolo en los casos típicos: juzgando, sin mayor conocimiento de la lengua, sobre si el asentimiento o el disentimiento del sujeto, subsiguiente a

una pregunta inesperada, ha sido promovido por la cosa que en esa ocasión estaba bajo examen. Señalando a la vez que se pregunta, se puede obtener una pista; si el objeto señalado es irrelevante, la respuesta irá acompañada, seguramente, de una mirada perpleja. Un indicio de irrelevancia puede estar en el hecho de que una pregunta no acompañada de ostensión dé lugar a que el indígena deje de prestar atención y parezca abstraído. Pero, dejando a un lado mecanismos hipotéticos, el hecho evidente es que, en virtud de las intuiciones no analizadas que sea, tendemos a obtener estos mínimos datos sobre las actitudes de los nativos sin la ayuda de un especial aparato lingüístico.

La pauta consistente en proponer oraciones en situaciones diversas vale únicamente para oraciones de una clase especial: aquellas que, como «Gavagai», «Rojo», «Eso hace daño», «Éste tiene la cara sucia», etc., sólo imponen asentimiento en presencia de ciertas circunstancias observables. Es una cuestión de *oraciones ocasionales* frente a *oraciones fijas*. Son las oraciones con las que nuestro lingüista ha de empezar y, también, aquellas a partir de las cuales podemos intentar una primera aproximación al concepto de significado.

La distinción entre oraciones ocasionales y oraciones fijas es definible en términos de la noción de asentimiento y disentimiento provocados que hemos supuesto disponible. Una oración es ocasional para un hombre si está preparado a asentir a, o a disentir de, ella solamente cuando la pregunta va acompañada de una estimulación que lo predisponga a ello.

No se trata de que el asentimiento o el disentimiento respecto a los enunciados fijos no pueda ser provocado de ese modo. Una estimulación visual fácilmente imaginable moverá a un buen conocedor de la ciudad a asentir al enunciado fijo «Hay casas de ladrillo en la calle de Los Olmos». En cierta ocasión la estimulación generada por un interferómetro movió a Michelson y a Morley a disentir del enunciado fijo «Existe una corriente de éter». Pero estos enunciados se diferencian de los ocasionales en que el sujeto, cuando es interrogado con posterioridad, puede insistir en su primer asentimiento o disentimiento aunque no medie ninguna estimulación específica; un enunciado ocasional, por el contrario, sólo suscita asentimiento o disentimiento si cada vez la pregunta va acompañada de la oportuna estimulación.

Definimos el *significado estimulativo afirmativo* de una oración ocasional S, para un hablante dado, como la clase de todas las estimulaciones que provocarían su asentimiento a S. Similarmente, pero en términos de disentimiento, podemos definir el significado estimulativo *negativo* de S. Finalmente, podemos definir el *significado estimulativo*, sin más, de S como el par ordenado de ambos. Podríamos distinguir grados de indecisión en el asentimiento y en el disenti-

o, por ejemplo, según el tiempo de reacción; y de forma fácilmente imaginable podríamos ampliar nuestra definición de significado estimulativo para incluir esta información; pero, con objeto de simplificar la exposición, no lo haremos.

Las distintas estimulaciones que reunimos en clases para integrar los significados estimulativos no deben ser tomadas como eventos particulares, fechados con exactitud, sino como tipos de eventos repetibles. Ha de poderse decir que la misma estimulación *ha ocurrido dos veces*, en vez de decir que han tenido lugar dos estimulaciones completamente similares. Para ver la necesidad de este enfoque, consideremos el significado estimulativo positivo de una oración ocasional S. Este significado es la clase  $\Sigma$  de todas las estimulaciones que *provocarían* el asentimiento a S. Si las estimulaciones se entendieran como eventos y no como tipos de eventos,  $\Sigma$  habría de ser una clase de eventos muchos de los cuales no han ocurrido ni ocurrirán pero que, si ocurrieran, provocarían asentimiento a S. Siempre que  $\Sigma$  contuviera un evento particular  $\sigma$ , realizado o no, tendría que contener todos los demás duplicados no realizados de  $\sigma$ ; pero *¿cuántos* son estos? Sin duda, es un irremediable sinsentido hablar de particulares no realizados y de su agrupamiento en clases. Las entidades no realizadas han de ser concebidas como universales porque, al carecer de especificaciones espacio-temporales y al ser semejantes en lo demás, es imposible distinguirlas entre sí.

Para nuestra presente tarea no es necesario determinar con exactitud cuándo hay que contar dos episodios de activación sensorial como recurrencias de la misma estimulación y cuándo como ocurrencias de estimulaciones diferentes. Está claro que en la práctica el lingüista nunca tendrá que preocuparse de los correlatos neurológicos de los episodios de estimulación. Siempre bastará con saber, por ejemplo, que el sujeto ha tenido una vislumbre fiable de un conejo. Esto es suficiente porque es razonable esperar que en circunstancias similares la conducta será la misma.

Los significados estimulativos, afirmativo y negativo, de un enunciado son mutuamente excluyentes. Hemos supuesto que el lingüista es capaz de reconocer el asentimiento y el disentimiento, y queremos interpretar estos últimos de manera que sea imposible decir de alguien que asiente y disiente en la misma ocasión al y del mismo enunciado ocasional. Es verdad que una cierta estimulación  $\sigma$  podría provocar, en cierto momento, el asentimiento de nuestro sujeto a S y que, más tarde, una recurrencia de  $\sigma$  podría provocar su disentimiento de S; pero en tal caso concluiríamos, simplemente, que el significado, para él, de S ha cambiado. Contaríamos a  $\sigma$  como elemento del significado estimulativo afirmativo que para él tenía S en la primera fecha, y como elemento del significado estimulativo negativo que para él tenía S en la segunda fecha. Una misma estimulación nunca



pertenecerá a la vez a la significación estimulativa afirmativa y a la significación negativa de S; es seguro que estas dos clases de estimulaciones son mutuamente excluyentes.

No obstante, los significados estimulativos afirmativo y negativo no se determinan uno a otro, pues el significado estimulativo negativo de S no abarca generalmente todas las estimulaciones que no provocarían el asentimiento a S. En general, por tanto, la comparación entre significados estimulativos íntegros puede ser una mejor base para la traducción que la mera comparación entre significados estimulativos afirmativos.

¿Qué decir, por último, de ese condicional fuerte, el «provocaría» de nuestra definición de significado estimulativo? El expediente es usado de manera tan indiscutida en sólidas ramas tradicionales de la ciencia que objetar su uso en un estudio tan inseguro como el presente sería, evidentemente, una pretensión fuera de lugar, algo así como un cumplido bien intencionado pero innecesario. Lo que el condicional fuerte define es una disposición; en este caso una disposición a asentir a S o a disentir de él ante estimulaciones diversas. Podemos suponer que esa disposición consiste en alguna sutil condición estructural, al modo de una alergia o de la solubilidad (en particular, se parece a la alergia en que no la comprendemos). Sea cual sea el estatuto ontológico de las disposiciones o el estatuto filosófico del discurso acerca de disposiciones, lo cierto es que sabemos bastante bien, en líneas generales, cómo establecer, a partir de comprobaciones juiciosas, muestras representativas y uniformidades observadas, una conjetura sobre la existencia de una determinada disposición.

## II. LA INESCRUTABILIDAD DE LOS TÉRMINOS

A la vista de la interdependencia de las oraciones, cabe preguntarse si podemos hablar razonablemente de significados, aunque sea de significados de enunciados completos y no de expresiones más breves, si no es en relación con los demás enunciados de una teoría inclusiva. Tal relatividad resultaría embarazosa porque, a su vez, el único acceso a la teoría viene dado por sus enunciados individualmente considerados. Ahora bien, la noción de significado estimulativo nos saca, en parte, del apuro. Esta noción aísla, para ciertos enunciados singulares, un tipo de significado empírico neto, y, aunque lo hace con independencia de la teoría, no por ello se pierde lo que el enunciado debe a ésta. En cierta medida, se trata de un instrumento para explorar el edificio de enunciados interconectados procediendo uno a uno. Algún expediente de este género es indispensable para iniciar la penetración en una cultura extraña, a la vez que es relevante para analizar nuestro propio conocimiento del mundo.

El punto de partida de nuestras consideraciones acerca del significado ha estado en las oraciones, si bien se ha tratado de oraciones de una clase especial y de una noción un tanto forzada de significado. Pues las palabras, cuando no son aprendidas como oraciones, lo son sólo, derivadamente, por abstracción de las funciones que desempeñan en las oraciones aprendidas. Con todo, antes de cualquier abstracción, hay oraciones de una sola palabra; y por fortuna estas últimas son, justamente, del tipo especial que ya estamos investigando; son oraciones ocasionales como «Blanco» y «Conejo». Además, tal sea en la medida en que se pueda decir que el concepto de significado estimulativo constituye, aunque en algún forzado sentido, un concepto de significado para oraciones ocasionales, se pueda decir también que constituye en particular un concepto de significado para términos generales como «Blanco» y «Conejo». Examinemos la aplicación de la noción de significado estimulativo a este último y convenientemente limitado ámbito de aplicación.

Afirmar la igualdad, para dos hablantes, del significado estimulativo de un término, o de dos términos para uno o dos hablantes, es afirmar una cierta igualdad en su aplicación: hay coincidencia tanto en las estimulaciones que provocan asentimiento como en las que provocan disentimiento. Ahora bien, ¿equivale eso a decir que el término o los términos tienen la misma *extensión*, es decir, que son verdaderos de los mismos objetos, para el hablante o hablantes en cuestión? Así podría parecer en el caso de «Conejo» y «Gavagai»; pero realmente la cosa es, en general, más complicada. Así, adaptando un ejemplo de Carnap, imaginemos un término general bárbaro aplicable a caballos y unicornios. Puesto que los unicornios no existen, la extensión de ese inclusivo término bárbaro es, sencillamente, la de «caballos». No obstante, nos gustaría de algún modo decir que el término, a diferencia de «caballo», también *sería* verdadero de los unicornios, si existieran. Pues bien, nuestro concepto de significado estimulativo nos ayuda realmente a dotar de sentido a esa determinación que queremos hacer respecto a objetos inexistentes, porque el significado estimulativo es, según la teoría del mismo, una cuestión de irritaciones de nuestras superficies sensoriales, no de caballos o unicornios. Cada estimulación causada por la observación de un unicornio es una combinación de impactos nerviosos que, en principio, no es menos real ni menos especificable que las causadas por la observación de un caballo. Incluso es posible provocar una estimulación de ese género mediante un artificio de cartón piedra. En la práctica también se puede hacer esto sin engaño, mediante descripciones y preguntas hipotéticas, siempre que se tenga un conocimiento suficiente del lenguaje; tales expedientes son maneras indirectas de hacer conjeturas acerca del significado estimulativo, si bien al margen de la definición del mismo.

Para términos como «Caballo», «Unicornio», «Blanco» y «Conejo» —términos generales para objetos externos observables— nuestro concepto de significado estimulativo parece proporcionar una relación de traducción razonablemente fuerte que va más allá de la mera coextensionalidad. Pero no es así; bien mirado, la relación ni siquiera alcanza la igualdad de extensión. Consideremos «Gavagai» de nuevo. ¿Quién sabe si los objetos a los que este término se aplica no son, después de todo, conejos sino simples estadios, o breves segmentos temporales, de conejos? En ambos casos, las mismas situaciones estimulativas que provocarían asentimiento a «Gavagai» provocarían asentimiento a «Conejo». O, quizás, «Gavagai» se aplica a cualquier parte no separada de conejos; y tampoco en este caso el significado estimulativo reflejaría diferencia alguna. Cuando, a partir de la igualdad de los significados estimulativos de «Gavagai» y «Conejo», el lingüista concluye que un gavagai es un conejo íntegro y duradero, está dando por sentado que el nativo es lo bastante semejante a nosotros para tener un término general breve para conejos y ninguno para estadios o partes de conejos.

Generalmente, podemos traducir algo (por ejemplo, «por morde») a un lenguaje dado aunque no haya nada en él que se corresponda con algunas de las sílabas componentes (por ejemplo, con «mor»). Precisamente de esta manera la oración ocasional «Gavagai» es traducible como diciendo que ahí hay un conejo, aunque ningún fragmento de «Gavagai», ni nada en el lenguaje nativo, se corresponda exactamente con el término «conejo». La sinonimia de «Gavagai» y «Conejo», en tanto oraciones, gira sobre consideraciones acerca de asentimientos provocados, las cuales trascienden todas las fronteras culturales; no ocurre lo mismo con la sinonimia de estas expresiones en tanto términos. Hacemos bien en escribir «Conejo» en vez de «conejo», para señalar que estamos considerando esa expresión en relación con sus sinónimos en tanto oración y no en relación con sus sinónimos en tanto término.

¿Podría superarse la supuesta indecisión entre conejos, estadios de conejos y partes integrantes de conejos mediante un pequeño suplemento de ostensión y de preguntas? Reflexionemos sobre esto. Cuando señalamos un conejo estamos señalando, también, un estadio de conejo y una parte integrante de un conejo. Cuando señalamos una parte de un conejo estamos, también, señalando un conejo y un estadio de conejo. Y sucede lo propio con la tercera alternativa. Nada que no pueda ser distinguido ya en el significado estimulativo mismo podrá serlo mediante ostensión, a menos que esta última vaya acompañada de preguntas sobre identidad y diferencia: ¿Es éste el mismo gavagai que aquél? ¿Hay aquí un gavagai, o dos? Tal interrogatorio exigiría del lingüista un dominio de la lengua indígena que nosotros por el momento no estamos en condiciones de justificar.

Más aún, supondría que, a semejanza del nuestro, el esquema conceptual indígena divide la realidad, de una forma o de otra, en una multiplicidad de objetos físicos distinguibles e identificables, sean conejos, estadios, o partes. Pero el enfoque indígena podría, después de todo, ser muy diferente al nuestro. El término «gavagai» podría ser el nombre propio de un universal recurrente, la cualidad de conejo, y *aun así* la oración ocasional «Gavagai» tendría el mismo significado estimulativo que posee bajo las otras alternativas sugeridas más arriba. Aún más, el punto de vista indígena podría ser tan ajeno al nuestro que hablar de objetos, incluso objetos abstractos como la conejidad, en relación con el mismo resultase carente de sentido. Los senderos por los que discurre el lenguaje de la jungla podrían ser completamente distintos del discurso occidental sobre esto y aquello, igual y diferente, uno y dos. Faltando algunos de estos usuales expedientes, no es posible decir, de modo significativo, que el indígena postula objetos; sustancias, tal vez, pero no objetos, concretos o abstractos. Y con todo, incluso en la perspectiva de esta actitud ontológica de naturaleza tan diferente, la oración ocasional «Gavagai» podría tener el mismo significado estimulativo que «(He aquí un) conejo». Las oraciones ocasionales y los significados estimulativos son moneda universal, en tanto que los términos, como expresiones que se aplican en algún sentido a objetos, son complementos provinciales de una cultura, como la nuestra, inclinada a la reificación.

¿Podemos siquiera imaginar alguna alternativa seria a nuestra pauta reificadora? Tal vez no, pues tendríamos que imaginarla en el proceso de traducción y lo que la traducción hace es imponer nuestra pauta. Quizá la noción misma de un tal contraste radical entre culturas carezca de sentido, salvo en uno puramente negativo: el fallo persistente en hallar análogos indígenas de nuestros familiares expedientes de la referencia objetiva, como los artículos, el predicado de identidad y las terminaciones de plural, que sean simples y convincentes. Sólo un fallo así podría hacernos percibir que el lenguaje indígena representa la materia del mundo de un modo inaccesible a nuestro lenguaje.

### III. ORACIONES OBSERVACIONALES

En las secciones primera y segunda hemos podido apreciar que, siempre que se limite a las oraciones ocasionales, la igualdad de significado estimulativo constituye, en algunos aspectos, una eficaz relación de sinonimia. Pero aun así limitado el significado estimulativo no cumple los requisitos implícitos en el habitual discurso acrítico sobre el significado. El problema está en que el asentimiento o el dis-asentimiento de un informante a una oración ocasional puede depen-

der sólo parcialmente de la estimulación actuante y hacerlo en amplia medida de información adicional propia que el lingüista desconoce. Cuando, primero, distinguimos entre oraciones ocasionales y oraciones fijas (sección I) y, después, pospusimos las segundas, excluimos todos los casos en los que el asentimiento o el disentimiento del informante podía depender por entero de información colateral, pero no los casos en los que su asentimiento, o disentimiento, depende principalmente de información de esa clase y sólo en muy escasa medida de la estimulación presente que lo provoca. Así, el asentimiento del nativo a «Gavagai» al vislumbrar algún movimiento entre la hierba puede deberse, más que nada, a la previa observación, cuando el lingüista estaba ausente, de conejos en los alrededores. Y hay oraciones ocasionales cuyos significados estimulativos no pueden ser considerados como sus «significados» por mucho esfuerzo de imaginación que se haga, porque el asentimiento a las mismas, aunque también requiere una estimulación, *siempre* depende en gran medida de información colateral. Un ejemplo es «Soltero»; el asentimiento a esta oración es provocado genuinamente por la visión de una cara; no obstante, se basa principalmente en información almacenada y no en la estimulación provocadora, salvo en la medida necesaria para reconocer al amigo soltero. La dificultad con «Soltero» radica en que su significado trasciende el mero aspecto físico de las personas que provocan el asentimiento y está relacionado con cuestiones que solamente pueden ser conocidas por otras vías. Es claro, entonces, que debemos procurar distinguir una subclase de las oraciones ocasionales, la de las *oraciones observacionales*, y reconocer que lo que he llamado significado estimulativo, constituye una razonable noción de significado para, a lo sumo, tales oraciones. Las oraciones ocasionales han sido definidas (sección I) como aquéllas respecto a las que se asiente o disiente sólo en presencia de una estimulación; lo que ahora exigimos de las oraciones observacionales, más particularmente, es que el asentimiento o el disentimiento sea provocado en todos los casos sin la ayuda de más información que la proporcionada por la estimulación misma.

Es curioso lo seguros que estamos de que cada asentimiento a «Soltero», o a un equivalente indígena, se basa en datos procedentes de dos fuentes: la estimulación presente y la información lateral. No carecemos de elaboradas, aunque asistemáticas, intuiciones sobre los usos de «Soltero» o de otras palabras de nuestro lenguaje. No obstante, no debemos enorgullecernos de esta clase de discurso fácil sobre significados y argumentos de sentido común, pese a toda su productividad, pues puede conducirnos casi sin darnos cuenta a las creencias más desesperadamente confusas y a controversias sin sentido.

Supongamos establecido que una determinada clase  $\Sigma$  compren

de exactamente las estimulaciones capaces de provocar directamente, sin el concurso de información lateral, el asentimiento al enunciado ocasional S. Supongamos igualmente que las estimulaciones comprendidas en otra clase  $\Sigma'$ , aptas también para provocar asentimiento a S, deben su eficacia más bien a cierta información lateral C muy extendida. ¿No podríamos también decir, en lugar de lo anterior, que al adquirir C los hombres han encontrado conveniente, de manera implícita, cambiar el *significado* mismo de S, en el sentido de que ahora los elementos de  $\Sigma'$  son aptos para provocar asentimiento directamente, al igual que los elementos de  $\Sigma$ ? Me parece que se pueden mantener las dos actitudes; ni la más completa perspicacia histórica revelaría distinción alguna, aunque revelara todas las etapas de la adquisición de C, porque el significado puede evolucionar *pari passu*. La distinción es ilusoria. Lo que hay objetivamente es una adaptación evolutiva a la naturaleza, reflejada en un conjunto evolutivo de disposiciones a dejarse provocar por estimulaciones a asentir, o a disentir, a, o de, enunciados ocasionales. Puede admitirse que estas disposiciones son impuras en el sentido de que incorporan conocimiento de cosas del mundo, pero lo hacen en una solución que no precipita nunca.

Las oraciones observacionales son oraciones ocasionales el asentimiento, o disentimiento, a las cuales es provocado sin la ayuda de información lateral. Pero la noción de ayuda por medio de información lateral ha resultado poco sólida. De todas maneras, la noción de oración observacional está en una posición mejor debido a un efecto estadístico de estabilización, un efecto que tal vez esté en mi mano explicar si por un momento sigo hablando acriticamente en términos de la dudosa noción de información lateral. Una parte de la información lateral relevante para una oración ocasional S puede estar muy extendida, otra parte puede no estarlo. A su vez, una porción de la información ampliamente extendida puede ser compartida por un grupo importante de personas y otra porción por otro grupo, de forma que pocas personas, si es que las hay, la conocerán completa. Por otra parte, el significado es social; incluso el individuo que usa una palabra de manera extravagante compartirá, probablemente, su desvío con algunos otros.

De cualquier modo, el efecto se pone de manifiesto llamativamente al comparar «Conejo» con «Soltero». El significado estimulativo de «Soltero» nunca será el mismo para dos hablantes que no sean hermanos siameses. Por el contrario, el significado estimulativo de «Conejo» será uniforme para casi todos los hablantes; excepciones como la del movimiento entre la hierba son raras. Entonces, un concepto que parece cumplir bastante bien el objetivo de la noción de oración observacional es, simplemente, el de *oración ocasional que posee un significado estimulativo intersubjetivo*.

¿Será entonces suficiente, para que una oración ocasional sea observacional, que haya *dos* personas para quienes aquella tenga el mismo significado estimulativo? No, como pone de relieve el ejemplo de los hermanos siameses. ¿Debe tener el mismo significado estimulativo para todos los miembros de la comunidad lingüística (suponiendo que ésta pudiera ser definida)? Seguramente, no. ¿Debe tener *exactamente* el mismo significado estimulativo para al menos dos hablantes? Quizá, no; considerando de nuevo el movimiento entre la hierba. Pero la cuestión es que estas preguntas aspiran a unos refinamientos fuera de lugar. Lo que aquí nos interesa son tendencias generales de conducta y lo que importa para la noción de oración observacional que aquí pretendemos construir es que para un número significativamente alto de hablantes los significados estimulativos tengan desviaciones significativamente pequeñas.

En un punto la variabilidad intersubjetiva del significado estimulativo de enunciados como «Soltero» ha sido expuesta incompletamente. El significado estimulativo de ese enunciado para una persona no sólo diferirá de su significado para otra persona sino que diferirá también del significado que para esta última tenga cualquier otro enunciado apropiado, en el mismo lenguaje o en otro.

El lingüista no está en condiciones de examinar y establecer *in extenso* un significado estimulativo indígena y, después, construir una oración castellana compleja cuyo significado estimulativo, para él, resulte equiparable, por medio de un exhaustivo agotamiento de casos, al significado indígena. Más bien tiene que extrapolar cualquier significado estimulativo indígena a partir de muestras relevantes, procurando adivinar la manera de pensar de su informante. Si la oración es tan poco observacional como «Soltero», simplemente no descubrirá líneas apropiadas de extrapolación; la traducción mediante significado estimulativo no producirá, entonces, un resultado erróneo, sencillamente no producirá resultado alguno. Esto es interesante porque lo que nos indujo a intentar definir las oraciones observacionales fue la consideración de que constituían la subclase de las oraciones ocasionales que parecían traducibles razonablemente por identidad de significados estimulativos. Vemos ahora que la limitación de este método de traducción a esta clase de oraciones es autorreforzante. Cuando una oración ocasional indígena no sea observacional, el lingüista se dará cuenta de que no podrá encontrar entre su propio repertorio de significados estimulativos de oraciones castellanas ninguno plausiblemente equiparable al que la oración indígena tiene para el informante.

Vimos que la noción de significado estimulativo no requiere una pluralidad de informantes. Lo que hay en general es el significado estimulativo de una oración para un hablante dado en un cierto momento de su vida (aunque al intentar delimitar ese significado el lin

gusta puede ver facilitada su tarea preguntando en momentos diversos y variando de informante). También vimos que la definición de oración observacional requiere puntos de referencia más amplios; en especial, requiere la comparación de diversos hablantes de la misma lengua. Pero en definitiva la consideración contenida en el párrafo anterior nos confirma que en realidad se puede prescindir de tal ampliación de horizontes: la traducción de oraciones ocasionales mediante el significado estimulativo se limitará de manera natural a las oraciones observacionales y, por tanto, no habrá necesidad de sacar a colación una y otra vez la definición de oración observacional.

La expresión «enunciado observacional» sugiere, a epistemólogos o metodólogos, los datos de la ciencia. A este respecto, nuestra versión no está fuera de lugar, pues, tal y como han sido definidos, nuestros enunciados observacionales son, precisamente, los enunciados ocasionales sobre los que es casi seguro que habrá un firme acuerdo por parte de todos los observadores bien situados. Por esto, ellos serán, justamente, los enunciados a los que un científico recurrirá en última instancia cuando sea requerido por colegas reticentes para que ordene sus datos y repita sus observaciones y experimentos.

#### IV. SINONIMIA INTRASUBJETIVA DE ORACIONES OCASIONALES

El significado estimulativo ha quedado definido para oraciones ocasionales en general, sin tener en cuenta la observacionalidad. No obstante, su parecido con lo que razonablemente podría ser llamado significado es menor cuando se aplica a oraciones no observacionales, como «Soltero». La traducción de «Bachelor» por «Soltero» no puede ser justificada sobre la base de la igualdad de los significados estimulativos; y tampoco puede serlo la sinonimia de «Soltero» y «Hombre que nunca se ha casado».

A pesar de eso, los significados estimulativos de «Soltero» y «Hombre que nunca se ha casado» son, curiosamente, idénticos para cada hablante. En todo momento un individuo será provocado por las mismas estimulaciones a asentir a «Soltero» y a «Hombre que nunca se ha casado»; y similarmente por lo que respecta al disentiimiento. Vemos de esta manera que, aunque el concepto de significado estimulativo está muy lejos del significado «auténtico» cuando se aplica a las oraciones ocasionales no observacionales «Soltero» y «Hombre que nunca se ha casado», la sinonimia se puede definir fielmente como la igualdad de significado estimulativo, tanto para estas oraciones como para las más escogidas oraciones observacionales, con tal de que nos limitemos a un hablante. Para cada hablante «Soltero» y «Hombre que nunca se ha casado» son sinónimos en un senti-



do definido (a saber, en tanto semejantes en significado estimulativo), sin tener por ello el mismo significado en ningún sentido aceptablemente definido de «significado» (pues en el caso de «Soltero» el significado estimulativo no es una base sólida para una definición aceptable). Puestas así las cosas, demos la bienvenida a la sinonimia y prescindamos del significado.

La restricción a un hablante no es obstáculo para afirmar que «Soltero» y «Hombre que nunca se ha casado» son sinónimas para el conjunto de la comunidad, en el sentido de que lo son para cada uno de sus miembros. Una extensión práctica al caso de dos lenguajes es posible si se dispone de un hablante bilingüe. Para un bilingüe, «Soltero» y «Bachelor» serán sinónimos en virtud de un criterio interno, a saber, la igualdad de significado estimulativo. Si tomamos a este hablante como muestra, podemos tratar «Soltero» y «Bachelor» como sinónimas para propósitos de traducción en las dos comunidades que él representa. Observando el grado de fluidez con el que se comunica en ambas comunidades, comparando otros individuos bilingües u observando cómo funcionan las traducciones, comprobaríamos si se trata de una muestra suficientemente buena.

Pero este uso de individuos bilingües no está al alcance del lingüista de la jungla que intenta penetrar en una cultura que ha permanecido aislada. Hasta ahora el único concepto disponible para la traducción radical es la igualdad de significado estimulativo, y aún este sólo para enunciados observacionales.

Las afinidades y diferencias entre la sinonimia intrasubjetiva y la traducción radical requieren una cuidadosa atención. La sinonimia intrasubjetiva, al igual que la traducción, puede valer para toda una comunidad. Es intrasubjetiva en el sentido de que cada sujeto conecta los sinónimos por medio de la igualdad, para él, de significado estimulativo; pero, no obstante, es comunitaria en el sentido de que las expresiones sinónimas en cuestión son conectadas mediante dicha igualdad por cada miembro de la comunidad. Obviamente, la sinonimia intrasubjetiva es en principio tan objetiva, tan susceptible de descubrimiento por el lingüista de campo, como lo es la traducción. Nuestro lingüista puede incluso descubrir que dos oraciones indígenas son intrasubjetivamente sinónimas sin por ello encontrar traducciones castellanas de las mismas —en resumen, sin comprenderlas— pues puede descubrir que ambas tienen el mismo significado estimulativo para su informante y no conocer ninguna oración castellana cuyo significado estimulativo, para él, sea presumiblemente el mismo. Así, para invertir la perspectiva, un marciano podría hallar que «Soltero» y «Hombre que nunca se ha casado» son sinónimas sin descubrir cuándo asentir a una u otra.

«Soltero» y «Sí» son dos oraciones ocasionales que podemos comparar instructivamente. Ni una ni otra son observacionales y, por

tanto, no son traducibles mediante identidad de significados estimulativos. Mal le iría al equivalente selvático de «Sí» (digamos, «Tak») si fuese traducido mediante el significado estimulativo. Las estimulaciones que, acompañando la pregunta del lingüista «¿Tak?», provocarían asentimiento a esta extraña oración, incluso por parte de todos los nativos, nunca habrían provocado el asentimiento de un castellano a «Sí», o a algo parecido, y ello a causa de su condición puramente verbal y de su exclusiva materialización en la lengua de la selva. «Tak» es precisamente lo que el lingüista busca como señal de asentimiento a cualquiera de las oraciones ocasionales indígenas que él pueda investigar, pero ella misma es inadecuada para ser investigada con los métodos que estamos considerando. Realmente, podemos suponer que, al igual que ocurre con «Soltero», no habrá dos hablantes, ni siquiera de la misma lengua, para quienes «Tak», o «Sí», tenga el mismo significado estimulativo; «Sí» sólo puede poseer el mismo significado estimulativo para hablantes que estén de acuerdo en cada cosa que cualquiera pueda decir, incluso impremeditadamente. Sin embargo, la igualdad de significado estimulativo es lo que define la sinonimia intrasubjetiva, no solamente entre «Soltero» y «Hombre que nunca se ha casado», sino también entre «Sí» y «Por supuesto» o «Ciertamente».

Repárese en que aún se mantienen las reservas formuladas en la sección II en relación con la coextensividad de términos. Aunque el maricano descubriera que «Soltero» y «Hombre que nunca se ha casado» son oraciones ocasionales sinónimas, no habría probado con ello que «soltero» y «hombre que nunca se ha casado» son términos generales coextensos. Por lo que él sabe, uno u otro de los términos podría aplicarse con exclusión del otro no a hombres sino a estadios temporales o partes de hombres, o incluso a un atributo abstracto (véase la sección II).

Cuando consideramos oraciones ocasionales y no términos, apreciamos que la sinonimia dentro de un lenguaje está en mejor situación que la traducción radical. La igualdad de significado estimulativo servirá como criterio de sinonimia intrasubjetiva de cualesquiera oraciones ocasionales, observacionales o no.

En realidad, debemos limitarnos a oraciones breves y sencillas. De lo contrario, la mera incapacidad de asimilar preguntas largas puede dar lugar, en términos de nuestras definiciones, a diferencias entre los significados estimulativos de oraciones que preferiríamos considerar como sinónimas. Puede ocurrir que una estimulación provoque asentimiento a la oración corta y no a la larga debido, precisamente, a la oscuridad de la segunda; en este caso no nos gustaría decir que para el sujeto el significado de la oración larga es diferente sino, sencillamente, que no la ha comprendido.

Con todo, no todas las oraciones serán cortas sino que algunas

contendrán a otras. Pienso en la ayuda que para ello prestan conjunciones como «o», «y», «pero», «si», «entonces», «que», etc., al regular la oración contenida como una cláusula de la oración continente. Pero también puede ocurrir en un nivel más elemental. Oraciones muy simples pueden contener sustantivos y adjetivos («rojo», «teja», «soltero», etc.) que también satisfacen los requisitos de las oraciones ocasionales, quedando sujetos, por tanto, a nuestro concepto de sinonimia. De esta manera, este concepto se aplica en pie de igualdad a oraciones que a veces ocurren como partes de otras. Es posible, entonces, cierta extensión de la sinonimia a oraciones ocasionales de mayor longitud que contienen a otras como partes, mediante el tipo de construcción que se expone a continuación.

Supongamos que  $R(S)$  es una oración ocasional que, aunque moderadamente corta, contiene como parte suya la oración ocasional  $S$ . Con  $R(\dots)$  podemos referirnos ahora al resultado de borrar la oración contenida en  $R(S)$ , y, siguiendo a Peirce, podemos hablar de ello como de un *rhema*. Diremos que un *rhema*  $R(\dots)$  es regular si cumple la siguiente condición: para cada  $S$  y  $S'$ , si  $S$  y  $S'$  son sinónimas y  $R(S)$  y  $R(S')$  son oraciones ocasionales idiomáticamente aceptables y suficientemente breves para nuestro concepto de sinonimia entonces  $R(S)$  y  $R(S')$  son sinónimas. Por ahora, este concepto de regularidad sólo tiene un sentido razonable para *rhemas* cortos, puesto que  $R(S)$  y  $R(S')$  deben ser, para  $S$  y  $S'$  convenientemente breves, suficientemente breves como para caer bajo nuestro concepto de sinonimia. No obstante, el propio concepto de regularidad invita de manera natural a una extensión: cuando los *rhemas*  $R_1(\dots)$  y  $R_2(\dots)$  sean regulares, admitiremos que también lo sea el *rhema* más largo  $R_1(R_2(\dots))$ . De este modo se hace posible calificar de regulares a *rhemas* cada vez más largos. A partir de aquí, podemos extender el concepto de sinonimia a diversas oraciones ocasionales largas, de la siguiente manera. Si  $R(\dots)$  es un *rhema* regular y  $S$  y  $S'$  son oraciones ocasionales cortas sinónimas en el sentido inicial, no extendido, y si  $R(S)$  y  $R(S')$  son, en alguna medida, combinaciones idiomáticamente aceptables, entonces, por extensión, podemos a su vez calificar de sinónimas a  $R(S)$  y  $R(S')$ , y ello aunque sean demasiado largas para ser sinónimas en el primer sentido. Ahora ya no hay limitación en la longitud, puesto que el *rhema* regular  $R(\dots)$  puede ser tan largo como queramos.

## V. FUNCIONES VERITATIVAS

En las secciones II y III justificamos la traducción radical de oraciones observacionales llevada a cabo mediante identificación de significados estimulativos. Pero hay un campo marcadamente diferente

que también se presta de manera inmediata a la traducción radical, a saber, el de las *funciones veritativas* como la negación, la conyunción lógica y la disyunción. Supongamos, como antes, que el asentimiento y el disentimiento fueran reconocibles en general. Las oraciones presentadas al nativo para aprobación o rechazo pueden ser, indiferentemente, ocasionales o fijas. Las ocasionales tendrán, naturalmente, que ir acompañadas de las oportunas estimulaciones provocadoras, tales que se pretende conseguir asentimiento o disentimiento; las fijas pueden ser presentadas en solitario. Ahora bien, por referencia al acuerdo y al desacuerdo podemos formular *criterios semánticos* para determinar si una expresión nativa dada ha de ser traducida como expresiva de la función veritativa en cuestión. Lo que sirve como criterio semántico de la negación es que ésta torna cualquier oración breve a la que estamos dispuestos a asentir en una oración de la que disentiremos, y viceversa. Para la conyunción tomamos el hecho de que la misma genera compuestos a los cuales, suponiendo que las oraciones componentes sean breves, estamos dispuestos a asentir cuando y sólo cuando estamos dispuestos a asentir a cada componente. El criterio de la disyunción es similar pero cambiando «asentir» por «disentir» en sus dos ocurrencias.

La razón de exigir brevedad en los componentes es simplemente que, como en la sección IV, si son largas, el sujeto puede confundirse. La identificación de una expresión indígena como negación, conyunción o disyunción no debe ser descartada sobre la base de la desviación de un sujeto respecto a nuestros criterios semánticos cuando es debida, tan sólo, a confusión. Adviértase que no se impone ningún límite a la longitud de los enunciados componentes a los que puede aplicarse la negación, la conyunción o la disyunción; se trata solamente de que los casos utilizados como prueba, a fin de reconocer por vez primera tales construcciones en una lengua extraña, siempre tendrán componentes breves.

Una vez que hallamos que una construcción nativa cumple alguno de estos tres criterios semánticos, ya no podemos pedir una comprensión más profunda de ello. Entonces estamos en condiciones de traducir la expresión nativa al castellano como «no», «y» u «o», según sea el caso, sin más salvedades que las rutinarias (es bien sabido que estas tres palabras castellanas no representan exactamente y sin ambigüedad a la negación, la conyunción y la disyunción).

En lógica se considera que una construcción que genera una oración a partir de otras oraciones expresa una función veritativa si cumple la condición de que el compuesto tenga un único «valor de verdad» (verdad o falsedad) para cada asignación de valores de verdad a los componentes. Obviamente, es posible formular criterios semánticos para todas las funciones veritativas de manera similar a como hicimos con la negación, la conyunción y la disyunción.

A veces, oímos hablar de pueblos de mentalidad prelógica que aceptan como verdaderas contradicciones manifiestas. Imaginemos forzando sin duda las intenciones de Levy-Bruhl, que alguien afirmó que estos indígenas admiten la verdad de cierto enunciado de la forma «*p ka bu p*», donde «*ka*» significa «*y*» y «*bu*» significa «*no*». Ahora bien, está a la vista que, si la traducción de «*ka*» como «*y*» y de «*bu*» como «*no*» sigue nuestros criterios semánticos, esta afirmación es absurda. ¿Y qué otros criterios podríamos aplicar sin incurrir en dogmatismo? A la inversa, afirmar sobre la base de un mejor diccionario que los nativos comparten nuestra lógica sería imponer esta última e incurrir en una petición de principio, en el caso de que aquí haya realmente algo que con sentido pueda ser pedido como principio. De todas maneras, yo me inclino por el diccionario mejor.

En el ámbito del castellano se puede ilustrar este punto mediante el tema de las lógicas alternativas. ¿Quién propone leyes lógicas heterodoxas está realmente contradiciendo nuestra lógica o sólo está sugiriendo nuevos e irrelevantes usos para algunos viejos vocablos familiares («*y*», «*o*», «*no*», «*todos*», etc.)? La cuestión carece de sentido, salvo desde la perspectiva de algún conjunto de criterios para traducir partículas lógicas. Dados los criterios ya expuestos, la respuesta es clara.

Oímos de vez en cuando que en su famosa libertad para resistematizar la ciencia o para forjar nuevos cálculos matemáticos el científico está obligado a respetar, al menos, la ley de contradicción. ¿Cuál ha de ser nuestra actitud ante esto? Huimos de la contradicción porque buscamos la verdad, pero ¿qué pensar de una revisión tan fundamental que considerara como verdaderas a las contradicciones? Antes que nada habría que sopesar cuidadosamente la utilidad del cambio. Las leyes de la lógica nos permiten deducir cualquier enunciado, indiscriminadamente, de una contradicción. Pero esa universal afirmación haría inútil la ciencia por ausencia de toda distinción. Así pues, una revisión que considere verdaderas las contradicciones habrá de ir acompañada de la revisión de otras leyes lógicas. Desde luego que todo esto puede hacerse; pero, una vez hecho, ¿cómo podemos afirmar que se ha logrado lo que se pretendía? Esta lógica heroicamente original cae bajo las consideraciones del párrafo anterior y, por tanto, tal vez pueda ser reconstruida, sencillamente, como la lógica de siempre, sólo que expresada en una mala notación.

Podemos, significativamente, contemplar el cambio de una ley lógica, la de tercero excluido, por ejemplo, o incluso la de contradicción. Pero esto es así porque mientras consideramos el cambio seguimos traduciendo *idénticamente*: «*y*» como «*y*», «*o*» como «*o*», etc. Después, una traducción más retorcida tal vez consiga anular el cambio de ley; o tal vez, por el contrario, se aprecie que el cambio da lugar a un sistema sustancialmente más fuerte que, además, no es en

modo alguno susceptible de ser traducido al viejo sistema. Pero, aun en este último caso, cualquier conflicto entre la vieja y la nueva lógica resulta ilusorio, pues sólo es posible porque traducimos por identidad.

En cualquier caso, una vez que hemos fijado nuestras traducciones mediante los repetidos criterios semánticos, hemos establecido completamente, al menos hasta donde alcanza la parte veritativo-funcional de la lógica, las leyes lógicas de un pueblo. En particular, queda determinada la clase de las *tautologías*, los compuestos veritativo-funcionales que son verdaderos en virtud, exclusivamente, de su estructura veritativo-funcional. Hay un conocido algoritmo tabular para determinar exactamente, en los casos de oraciones en las que las conectivas lógicas están excesivamente repetidas y superpuestas, qué asignaciones de valores de verdad a las oraciones componentes últimas hacen verdadero a todo el compuesto; las tautologías son los compuestos que resultan verdaderos para toda asignación.

Es un lugar común de la epistemología (y, por esa razón, ocasionalmente objetado) que solamente dos ámbitos muy opuestos del conocimiento disfrutaban de una certeza inamovible. Uno es el conocimiento de lo que se hace directamente presente a la experiencia sensible, el otro es el conocimiento de la verdad lógica. Es notable que estos sean, aproximadamente, los dos dominios donde hemos dotado a la traducción radical de un pleno sentido conductista. Un dominio donde la traducción parecía no ofrecer problemas era el de las oraciones observacionales. El otro, el de las funciones veritativas y, por tanto, también en cierto sentido el de las tautologías, dado que éstas son las verdades para las que solamente son relevantes las funciones veritativas.

Ahora bien, las funciones veritativas y las tautologías son tan sólo las más simples de las funciones y las verdades lógicas. ¿Podemos ir más allá? Las funciones lógicas que ahora se presentan como candidatos naturales son las oraciones *categorías*, representadas tradicionalmente por *A*, *E*, *I*, y *O*, y formadas habitualmente en castellano con las construcciones «todos son» («Todos los conejos son tímidos»), «ninguno es», «algunos son» y «algunos no son». A primera vista, un criterio semántico para *A* podría ser el siguiente: el compuesto impone asentimiento (a un hablante dado) si y sólo si el significado estimulativo positivo (para él) del primer componente es un subconjunto del significado estimulativo positivo del segundo. Es bastante obvio el modo en que este criterio puede ser adaptado a *E*, *I*, y *O*; lo malo es que, según lo visto en la sección II, la línea de razonamiento es equivocada en su conjunto. Tomemos, por ejemplo, *A*. Si «hipoide» es un término general aplicable a los caballos y a los unicornios, entonces, al no haber unicornios, todos los hipoidees son caballos, y, sin embargo, el significado estimulativo positivo de «Hipoidees

de» incluye patrones de estimulación, los correspondientes a «Unicornio», que no pertenecen al significado estimulativo positivo de «Caballo». Por esta razón, porque va más allá de la extensión, el criterio semántico sugerido no vale para «Todos los S son P». Y falla en un punto aún más serio, aunque de tipo opuesto, pues, mientras que los estadios de conejos no son conejos, vimos en la sección II que en un punto a significado estimulativo no hay ninguna diferencia.

La dificultad es fundamental. La verdad de los enunciados categóricos depende de los objetos de los que los términos componentes son verdaderos, por muy externos e inferenciales que sean; y qué objetos sean esos es algo que no viene determinado unívocamente por los significados estimulativos. En realidad, los enunciados categóricos, al igual que las desinencias de plural y la identidad, sólo adquieren sentido en relación a un esquema conceptual que postule la existencia de objetos; en contraste con esto, y conforme a lo subrayado en la sección II, los significados estimulativos pueden ser exactamente los mismos tanto para personas que posean esquemas de ese género como para personas tan ajenas a ello como se quiera. De lo que consideramos como lógica, la parte veritativo-funcional es, al parecer, la única que en una lengua extranjera somos capaces de atrapar con criterios conductistas.

## VI. HIPÓTESIS ANALÍTICAS

¿Cómo extiende, entonces, nuestro lingüista la traducción radical más allá de los límites de las oraciones observacionales y las funciones de verdad? En líneas generales procede de la manera que se expone a continuación. Segmenta las preferencias oídas, en trozos recurrentes manejablemente cortos y, a partir de ahí, compila una relación de «palabras» nativas. A título de hipótesis pone en ecuación diferentes miembros de esa relación con palabras y expresiones castellanas, cuidando que sean respetadas las traducciones ya establecidas de oraciones observacionales completas. Estas ecuaciones conjeturales pueden ser llamadas *hipótesis analíticas* de traducción. El lingüista tendrá que hacer uso de estas hipótesis no sólo para las palabras nativas sino también para las construcciones sintácticas o modos de reunir palabras, porque no cabe presumir que el lenguaje nativo siga el orden que las palabras tienen en castellano. El conjunto de estas hipótesis analíticas de traducción constituye una gramática y un diccionario jungla-castellano que son aplicados por el lingüista a oraciones para cuya traducción no hay evidencia independiente.

Las hipótesis analíticas de traducción no dependen exclusivamente, en cuanto a su evidencia, de traducciones previas de enunciados observacionales. También pueden ser parcialmente contrastadas

comprobando su conformidad con la sinonimia intrasubjetiva (sección IV) que pueda darse entre enunciados ocasionales. Por ejemplo, si las hipótesis analíticas llevan a traducir los enunciados  $S_1$  y  $S_2$  como «He aquí un soltero» y «He aquí un hombre que nunca se ha casado», respectivamente, debemos esperar entonces que para cada motivo el significado estimulativo de  $S_1$  sea el mismo que el de  $S_2$ .

Además de a la luz de la traducción de oraciones ocasionales, las hipótesis analíticas de traducción pueden, algunas veces, ser comprobadas parcialmente a la luz de la traducción de oraciones fijas. Estas últimas difieren de las ocasionales sólo en que el asentimiento y el disasentimiento respecto a ellas pueden producirse sin que medie una estimulación provocadora, no en que jamás medie una estimulación. El concepto de asentimiento provocado es razonablemente aplicado, para un hablante dado y por una vez, a la oración fija «Algunos conejos son negros», siempre que consigamos presentarle un ejemplar antes de que sepa que existen conejos negros. Un hablante puede, incluso, ser provocado repetidamente a asentir a algunas oraciones fijas; cada año se le puede realmente provocar a asentir a «El azafrán ha florecido», y cada día a «*The Times* ha llegado». De esta manera se reducen distancias entre oraciones fijas y ocasionales, lo que no obsta a la permanencia del límite definido hacia la mitad de la sección I. Así, el lingüista puede también evaluar sus hipótesis analíticas de traducción comparando las traducciones que se desprenden de las mismas con las traducciones originales que pueden ser llevadas a cabo a partir de asentimientos y disasentimientos provocados.

Las oraciones fijas pueden aportar una pequeña prueba adicional a las hipótesis analíticas al margen de asentimientos y disasentimientos provocados. Si, por ejemplo, las hipótesis analíticas apuntan a una oración castellana más bien trivial como traducción de una oración nativa  $S$ , el lingüista se tranquilizará si encuentra que  $S$  también impone, sin necesidad de provocación, un asentimiento general.

No es necesario que en la práctica las hipótesis analíticas de traducción adopten forma de ecuación. No es necesario que cada palabra indígena sea directamente puesta en ecuación con alguna palabra o expresión castellana. El traductor puede especificar ciertos contextos en los que una palabra debe ser traducida de una manera y otros en los que debe ser traducida de otra. Puede añadir a la forma ecuacional indicaciones semánticas suplementarias *ad libitum*. «En mal estado (*dicho de un huevo*)» es una definición lexicográfica tan buena como «podrido», a pesar de la intrusión de observaciones complementarias. Las instrucciones de traducción que tienen que ver con inflexiones gramaticales —por tomar un caso extremo— se pueden hacer depender de ecuaciones de palabras y expresiones en combinación inextricable con elementos no ecuacionales, pues el objetivo no es la traducción de palabras o construcciones sueltas sino la del dis-



curso coherente. Las hipótesis a las que llega el lingüista y las instrucciones que formula son hipótesis e instrucciones concomitantes acerca de la traducción del discurso coherente y pueden ser presentadas en cualquier forma que resulte clara y práctica, ecuacional o no.

No obstante, conviene prestar una particular atención a la forma más simple de hipótesis analítica, la que iguala directamente una palabra nativa o una construcción con un hipotético equivalente castellano. La razón de ello es que las hipótesis han de ser concebidas por el lingüista y el caso típico de concepción es aquél en el que éste percibe un paralelismo funcional entre un fragmento de una oración nativa S y una palabra de la traducción castellana de S. Solamente así podemos explicar que alguien decida traducir radicalmente al castellano una locución indígena como una desinencia de plural, como el predicado de identidad «=», como una cópula categórica o como cualquier otra parte de nuestro propio aparato de la referencia objetiva; porque, según ha sido subrayado en anteriores páginas, el examen, por muy detallado que sea, de los significados estimulativos o de otros aspectos de la conducta no puede determinar, ni siquiera si el nativo posee un esquema conceptual tan inclinado como el nuestro a postular la existencia de objetos. Únicamente por medio de esa abierta proyección de sus propios hábitos lingüísticos puede el traductor radical encontrar términos en la lengua indígena o, después de encontrados, emparejarlos con los de su propia lengua. Los significados estimulativos nunca bastarán ni siquiera para determinar qué palabras son términos, si es que los hubiera; menos aún, para determinar qué términos son coextensos.

En realidad, el lingüista que se toma el lenguaje de la jungla lo suficientemente en serio como para emprender la elaboración de un diccionario y una gramática completos no actuará exactamente como hemos imaginado. Desdeñando los paralelismos entre el castellano y la lengua indígena, se sumergirá en esta última hasta llegar a hablarla como un nativo. Desde el comienzo el aprendizaje puede desarrollarse tan exento de cualquier apelación a otras lenguas como queremos suponer; puede ser, virtualmente, un duplicado acelerado del aprendizaje infantil. Cuando, finalmente, el lingüista se aplica a la traducción y a elaborar un diccionario jungla-castellano y la correspondiente gramática, puede hacerlo como bilingüe. En adelante, sus dos personalidades asumen los papeles que en las anteriores páginas se repartían el lingüista y su informante. Iguala «Gavagai» con «Conejo» al apreciar que ambas oraciones poseen, para él, el mismo significado estimulativo. Y es fácil imaginar que seguirá el mismo procedimiento para traducir oraciones no observacionales del tipo de «Soltero»; en este punto pone de manifiesto sus ventajas la situación intrasubjetiva (cf. la sección IV). Cuando el lingüista obtiene buenos resultados con otras traducciones más abstrusas, sin duda pone en

juego, esencialmente, el método de las hipótesis analíticas, con la diferencia de que ahora las proyecta desde el previo dominio por separado de ambas lenguas y no las utiliza para llegar a dominar la de la jungla. Ahora bien, aunque la traducción llevada a cabo por un bilingüe es la que hace mayor justicia a la lengua de los nativos, la reflexión sobre la misma nos revela muy poco sobre la naturaleza del significado, porque el traductor bilingüe avanza en su trabajo en virtud de la comunicación que se lleva a cabo en el seno de una personalidad desdoblada, y este método solamente tendría para nosotros un sentido operativo en la medida en que consiguiéramos exteriorizarlo. Por lo tanto, sigamos pensando en los términos de nuestro plan primitivo, que incluía al informante nativo como colaborador vivo, antes que permitir que el lingüista empiece por devorarlo.

## VII. UN PUÑADO DE SIGNIFICADO

Una vez terminado, el manual jungla-castellano debe ser considerado como un manual para la traducción de oraciones por oraciones. Cualesquiera que sean los detalles de sus indicaciones sobre la traducción de palabras y sobre paradigmas sintácticos, el logro esencial del lingüista es una *correlación semántica* infinita de oraciones: la implícita especificación de una oración castellana para cada una de las infinitas oraciones posibles de la jungla. No es necesario que a cada oración de la jungla le corresponda una sola oración castellana, pero sí lo es que las diversas alternativas sean sinónimas bajo cualquier criterio aceptable de sinonimia intrasubjetiva entre oraciones castellanas; y lo mismo vale de la traducción inversa. Aunque la concepción y formulación de esa correlación semántica de oraciones depende de la fragmentación de las mismas en palabras, la evidencia que pueda haber será evidencia en favor de oraciones. Según lo indicado en la sección VI, esta consiste en diversas concordancias o conformidades en torno al significado estimulativo, las sinonimias intrasubjetivas y otros puntos relacionados con el asentimiento y el disentimiento provocado y no provocado.

A pesar de que la correlación semántica agota la totalidad de las oraciones indígenas, ella misma no está determinada en toda su extensión por su soporte evidencial. Debido a ello, incontables correlaciones alternativas son igualmente compatibles con esa evidencia. Si el lingüista arriba a la suya sin experimentar la sensación de que su opción haya sido demasiado arbitraria, ello se debe a lo limitado que está en cuanto a las correlaciones que puede manejar. A causa de su limitud no tiene las manos libres para asignar a cada una de las infinitas oraciones de la jungla cualquier oración castellana que sea compatible con la evidencia pertinente; ha de hacer estas asignaciones de

un modo manejablemente sistemático y sin salirse de un manejablemente limitado conjunto de fragmentos verbales repetibles. Al lingüista le es indispensable proceder palabra a palabra a la hora de especificar su correlación semántica, e incluso a la hora de concebirla.

El método de la segmentación no sólo limita las posibilidades de una eventual correlación semántica sino que, además, contribuye a definir los objetivos del lingüista en relación con la traducción. Concederá gran valor a los paralelismos estructurales, es decir, a la correspondencia entre las partes del enunciado indígena, según el mismo lo segmenta, y las partes de su traducción castellana. A igualdad de otros factores, la traducción más literal es considerada la mejor traducción<sup>1</sup>. La tendencia a la traducción literal está asegurada por un factor técnico. En efecto, la verdadera utilidad de la segmentación no es otra que permitir llevar a cabo traducciones largas a partir de correspondencias entre trozos cortos. Después, la tendencia termina por convertirse en un objetivo —y un objetivo que incluso varía en los detalles según la segmentación que haya sido adoptada en la práctica—.

Por medio de sus hipótesis analíticas nuestro lingüista de campo formula implícitamente (y, en realidad, alcanza) la gran hipótesis sintética que su correlación semántica general constituye. La evidencia que hay en favor de su correlación semántica es la misma que hay para sus hipótesis analíticas. Cronológicamente, las hipótesis analíticas son formuladas antes de que se disponga de toda esa evidencia; después de formuladas, la evidencia que vaya siendo obtenida será percibida como la corroboración pragmática de un diccionario de uso. En cualquier caso, la traducción de un vasto dominio de oraciones indígenas, si bien amparada por la correlación semántica, nunca podrá ser corroborada o sustentada si no es a modo de voladizo: es, sencillamente, lo que ocurre cuando las hipótesis analíticas son aplicadas más allá de la zona cubierta por la evidencia. El hecho de que estas traducciones inverificables vayan adelante sin tropiezos no debe ser tomado como evidencia pragmática de un buen trabajo lexicográfico, pues el tropiezo es imposible.

Debemos reconocer, por consiguiente, que las hipótesis analíticas y la gran hipótesis sintética que ellas componen son hipótesis sólo en un sentido incompleto. Un caso muy diferente es el de la traducción de «Gavagai» como «He aquí un conejo» mediante la igualdad de significado estimulativo. Esta es una hipótesis genuina que, aunque formulada a partir de la observación de casos paradigmáticos, puede ser errónea. «Gavagai» y «He aquí un conejo» poseen significados estimulativos para ambos hablantes y estos significados son

<sup>1</sup> De aquí, también, el concepto de sinonimia estructural de Carnap. Véase su *Meaning and Necessity*, Chicago, 1947, secciones 14-16.

Iguales o diferentes con independencia de nuestras conjeturas. Por el contrario, no es posible dotar de sentido a la sinonimia entre palabras como las que constituyen el objeto típico de las hipótesis analíticas. No se trata de que no podemos estar seguros de que las hipótesis analíticas son correctas, sino de que ni siquiera hay, como sí ocurre en el caso de «Gavagai», una materia objetiva sobre la que acertar o equivocarse.

La traducción radical progresa en todos sus frentes y las hipótesis analíticas resultan indispensables. Además, no son caprichosas; precisamente un poco más arriba hemos esbozado el modo en que reciben apoyo. ¿No podemos, entonces, decir que, después de todo, estos modos de concebir y respaldar las hipótesis analíticas *proporcionan* un sentido a la igualdad de significado entre las expresiones que tales hipótesis ponen en ecuación? No. Solamente podríamos afirmar esto si no fuera el caso de que dos conjuntos de hipótesis analíticas pueden chocar entre sí en el plano semántico y, no obstante, estar apoyadas por igual por toda la evidencia teóricamente accesible (incluyendo cualesquiera consideraciones sobre simplicidad).

Esta imposibilidad de definir la sinonimia por referencia al método de las hipótesis analíticas no es, formalmente, diferente de la imposibilidad de definir la verdad por referencia al método científico. También las consecuencias son paralelas. De la misma manera que sólo podemos hablar significativamente de la verdad de un enunciado en los términos de una teoría o esquema conceptual, sólo podemos, en general, hablar significativamente de sinonimia entre palabras y expresiones de dos lenguas en términos de un determinado sistema de hipótesis analíticas.

El método de las hipótesis analíticas es un procedimiento para que el lenguaje propio nos catapulte al lenguaje de la jungla. Es un procedimiento para injertar brotes exóticos en el viejo arbusto familiar hasta que sólo sea visible lo exótico. Las oraciones indígenas que no son semánticamente neutrales son traducidas tentativamente por oraciones castellanas sobre la base, en realidad, de una aparente analogía funcional en ambos lenguajes. Puesto que no son únicas, estas relaciones de analogía no pueden ser consideradas como los significados. Y, de todos modos, las analogías van siendo más débiles a medida que avanzamos hacia las oraciones más alejadas de la observación, es decir, hacia las oraciones teóricas. Puestas así las cosas, ¿quién se arriesgaría a traducir «Los neutrinos carecen de masa» a la lengua de la jungla? Si alguien lo hace, será acuñando nuevas palabras indígenas o desvirtuando el uso de las viejas. Y, probablemente, alegará en su descargo que los nativos carecen de los conceptos necesarios, así como que saben demasiado poca física. Y tendrá razón. No obstante, hay otra manera de presentar la cuestión: las hipótesis analíticas son, en el mejor de los casos, expedientes por medio de los

cuales ponemos de manifiesto, indirectamente, ciertas analogías entre las oraciones que han sido realmente traducidas y las que no, extendiendo así los límites iniciales de la traducción; y «Los neutrinos carecen de masa» está fuera de lugar allí donde los efectos de las hipótesis analíticas que somos capaces de idear son demasiado borrosos para servir de algo.

La integración de ambos en el continuo lingüístico que era el bajo-alemán facilitó la traducción del frisón al inglés (sección I), y su integración en un continuo de evolución cultural facilitó la traducción del húngaro al inglés. Al facilitar la traducción, estas continuidades despiertan la ilusión de una materia objetiva: la ilusión de que esas oraciones nuestras tan fácilmente intertraducibles son diversas encarnaciones verbales de alguna proposición o significación intercultural, cuando en realidad sería mejor considerarlas como meras variantes de una única expresión verbal intracultural. Sólo la discontinuidad de la traducción radical pone a prueba nuestras significaciones: las contrapone a sus encarnaciones verbales o, más propiamente, no encuentra nada en éstas.

Las oraciones observacionales se dejan mondar bien; sus significados, los significados estimulativos, surgen íntegros y libres de toda contaminación verbal residual. Las oraciones teóricas, como «Los neutrinos carecen de masa», la ley de entropía o la de constancia de la velocidad de la luz, están en el extremo opuesto. No es imaginable una situación estimulativa que provoque asentimiento o disasentimiento a una de ellas y que no incluya estimulación verbal procedente del interior del lenguaje. Las oraciones de este extremo teórico, y también otras que se encuentran entre ambos extremos, carecen de significado lingüísticamente neutral.

Es una afirmación trivial decir que no podemos conocer el significado de un enunciado extranjero salvo en la medida en que podemos ofrecer una traducción suya a nuestra propia lengua. Pero eso no es todo: sólo relativamente a un manual de traducción, en gran parte arbitrario, es posible, de muchos enunciados extranjeros, decir que comparten el significado de los correspondientes enunciados castellanos, y aun esto en un sentido muy parroquial de significado, a saber, su uso en castellano. Dejando a un lado los significados estimulativos de los enunciados observacionales, la mayor parte del discurso sobre el significado requiere la referencia tácita a una determinada lengua de manera similar a como el discurso sobre la verdad supone una referencia tácita a nuestro propio sistema del mundo, el mejor que por el momento podemos aglutinar.

Habiendo tan poca cosa en cuanto a significados neutrales (salvo los significados estimulativos) que sea relevante para la traducción radical, no es posible saber cuánto de nuestro éxito con las hipótesis analíticas se debe a una auténtica afinidad entre nuestra manera de

ver las cosas y la de los indígenas, y cuánto se debe al talento lingüístico o a la coincidencia afortunada. Ni siquiera estoy seguro de que la cuestión tenga sentido. Podemos admirarnos tanto de la inescrutabilidad de la mente de los indígenas como del estrecho parecido entre ellos y nosotros, según que hayamos sido incapaces de encontrar una buena traducción o, por el contrario, hayamos hecho un trabajo más detallado, hasta leer nuestros propios usos lingüísticos provinciales en el lenguaje de la jungla.

Usener, Cassirer, Sapir y, más recientemente, B. L. Whorf han subrayado que diferencias profundas en el lenguaje comportan diferencias esenciales en el modo de pensar o de ver el mundo. Yo prefiero no plantear la cuestión de manera que pareciera sugerir que ciertas proposiciones filosóficas son afirmadas en una cultura y negadas en otra. De lo que realmente se trata es de la dificultad o indeterminación para establecer una correlación. A medida que dejamos atrás las oraciones que tienen un claro condicionamiento directo a estímulos no verbales y a medida que abandonamos el terreno conocido, hay menos base para comparar dos traducciones y tiene menos sentido decir que una es una buena traducción y que la otra es mala.



## IV. SIGNIFICADO Y VERDAD

[The following text is extremely blurry and illegible. It appears to be a series of paragraphs of text, but the content cannot be transcribed accurately.]



SECRETADO Y VERDAD

La filosofía del lenguaje contemporánea ha heredado de la obra seminal de G. Frege dos ideas que alguna de las propuestas de explicación del significado que se recogen en este apartado explotan de manera intensiva. Una de ellas es el denominado *Principio estructural o contextual*, que afirma que el significado de cualquier constituyente oracional viene dado por su contribución a las estructuras oracionales en cuyo contexto puede figurar. Con ello se atribuye, por un lado, prioridad semántica a la noción de significado oracional y, por otro, se adscribe como tarea a la teoría del significado para un lenguaje particular la de proporcionar los significados de cada una de sus oraciones en virtud de la contribución sistemática a los mismos de sus componentes y estructura. La otra consiste en la afirmación de que el significado de una oración viene dado por sus condiciones de verdad; la conjunción de estas dos ideas —principio estructural y significado como condiciones de verdad— proporciona un rasgo único de las oraciones —el concepto de verdad— en términos del cual derivar los significados de las oraciones de un lenguaje de acuerdo con las exigencias anteriormente enunciadas.

En «La concepción semántica de la verdad y los fundamentos de la semántica» —una versión informal de un artículo anterior más técnico publicado en 1933— Alfred Tarski muestra cómo es posible, en contra de lo que mantenían la mayor parte de los positivistas lógicos, tratar de manera rigurosa las nociones semánticas presentando una *definición satisfactoria*, esto es: *materialmente adecuada y formalmente correcta*, del concepto de verdad. En «Verdad y significado» Donald Davidson, teniendo presente la conexión fregeana entre significado y verdad, mantiene que la mejor manera de proporcionar una explicación sistemática de los significados de un lenguaje consiste en desarrollar una definición recursiva de la verdad *à la* Tarski para ese lenguaje: «extraer un concepto rico (aquí algo razonablemente cercano a la traducción) a partir de finos trocitos de evidencia (aquí los valores de verdad de las oraciones) imponiendo estructura formal a bastantes trozos».

El proyecto de Davidson es, sin duda, grandioso a la vez que deficiente. Desde Frege era casi un asunto de dogma en filosofía del lenguaje que, mientras que era posible pasar del sentido a la referencia, «no había camino de vuelta —por decirlo con palabras de Russell— de la referencia del sentido» (justamente lo que el programa de Davidson propone). No es extraño, pues, que pronto surgieran respuestas y modificaciones. Peter Strawson mantiene en «Significado y verdad» —un título bastante revelador— que la «pugna homérica» entre los teóricos de la semántica formal (entre los que se cuenta Da-

vidson) y los de la intención comunicativa se resuelve a favor de estos últimos, dado que la noción de enunciado verdadero no es una noción neutral, sino que ha de analizarse en términos de los que un hablante *quiere decir* mediante tal enunciado. El propio Davidson consciente de la existencia de teorías extensionalmente equivalentes pero no igualmente interpretativas, modifica sus posiciones iniciales de «Verdad y significado» en «Interpretación radical» presentando las constricciones empíricas que (además de las formales) una teoría adecuada debe respetar. Cabe preguntarse si la doctrina de la interpretación radical es capaz de proporcionar teoremas interpretativos dejando a salvo lo que se proclama como mayor mérito de la teoría su carácter extensional. «¿Qué es una teoría del significado?» es la primera de una serie de dos conferencias [la segunda se publicó en 1976 en G. Evans y John McDowell (eds.), *Truth and Meaning*, Clarendon Press, Oxford] en las que Michael Dummett intenta establecer las condiciones generales que tal teoría debe cumplir. Para él las teorías «modestas» como las de Davidson —contrariamente a las teorías «vigorosas» como las que él propone— suponen que ya entendemos ciertos conceptos sobre los que descansa el significado de los demás. Su método consiste en derivar el significado de cada oración de un lenguaje a partir de un conjunto de axiomas de acuerdo con determinadas condiciones. Pero no nos dice nada sobre en qué consiste el significado de los axiomas, o sobre el significado de «es verdadera» en los teoremas. Esto lleva de manera natural al holismo. Pero una teoría adecuada del significado debe ser, según Dummett, atomista o, cuando menos, molecular, y debe poder determinar no meramente lo que se tiene que saber para conocer el significado de una expresión dada, sino en qué consiste tener tal conocimiento. Y esto es algo que, de acuerdo con Dummett, propuestas como la de Davidson simplemente no pueden hacer.

## LA CONCEPCIÓN SEMÁNTICA DE LA VERDAD Y LOS FUNDAMENTOS DE LA SEMÁNTICA \*

ALFRED TARSKI

Este trabajo consta de dos partes: la primera es de carácter expositivo, y la segunda es más bien polémica.

En la primera parte me propongo resumir de manera no formal los principales resultados de mis investigaciones concernientes a la definición de la verdad y al problema, más general, de los fundamentos de la semántica. Estos resultados están incorporados en una obra publicada hace varios años<sup>1</sup>. Aunque mis investigaciones conciernen a conceptos de los que se ha ocupado la filosofía clásica, se las conoce comparativamente poco en los círculos filosóficos a causa de su carácter estrictamente técnico. Por esta razón espero que se me acusará por retomar el asunto<sup>2</sup>.

Desde que apareció mi obra, mis investigaciones han suscitado varias objeciones de valor desigual; algunas de ellas fueron publicadas y otras fueron formuladas en discusiones públicas y privadas en que he tomado parte<sup>3</sup>. En la segunda parte de este trabajo expondré mis opiniones acerca de estas objeciones. Espero que las observaciones que formularé al respecto no sean consideradas de carácter pura-

\* Versión castellana de Mario Bunge, Emilio O. Colombo, Estela Arias y Lilia Fornari.

<sup>1</sup> Compárese Tarski (2) (véase la bibliografía al final de este trabajo). Esta obra puede consultarse para encontrar una presentación más detallada y formal del asunto de que trata esta memoria, y en particular de los tópicos incluidos en las secciones 6 y 7 a 13. También contiene referencias a mis primeras publicaciones sobre los problemas semánticos (una comunicación en polaco, 1930; el artículo Tarski (1) en francés, 1931; una comunicación en alemán, 1932; y un libro en polaco, 1933). La parte expositiva del presente trabajo se relaciona con Tarski (3). Mis investigaciones sobre la noción de verdad y sobre la semántica teórica han sido reseñadas o discutidas por Hofstadter (1), Juhos (1), Kokoszynska (1) y (2), Kotarbinski (2), Scholz (1), Weinberg (1) y otros.

<sup>2</sup> Puede esperarse que aumente el interés por la semántica teórica, de resultados de la reciente publicación de la importante obra de Carnap (2).

<sup>3</sup> Esto se aplica, en particular, a las discusiones públicas durante el I Congreso Internacional para la Unidad de la Ciencia (París, 1935) y la Conferencia de Congresos Internacionales para la Unidad de la Ciencia (París, 1937); cfr., por ejemplo, Neurath (1) y Gonseth (1).

mente polémico, sino que se encuentre en ellas algunas contribuciones constructivas al asunto.

En la segunda parte de este trabajo hago amplio uso de materiales gentilmente puestos a mi disposición por la Dra. Marja Koszowska (Universidad de Lwów). He contraído una deuda de gratitud con los profesores Ernest Nagel (Universidad de Columbia) y David Rynin (Universidad de California), quienes me han ayudado a preparar el texto final y me han hecho varias observaciones críticas.

## I. EXPOSICIÓN

1. *El problema principal: una definición satisfactoria de la verdad.* Nuestro discurso tendrá como centro la noción <sup>4</sup> de *verdad*. El problema principal es el de dar una *definición satisfactoria* de esta noción, es decir, una definición que sea *materialmente adecuada* y *formalmente correcta*. Pero semejante formulación del problema no puede, por su generalidad, considerarse inequívoca; requiere, pues, algunos comentarios adicionales.

Con el fin de evitar toda ambigüedad, debemos comenzar por especificar las condiciones en que la definición de verdad será considerada adecuada desde el punto de vista material. La definición descarta no se propone especificar el significado de una palabra familiar que se usa para denotar una noción nueva; por el contrario, se propone asir el significado real de una noción vieja. Por consiguiente, debemos caracterizar esta noción con la suficiente precisión para que cualquiera pueda determinar si la definición desempeña realmente su tarea.

En segundo lugar, debemos determinar de qué depende la corrección formal de la definición. Por esto, debemos especificar las palabras o conceptos que deseamos usar al definir la noción de verdad; y también debemos dar las reglas formales a que debiera someterse la definición. Hablando con mayor generalidad, debemos describir la estructura formal del lenguaje en que se dará la definición.

El tratamiento de estos puntos ocupará una considerable porción de la primera parte de este trabajo.

---

<sup>4</sup> Las palabras «noción» y «concepto» se usan en este trabajo con toda la vaguedad y ambigüedad con que figuran en la literatura filosófica. De modo que unas veces se refieren simplemente a un término. A veces no tiene importancia determinar cuál de estas interpretaciones se tiene en cuenta y en ciertos casos tal vez ninguna de ellas se aplica adecuadamente. Si bien en principio comparto la tendencia a evitar estos términos en toda discusión exacta, no he considerado necesario hacerlo así en esta presentación informal.

2. *La extensión del término «verdadero».* Comenzaremos por hacer algunas observaciones acerca de la extensión del concepto de verdad que aquí consideramos.

El predicado «verdadero» se usa con referencia a fenómenos psicológicos, tales como juicios o creencias, otras veces en relación con ciertos objetos físicos —a saber, expresiones lingüísticas y, específicamente, oraciones [*sentences*]— y a veces con ciertos entes ideales llamados «proposiciones». Por «oración» entenderemos aquí lo que en gramática se llama usualmente «oración enunciativa»; en lo que respecta al término «proposición», su significado es, notoriamente, tema de largas disputas de varios filósofos y lógicos, y parece que nunca se lo ha tornado bastante claro e inequívoco. Por diversas razones, lo más conveniente parece *aplicar el término «verdadero» a las oraciones*; es lo que haremos<sup>5</sup>.

Por consiguiente, siempre debemos relacionar la noción de verdad, así como la de oración con un lenguaje específico; pues es obvio que la misma expresión que es una oración verdadera en un lenguaje puede ser falsa o carente de significado en otro.

Desde luego, el hecho de que en este lugar nos interese primariamente la noción de verdad de las oraciones no excluye la posibilidad de extender subsiguientemente esta noción a otras clases de objetos.

3. *El significado del término «verdadero».* El problema del significado (o intensidad) del concepto de verdad plantea dificultades mucho más graves.

La palabra «verdad», como otras palabras del lenguaje cotidiano, ciertamente no es inequívoca. Y no me parece que los filósofos que han tratado este concepto hayan ayudado a disminuir su ambigüedad. En las obras y discusiones de filósofos encontramos muchas concepciones diferentes de la verdad y de la falsedad; debemos indicar cuál de ellas constituirá la base de nuestra discusión.

Quisiéramos que nuestra definición hiciese justicia a las intuiciones vinculadas con la *concepción aristotélica clásica de la verdad*, intuiciones que encuentran su expresión en las conocidas palabras de la *Metafísica* de Aristóteles:

*Decir de lo que es que no es, o de lo que no es que es, es falso, mientras que decir de lo que es que es, o de lo que no es que no es, es verdadero.*

<sup>5</sup> Para nuestros fines es más conveniente entender por «expresiones», «frases», etc., no inscripciones individuales, sino clases de inscripciones de forma similar (por consiguiente, no cosas físicas individuales, sino clases de tales cosas).

Si quisiéramos adaptarnos a la terminología filosófica moderna quizá podríamos expresar esta concepción mediante la familiar fórmula:

*La verdad de una oración consiste en su acuerdo (o correspondencia) con la realidad.*

(Se ha sugerido el término «teoría de la correspondencia» para designar una teoría de la verdad que se base en esta última formulación.)

En cambio, si decidimos extender el uso popular del término «designa» aplicándolo no sólo a nombres, sino también a oraciones; y si acordamos hablar de los designados [*designata*] de las oraciones como de «estados de cosas», posiblemente podríamos usar, para los mismos fines, la oración siguiente:

*Una oración es verdadera si designa un estado de cosas existente.*<sup>6</sup>

Sin embargo, todas estas formulaciones pueden conducir a diversos equívocos, pues ninguna de ellas es suficientemente precisa y clara (aunque esto se aplica mucho menos a la formulación aristotélica original que a cualquiera de las otras); en todo caso, ninguna de ellas puede considerarse una definición satisfactoria de la verdad. De nosotros depende que busquemos una expresión más precisa de nuestras intuiciones.

4. *Un criterio de adecuación material de la definición*<sup>7</sup>. Empecemos con un ejemplo concreto. Consideremos la oración «la nieve es blanca». Nos preguntamos en qué condiciones esta oración es verdadera o falsa. Parece claro que, si nos basamos sobre la concepción clásica de la verdad, diremos que la oración es verdadera si la nieve es blanca, y falsa si la nieve no es blanca. Por consiguiente, si la definición

<sup>6</sup> Para la formulación aristotélica, véase Aristóteles (1), Γ, 7, 27. Las otras dos formulaciones son muy comunes en la literatura, pero no sé a quiénes se deben. Pueden encontrarse un tratamiento crítico de varias concepciones de la verdad, p. ej., en Kotarbinski (1) (en polaco solamente por ahora), pp. 123 ss., y Russell (1), pp. 362 ss.

<sup>7</sup> En lo que respecta a la mayoría de las observaciones contenidas en las secciones 4 y 8, reconozco mi deuda con S. Lesniewski, quien las desarrolló en sus clases inéditas en la Universidad de Varsovia (en 1910 y años posteriores). Sin embargo, Lesniewski no anticipó la posibilidad de un desarrollo riguroso de la teoría de la verdad, y menos aun de una definición de esta noción; por consiguiente, si bien señaló equivalencias de la forma (V) como premisas de la antinomia del mentiroso, no las concibió como condiciones suficientes para un uso adecuado (o definición) de la noción de verdad. Tampoco se le deben las observaciones de la sección 8 respecto de la presencia de una premisa empírica en la antinomia del mentiroso, y la posibilidad de eliminar dicha premisa.

ción de verdad ha de conformarse a nuestra concepción, debe implicar la siguiente equivalencia:

*La oración «la nieve es blanca» es verdadera si, y sólo si, la nieve es blanca.*

Obsérvese que la oración «*la nieve es blanca*» figura entre comillas en el primer miembro de esta equivalencia, y sin comillas en el segundo miembro. En el segundo miembro tenemos la oración misma, y en el primero el nombre de la oración. Empleando la terminología lógica medieval, también podríamos decir que en el segundo miembro las palabras «*la nieve es blanca*» figuran en *suppositio formalis*, y en el primero en *suppositio materialis*. Apenas hace falta explicar por qué debemos poner el nombre de la oración, y no la oración misma, en el primer miembro de la equivalencia. En primer lugar, desde el punto de vista de la gramática de nuestro lenguaje, una expresión de la forma «*X es verdadera*» no se convertirá en una oración significativa si en ella reemplazamos «*X*» por una oración o por cualquier otra cosa que no sea un nombre, ya que el sujeto de una oración sólo puede ser un nombre o una expresión que funcione como nombre. En segundo lugar, las convenciones fundamentales que regulan el uso de cualquier lenguaje requieren que, toda vez que nos pronunciemos acerca de un objeto, sea el nombre del objeto el que se emplee y no el objeto mismo. Por consiguiente, si deseamos decir algo acerca de una oración —por ejemplo, que es verdadera— debemos usar el nombre de esa oración y no la oración misma <sup>8</sup>.

Puede agregarse que el poner una oración entre comillas no es, de ningún modo, la única manera de formar su nombre. Por ejemplo, suponiendo el orden usual de las letras de nuestro alfabeto, podemos usar la siguiente expresión como nombre (descripción) de la oración «*la nieve es blanca*».

*La oración constituida por cuatro palabras, la primera de las cuales consiste en las letras 13.<sup>a</sup> y 1.<sup>a</sup>, la segunda en las letras 16.<sup>a</sup>, 10.<sup>a</sup>, 25.<sup>a</sup>, y 6.<sup>a</sup>, la tercera en las letras 6.<sup>a</sup> y 22.<sup>a</sup>, y la cuarta en las letras 2.<sup>a</sup>, 13.<sup>a</sup>, 1.<sup>a</sup>, 16.<sup>a</sup>, 3.<sup>a</sup> y 1.<sup>a</sup> del alfabeto castellano.*

Generalicemos ahora el procedimiento que acabamos de aplicar. Consideremos una oración arbitraria; la reemplazaremos por la letra «*p*». Formemos el nombre de esta oración y reemplacémoslo por otra letra, por ejemplo, «*X*». Nos preguntamos cuál es la relación ló-

<sup>8</sup> En relación con diversos problemas lógicos y metodológicos envueltos en este trabajo, el lector puede consultar Tarski (6).



gica que existe entre las dos oraciones «*X es verdadera*» y «*p*». Está claro que, desde el punto de vista de nuestra concepción básica de la verdad, estas oraciones son equivalentes. En otras palabras, vale la siguiente equivalencia:

(V) *X es verdadera si, y sólo si, p.*

Llamaremos «*equivalencia de la forma (V)*» a toda equivalencia de esta clase (en la que «*p*» sea reemplazada por cualquier oración del lenguaje a que se refiere la palabra «*verdadero*», y «*X*» sea reemplazada por un nombre de esta oración).

Por fin podemos formular de manera precisa las condiciones en que consideraremos el uso y la definición del término «*verdadero*» como adecuado desde el punto de vista material: deseamos usar el término «*verdadero*» de manera tal que puedan enunciarse todas las equivalencias de la forma (V), y llamaremos «*adecuada*» a una definición de la verdad si de ella se siguen todas estas equivalencias.

Debemos subrayar que ni la expresión (V) misma (que no es una oración sino sólo un esquema de oración), ni caso particular alguno de la forma (V) pueden considerarse como una definición de la verdad. Sólo podemos decir que toda equivalencia de la forma (V), obtenida reemplazando «*p*» por una oración particular, y «*X*» por un nombre de esta oración, puede considerarse una definición parcial de la verdad, que explica en qué consiste la verdad de esta oración individual. La definición general debe ser, en cierto sentido, una conjunción lógica de todas estas definiciones parciales.

(La última observación exige algunos comentarios. Un lenguaje puede admitir la construcción de infinitas oraciones; por lo tanto, el número de definiciones parciales de la verdad referentes a oraciones de dicho lenguaje también será infinito. De modo que, para darle a nuestra observación un sentido preciso, tendríamos que explicar qué se entiende por «*conjunción lógica*» de infinitas oraciones; pero esto nos llevaría muy lejos en la consideración de problemas técnicos de la lógica moderna.)

5. *La verdad como concepto semántico.* Propongo el nombre de «*concepción semántica de la verdad*» para designar la concepción de la verdad que se acaba de exponer.

La *semántica* es una disciplina que —para decirlo sin gran precisión— *se ocupa de ciertas relaciones entre las expresiones de un lenguaje y los objetos (o «estados de cosas») a que se «refieren» esas expresiones.* Como ejemplos típicos de conceptos semánticos mencionemos los de *designación, satisfacción y definición*, tal como figurarían en los ejemplos siguientes:

*La expresión «el padre de este país» designa (denota) a George Washington; la nieve satisface la función proporcional [sentential] (la condición) «x es blanca»;*

*la ecuación « $2 \cdot x = 1$ » define (determina unívocamente) el número  $1/2$ .*

Mientras que las palabras «*designa*», «*satisface*» y «*define*» expresan relaciones (entre ciertas expresiones y los objetos a que se «refieren» estas expresiones), la palabra «*verdadero*» posee una naturaleza lógica diferente: expresa una propiedad (o denota una clase) de ciertas expresiones, a saber, de oraciones. Sin embargo, se ve fácilmente que todas las formulaciones que se dieron anteriormente (cfr. las secciones 3 y 4) y que tenían por finalidad explicar el significado de esta palabra, no se referían a las oraciones mismas sino a objetos «acerca de los que hablan» estas oraciones, o posiblemente a «estados de cosas» descritas por ellas. Más aún, resulta que la manera más simple y natural de obtener una definición exacta de verdad es la que acarrea el uso de otras nociones semánticas, p. ej., la noción de satisfacción. Por estas razones incluimos el concepto de verdad que aquí tratamos entre los conceptos semánticos, y el problema de definir la verdad resulta estar estrechamente relacionado con el problema más general de echar los fundamentos de la semántica teórica.

Acaso valga la pena decir que la semántica, tal como se la concibe en este trabajo (y en trabajos anteriores del autor), es una disciplina sobria y modesta que no tiene pretensiones de ser una panacea universal para curar todos los males y las enfermedades de la humanidad, sean imaginarios o reales. No se encontrará en la semántica remedio alguno para la caries dental, el delirio de grandezas o los conflictos de clase. Tampoco es la semántica un artificio para establecer que todos, con excepción del que habla y sus amigos, dicen disparates.

Desde la antigüedad hasta nuestros días, los conceptos semánticos han desempeñado un importante papel en las discusiones de los filósofos, lógicos y filólogos. Sin embargo, estos conceptos se han tratado durante mucho tiempo con cierta sospecha. Desde el punto de vista histórico, esta sospecha está completamente justificada. Pues, aunque el significado de los conceptos semánticos, tal como se los usa en el lenguaje cotidiano, parece bastante claro e inteligible, todas las tentativas de caracterizar este significado de manera general y exacta han fracasado. Y, lo que es peor, varios argumentos que implicaban estos conceptos, y que por lo demás parecían correctos y estar basados sobre premisas aparentemente obvias, conducían con frecuencia a paradojas y antinomias. Baste mencionar aquí la *antinomía del mentiroso*, la *antinomía de la definibilidad* (mediante un nú-

gica que existe entre las dos oraciones «*X es verdadera*» y «*p*». Está claro que, desde el punto de vista de nuestra concepción básica de la verdad, estas oraciones son equivalentes. En otras palabras, vale la siguiente equivalencia:

(V) *X es verdadera si, y sólo si, p.*

Llamaremos «*equivalencia de la forma (V)*» a toda equivalencia de esta clase (en la que «*p*» sea reemplazada por cualquier oración del lenguaje a que se refiere la palabra «*verdadero*», y «*X*» sea reemplazada por un nombre de esta oración).

Por fin podemos formular de manera precisa las condiciones en que consideraremos el uso y la definición del término «*verdadero*» como adecuado desde el punto de vista material: deseamos usar el término «*verdadero*» de manera tal que puedan enunciarse todas las equivalencias de la forma (V), y llamaremos «*adecuada*» a una *definición de la verdad si de ella se siguen todas estas equivalencias*.

Debemos subrayar que ni la expresión (V) misma (que no es una oración sino sólo un esquema de oración), ni caso particular alguno de la forma (V) pueden considerarse como una definición de la verdad. Sólo podemos decir que toda equivalencia de la forma (V), obtenida reemplazando «*p*» por una oración particular, y «*X*» por un nombre de esta oración, puede considerarse una definición parcial de la verdad, que explica en qué consiste la verdad de esta oración individual. La definición general debe ser, en cierto sentido, una conjunción lógica de todas estas definiciones parciales.

(La última observación exige algunos comentarios. Un lenguaje puede admitir la construcción de infinitas oraciones; por lo tanto, el número de definiciones parciales de la verdad referentes a oraciones de dicho lenguaje también será infinito. De modo que, para darle a nuestra observación un sentido preciso, tendríamos que explicar qué se entiende por «conjunción lógica» de infinitas oraciones; pero esto nos llevaría muy lejos en la consideración de problemas técnicos de la lógica moderna.)

5. *La verdad como concepto semántico.* Propongo el nombre de «*concepción semántica de la verdad*» para designar la concepción de la verdad que se acaba de exponer.

La *semántica* es una disciplina que —para decirlo sin gran precisión— *se ocupa de ciertas relaciones entre las expresiones de un lenguaje y los objetos (o «estados de cosas») a que se «refieren» esas expresiones*. Como ejemplos típicos de conceptos semánticos mencionemos los de *designación, satisfacción y definición*, tal como figuran en los ejemplos siguientes:

*La expresión «el padre de este país» designa (denota) a George Washington; la nieve satisface la función proporcional [sentencial] (la condición) «x es blanca»;*

*la ecuación « $2 \cdot x = 1$ » define (determina unívocamente) el número  $1/2$ .*

Mientras que las palabras «*designa*», «*satisface*» y «*define*» expresan relaciones (entre ciertas expresiones y los objetos a que se «refieren» estas expresiones), la palabra «*verdadero*» posee una naturaleza lógica diferente: expresa una propiedad (o denota una clase) de ciertas expresiones, a saber, de oraciones. Sin embargo, se ve fácilmente que todas las formulaciones que se dieron anteriormente (cfr. las secciones 3 y 4) y que tenían por finalidad explicar el significado de esta palabra, no se referían a las oraciones mismas sino a objetos «acerca de los que hablan» estas oraciones, o posiblemente a «estados de cosas» descritas por ellas. Más aún, resulta que la manera más simple y natural de obtener una definición exacta de verdad es la que acerca el uso de otras nociones semánticas, p. ej., la noción de satisfacción. Por estas razones incluimos el concepto de verdad que aquí tratamos entre los conceptos semánticos, y el problema de definir la verdad resulta estar estrechamente relacionado con el problema más general de echar los fundamentos de la semántica teórica.

Acaso valga la pena decir que la semántica, tal como se la concibe en este trabajo (y en trabajos anteriores del autor), es una disciplina sobria y modesta que no tiene pretensiones de ser una panacea universal para curar todos los males y las enfermedades de la humanidad, sean imaginarios o reales. No se encontrará en la semántica remedio alguno para la caries dental, el delirio de grandezas o los conflictos de clase. Tampoco es la semántica un artificio para establecer que todos, con excepción del que habla y sus amigos, dicen disparates.

Desde la antigüedad hasta nuestros días, los conceptos semánticos han desempeñado un importante papel en las discusiones de los filósofos, lógicos y filólogos. Sin embargo, estos conceptos se han tratado durante mucho tiempo con cierta sospecha. Desde el punto de vista histórico, esta sospecha está completamente justificada. Pues, aunque el significado de los conceptos semánticos, tal como se los usa en el lenguaje cotidiano, parece bastante claro e inteligible, todas las tentativas de caracterizar este significado de manera general y exacta han fracasado. Y, lo que es peor, varios argumentos que implicaban estos conceptos, y que por lo demás parecían correctos y estar basados sobre premisas aparentemente obvias, conducían con frecuencia a paradojas y antinomias. Baste mencionar aquí la *antinomia del mentiroso*, la *antinomia de la definibilidad* (mediante un nú-

mero finito de palabras) de Richard, y la *antinomia de los términos heterólogos*, de Grelling y Nelson <sup>9</sup>.

Creo que el método esbozado en este trabajo ayuda a superar estas dificultades y asegura la posibilidad de lograr un uso coherente de los conceptos semánticos.

6. *Lenguajes con una estructura especificada*. A causa de la posible aparición de antinomias, el problema de especificar la estructura formal y el vocabulario de un lenguaje en que hayan de darse definiciones de conceptos semánticos se hace especialmente agudo. Nos ocuparemos ahora de este problema.

Hay ciertas condiciones generales en las cuales se considera *exclusivamente especificada* la estructura de un lenguaje. Para especificar la estructura de un lenguaje debemos, por ejemplo, caracterizar inequívocamente la clase de palabras o expresiones que hayan de considerarse *significativas* [*meaningful*]. En particular, debemos indicar todas las palabras que hayamos decidido usar sin definir las, y que se llaman «*términos indefinidos* (o *primitivos*)»; y debemos dar las llamadas *reglas de definición* para introducir *términos definidos* o nuevos. Más aún, debemos establecer criterios para distinguir, dentro de la clase de expresiones, aquellas que llamaremos «*oraciones*» [*sentences*]. Por último, debemos formular las condiciones en que puede afirmarse una oración del lenguaje. En particular, debemos indicar todos los *axiomas* (u *oraciones primitivas*), esto es, oraciones que hayamos decidido afirmar sin prueba; y debemos dar las llamadas *reglas de inferencia* (o *reglas de prueba*) mediante las cuales podemos deducir nuevas oraciones afirmadas a partir de otras oraciones afirmadas previamente. Los axiomas, así como las oraciones que se deducen de ellos mediante las reglas de inferencia, se denominan «*teoremas*» u «*oraciones comprobables*».

Si, al especificar la estructura de un lenguaje, nos referimos exclusivamente a la forma de las expresiones que comprenden, se dirá que el lenguaje está *formalizado*. En tal lenguaje, los teoremas son las únicas oraciones que pueden afirmarse.

En la actualidad, los únicos lenguajes que poseen una estructura especificada son los lenguajes formalizados de los diversos sistemas de lógica deductiva, posiblemente enriquecidos mediante ciertos términos no lógicos. Sin embargo, el campo de aplicación de estos lenguajes es bastante amplio; teóricamente podemos desarrollar en

<sup>9</sup> La antinomia del mentiroso (atribuida a Eubúlides o Epiménides) se trata en las secciones 7 y 8. Para la antinomia de la definibilidad (debida a J. Richard) véase, p. ej., Hilbert-Bernays (1), vol. 2, pp. 263 ss.; para la antinomia de los términos heterólogos, véase Grelling-Nelson (1), p. 307.

los varias ramas de la ciencia, por ejemplo, la matemática y la física teórica.

(En cambio, podemos imaginar la construcción de lenguajes que tienen una estructura exactamente especificada sin estar formalizados. En un lenguaje de este tipo la afirmabilidad [*assertability*] de las oraciones, por ejemplo, puede no depender siempre de su forma sino de otros factores, de índole no lingüística. Sería interesante e importante construir realmente un lenguaje de este tipo, y más particularmente un lenguaje que resultara suficiente para el desarrollo de una amplia rama de la ciencia empírica; pues esto justificaría la esperanza de que los lenguajes de estructura especificada terminaran por reemplazar el lenguaje cotidiano en el discurso científico.)

*El problema de la definición de la verdad adquiere un significado preciso y puede resolverse en forma rigurosa solamente para aquellos lenguajes cuya estructura se ha especificado exactamente.* Para otros lenguajes —por ejemplo, para todos los lenguajes naturales o «hablados»— el significado del problema es más o menos vago, y su solución sólo puede tener un carácter aproximado. *Grosso modo*, la aproximación consiste en reemplazar un lenguaje natural (o un trozo del mismo en que estemos interesados) por otro cuya estructura se especifica exactamente, y que difiere del lenguaje dado «tan poco como sea posible».

7. *La antinomia del mentiroso.* Para descubrir algunas de las condiciones más específicas que deben satisfacer los lenguajes en que (o para los cuales) haya de darse la definición de la verdad, es aconsejable comenzar con el tratamiento de la antinomia que implica directamente la noción de verdad, a saber, la antinomia del mentiroso.

Para obtener esta antinomia en una forma clara <sup>10</sup>, consideremos la oración siguiente:

*la oración impresa en la página 283, líneas 30-31, de este trabajo, no es verdadera.*

Para abreviar reemplazaremos la oración que acabamos de enunciar por la letra «S».

De acuerdo con nuestra convención concerniente al uso adecuado del término «verdadero», afirmamos la siguiente equivalencia de la forma (V):

(1) «S» es verdadera si, y sólo si, la oración impresa en la página 283, líneas 30-31, de este trabajo, no es verdadera.

<sup>10</sup> Debida al profesor J. Lukasiewicz (Universidad de Varsovia).

Por otra parte, teniendo presente el significado del símbolo «s», establecemos empíricamente el siguiente hecho:

- (2) «s» es idéntica a la oración impresa en la página 283, líneas 30-31 de este trabajo.

Ahora bien, por una ley familiar de la teoría de la identidad (ley de Leibniz), se sigue de (2) que en (1) podemos reemplazar la expresión «la oración impresa en la página 283, líneas 30-31, de este trabajo» por el símbolo «“s”». Obtenemos así lo que sigue:

- (3) «s» es verdadera si, y sólo si, «s» no es verdadera.

De esta manera, hemos llegado a una contradicción evidente.

A mi juicio, sería erróneo y peligroso, desde el punto de vista del progreso científico, despreciar la importancia de esta y otras antinomias, tratándolas como bromas o sofistequerías. Es un hecho que estamos en presencia de un absurdo, que nos hemos visto obligados a afirmar una oración falsa (puesto que (3), como equivalencia entre dos oraciones contradictorias, es necesariamente falsa). Si tomamos en serio nuestro trabajo no podemos tolerar este hecho. Debemos descubrir su causa, es decir, debemos analizar las premisas sobre las que se basa la antinomia; luego debemos rechazar por lo menos una de esas premisas, y debemos investigar las consecuencias que esto tiene para el dominio íntegro de nuestra investigación.

Debemos insistir en que las antinomias han desempeñado un papel prominente en el establecimiento de los fundamentos de las modernas ciencias deductivas. Y, así como las antinomias de la teoría de las clases —y en particular la antinomia de Russell (de la clase de todas las clases que no son miembros de sí mismas)— fueron el punto de partida de las tentativas exitosas por formalizar coherentemente la lógica y la matemática, por su parte la antinomia del mentiroso y otras antinomias semánticas dan origen a la construcción de la semántica teórica.

8. *La incoherencia [inconsistency] de los lenguajes semánticamente cerrados.* Analizando las suposiciones que conducen a la antinomia del mentiroso, observamos las siguientes:

(I) Hemos supuesto, implícitamente, que el lenguaje en que se construye la antinomia contiene, además de sus expresiones, los nombres de estas expresiones, así como términos semánticos tales como el término «verdadero» referido a oraciones de este lenguaje. También hemos supuesto que todas las oraciones que determinan el uso adecuado de este término pueden afirmarse en el lenguaje. Un lenguaje que goza de estas propiedades se llamará «semánticamente cerrado».

(II) Hemos supuesto que en este lenguaje valen las leyes ordinarias de la lógica.

(III) Hemos supuesto que podemos formular y afirmar en nuestro lenguaje una premisa empírica, tal como el enunciado (2) que figuraba en nuestro argumento.

Resulta que la suposición (III) no es esencial, pues es posible reconstruir la antinomia del mentiroso sin su ayuda <sup>11</sup>. En cambio, se demuestra que las suposiciones (I) (II) son esenciales. Puesto que todo lenguaje que satisface ambas suposiciones es incoherente [*inconsistent*], debemos rechazar al menos una de ellas.

Sería superfluo subrayar en este punto las consecuencias del rechazo de la suposición (II), esto es, del cambio de nuestra lógica (suponiendo que esto fuera posible) aunque sólo fuera en sus partes más elementales y fundamentales. Por esto consideraremos solamente la posibilidad de rechazar la suposición (I). Decidiremos *no usar lenguaje alguno que sea semánticamente cerrado* en el sentido dado anteriormente.

Esta restricción sería, desde luego, inaceptable para quienes —por razones que no son claras para mí— creen que hay un solo lenguaje «genuino» (o, al menos, que todos los lenguajes «genuinos» son mutuamente traducibles). Sin embargo, esta restricción no afecta a las necesidades o a los intereses de la ciencia de una manera esencial. Los lenguajes (sea los formalizados o —lo que ocurre con mayor frecuencia— los trozos del lenguaje cotidiano) que se usan en el discurso científico no tienen por qué ser semánticamente cerrados. Esto es obvio en el caso en que los fenómenos lingüísticos y, en particular, las nociones semánticas, no intervienen de manera alguna en el asunto de una ciencia; pues en tal caso el lenguaje de esta ciencia no necesita ser provisto de términos semánticos. Sin embargo, veremos en la próxima sección cómo puede prescindirse de lenguajes semánticamente cerrados incluso en aquellas discusiones científicas que acarrean esencialmente nociones semánticas.

<sup>11</sup> Esto puede hacerse, a grandes rasgos, de la siguiente manera. Sea *S* un enunciado cualquiera que comience con las palabras «*Todo enunciado*». Correlacionamos con *S* un nuevo enunciado *S\** sometiendo a *S* a las siguientes modificaciones: reemplazamos en *S* la primera palabra, «*Todo*», por «*El*»; y después de la segunda palabra, «*enunciado*», insertamos toda la frase *S* entre comillas. Convengamos en llamar «(auto) aplicable» o «no (auto) aplicable» al enunciado *S*, según que el enunciado correlacionado *S\** sea verdadero o falso. Consideremos ahora el enunciado siguiente:

*Todo enunciado es no aplicable.*

Es fácil comprobar que el enunciado que acaba de formularse debe ser a la vez aplicable y no aplicable; por consiguiente, constituye una contradicción. Puede no ser del todo claro en qué sentido esta formulación de la antinomia no envuelve una premisa empírica; pero no me detendré más en este punto.



Se presenta el problema de la posición que ocupa el lenguaje cotidiano a este respecto. A primera vista parecería que este lenguaje satisficiera las suposiciones (I) y (II), y que por ello es incoherente. Pero en realidad el caso no es tan simple. Nuestro lenguaje cotidiano no es, ciertamente, un lenguaje que posea una estructura exactamente especificada. No sabemos con precisión cuáles expresiones son oraciones, y sabemos aún menos cuáles oraciones pueden tomarse como afirmables. De manera que el problema de la coherencia carece de sentido exacto respecto de este lenguaje. En el mejor de los casos sólo podemos arriesgarnos a conjeturar que un lenguaje cuya estructura ha sido especificada exactamente, y que se parece a nuestro lenguaje cotidiano tanto como sea posible, es incoherente.

9. *Lenguaje-objeto y metalenguaje.* Puesto que hemos acordado no emplear lenguajes semánticamente cerrados, debemos usar dos lenguajes diferentes al tratar el problema de la definición de la verdad y, en general, todos los problemas semánticos. El primero de estos lenguajes es el lenguaje acerca del que «se habla», y que es el tema de toda la discusión; la definición de la verdad que estamos buscando se aplica a las oraciones de este lenguaje. El segundo es el lenguaje en que «hablamos acerca del» primer lenguaje, y en cuyos términos deseamos, en particular, construir la definición de verdad para el primer lenguaje. Denominaremos *lenguaje-objeto* al primer lenguaje y *metalenguaje* al segundo.

Obsérvese que estos términos, «lenguaje-objeto» y «metalenguaje», sólo tienen un sentido relativo. Por ejemplo, si nos interesa la noción de verdad aplicada a oraciones, este último se convierte automáticamente en el lenguaje objeto de nuestra discusión; y para definir la verdad para este lenguaje, debemos ir a un nuevo metalenguaje, a un metalenguaje, por así decir, de un nivel superior. De esta manera llegamos a toda una jerarquía de lenguajes.

El vocabulario del metalenguaje está determinado, en gran parte, por las condiciones enunciadas anteriormente, en las que se considerará materialmente adecuada una definición de la verdad. Recordemos que esta definición debe implicar todas las equivalencias de la forma (V):

(V)  $X$  es verdadera si, y sólo si,  $p$ .

La definición misma, y todas las equivalencias implicadas por ella, han de formularse en el metalenguaje. En cambio, el símbolo « $p$ » que figura en (V) representa una oración arbitraria de nuestro lenguaje-objeto. Por consiguiente, toda oración que figure en el lenguaje-objeto también debe figurar en el metalenguaje; en otras palabras, el metalenguaje debe contener el lenguaje-objeto como parte

de él. Esto es al menos necesario para probar que la definición es adecuada, aun cuando la definición misma puede formularse a veces en un metalenguaje menos amplio que no satisface esta condición.

(La condición en cuestión puede modificarse un tanto, pues basta suponer que el lenguaje-objeto puede traducirse al metalenguaje; esto requiere cierto cambio de la interpretación del símbolo « $p$ » en (V). En todo lo que sigue ignoraremos la posibilidad de esta modificación.)

Más aún, el símbolo « $X$ » que figura en (V) representa el nombre de la oración representada por « $p$ ». Vemos, pues, que el metalenguaje debe tener la riqueza suficiente para dar la posibilidad de construir un nombre para cada una de las frases del lenguaje objeto.

Además, el metalenguaje debe contener, obviamente, términos de carácter lógico general, tal como la expresión «si y sólo si»<sup>12</sup>.

Es deseable que el metalenguaje no contenga términos indefinidos, a excepción de los involucrados explícita o implícitamente en las observaciones precedentes (es decir, términos del lenguaje-objeto), de los términos referentes a la forma de las expresiones del lenguaje objeto, de los términos que se usan para construir nombres de estas expresiones, y de los términos lógicos. En particular, deseamos que los términos semánticos (referentes al lenguaje-objeto) se introduzcan en el metalenguaje sólo por definición. Pues, si se satisface este postulado, la definición de la verdad, o de cualquier otro concepto semántico, cumplirá lo que esperamos intuitivamente de toda definición; es decir, explicará el significado del término que se define en términos cuyos significados parecen completamente claros e inequívocos. Más aún, tendremos entonces una garantía de que el uso de conceptos semánticos no nos complicará en contradicciones.

No tendremos otros requisitos que imponer a la estructura formal del lenguaje-objeto y del metalenguaje; suponemos que es semejante a la de otros lenguajes formalizados conocidos en la actualidad. En particular, suponemos que en el metalenguaje se observan las habituales reglas formales de definición.

10. *Condiciones de una solución positiva del problema principal.* Ahora ya tenemos una idea clara, tanto de las condiciones de adecuación material a que se sujetará la definición de la verdad como

<sup>12</sup> Los términos «lógica» y «lógico» se usan en este trabajo en un sentido amplio, que se ha tornado casi tradicional en las últimas décadas; la lógica comprende —según se supone aquí— toda la teoría de las clases y relaciones (esto es, la teoría matemática de los conjuntos). Por muchas y diferentes razones, me inclino personalmente a usar el término «lógica» en un estudio mucho más estrecho, a saber, de manera que sólo se aplique a lo que a veces se llama la «lógica elemental», es decir, al cálculo proposicional y al cálculo (restringido) de predicados.

de la estructura formal del lenguaje en que haya de construirse esta definición. En estas circunstancias, el problema de definir la verdad adquiere el carácter de un problema determinado de naturaleza puramente deductiva.

Sin embargo, la solución del problema no es en manera alguna obvia, y no la daría en detalle sin usar toda la maquinaria de la lógica contemporánea. En este lugar me limitaré a esbozar la solución y a tratar algunos de los puntos de mayor interés general comprendidos en ella.

La solución resulta ser unas veces positiva y otras negativa. Esto depende de ciertas relaciones formales entre el lenguaje objeto y su metalenguaje; o, más específicamente, del hecho de si el metalenguaje en su parte lógica es «*esencialmente más rico*» que el lenguaje objeto, o no. No es fácil dar una definición general y precisa de esta noción de «riqueza esencial». Si nos limitamos a los lenguajes que se basan sobre la teoría lógica de los tipos, la condición para que el metalenguaje sea «esencialmente más rico» que el lenguaje-objeto es que contenga variables de un tipo lógico superior al de las del lenguaje objeto.

Si no se satisface la condición de «riqueza esencial», usualmente puede demostrarse que es posible formular una interpretación del metalenguaje en el lenguaje-objeto; es decir, cualquier término dado del metalenguaje puede correlacionarse con un término bien determinado del lenguaje-objeto, de manera tal que las oraciones afirmables [*assertible*] de uno de los lenguajes resulten correlacionadas con oraciones afirmables del otro. De resultados de esta interpretación, la hipótesis de que en el metalenguaje se ha formulado una definición satisfactoria de verdad implica la posibilidad de reconstruir, en ese lenguaje, la antinomia del mentiroso; y esto nos obliga, a su vez, a rechazar la hipótesis en cuestión.

(El hecho de que el metalenguaje, en su parte no lógica, sea comúnmente más amplio que el lenguaje-objeto, no afecta a la posibilidad de interpretar el primero en el segundo. Por ejemplo, los nombres de las expresiones del lenguaje-objeto figuran en el metalenguaje, aunque en su mayor parte no figuran en el lenguaje-objeto; sin embargo, es posible interpretar estos nombres en términos del lenguaje-objeto.)

Vemos, pues, que la condición de «riqueza esencial» es necesaria para que sea posible dar una definición satisfactoria de la verdad en el metalenguaje. Si queremos desarrollar la teoría de la verdad en un metalenguaje que no satisfaga esta condición, debemos abandonar la idea de definir la verdad con la sola ayuda de los términos que hemos señalado anteriormente (en la sección 8). Debemos incluir entonces el término «*verdadero*», o algún otro término semántico, en la lista de los términos indefinidos del metalenguaje, expresando las propie-

dades fundamentales de la noción de verdad en una serie de axiomas. No hay nada que sea esencialmente incorrecto en semejante procedimiento axiomático, y puede resultar útil para diversos fines<sup>13</sup>.

Sucede, sin embargo, que puede evitarse este procedimiento. Pues la condición de «riqueza esencial» del metalenguaje resulta ser, no sólo necesaria, sino también suficiente para construir una definición satisfactoria de la verdad; es decir, si el metalenguaje satisface esta condición, en él puede definirse la noción de verdad. Indicaremos ahora, en términos generales, cómo puede llevarse a cabo esta construcción.

11. *La construcción de la definición (bosquejo)*<sup>14</sup>. A partir de la definición de otra noción semántica, la de *satisfacción*, puede obtenerse en forma muy sencilla una definición de verdad.

La de satisfacción es una relación entre objetos arbitrarios y ciertas expresiones llamadas «funciones proposiciones» [*sentential functions*]. Éstas son expresiones tales como «*x es blanca*», «*x es mayor que y*», etc. Su estructura formal es análoga a la de las proposiciones; sin embargo, pueden contener variables de las llamadas libres (tales como «*x*» e «*y*» en «*x es mayor que y*») que pueden figurar en enunciados.

Al definir la noción de función proposicional en los lenguajes formalizados, comúnmente aplicamos lo que se llama «procedimiento recursivo»; es decir, primero describimos funciones proposicionales de la estructura más simple (lo que comúnmente no ofrece dificultades) y luego indicamos las operaciones mediante las cuales pueden construirse funciones compuestas a partir de otras más simples. Una operación de este tipo puede consistir, por ejemplo, en formar la disyunción o la conjunción lógica de dos funciones dadas, es decir, en combinarlas por las palabras «o» o «y». Una oración [*sentence*] puede definirse ahora simplemente como una función proposicional que no contiene variables libres.

En lo que respecta a la noción de satisfacción, podríamos tratar de definirla diciendo que ciertos objetos *satisfacen* una función dada si ésta se convierte en una oración verdadera cuando reemplazamos sus variables libres por nombres de los objetos dados. En este sentido, por ejemplo, la nieve satisface la función proposicional «*x es blanca*», ya que la oración «*la nieve es blanca*» es verdadera. Pero,

<sup>13</sup> Véase, sin embargo Tarski (3), pp. 5 ss.

<sup>14</sup> El método de construcción que esbozaremos puede aplicarse —mediando cambios apropiados— a todos los lenguajes formalizados que se conocen en la actualidad; sin embargo, no se sigue que no podría construirse un lenguaje al que no pudiera aplicarse este método.

aparte de otras dificultades, no podemos emplear este método porque deseamos usar la noción de satisfacción para definir la verdad.

Para obtener una definición de satisfacción debemos aplicar nuevamente un procedimiento recurrente. Indicamos cuáles son los objetos que satisfacen las funciones proposicionales más simples, y luego enunciamos las condiciones en que los objetos dados satisfacen una función compuesta (suponiendo que sabemos cuáles son los objetos que satisfacen las funciones simples a partir de las cuales se construye la compuesta). Así, por ejemplo, decimos que ciertos números satisfacen la disyunción lógica «*x es mayor que y o x es igual a y*» si satisfacen por lo menos una de las funciones «*x es mayor que y*» o «*x es igual a y*».

Una vez obtenida la definición general de satisfacción, observamos que también se le aplica automáticamente a las funciones proposicionales especiales que no contienen variables libres, es decir, a las oraciones. Resulta que para una oración hay sólo dos casos posibles: una oración o bien es satisfecha por todos los objetos, o no es satisfecha por objeto alguno. Por consiguiente, llegamos a una definición de la verdad y de la falsedad diciendo simplemente que *una oración es verdadera si es satisfecha por todos los objetos, y falsa en caso contrario*<sup>15</sup>.

(Puede parecer extraño que hayamos elegido un rodeo para definir la verdad de una oración, en lugar de tratar de aplicar, por ejemplo, un procedimiento directo de recurrencia. La razón de esto es que las oraciones compuestas se construyen a partir de funciones proposicionales sencillas, pero no siempre a partir de oraciones simples; por consiguiente, no se conoce ningún método general de recurrencia que se aplique específicamente a las oraciones.)

Este tosco esbozo no aclara dónde y cómo está implicada la suposición

<sup>15</sup> Al llevar a la práctica esta idea surge cierta dificultad técnica. Una función proposicional puede contener un número arbitrario de variables libres; y la naturaleza lógica de la noción de satisfacción varía con este número. Así, por ejemplo, la noción en cuestión, aplicada a funciones de una variable, es una relación binaria entre estas funciones y objetos singulares; aplicada a funciones de dos variables se convierte en una relación ternaria entre funciones y pares de objetos; y así sucesivamente. Por consiguiente, estrictamente hablando no se nos presenta una sola noción de satisfacción sino infinitas nociones; y resulta que estas nociones no pueden definirse independientemente entre sí, sino que deben introducirse simultáneamente.

Para vencer esta dificultad empleamos la noción matemática de sucesión infinita (o, posiblemente, de sucesión finita con un número arbitrario de términos). Convenimos en considerar la satisfacción, no como una relación de orden superior entre funciones proposicionales y un número indefinido de objetos, sino como una relación binaria entre funciones y sucesiones de objetos. Con esta suposición, la formulación de una definición general y precisa de satisfacción ya no presenta dificultades; y un enunciado verdadero puede definirse ahora como aquel que es satisfecho por toda sucesión.

ción de la «riqueza esencial» del metalenguaje; esto no se aclara sino cuando se lleva a cabo la construcción de manera detallada y formal<sup>16</sup>.

12. *Consecuencias de la definición.* La definición de verdad esbozada precedentemente tiene muchas consecuencias interesantes.

En primer lugar, la definición resulta ser no sólo formalmente correcta, sino también materialmente adecuada (en el sentido establecido en la sección 4); en otras palabras, implica todas las equivalencias de la forma (V). A este respecto, es importante señalar que las condiciones de adecuación material de la definición determinan únicamente la extensión del término «verdadero». Por esto, toda definición de la verdad que sea materialmente adecuada es necesariamente equivalente a la que hemos construido. La concepción semántica de la verdad no nos da, por así decir, ninguna posibilidad de elección entre diversas definiciones no equivalentes de esta noción.

Más aún, de nuestra definición podemos deducir varias leyes de naturaleza general. En particular, con su ayuda podemos probar las leyes de contradicción y del tercero excluido, tan características de la concepción aristotélica de la verdad. Estas leyes semánticas no debieron identificarse con las leyes lógicas de contradicción y del tercero excluido, relacionadas con ellas; las leyes lógicas pertenecen al cálculo proposicional, es decir, a la parte más elemental de la lógica, y no incluyen para nada el término «verdadero».

Aplicando la teoría de la verdad a los lenguajes formalizados de cierta clase muy amplia de disciplinas matemáticas, se obtienen otros resultados importantes; sólo se excluyen de esta clase disciplinas de un carácter elemental y de una estructura lógica muy elemental. Resulta que, para una disciplina de esta clase, la noción de verdad nunca coincide con la de comprobabilidad [provability]; pues todas las oraciones comprobables son verdaderas, pero hay oraciones verdaderas que no son comprobables<sup>17</sup>. Se sigue, entonces, que toda disciplina

<sup>16</sup> Para definir por recurrencia la noción de satisfacción, debemos aplicar cierta forma de la definición por recurrencia que no se admite en el lenguaje-objeto. Luego, la «riqueza esencial» del metalenguaje puede consistir simplemente en admitir este tipo de definición. En cambio, se conoce un método general que haga posible la eliminación de todas las definiciones por recurrencia, reemplazándolas por definiciones normales, explícitas. Si tratamos de aplicar este método a la definición de satisfacción, vemos que, o bien debemos introducir en el metalenguaje variables de tipo lógico superior al de las que figuran en el lenguaje-objeto, o bien debemos suponer axiomáticamente, en el metalenguaje, la existencia de clases más amplias que todas aquellas cuya existencia puede establecerse en el lenguaje-objeto. Véase a este respecto Tarski (2), pp. 193 ss., y Tarski (5), p. 110.

<sup>17</sup> A causa del desarrollo de la lógica moderna, la noción de prueba matemática ha

de este tipo es coherente pero incompleta; es decir, de dos oraciones contradictorias cualesquiera, a lo sumo una es comprobable y, lo que es más, existe un par de oraciones contradictorias ninguna de las cuales es comprobable<sup>18</sup>.

13. *Extensión de los resultados a otras nociones semánticas.* La mayor parte de los resultados obtenidos en las secciones anteriores al tratar la noción de verdad pueden extenderse, mediando cambios apropiados, a otras nociones semánticas; por ejemplo, a la noción de satisfacción (implicada en nuestra discusión precedente) y a las de *designación y descripción*.

Cada una de estas nociones puede analizarse siguiendo las líneas

---

sufrido una simplificación de grandes alcances. Un enunciado de una disciplina formalizada dada es comprobable si puede obtenerse a partir de los axiomas de esta disciplina por la aplicación de ciertas reglas de inferencia sencillas y puramente formales tales como las de separación y sustitución. Por consiguiente, para mostrar que todos los enunciados comprobables son verdaderos, basta probar que todos los enunciados aceptados como axiomas son verdaderos, y que las reglas de inferencia, cuando se las aplica a enunciados verdaderos, producen nuevos enunciados verdaderos; y por lo común esto no ofrece dificultades.

En cambio, a causa de la naturaleza elemental de la noción de comprobabilidad una definición precisa de esta noción sólo requiere medios lógicos bastante simples. En la mayoría de los casos, los artificios lógicos disponibles en la disciplina formalizada (con la que está relacionada la noción de comprobabilidad) son más que suficientes para estos fines. Sabemos, sin embargo, que en lo que respecta a la definición de la verdad vale justamente lo contrario. Por consiguiente, en general las nociones de verdad y de comprobabilidad no pueden coincidir; y, puesto que todo enunciado comprobable es verdadero, debe haber enunciados verdaderos que no son comprobables.

<sup>18</sup> La teoría de la verdad nos da, pues, un método general para efectuar pruebas de coherencia [*consistency*] en las disciplinas metamáticas formalizadas. Es fácil advertir, sin embargo, que una prueba de coherencia obtenida por este método puede poseer algún valor intuitivo —esto es, puede convencernos, o reforzar nuestra creencia, de que la disciplina en cuestión es realmente coherente— tan sólo en el caso de que logremos definir la verdad en términos de un metalenguaje que no contenga como parte al lenguaje-objeto (ver a este respecto una observación en la sección 9). Pues sólo en este caso pueden ser intuitivamente más simples y obvias las suposiciones deductivas del metalenguaje que las del lenguaje objeto, aun cuando se satisfaga formalmente la condición de «riqueza esencial». Cf. también Tarski (3), p. 7.

La incompletitud de una amplia clase de disciplinas formalizadas constituye el contenido esencial de un teorema fundamental de K. Gödel; cfr. Gödel (1), pp. 187 ss. La explicación del hecho de que la teoría de la verdad conduce tan directamente al teorema de Gödel es bastante simple. Al deducir el resultado de Gödel a partir de la teoría de la verdad hacemos un uso esencial del hecho de que la definición de verdad no puede darse en un lenguaje que sea sólo tan «rico» como el lenguaje-objeto (cf. nota 17); sin embargo, al establecer este hecho se aplica un método de razonamiento que está estrechamente relacionado con el usado (por primera vez) por Gödel. Puede añadirse que Gödel fue obviamente guiado, en su prueba, por ciertas consideraciones intuitivas concernientes a la noción de verdad, aun cuando esta noción no figure explícitamente en la prueba; cfr. Gödel (1), pp. 174 ss.

generales del análisis de la verdad. De esta manera pueden establecerse criterios para un uso adecuado de estas nociones; puede mostrarse que cada una de estas nociones, cuando se la usa en un lenguaje semánticamente cerrado de acuerdo con estos criterios, conduce necesariamente a una contradicción<sup>19</sup>; vuelve a tornarse indispensable una distinción entre el lenguaje-objeto y el metalenguaje; y en todos los casos la «riqueza esencial» del metalenguaje resulta ser una condición necesaria y suficiente para lograr una definición satisfactoria de la noción en cuestión. Por consiguiente, los resultados obtenidos al discutir una noción semántica particular se aplican al problema general de los fundamentos de la semántica teórica.

Dentro de la semántica teórica podemos definir y estudiar algunas otras nociones, cuyo contenido intuitivo es más complicado y cuyo origen semántico es menos evidente; nos referimos, por ejemplo, a las importantes nociones de *consecuencia*, *sinonimia* y *significado*<sup>20</sup>.

En este trabajo nos hemos ocupado de la teoría de nociones semánticas vinculadas con un lenguaje objeto individual (aun cuando en nuestros argumentos no han figurado propiedades específicas de este lenguaje). Sin embargo, también podríamos considerar el problema de desarrollar una *semántica general* que se aplique a una amplia clase de lenguajes objeto. Una parte considerable de nuestras observaciones previas puede extenderse a este problema general; sin embargo, a este respecto surgen ciertas dificultades nuevas que no discutiremos en este lugar. Sólo observaré que el método axiomático (mencionado en la sección 10) puede resultar el más apropiado para el tratamiento del problema<sup>21</sup>.

<sup>19</sup> Las nociones de designación y definición llevan directamente a las antinomias de Grelling-Nelson y de Richard (cfr. nota 9). Para obtener una antinomia a partir de la noción de satisfacción, construimos la siguiente expresión:

*La función proposicional X no satisface a X.*

Surge una contradicción cuando consideramos la cuestión de si esta expresión, que es claramente una función proposicional, se satisface a sí misma o no.

<sup>20</sup> Todas las nociones mencionadas en esta sección pueden definirse en términos de satisfacción. Podemos decir, p. ej., que un término dado designa un objeto dado si este objeto satisface la función proposicional «*x es idéntico a T*», donde «*T*» representa el término dado. Análogamente, se dirá que una función proposicional define un objeto dado si este último es el único objeto que satisface esta función. Para una definición de consecuencia, véase Tarski (4), y para la sinonimia, Carnap (2).

<sup>21</sup> La semántica general es el tema de Carnap (2). A este respecto véanse también observaciones de Tarski (2), pp. 388 ss.



## II. OBSERVACIONES POLÉMICAS

14. *La concepción semántica de la verdad ¿es la «correcta»? Comenzaré la parte polémica de este trabajo haciendo algunas observaciones generales.*

Espero que nada de lo que aquí se diga se interprete como una pretensión de que la concepción semántica de la verdad es la «correcta» o aun la «única posible». No tengo la menor intención de contribuir de manera alguna a esas discusiones interminables y a menudo violentas sobre el asunto: «¿Cuál es la concepción correcta de la verdad?»<sup>22</sup>. Confieso que no entiendo de qué se trata en esas disputas pues el problema mismo es tan vago que no es posible alcanzar una solución determinada. En efecto, me parece que nunca se ha aclarado el sentido en que se usa la oración «la concepción correcta». En la mayoría de los casos se tiene la impresión de que la oración se usa en un sentido casi místico que se funda en la creencia de que toda palabra tiene un solo significado «real» (idea de tipo platónico o aristotélico), y que todas las concepciones rivales realmente intentan captar este significado único; pero, puesto que se contradicen entre sí, solo una de las tentativas puede tener éxito, y por lo tanto una sola de las concepciones es la «correcta».

Las disputas de este tipo no se restringen, en modo alguno, a la noción de verdad. Se producen en todos los dominios en que se usa el lenguaje común, con su vaguedad y ambigüedad, en lugar de una terminología exacta, científica; y carecen siempre de sentido y son, por ello, vanas.

Me parece evidente que el único enfoque racional de estos problemas es el siguiente: debiéramos aceptar el hecho de que no nos enfrentamos con un concepto sino con diversos conceptos diferentes denotados por una palabra; debiéramos tratar de aclarar estos conceptos todo lo posible (mediante la definición, o un procedimiento axiomático, o de alguna otra manera); para evitar más confusiones, debiéramos convenir en usar diferentes términos para designar los diferentes conceptos; y luego podremos emprender un estudio tranquilo y sistemático de todos estos conceptos que exhiba sus principales propiedades y relaciones mutuas.

Para referirnos específicamente a la noción de verdad, sin duda acontece que en las discusiones filosóficas —y quizá también en el uso cotidiano— pueden encontrarse algunas concepciones incipientes de esta noción que difieren esencialmente de la clásica (y de la cual la concepción semántica no es sino una forma modernizada). En efecto, en la literatura se han discutido varias concepciones de esta

<sup>22</sup> Cfr. varias citas en Ness (1), pp. 13 ss.

clase; por ejemplo, la concepción pragmatista, la teoría de la coherencia, etc.

Me parece que ninguna de estas concepciones ha sido formulada, hasta ahora, de una manera inteligible e inequívoca. Sin embargo, esto puede cambiar; puede venir una época en que nos veamos frente a varias concepciones de la verdad, incompatibles pero igualmente claras y precisas. Se hará entonces necesario abandonar el uso ambiguo del término «verdadero», introduciendo en su lugar diversos términos, cada uno de los cuales denote una noción diferente. Personalmente, no me sentiría herido si un futuro congreso mundial de «teóricos de la verdad» decidiera, por mayoría de votos, reservar la palabra «verdad» para una de las concepciones no clásicas y sugiriera otra palabra, por ejemplo «ferdad», para designar la concepción que aquí consideramos. Pero no puedo imaginar que nadie pueda presentar argumentos sólidos en sostén de la tesis de que la concepción semántica es «equivocada» y debe abandonarse por entero.

15. *Corrección formal de la definición de la verdad que se ha sugerido.* Las objeciones específicas que se han formulado a mis investigaciones pueden dividirse en varios grupos, que discutiremos por separado.

Creo que prácticamente todas estas objeciones se aplican, no a la definición especial que he propuesto, sino a la concepción semántica de la verdad en general. Aun aquellas que se formularon contra la definición propuesta podrían referirse a cualquier otra definición que se conformara a esta concepción.

Esto se aplica, en particular, a aquellas objeciones que concierne a la corrección formal de la definición. He oído unas pocas objeciones de esta clase; sin embargo, dudo mucho que cualquiera de ellas pueda ser tratada seriamente.

Como ejemplo típico citaré en sustancia una de estas objeciones<sup>23</sup>. Al formular la definición usamos necesariamente conectivas proposicionales, es decir, expresiones tales como «si..., entonces», «o», etc. Ellas aparecen en el *definiens*; y una de ellas, a saber, la frase «si, y sólo si», se emplea comúnmente para combinar el *definiendum* con el *definiens*. Sin embargo, es bien sabido que el significado de las conectivas proposicionales se explica en lógica con ayuda de las palabras «verdadero» y «falso»; por ejemplo, decimos que una equivalencia, es decir, un enunciado de la forma «*p* si, y sólo si *q*», es verdadero si sus dos miembros —esto es, las oraciones representadas por «*p*» y «*q*»— son verdaderos, o son falsos. Por lo tanto, la definición de la verdad implica un círculo vicioso.

<sup>23</sup> No citaremos los nombres de las personas que han formulado objeciones, a menos que dichas objeciones hayan sido publicadas.

Si esta objeción fuera válida no sería posible ninguna definición formalmente correcta de la verdad; pues no podemos formular ninguna oración compuesta sin usar conectivas proposicionales u otros términos lógicos definidos con su ayuda. Afortunadamente, la situación no es tan grave.

Sin duda, un desarrollo estrictamente deductivo de la lógica es precedido a menudo por ciertas declaraciones que explican en qué condiciones se consideran verdaderas o falsas oraciones de la forma «*si p, entonces q*», etc. (Tales explicaciones se dan a menudo esquemáticamente, mediante las llamadas tablas de verdad.) Sin embargo, esas declaraciones están fuera del sistema de la lógica, y no deberían considerarse como definiciones de los términos en cuestión. No se formulan en el lenguaje del sistema, sino que constituyen consecuencias especiales de la definición de la verdad que se da en el metalenguaje. Más aún, esas declaraciones no influyen de manera alguna en el desarrollo deductivo de la lógica. Pues en tal desarrollo no tratamos la cuestión de si una oración dada es verdadera: sólo nos interesa el problema de si es comprobable<sup>24</sup>.

En cambio, desde el momento en que nos encontramos dentro del sistema deductivo de la lógica —o de cualquier disciplina basada sobre la lógica, tal como la semántica— tratamos las conectivas proposicionales como términos indefinidos, o bien las definimos mediante otras conectivas proposicionales, pero nunca mediante términos semánticos tales como «verdadero» o «falso». Por ejemplo, si convenimos en considerar las expresiones «no» y «si... entonces» (y posiblemente también «si y sólo si») como términos indefinidos, podemos definir «o» diciendo que una oración de la forma «*p o q*» es equivalente a la oración correspondiente de la forma «*si no p, entonces q*». La definición puede formularse, por ejemplo, de la manera siguiente:

*(p o q) si, y sólo si (si no p, entonces q).*

Obviamente, esta definición no contiene términos semánticos.

<sup>24</sup> Debe subrayarse, sin embargo, que en lo que respecta a la cuestión de un presunto círculo vicioso la situación no cambiaría aun cuando adoptáramos un punto de vista diferente, tal como el de Carnap (2); esto es, si consideráramos la especificación de las condiciones en que son verdaderas las oraciones de un lenguaje como parte esencial de la descripción de ese lenguaje. En cambio, puede observarse que el punto de vista representado en el texto no excluye la posibilidad de usar tablas de verdad en un desarrollo deductivo de la lógica. Sin embargo, estas tablas deben considerarse meramente como un instrumento formal para verificar la comprobabilidad de ciertas oraciones; y los símbolos «V» y «F» que figuran en ellas, y que usualmente se consideran abreviaturas de «verdadero» y «falso», no deberían interpretarse en ninguna forma intuitiva.

Sin embargo, un círculo vicioso surge en la definición sólo cuando el *definiens* contiene, ya el término que se define, ya otros términos definidos con su ayuda. Vemos así claramente que el uso de las conectivas proposicionales en la definición del término semántico «verdadero» no acarrea círculo alguno.

Mencionaré otra objeción que encontré en la literatura y que también parece concernir a la corrección formal, si no de la definición misma de verdad, al menos a los argumentos que conducen a esta definición<sup>25</sup>.

El autor de esta objeción se equivoca al considerar el esquema (V) de la sección 4 como una definición de la verdad. Objeta a esta presunta definición que está afectada de una «brevedad inadmisibles, es decir, incompletitud», que «no nos da un medio para decidir si por «equivalencia» se entiende una relación lógico-formal, o bien no lógica y también estructuralmente no descriptible». Para eliminar este «defecto» sugiere completar (V) de una de las dos maneras siguientes:

(V') *X es verdadera si, y sólo si, p es verdadera.*

(V'') *X es verdadera si, y sólo si, se da p (es decir, si ocurre lo que declara p).*

Luego discute estas dos nuevas «definiciones», que estarían libres del «defecto» formal de la vieja, pero que resultan insatisfactorias por otras razones, de índole no formal.

Esta nueva objeción parece surgir de una incompreensión relativa a la naturaleza de las conectivas proposicionales (por lo cual está de alguna manera relacionada con la que tratamos anteriormente). El autor de la objeción no parece advertir que la frase «*si y sólo si*» (contrastariamente a oraciones tales como «*son equivalentes*», o «*es equivalente a*») no expresa una relación entre oraciones, puesto que no combina nombres de oraciones.

En general, todo el argumento se funda sobre una obvia confusión entre oraciones y sus nombres. Baste señalar que —a diferencia de (V)— los esquemas (V') y (V'') no dan ninguna expresión significativa si en ellos sustituimos «*p*» por una oración; pues las oraciones «*p es verdadera*» y «*se da p*» (es decir, «*lo que declara p ocurre*») pier-

<sup>25</sup> Cfr. Juhos (1). Debo admitir que no entiendo claramente las objeciones de Juhos y que no sé cómo clasificarlas; por esto me limito a ciertos puntos de carácter formal. Von Juhos parece ignorar mi definición de la verdad; sólo se refiere a una presentación informal en Tarski (3), en la que la definición no aparece para nada. Si conociera la definición real tendría que cambiar su argumento. Sin embargo, no dudo de que también en esta definición descubriría algunos «defectos». Pues él cree que ha probado que «por razones de principio es imposible dar tal definición».

den significado si se reemplaza «*p*» por una oración, y no por el nombre de una oración (cf. la sección 4) <sup>26</sup>.

Mientras que el autor de la objeción considera el esquema (V) como «inadmisiblemente breve», por mi parte me inclino a considerar los esquemas (V') y (V'') como «inadmisiblemente largos». Y hasta creo que puedo probar rigurosamente esta afirmación sobre la base de la siguiente definición: Se dice que una expresión es «inadmisiblemente larga» si (I) no es significativa y (II) se ha obtenido a partir de una expresión significativa insertándole palabras superfluas.

16. *Redundancia de términos semánticos; su posible eliminación.* La objeción que me propongo discutir ahora no concierne ya a la corrección formal de la definición, pero con todo trata de ciertos rasgos formales de la concepción semántica de la verdad.

Hemos visto que esta concepción consiste, en esencia, en considerar la oración «*X es verdadera*» como equivalente a la oración denotada por «*X*» (donde «*X*» representa un nombre de una oración del lenguaje-objeto). Por consiguiente, el término «verdadero», cuando aparece en una oración simple de la forma «*X es verdadera*», puede eliminarse fácilmente, y la oración misma, que pertenece al metalenguaje, puede reemplazarse por una oración equivalente del lenguaje-objeto; y lo mismo se aplica a oraciones compuestas siempre que el término «verdadero» figure en ellas exclusivamente como parte de expresiones de la forma «*X es verdadera*».

Por este motivo, algunos han insistido en que el término «verdadero», en el sentido semántico, siempre puede eliminarse, y que por esta razón la concepción semántica de la verdad es del todo estéril e inútil. Y, puesto que las mismas consideraciones se aplican a otras nociones semánticas, se ha sacado la conclusión de que la semántica en su conjunto es un juego puramente verbal y, en el mejor de los casos, sólo un pasatiempo inofensivo.

Pero la cosa no es tan simple <sup>27</sup>. No siempre puede efectuarse esta clase de eliminación. No puede hacerse en el caso de los enunciados universales que expresan el hecho de que todos los enunciados de cierto tipo son verdaderos, o que todas las oraciones verdaderas tienen cierta propiedad. Por ejemplo, en la teoría de la verdad podemos probar el siguiente enunciado:

<sup>26</sup> Las oraciones «*p es verdadera*» y «*ocurre p*» [“*p is the case*”] (o, mejor, «*es verdad que p*» y «*ocurre que p*») se usan a veces en tratamientos informales, principalmente por razones estilísticas; pero se las considera sinónimas de la oración representada por «*p*». En cambio, en la medida en que entiendo la situación, las oraciones en cuestión no pueden ser usadas por Juhos como sinónimas de «*p*»; pues de lo contrario la sustitución de (V) por (V') o (V'') no constituirían ningún «adelanto».

<sup>27</sup> Cfr. la discusión de este problema en Kokoszynska (1), pp. 161 ss.

*Todas las consecuencias de los enunciados verdaderos son verdaderas.*

Sin embargo, no podemos librarnos en este caso de la palabra «verdadera» en la forma sencilla que se ha puesto.

Además, aun en el caso de los enunciados particulares que tienen la forma «*X es verdadera*», semejante eliminación sencilla no puede hacerse siempre. En efecto, la eliminación es posible sólo en aquellos casos en que el nombre del enunciado del que se dice que es verdadero figura en una forma que nos permite reconstruir el enunciado mismo. Por ejemplo, nuestro conocimiento histórico actual no nos da posibilidad de eliminar la palabra «verdadera» de la siguiente oración:

*La primera oración escrita por Platón es verdadera.*

Por supuesto, desde que tenemos una definición de la verdad, y desde que toda definición permite reemplazar el *definiendum* por su *definiens*, siempre es teóricamente posible eliminar el término «verdadero» en un sentido semántico. Pero ésta no sería la eliminación simple aludida anteriormente, y no daría como resultado la sustitución de un enunciado del metalenguaje por un enunciado del lenguaje-objeto.

Sin embargo, si alguien insistiera en que —a causa de la posibilidad teórica de eliminar la palabra «verdadero» sobre la base de su definición— el concepto de verdad es estéril, debe aceptar la conclusión de que todas las nociones definidas son estériles. Pero este resultado es tan absurdo y tan irrazonable históricamente, que no es necesario comentarlo. Por mi parte, me inclino más bien a concordar con quienes sostienen que los momentos de mayor avance creador de la ciencia coinciden con frecuencia con la introducción de nuevas nociones por medio de definiciones.

17. *Conformidad de la concepción semántica de la verdad con los usos filosófico y vulgar.* Se ha suscitado la cuestión de si la concepción semántica de la verdad puede considerarse como una forma precisa de la vieja concepción clásica de esta noción.

En la primera parte de este trabajo se citaron varias formulaciones de la concepción clásica (sección 3). Debo repetir que, a mi juicio, ninguna de ellas es bastante precisa y clara. Por consiguiente, la única manera segura de resolver la cuestión sería confrontar a los autores de aquellos enunciados con nuestra nueva formulación, y preguntarles si ella concuerda con sus intenciones. Desgraciadamente, este método no es practicable, porque dichos autores murieron hace algún tiempo.

En lo que a mí respecta, no tengo duda alguna de que nuestra formulación se conforma al contenido intuitivo de la de Aristóteles. Estoy menos seguro respecto de las formulaciones posteriores de la concepción clásica, pues son, por cierto, muy vagas<sup>28</sup>.

Más aún, se han expresado algunas dudas acerca de si la concepción semántica refleja la noción de verdad en su uso vulgar y cotidiano. Me doy cuenta (como ya lo he señalado) de que el sentido vulgar de la palabra «verdadero» —como el de cualquier otra palabra del lenguaje cotidiano— es hasta cierto punto vago, y que su uso es más o menos fluctuante. Por lo tanto, el problema de asignarle a esta palabra un significado fijo y exacto queda relativamente muy especificado, y toda solución de este problema implica necesariamente cierta desviación respecto de la práctica del lenguaje cotidiano.

A pesar de todo esto, creo que la concepción semántica se conforma en medida considerable al uso vulgar, aunque me apresuro a admitir que puedo estar equivocado. Y, lo que es más pertinente, creo que la cuestión suscitada puede resolverse científicamente, aunque desde luego no mediante un procedimiento deductivo, sino con ayuda del método estadístico de la encuesta. De hecho, semejante investigación se ha llevado a cabo, y de algunos de sus resultados se ha informado a congresos y han sido en parte publicados<sup>29</sup>.

Desearía subrayar que, en mi opinión, semejantes investigaciones deben llevarse a cabo con el máximo cuidado. Por ejemplo, si le preguntáramos a un muchacho de escuela secundaria, o a un adulto inteligente sin preparación filosófica especial, si considera que una oración es verdadera si concuerda con la realidad, o si designa una situación existente, puede resultar simplemente que no comprenda la pregunta; por consiguiente su respuesta, cualquiera que sea, carecerá de valor para nosotros. Pero su respuesta a la pregunta acerca de si admitiría que la oración «*está nevando*» pueda ser verdadera aun cuando no esté nevando, o falsa aunque esté nevando, sería, naturalmente, muy importante para nuestro problema.

Por esto, nada me sorprendió (en una discusión dedicada a estos problemas) enterarme de que en un grupo de personas preguntadas sólo el 15 % concordó en que «verdadero» significa para ellos «concordante con la realidad», en tanto que el 90 % convino en que una oración tal como «*está nevando*» es verdadera si, y sólo si, está nevando. De modo que una gran mayoría de esas personas parecían

<sup>28</sup> La mayoría de los autores que han discutido mi obra sobre la noción de verdad opinan que mi definición se conforma a la concepción clásica de esta noción; véase, p. ej., Kortabinski (2) y Scholz (1).

<sup>29</sup> Cfr. Ness (1). Desgraciadamente, los resultados de la parte de la investigación de Ness que es particularmente importante para nuestro problema no se tratan en su libro; cfr. p. 148, nota 1.

rechazar la concepción clásica de la verdad en su formulación «filosófica», aceptando en cambio la misma concepción cuando se la formulaba en palabras sencillas (haciendo a un lado la cuestión de si se justifica en este lugar el uso de la oración «la misma concepción»).

18. *La definición en su relación con «el problema filosófico de la verdad» y con varias corrientes gnoseológicas.* He oído la observación de que la definición formal de la verdad no tiene nada que ver con «el problema filosófico de la verdad»<sup>30</sup>. Sin embargo, nadie me ha enseñado jamás, en forma inteligible, en qué consiste este problema. Se me ha informado, a este respecto, que mi definición, aunque enuncia condiciones necesarias y suficientes para que una frase sea verdadera, en realidad no aprehende la «esencia» de este concepto. Como nunca he logrado entender lo que es la «esencia» de un concepto, permítaseme abandonar la discusión en este punto.

En general, no creo que exista algo así como «el problema filosófico de la verdad». Creo, en cambio, que hay varios problemas inteligibles e interesantes (pero no necesariamente filosóficos) concernientes a la noción de verdad, pero creo también que pueden formularse exactamente y resolverse, posiblemente, sólo sobre la base de una concepción precisa de esta noción.

Si bien por una parte la definición de la verdad ha sido criticada por no ser suficientemente filosófica, por la otra se le han opuesto una serie de objeciones que la acusan de graves implicaciones filosóficas, todas ellas de naturaleza muy indeseable. Discutiré ahora una objeción especial de este tipo; trataré otro grupo de objeciones de esta clase en la próxima sección.

Se ha sostenido que —a causa de que una oración tal como «la nieve es blanca» se considera semánticamente verdadera si la nieve es de hecho blanca (el subrayado es del crítico)— la lógica se encuentra envuelta en un realismo extremadamente acrítico<sup>31</sup>.

Si yo tuviera la oportunidad de discutir esa objeción con su autor, diría dos cosas. En primer lugar, le pediría que eliminase las palabras «de hecho», que no figuran en la formulación original y que son equívocas, aun cuando no afectan el contenido. Pues estas palabras producen la impresión de que la concepción semántica de la verdad tiene por finalidad establecer las condiciones en que tenemos la garantía de poder afirmar cualquier oración, y en particular cualquier oración

<sup>30</sup> Aunque he oído esta opinión varias veces, sólo una vez la he visto escrita y, lo que por cierto es curioso, en una obra que no tiene un carácter filosófico: en Hilbert-Bernays (1), vol. II, p. 269 (dónde, dicho sea de pasada, no se la expresa como objeción). En cambio, no he encontrado ninguna observación a este respecto en el tratamiento de mi obra por los filósofos profesionales (cfr. nota 1).

<sup>31</sup> Cfr. Gonsseth (1), pp. 187 ss.



empírica. Pero bastará reflexionar brevemente para ver que esta impresión no es sino una ilusión; y creo que el autor de la objeción es víctima de la ilusión que él mismo creó.

En efecto, la definición semántica de la verdad nada implica respecto de las condiciones en que puede afirmarse una oración tal como (1),

*La nieve es blanca.*

Sólo implica que, siempre que afirmamos o rechazamos esta oración debemos estar listos para afirmar o rechazar la oración correlacionada (2),

*La oración «la nieve es blanca» es verdadera.*

De manera que podemos aceptar la concepción semántica de la verdad sin abandonar ninguna actitud gnoseológica que podamos haber tenido; seguimos siendo realistas ingenuos, realistas críticos o idealistas, empiristas o metafísicos: lo que hayamos sido antes. La concepción semántica es completamente neutral respecto de todas esas posiciones.

En segundo lugar, yo trataría de obtener alguna información respecto de la concepción de la verdad que, en opinión del autor de la objeción, no envuelva a la lógica en el más ingenuo de los realismos. Diría que esta concepción debe ser incompatible con la semántica. Por ejemplo, debe haber oraciones que son verdaderas en una de estas concepciones sin ser verdaderas en la otra. Supongamos, v.g., que la oración (1) es de esta clase. La verdad de esta oración está determinada, en la concepción semántica, por una equivalencia de la forma (V):

*La oración «la nieve es blanca» es verdadera si, y sólo si, la nieve es blanca.*

Por consiguiente, en la nueva concepción debemos rechazar esta equivalencia, y por lo tanto, debemos aceptar su negación:

*La oración «la nieve es blanca» es verdadera si, y sólo si, la nieve no es blanca (o quizá, la nieve no es, de hecho, blanca).*

Esto suena a paradoja. No considero absurda semejante consecuencia de la nueva concepción; pero temo un poco que alguien, en el futuro, pueda acusarla de envolver a la lógica en un «irrealismo extremadamente artificioso». En todo caso, me parece importante advertir que toda concepción de la verdad incompatible con la semántica tiene consecuencias de este tipo.

Me he detenido un tanto en esta cuestión, no porque me parezca importante la objeción que hemos tratado, sino porque al discutirla han surgido ciertos puntos que debieran tomar en cuenta todos aquellos que, por diversas razones gnoseológicas, se inclinan a rechazar la concepción semántica de la verdad.

19. *Los supuestos elementos metafísicos de la semántica.* La concepción semántica de la verdad ha sido acusada varias veces de envolver ciertos elementos metafísicos. Se han hecho objeciones de esta clase no sólo a la teoría de la verdad, sino a todo el dominio de la semántica teórica<sup>32</sup>.

No me propongo tratar el problema general de si es objetable la introducción de un elemento metafísico en la ciencia. El único punto que me interesará en este lugar será si, y en qué sentido, está envuelta la metafísica en el tema de nuestra discusión.

Toda la cuestión depende, evidentemente, de lo que se entienda por «metafísica». Por desgracia, esta noción es extremadamente vaga y equívoca. Cuando se escuchan discusiones sobre este tema, a veces se tiene la impresión de que el término «metafísico» ha perdido todo significado objetivo, usándose tan sólo como una especie de jerga filosófica profesional.

Para algunos, la metafísica es una teoría general de los objetos (ontología), una disciplina que debe desarrollarse de una manera puramente empírica, y que difiere de otras ciencias empíricas tan sólo por su generalidad. No sé si realmente existe semejante disciplina (algunos cínicos pretenden que en filosofía es habitual bautizar niños no nacidos); pero creo que, en todo caso, la metafísica así entendida no puede ser objetada por nadie, y apenas tiene conexiones con la semántica.

Pero la mayoría de las veces, el término «metafísico» se usa como directamente opuesto —en uno u otro sentido— al término «empírico»; en todo caso, es usado de esta manera por quienes se inquietan con el pensamiento de que pueda haberse introducido algún elemento metafísico en la ciencia. Esta concepción general de la metafísica toma varias formas más específicas.

Por ejemplo, algunos consideran que es sintomático de la presencia de un elemento metafísico en una ciencia cuando se emplean métodos de investigación que no son deductivos ni empíricos. Pero en el desarrollo de la semántica no pueden encontrarse vestigios de este automa (a menos que estén envueltos algunos elementos metafísicos

<sup>32</sup> Véase Nagel (1) y Nagel (2), pp. 471 ss. Una observación dirigida, tal vez, en la misma dirección, se encuentra también en Weinberg (1), p. 77; véase, sin embargo, sus observaciones anteriores, pp. 75 ss.

en el lenguaje-objeto a que se refieren las nociones semánticas). En particular, la semántica de los lenguajes formalizados se construye de manera puramente deductiva.

Otros sostienen que el carácter metafísico de una ciencia depende principalmente de su vocabulario y, más específicamente, de sus términos primitivos. Así, por ejemplo, se dice que un término es metafísico si no es lógico ni matemático, y si no está asociado con un procedimiento empírico que nos permita decidir si una cosa es denotada por este término, o no. Con respecto a esta opinión sobre la metafísica, baste recordar que un metalenguaje sólo incluye tres clases de términos indefinidos: (I) términos tomados de la lógica, (II) términos del lenguaje-objeto correspondiente, y (III) nombres de expresiones del lenguaje-objeto. Es, pues, obvio que en el metalenguaje no figuran términos indefinidos de índole metafísica (a menos, nuevamente, que tales términos aparezcan en el propio lenguaje-objeto).

Hay, sin embargo, quienes creen que, aun cuando no figuren términos metafísicos entre los términos primitivos de un lenguaje, pueden introducirse por definición; a saber, mediante aquellas definiciones que no nos proveen de criterios generales para decidir si un objeto cae dentro del concepto definido. Se arguye que el término «verdadero» es de esta clase, ya que ningún criterio universal de verdad se deduce en forma inmediata de la definición de este término, y ya que se cree generalmente (y en cierto sentido hasta pudo probarse) que jamás se encontrará semejante criterio. Este comentario sobre el carácter real de la noción de verdad parece perfectamente justo. Sin embargo, debe advertirse que la noción de verdad no difiere, a este respecto, de muchas nociones de la lógica, de la matemática, y de las partes teóricas de diversas ciencias empíricas, p. ej., de la física teórica.

En general, es preciso decir que si el término «metafísico» se emplea en un sentido tan amplio que abarque ciertas nociones (o métodos) de la lógica, de la matemática o de las ciencias empíricas, se aplicará *a fortiori* a aquellas de la semántica. En efecto, como ya lo sabemos por la Parte I de este trabajo, al desarrollar la semántica de un lenguaje usamos todas las nociones de este lenguaje, y aplicamos un aparato lógico aun más poderoso que el que se usa en el lenguaje mismo. Por otra parte, puedo resumir los argumentos expuestos anteriormente, afirmando que en ninguna de las interpretaciones del término «metafísico» que me son familiares o más o menos inteligibles, envuelve la semántica términos metafísicos que le sean peculiares.

Haré una última observación en relación con este grupo de objeciones. La historia de la ciencia muestra muchos ejemplos de conceptos que fueron juzgados metafísicos (en un sentido vago, pero en

todo caso despectivo de este término) antes que fuera precisado su sentido; pero una vez que recibieron una definición rigurosa, formal, se evaporó la desconfianza que se les tenía. Como ejemplos típicos podemos mencionar los conceptos de números negativos e imaginarios en la matemática. Espero que el concepto de verdad y otros conceptos semánticos tengan un destino similar; y me parece, por lo tanto, que quienes han desconfiado de dichos conceptos a causa de sus presuntas implicaciones metafísicas debieran acoger con agrado el hecho de que se dispone ahora de definiciones precisas de ellos. Si a consecuencia de esto los conceptos semánticos perdiesen interés filosófico, no harían sino compartir el destino de muchos otros conceptos científicos, lo que no es de lamentar.

20. *Aplicabilidad de la semántica a las ciencias empíricas especiales.* Llegamos a las objeciones del último y acaso del más importante de los grupos. Se han expresado algunas fuertes dudas acerca de si las nociones semánticas tienen o pueden encontrar aplicaciones en varios dominios de la actividad intelectual. En su mayoría, estas dudas han concernido a la aplicabilidad de la semántica al campo de la ciencia empírica, sea a las ciencias especiales o a la metodología general de este campo; aunque se ha expresado un escepticismo similar con respecto a las posibles aplicaciones de la semántica a las ciencias matemáticas y a su metodología.

Creo que es posible calmar un tanto estas dudas, y que no carece de fundamento cierto optimismo respecto del valor potencial de la semántica para varios dominios del pensamiento.

Para justificar este optimismo, creo que basta subrayar dos puntos bastante obvios. En primer lugar, el desarrollo de una teoría que formula una definición precisa de una noción y establece sus propiedades generales provee, *eo ipso*, de una base más firme para todas las discusiones en que se halle envuelta dicha noción; por esto, no puede ser indiferente para nadie que use esa noción y desee hacerlo de manera consciente y coherente. En segundo lugar, las nociones semánticas están de hecho comprendidas en varias ramas de la ciencia, y en particular de la ciencia empírica.

El hecho de que en la investigación empírica sólo tratemos con lenguajes naturales, y que la semántica teórica se aplique a estos lenguajes sólo con cierta aproximación, no afecta esencialmente al problema. Sin embargo, tiene sin duda la consecuencia de que el progreso de la semántica tendrá una influencia retardada y algo limitada a este campo. Esta situación no difiere esencialmente de la que se presenta cuando aplicamos las leyes de la lógica a las discusiones de la vida diaria o, en general, cuando intentamos aplicar la ciencia teórica a los problemas empíricos.

En la psicología, la sociología y prácticamente en todas las huma-

nidades están envueltas, en mayor o menor grado, nociones semánticas. Así, por ejemplo, un psicólogo define el llamado cociente de inteligencia en términos del número de respuestas *verdaderas* (correctas) y *falsas* (incorrecta) que da una persona a ciertas preguntas; para un historiador de la cultura, puede ser de gran importancia el dominio de los objetos para los cuales una raza humana, en etapas sucesivas de su desenvolvimiento, posee *designaciones* adecuadas; un estudioso de la literatura puede estar intensamente interesado en el problema de si un autor dado siempre usa dos palabras dadas con el mismo *significado*. Los ejemplos de este tipo pueden multiplicarse indefinidamente.

El dominio más natural y promisorio para la aplicación de la semántica teórica es, claramente, la lingüística, esto es, el estudio empírico de los lenguajes naturales. Ciertas partes de esta ciencia se llaman incluso «semántica», a veces con un calificativo. Ocasionalmente se le da este nombre a ese trozo de la gramática que intenta clasificar todas las palabras de un lenguaje en partes de la oración, según lo que significan o designan las palabras. A veces se llama «semántica histórica» al estudio de la evolución de los significados en el desarrollo histórico de un lenguaje. En general, la totalidad de las investigaciones sobre relaciones semánticas que figuran en un lenguaje natural se denomina «semántica descriptiva». La relación entre la semántica teórica y la descriptiva es análoga a la que existe entre la matemática pura y la aplicada, o quizás a la que existe entre la física teórica y la experimental; el papel que desempeñan los lenguajes formalizados en la semántica puede compararse *grosso modo* al de los sistemas aislados en física.

Acaso sea innecesario decir que la semántica no puede encontrar aplicación directa alguna en las ciencias naturales tales como la física, la biología, etc.; pues en ninguna de estas ciencias tratamos con fenómenos lingüísticos, y aun menos con relaciones semánticas entre expresiones lingüísticas y objetos a que se refieren estas expresiones. En la sección siguiente veremos, sin embargo, que la semántica puede tener una especie de influencia indirecta sobre aquellas ciencias en que no intervienen directamente las nociones semánticas.

21. *Aplicabilidad de la semántica a la metodología de las ciencias empíricas.* Además de la lingüística, otro importante dominio de posibles aplicaciones de la semántica es la metodología de la ciencia; este término se usará aquí en un sentido amplio, que abarque la teoría de la ciencia en general. Independientemente de si la ciencia se concibe meramente como un sistema de enunciados o como una totalidad de ciertos enunciados y actividades humanas, el estudio del lenguaje científico constituye una parte esencial del tratamiento metodológico de una ciencia. Y me parece claro que cualquier tendencia a

eliminar las nociones semánticas (tales como las de verdad y designación) de esta discusión la haría fragmentaria e inadecuada<sup>33</sup>. Más aún, tal tendencia no tiene razón de ser hoy día, cuando se han superado las principales dificultades que presenta el uso de los términos semánticos. La semántica del lenguaje científico debiera incluirse simplemente como parte de la metodología de la ciencia.

No me inclino, de modo alguno, a encargar a la metodología y, en particular, a la semántica —sea teórica o descriptiva— la tarea de aclarar los significados de todos los términos científicos. Esta tarea se deja a las ciencias que usan los términos, y en realidad es cumplida por ellas (de la misma manera en que, p. ej., la tarea de aclarar el significado del término «verdadero» se deja a la semántica, la que la lleva a cabo). Sin embargo, puede haber ciertos problemas especiales de esta clase, en que es deseable un enfoque metodológico, o incluso en que éste es necesario (quizás el problema de la noción de causalidad sea un buen ejemplo de esto); y en una discusión metodológica de semejantes problemas, las nociones semánticas pueden desempeñar un papel esencial. Así, pues, la semántica puede tener alguna influencia sobre cualquiera de las ciencias.

Se presenta el problema de si la semántica puede ayudar a resolver problemas generales y, por decirlo así, clásicos de la metodología. Trataré con algún detalle un aspecto especial, aunque muy importante, de esta cuestión.

Uno de los principales problemas de la metodología de la ciencia empírica consiste en establecer las condiciones en que puede considerarse aceptable una teoría o una hipótesis empírica. Esta noción de aceptabilidad debe hacerse relativa a una etapa dada del desarrollo de una ciencia (o a un cierto cúmulo de conocimiento). En otras palabras, podemos considerarla provista de un coeficiente dependiente del tiempo; pues una teoría aceptable hoy, puede ser insostenible mañana como resultado de nuevos descubrimientos científicos.

Parece *a priori* muy plausible que la aceptabilidad de una teoría depende de alguna manera de la verdad de sus enunciados, y que por consiguiente un metodólogo, en sus (hasta ahora bastante infructuosos) intentos de precisar la noción de aceptabilidad, puede esperar alguna ayuda de la teoría semántica de la verdad. Por consiguiente, nos preguntamos: ¿Hay algún postulado que pueda imponerse razonablemente a las teorías aceptables y que envuelva la noción de verdad? Y, en particular, nos preguntamos si es razonable el siguiente postulado:

<sup>33</sup> Esta tendencia era evidente en obras anteriores de Carnap (véase, p. ej., Carnap (1), especialmente Parte V) y en escritos de otros miembros del Círculo de Viena. Con respecto a este respecto Kokoszynska (1) y Weinberg (1).

*Una teoría aceptable no puede contener (o implicar) enunciados falsos alguno.*

La respuesta a esta última pregunta es claramente negativa. Pues, ante todo, estamos prácticamente seguros —sobre la base de nuestra experiencia histórica— que toda teoría empírica aceptada hoy será tarde o temprano rechazada o reemplazada por otra teoría. También es muy probable que la nueva teoría sea incompatible con la vieja, es decir, implicará un enunciado contradictorio con uno de los enunciados contenidos en la vieja teoría. Por lo tanto, al menos una de las dos teorías debe incluir enunciados falsos, pese al hecho de que cada una de ellas es aceptada en cierto momento. En segundo lugar, el postulado en cuestión difícilmente podría ser satisfecho en la práctica; pues no conocemos, y es muy improbable que los encontremos, criterios de verdad que nos permitan mostrar que ningún enunciado de una teoría empírica es falso.

El postulado en cuestión podría considerarse, a lo sumo, como la expresión de un ideal de teorías sucesivamente más adecuadas en un dominio dado de la investigación; pero a esto apenas se le puede dar un significado preciso.

Sin embargo, me parece que hay un importante postulado que puede imponerse razonablemente a las teorías empíricas aceptables, y que envuelve la noción de verdad. Está estrechamente relacionado con el que acabamos de tratar, pero es esencialmente más débil. Recordando que la noción de aceptabilidad está dotada de un coeficiente temporal, podemos darle a este postulado la siguiente forma:

*Tan pronto como logramos mostrar que una teoría empírica contiene (o implica) frases falsas, ya no puede considerarse aceptable.*

En apoyo de este postulado quisiera hacer las siguientes observaciones.

Creo que todo el mundo concuerda en que una de las razones que pueden obligarnos a rechazar una teoría empírica es la prueba de su incoherencia [*inconsistency*]: una teoría se torna insostenible si logramos deducir de ella dos frases contradictorias. Ahora podemos preguntar cuáles son los motivos usuales para rechazar una teoría por tales motivos. Quienes están familiarizados con la lógica moderna se inclinan a responder a esta cuestión de la siguiente manera: Una conocida ley lógica muestra que una teoría que nos permite deducir dos frases contradictorias también nos permite deducir cualquier enunciado; por consiguiente, tal teoría es trivial y carece de interés científico.

Tengo algunas dudas de que esta respuesta contenga un análisis adecuado de la situación. Creo que las personas que no conocen lógica

La moderna se inclinan tan poco a aceptar una teoría incoherente como quienes están totalmente familiarizados con ella; y probablemente esto se aplique incluso a quienes consideran (como aún ocurre en algunos) que la ley lógica sobre la que se basa el argumento es altamente controvertible y casi paradójica. No creo que cambiara nuestra actitud para con una teoría incoherente aun cuando decidiéramos, por alguna razón, debilitar nuestro sistema lógico privándonos de la posibilidad de deducir todo enunciado a partir de dos enunciados contradictorios cualesquiera.

Me parece que la auténtica razón de nuestra actitud es diferente, o al menos sabemos (aunque sólo sea intuitivamente) que una teoría incoherente debe contener ciertos enunciados falsos; y no nos inclinamos a considerar como aceptable ninguna teoría acerca de la cual se haya demostrado que contiene enunciados de esa clase.

Hay varios métodos para mostrar que una teoría dada incluye enunciados falsos. Algunos se fundan sobre propiedades puramente lógicas de la teoría en cuestión; el método que acabamos de tratar (esto es, la prueba de la incoherencia) no es el único método de este tipo, pero es el más simple y el que se aplica con mayor frecuencia en la práctica. Con ayuda de ciertas suposiciones referentes a la verdad de los enunciados empíricos, podemos obtener métodos que tienen la misma finalidad pero que no son de naturaleza puramente lógica. Si decidimos aceptar el postulado general sugerido más arriba, una aplicación exitosa de cualquiera de estos métodos tornará insostenible a la teoría.

22. *Aplicaciones de la semántica a la ciencia deductiva.* En lo que respecta a la aplicabilidad de la semántica a las ciencias matemáticas y a su metodología, esto es, a la matemática, estamos en una posición mucho más favorable que en el caso de las ciencias empíricas. Pues, en lugar de proponer razones que justifiquen algunas esperanzas para el futuro (haciendo así una especie de propaganda en favor de la semántica), podemos señalar resultados concretos que ya se han alcanzado.

Siguen expresándose dudas acerca de si la noción de enunciado verdadero —a diferencia de la de enunciado comprobable— puede tener importancia para las disciplinas matemáticas y desempeña algún papel en las discusiones acerca de la metodología de la matemática. Me parece, sin embargo, que precisamente esta noción de enunciado verdadero constituye una valiosísima contribución de la semántica a la metamatemática. Ya poseemos una serie de interesantes resultados metamatemáticos obtenidos con ayuda de la teoría de la verdad. Estos resultados conciernen a las relaciones mutuas entre la noción de verdad y la de comprobabilidad; establecen nuevas propiedades de esta última noción (que, como es sabido, es una de las



nociones básicas de la metamatemática), y echan alguna luz sobre los problemas fundamentales de la coherencia y de la completitud. Los más importantes de estos resultados ya fueron considerados brevemente en la sección 12<sup>34</sup>.

Más aun, aplicando el método semántico podemos definir adecuadamente diversas nociones metamatemáticas de importancia que hasta ahora se han usado solamente en forma intuitiva; tales la noción de definibilidad o la de modelo de un sistema axiomático. De esta manera podemos encarar un estudio sistemático de estas nociones. En particular, las investigaciones sobre la definibilidad ya han producido algunos resultados interesantes, y prometen más para el futuro<sup>35</sup>.

Hemos tratado las aplicaciones de la semántica a la metamatemática y no a la matemática propiamente dicha. Pero esta distinción entre matemática y metamatemática no tiene gran importancia. Pues la propia metamatemática es una disciplina deductiva y, por consiguiente, desde cierto punto de vista, es parte de la matemática; y es bien sabido que —a causa del carácter formal del método deductivo— los resultados que se obtienen en una disciplina deductiva pueden extenderse automáticamente a cualquier otra disciplina en que la disciplina dada encuentre una interpretación. Así, por ejemplo, todos los resultados metamatemáticos pueden interpretarse como resultados de la teoría de los números. Tampoco desde el punto de vista práctico existe una nítida línea divisoria entre la metamatemática y la matemática propiamente dicha; por ejemplo, las investigaciones sobre la definibilidad podrían incluirse en cualquiera de estos dominios.

### 23. *Observaciones finales.* Deseo concluir esta discusión con al

<sup>34</sup> Para otros resultados obtenidos con ayuda de la teoría de la verdad, véase Gödel (2); Tarski (2), pp. 401 ss.; y Tarski (5), pp. 111 ss.

<sup>35</sup> Un objeto —p. ej., un número o un conjunto de números— se dice definible (en cierto formalismo) si existe una función proposicional que lo define; cf. nota 20. Por consiguiente, el término «definible», aunque de origen matemático (semántico), es puramente matemático en lo que respecta a su extensión, puesto que expresa una propiedad (denota una clase) de objetos matemáticos. Por consiguiente, la noción de definibilidad puede redefinirse en términos puramente matemáticos, aunque no dentro de la disciplina formalizada a que se refiere esta noción; con todo, la idea fundamental de la definición no cambia. Ver a este respecto, y también para mayores referencias bibliográficas, Tarski (1). En la literatura pueden encontrarse varios otros resultados concernientes a la definibilidad; p. ej., en Hilbert-Bernays (1), vol. I, pp. 354 ss., 400 ss., 456 ss., etc., y en Lindenbaum-Tarski (1). Obsérvese que el término «definible» se usa a veces en otro sentido, metamatemático pero no semántico; esto ocurre, por ejemplo, cuando decimos que un término es definible en otros términos (sobre la base de un sistema axiomático dado). Para una definición de modelo de un sistema axiomático, véase Tarski (4).

gunas observaciones generales y más bien libres acerca de la cuestión de la evaluación de las conquistas de la ciencia en términos de su aplicabilidad. Debo confesar que tengo varias dudas a este respecto.

Por ser matemático (y también lógico, y acaso filósofo de cierta especie), he tenido oportunidad de asistir a muchas discusiones entre especialistas en matemática, donde el problema de la aplicación es especialmente agudo, y he observado en varias ocasiones el siguiente fenómeno: si un matemático desea disminuir la importancia de la obra de uno de sus colegas, digamos A, el mejor método que encuentra para hacerlo es preguntarle a qué pueden aplicarse sus resultados. El interrogado, puesto entre la espada y la pared, termina por desenterrar las investigaciones de otro matemático, B, como el lugar de las aplicaciones de sus propios resultados. Si a su vez B es sometido al mismo interrogatorio, se referirá a otro matemático C. Después de unos pocos pasos de esta clase se vuelve a hacer referencia a las investigaciones de A, cerrándose así la cadena.

Hablando más seriamente, no negaré que el valor de la obra de un hombre pueda aumentar por sus implicaciones para la investigación de otros y para la práctica. Creo, sin embargo, que es contrario al progreso de la ciencia medir la importancia de investigación alguna exhaustiva o primordialmente en términos de su utilidad y aplicabilidad. Sabemos, por la historia de la ciencia, que muchos resultados y descubrimientos de importancia hubieron de esperar siglos hasta recibir aplicación en algún campo. Y, en mi opinión, hay otros factores de importancia que no pueden dejarse de lado al determinar el valor de una obra científica. Me parece que hay un dominio especial de necesidades humanas muy profundas e intensas, que están relacionadas con la investigación científica, y que son en muchos aspectos similares a las necesidades estéticas y acaso religiosas. Y también me parece que la satisfacción de estas necesidades debiera considerarse como una importante tarea de la investigación. Por consiguiente, creo que la cuestión del valor de una investigación cualquiera no puede contestarse adecuadamente sin tener en cuenta la satisfacción intelectual que producen los resultados de esa investigación a quienes la comprenden y estiman. Acaso sea impopular y anticuado decirlo, pero no creo que un resultado científico que nos dé una mejor comprensión del mundo y lo haga más armonioso a nuestros ojos deba tenerse en menos que, por ejemplo, una invención que reduzca el costo de la pavimentación de los caminos o mejore las instalaciones sanitarias del hogar.

Está claro que las observaciones que acabo de hacer son inútiles si se usa la palabra «aplicación» en un sentido muy amplio y liberal. No es menos obvio, quizá, que nada se deduce, de estas observaciones generales, que concierna a los tópicos específicos que se han tratado en este trabajo; y realmente no sé si la investigación semántica puede

ganar o perder con la introducción del patrón de valor que he sugerido.

#### BIBLIOGRAFÍA

Sólo se da la lista de los libros y artículos a que se hace referencia en este trabajo.

- ARISTÓTELES (1): *Metaphysica (Works*, vol. VIII, traducción inglesa de W. D. Ross], Oxford, 1908. [Versión castellana: Gredos, Madrid, 1970.]
- CARNAP, R. (1): *Logical Syntax of Language*, Londres/Nueva York, 1937.
- CARNAP, R. (2): *Introduction to Semantics*, Cambridge, 1942.
- GÖDEL, K. (1): «Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme, I», en *Monatshefte für Mathematik und Physik*, vol. XXXVIII, 1931, pp. 173-198. [Versión castellana: *Cuadernos Teorema*, Valencia, 1980.]
- GÖDEL, K. (2): «Über die Länge von Beweisen», en *Ergebnisse eines mathematischen Kolloquiums*, vol. VII, 1936, pp. 23-24. [Versión castellana en K. Gödel, *Obras completas*, Alianza, Madrid, 1980.]
- GONSETH, F. (1): «Le Congrès Descartes. Questions de Philosophie Scientifique», en *Revue Thomiste*, vol. XLIV, 1938, pp. 183-193.
- GRELLING, K., y NELSON, L. (1): «Bemerkungen zu den Paradoxien von Russell und Burali-Forti», en *Abhandlungen der Fries'schen Schule*, vol. II (nueva serie), 1908, pp. 301-334.
- HOFSTADTER, A. (1): «On Semantic Problems», en *The Journal of Philosophy*, vol. XXXV, 1938, pp. 225-232.
- HILBERT, D., y BERNAYS, P. (1): *Grundlagen der Mathematik*, 2 vols., Berlin, 1934-1939.
- JUHOS, B. von (1): «The Truth of Empirical Statements», en *Analysis*, vol. IV, 1937, pp. 65-70.
- KOKOSZYNSKA, M. (1): «Über den absoluten Wahrheitsbegriff und einige andere semantische Begriffe», en *Erkenntnis*, vol. VI, 1936, pp. 143-165.
- KOKOSZYNSKA, M. (2): «Syntax, Semantik und Wissenschaftslogik», en *Actes du Congrès International de Philosophie Scientifique*, vol. III, pp. 9-14, Paris, 1936.
- KOTARBINSKI, T. (1): *Elementy teorii poznania, logiki formalnej i metodologii nauk (Elementos de teoría del conocimiento, lógica formal y metodología de las ciencias)*, en polaco, Lwów, 1929.
- KOTARBINSKI, T. (2): «W sprawie pojęcia prawdy» («Sobre el concepto de verdad»), en *Przegląd filozoficzny*, vol. XXXVII, 1934, pp. 85-91. [Traducción castellana en *Cuadernos de Epistemología*, n.º 9, Buenos Aires, 1959.]
- LINDENBAUM, A., y TARSKI, A. (1): «Über die Beschränktheit der Ausdrucksmittel deduktiver Theorien», en *Ergebnisse eines mathematischen Kolloquiums*, vol. VII, 1936, pp. 15-23.
- NAGEL, E. (1): Reseña de Hofstadter (1), en *The Journal of Symbolic Logic*, vol. III, 1938, p. 90.
- NAGEL, E. (2): Reseña de Carnap (2), en *The Journal of Philosophy*, vol. XXXIX, 1942, pp. 468-473.
- NESS, A. (1): «"Truth" as Conceived by Those who are not Professional Philosophers», en *Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo, II. Hist.-Filos. Klasse*, vol. IV, Oslo, 1938.
- NEURATH, O. (1): «Erster Internationaler Kongress für Einheit der Wissenschaft in Paris 1935», en *Erkenntnis*, vol. V, 1935, pp. 377-406.

- MOSKILL, B. (1): *An Inquiry into Meaning and Truth*, Nueva York, 1940.
- REIDOLZ, H. (1): Reseña de *Studia Philosophica*, vol. I, en *Deutsche Literaturzeitung*, vol. LVIII, 1937, pp. 1914-1917.
- ŁAŚKI, A. (1): «Sur les ensembles définissables des nombres réels.», en *Fundamenta Mathematicae*, vol. XVII, 1931, pp. 210-239.
- ŁAŚKI, A. (2): «Der Wahrheitsbegriff in der formalisierten Sprachen» (traducción alemana de un libro en polaco, 1933), en *Studia Philosophica*, vol. I, 1935, pp. 261-405 [incluido en *Logic, Semantics, Methamathematics*, Oxford, 1956.]
- ŁAŚKI, A. (3): «Grundlegung der wissenschaftlichen Semantik», en *Actes du Congrès International de Philosophie Scientifique*, vol. III, París, 1936, pp. 1-8.
- ŁAŚKI, A. (4): «Über den Begriff der logischen Folgerung», en *Actes du Congrès International de Philosophie Scientifique*, vol. VII, París, 1937, pp. 1-11.
- ŁAŚKI, A. (5): «On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth», en *The Journal of Symbolic Logic*, vol. IV, 1939, pp. 105-112.
- ŁAŚKI, A. (6): *Introduction to Logic*, Nueva York, 1941. [Traducción castellana: *Introducción a la lógica y a la metodología de las ciencias deductivas*, Buenos Aires, 1951.]
- WEINBERG, J. (1): Reseña de *Studia Philosophica*, vol. I, en *The Philosophical Review*, vol. XLVII, pp. 70-77.

## VERDAD Y SIGNIFICADO \*

DONALD DAVIDSON

Se concede por muchos filósofos del lenguaje, y recientemente incluso por algunos lingüistas, que una teoría satisfactoria del significado debe dar una explicación de cómo los significados de las oraciones dependen de los significados de las palabras. A menos que tal explicación pueda proporcionarse para un lenguaje particular, se arguye, no se habrá explicado el hecho de que podemos aprender el lenguaje: no se habrá explicado el hecho de que, al dominar un vocabulario finito y un conjunto finitamente enunciado de reglas, estamos preparados para producir y entender cualquiera de un número potencialmente infinito de oraciones. No discuto estas vagas afirmaciones, en las que percibo más que una porción seminal de verdad.<sup>1</sup> En su lugar quiero plantear qué es para una teoría dar una explicación del género bosquejado.

Una propuesta es comenzar asignando a cada palabra (u otro rasgo sintáctico significativo) de la oración alguna entidad como su significado; así podríamos asignar Teeteto a «Teeteto» y la propiedad de volar a «vuela» en la oración «Teeteto vuela». Surge entonces el problema de cómo a partir de esos significados se genera el significado de la oración. Contemplando la concatenación como un fragmento significativo de la sintaxis, podríamos asignarle a ella la relación de participar en o instanciar; es, sin embargo, obvio que tenemos aquí el comienzo de un regreso al infinito. Frege buscaba evitar el regreso diciendo que las entidades que corresponden a predicados (por ejemplo) eran «insaturadas» o «incompletas» en con-

\* Versión castellana de Luis M. Valdés Villanueva.

Una primera versión de este artículo fue leída en la reunión de la *Eastern Division* de la *American philosophical Association* en diciembre de 1966; el tema principal se remonta a un artículo no publicado presentado en la *Pacific Division* de la *American Philosophical Association* en 1953. Las formulaciones presentes deben mucho a John Wallace, con el que he discutido estos asuntos desde 1962. Mi investigación fue apoyada por la *National Science Foundation*.

<sup>1</sup> En otras partes he defendido insistentemente que es una condición necesaria para un lenguaje ha de ser aprendible, que tenga solamente un número finito de primitivos semánticos: véase «Theories of Meaning and learnable Languages», en *Proceedings of the 1964 International Congress for logic, methodology and philosophy of science*, North Holland Publishing Company, Amsterdam, 1965, pp. 383-394.

traste con las entidades que corresponden a nombres. Pero esta doctrina parece rotular una dificultad más bien que resolverla.

Lo que quiero decir se pondrá de manifiesto si pensamos por un momento en los términos singulares complejos, a los que juntamente con las oraciones se aplica la teoría de Frege. Considérese la expresión «el padre de Annette»; ¿cómo depende el significado del todo del significado de las partes? La respuesta parecería ser que el significado de «el padre de» es tal que, cuando esta expresión es prefijada a un término singular, el resultado se refiere al padre de la persona a la que se refiere el término singular. ¿Qué papel desempeña, en esta explicación, la entidad insaturada o incompleta por la que está «el padre de»? Todo lo que podemos pensar en decir es que esta entidad «proporciona» o «da» al padre de  $x$  como valor cuando el argumento es  $x$ , o quizás que esta entidad pone en correspondencia personas con sus padres. Puede no estar claro si la entidad por la que se dice que está «el padre de» realiza alguna función explicativa genuina en la medida en que permanecemos aferrados a expresiones individuales; pensemos pues, en cambio, en la clase infinita de expresiones formadas al escribir cero o más veces «el padre de» delante de «Annette». Es fácil proporcionar una teoría que diga, para uno de esos términos singulares arbitrarios, a qué se refiere: si el término es «Annette», se refiere a Annette, mientras que, si el término es complejo, constando de «el padre de» prefijado a un término singular  $t$ , entonces se refiere al padre de la persona a la que  $t$  se refiere. Es obvio que ninguna entidad correspondiente a «el padre de» es, o necesita ser, mencionada al enunciar esta teoría.

Sería inapropiado quejarse de que esta pequeña teoría *usa* las palabras «el padre de» al dar la referencia de expresiones que contienen esas palabras. Pues la tarea era dar el significado de todas las expresiones de un cierto conjunto infinito sobre las bases del significado de las partes; no estaba en el trato dar también los significados de las partes atómicas. Por otro lado, es ahora evidente que una teoría satisfactoria del significado de las expresiones complejas puede no exigir entidades como significado de todas las partes. Nos incumbe entonces reformular nuestra exigencia impuesta a una teoría satisfactoria del significado de modo que no sugiera que las palabras individuales tienen que tener significados en absoluto, en algún sentido que trascienda al hecho de que tienen un efecto sistemático sobre el significado de las oraciones en las que ocurren. En efecto, para el caso que traemos entre manos podemos hacerlo aún mejor al enunciar el criterio de éxito: lo que queríamos, y lo que obtuvimos, es una teoría que entraña toda oración de la forma « $t$  se refiere a  $x$ » donde « $t$ » es reemplazada por una descripción estructural<sup>2</sup> de un tér-

<sup>2</sup> Una «descripción estructural» de una expresión describe la expresión como una

mino singular, y « $x$ » es reemplazada por ese término mismo. Además nuestra teoría cumple esto sin apelar a ningún concepto semántico más allá del básico «se refiere a». Finalmente, la teoría sugiere claramente un procedimiento efectivo para determinar, para cualquier término singular de su universo, a qué se refiere ese término.

Una teoría con tales méritos evidentes merece una aplicación más amplia. El dispositivo propuesto por Frege para este fin tiene una brillante simplicidad: contar los predicados como una clase especial de expresiones funcionales, y las oraciones como un caso especial de términos singulares complejos. Ahora bien, se vislumbra sin embargo una dificultad si continuamos en nuestra presente trayectoria (implícita) de identificar el significado de un término singular con su referencia. La dificultad se sigue de hacer dos suposiciones razonables: que términos singulares lógicamente equivalentes tienen la misma referencia; y que un término singular no cambia su referencia si un término singular contenido es reemplazado por otro con la misma referencia. Pero supongamos ahora que « $R$ » y « $S$ » abrevian cualesquiera dos oraciones con el mismo valor de verdad. Entonces las cuatro oraciones siguientes tienen la misma referencia:

- (1)  $R$
- (2)  $\hat{x}(x=x.R)=\hat{x}(x=x)$
- (3)  $\hat{x}(x=x.S)=\hat{x}(x=x)$
- (4)  $S$

Pues (1) y (2) son lógicamente equivalentes, como lo son (3) y (4), mientras que (3) difiere solamente de (2) en que contiene el término singular « $\hat{x}(x=x.S)$ », donde (2) contiene « $\hat{x}(x=x.R)$ » y éstos se refieren a la misma cosa si  $S$  y  $R$  tienen el mismo valor de verdad. Por consiguiente, cualesquiera dos oraciones tienen la misma referencia si tienen el mismo valor de verdad<sup>3</sup>. Y, si el significado de una oración es aquello a lo que se refiere, todas las oraciones que tienen el mismo valor de verdad tienen que ser sinónimas —un resultado inteligible—.

Aparentemente tenemos que abandonar el presente enfoque en tanto que conduciéndonos a una teoría del significado. Éste es el punto natural al que volver en busca de ayuda a la distinción entre significado y referencia. El problema, se nos dice, es que las cuestiones

concatenación de elementos extraídos de una lista finita (por ejemplo, de palabras o letras).

<sup>3</sup> El argumento es esencialmente de Frege. Véase A. Church, *Introduction to Mathematical Logic*, vol. I, Princeton, 1956, pp. 24-25. Vale la pena quizás mencionar que el argumento no depende de ninguna identificación particular de las entidades a las que se supone que se refieren las oraciones.

de referencia se establecen, en general, por medio de hechos extralingüísticos; las cuestiones de significado no, y los hechos pueden mostrar las referencias de expresiones que no son sinónimas. Si queremos una teoría que dé el significado (como distinto de la referencia) de cada oración, tenemos que comenzar con el significado (como distinto de la referencia) de las partes.

Hasta aquí hemos estado siguiendo las huellas de Frege; gracias a él el sendero es bien conocido e incluso está bien gastado. Pero ahora, me gustaría sugerir, hemos alcanzado un *impasse*: el cambio de referencia a significado no lleva a ninguna explicación provechosa de cómo los significados de las oraciones dependen de los significados de las palabras (u otros rasgos estructurales) que las componen. Pregúntese, por ejemplo, por el significado de «Teeteto vuela». Una respuesta fregeana podría decir algo parecido a esto: dado el significado de «Teeteto» como argumento, el significado de «vuela» proporciona el significado de «Teeteto vuela» como valor. La vacuidad de esta respuesta es obvia. Queríamos saber cuál es el significado de «Teeteto vuela»; no constituye ningún progreso el que se nos diga que es el significado de «Teeteto vuela». Todo esto ya lo sabíamos antes que teoría alguna estuviera a la vista. En la explicación no genuina que se acaba de dar, hablar de la estructura de la oración y del significado de las palabras era ocioso, pues esto no desempeñaba ningún papel en la descripción dada del significado de la oración.

El contraste que hay aquí entre una explicación real y otra aparentemente será aún más claro si reclamamos una teoría, análoga a la teoría en miniatura para los términos singulares que acabamos de bosquejar, pero diferente por el hecho de que trata de significados en lugar de referencias. Lo que la analogía exige es una teoría que tenga como consecuencias todas las oraciones de la forma «*o* significa *m*», donde «*o*» es reemplazada por una descripción estructural de una oración y «*m*» es reemplazada por un término singular que se refiere al significado de esa oración; una teoría que, por otra parte, proporciona un método efectivo para llegar al significado de una oración arbitraria descrita estructuralmente. Claramente si estos criterios han de ser cumplidos <sup>4</sup> es esencial algún modo más articulado de referirse a los significados. Los significados como entidades, o el concepto relacionado de sinonimia, nos permite formular la regla siguiente que pone

<sup>4</sup> Puede pensarse que Church, en «A Formulation of the Logic of Sense and Denotation», en *Structure, Method and Meaning: Essays in Honor of H. M. Sheffer* (editado por Henle, Kallen y Langer), Liberal Arts Press, Nueva York, 1951, pp. 3-24, ha dado una teoría de significado que hace uso esencial de significados como entidades. Pero esto no es el caso: las lógicas del sentido y la denotación de Church se interpretan como siendo sobre significados, pero no mencionan expresiones y así no pueden ser, desde luego, teorías del significado en el sentido que está ahora en discusión.



en relación oraciones y sus partes: son sinónimas aquellas oraciones cuyas partes correspondientes son sinónimas («correspondientes» necesita aquí, desde luego, aclararse detalladamente). Y los significados como entidades pueden, en teorías tales como la de Frege, actuar de vez en cuando como referencias, perdiendo entonces su *status* como entidades distintas de las referencias. Paradójicamente, la única cosa que los significados no parecen hacer es engrasar las ruedas de una teoría del significado —al menos en la medida en que exigimos de tal teoría que dé de manera no trivial el significado de toda oración del lenguaje—. Mi objeción a los significados en la teoría de identidad no es que son abstractos o que sus condiciones de identidad son oscuras, sino que no tienen un uso demostrado.

Éste es el lugar para descartar otro prometedor pensamiento. Supóngase que tenemos una teoría satisfactoria de la sintaxis para nuestro lenguaje, que conste de un método efectivo de decir, para una expresión arbitraria, si es o no significativa independientemente (*i. e.*, si es una oración), y supongamos como algo usual que esto comporta contemplar cada oración como compuesta, de maneras admisibles, a partir de elementos extraídos de un *stock* finito fijo de elementos sintácticos atómicos (aproximadamente, palabras). El pensamiento prometedor es que la sintaxis, así concebida, proporcionará la semántica cuando se añade un diccionario que da el significado de cada átomo sintáctico. Las esperanzas se frustran, sin embargo, si la semántica ha de comprender una teoría del significado en nuestro sentido, pues el conocimiento de las características estructurales que contribuyen a la significatividad en una oración, más el conocimiento del significado de las partes últimas, no equivale al conocimiento de lo que significa la oración. Este punto es fácilmente ilustrado por las oraciones de creencia. Su sintaxis carece relativamente de problemas. Con todo, añadir un diccionario deja intacto el problema semántico estándar, que es que no podemos dar cuenta ni tan siquiera de las condiciones de verdad de tales oraciones sobre la base de lo que sabemos del significado de las palabras que están en ellas. La situación no se altera radicalmente refinando el diccionario para indicar qué significado o significados comporta una expresión ambigua en cada uno de sus posibles contextos; el problema de las oraciones de creencia persiste después de que las ambigüedades se resuelven.

El hecho de que la sintaxis recursiva con diccionario añadido no sea necesariamente semántica recursiva ha sido oscurecido en algún escrito reciente sobre lingüística mediante la intromisión de criterios semánticos en la discusión de pretendidas teorías sintácticas. El asunto se reduciría a una inocua diferencia sobre terminología si los criterios semánticos fuesen claros; pero no lo son. Mientras que hay acuerdo en que la tarea central de la semántica es dar la interpretación semántica (el significado) de toda oración del lenguaje, no se en-

contrará en parte alguna de la literatura lingüística, al menos que yo sepa, una explicación franca de cómo una teoría realiza esta tarea, o de cómo decir cuándo ha sido cumplida. El contraste con la sintaxis es chocante. La principal tarea de una sintaxis modesta es caracterizar *significatividad* (u oracionalidad). Podemos tener tanta confianza en la corrección de tal caracterización como tenemos en la representatividad de nuestra muestra y en nuestra capacidad de decir cuándo expresiones particulares son significativas (oraciones). ¿Qué tarea y tests claros y análogos existen para la semántica?<sup>5</sup>

Decidimos hace un rato no suponer que las partes de las oraciones tienen significados excepto en el sentido ontológicamente neutral de hacer una contribución sistemática al significado de las oraciones en las que aparecen. Puesto que postular significados no ha producido ninguna ganancia, volvamos a esa intuición. Una dirección hacia la que apunta es una cierta visión holista del significado. Si las oraciones dependen para su significado de su estructura, y si entendemos el significado de cada elemento de la estructura solamente como una abstracción a partir de la totalidad de las oraciones en las que desempeña un papel, entonces podemos dar el significado de cualquier oración (o palabra) dando solamente el significado de toda oración (y palabra) del lenguaje. Frege dijo que solamente en el contexto de una oración tiene significado una palabra; podría haber añadido, en la misma vena, que solamente en el contexto del lenguaje tiene significado una oración (y, por lo tanto, una palabra).

Este grado de holismo estaba ya implícito en la sugerencia de que una teoría adecuada del significado tenía que entrañar *todas* las oraciones de la forma «*o* significa *m*». Pero ahora, no habiendo encontrado más ayuda en los significados de las oraciones que en los significados de las palabras, preguntémonos si podemos deshacernos de los molestos términos singulares que se supone reemplazan a «*m*» y referirnos a significados. De algún modo, nada podría ser más fácil: escribamos solamente «*o* significa que *p*», e imaginemos «*p*» reemplazada por una oración. Las oraciones, como hemos visto, no pueden nombrar significados, y las oraciones con «que» como prefijo no son en

<sup>5</sup> Para una formulación reciente e instructiva del papel de la semántica en la lingüística, véase Noam Chomsky, «Topics in the Theory of Generative Grammar», en *Current Trends in Linguistics* (editado por Thomas A. Sebeok), vol. III, La Haya, 1966. En este artículo Chomsky (1) subraya la importancia central de la semántica en la teoría lingüística, (2) argumenta a favor de la superioridad de las gramáticas transformativas sobre las gramáticas de estructura de frase sobre la base, en gran medida, de que, aunque las gramáticas de estructura de frase pueden ser adecuadas para definir oracionalidad para (al menos) algunos lenguajes naturales, son inadecuadas como fundamento para semántica y (3) hace comentarios repetidamente sobre el «estado más bien primitivo» de los conceptos de la semántica y observa que la noción de interpretación semántica «opone aún resistencia a cualquier análisis profundo».

absoluto nombres, a menos que lo decidamos así. Sin embargo, parece ahora como si estuviésemos en dificultades bajo otra consideración, pues parece razonable esperar que, al luchar con la lógica de la expresión aparentemente no extensional «significa que», nos encontremos con problemas tan difíciles como, o quizás idénticos a los problemas que nuestra teoría se está esforzando en resolver.

El único modo que conozco de tratar con esta dificultad es simple y radical. La ansiedad de que estamos enmarañados en lo intensional surge de usar las palabras «significa que» como relleno entre la descripción de la oración y la oración, pero puede ser que el éxito de nuestra aventura dependa no del relleno, sino de lo que éste rellena. La teoría habrá realizado su trabajo si proporciona, para toda oración *o* del lenguaje que se está estudiando, una oración apareada con la anterior (para reemplazar a «*p*») que, en algún sentido que todavía ha de clarificarse, «dé el significado» de *o*. Una candidata obvia para oración a aparear es justamente *o* misma, si el lenguaje objeto está contenido en el metalenguaje; en el caso contrario una traducción de *o* en el metalenguaje. Como paso intrépido final, intentemos tratar la posición ocupada por «*p*» extensionalmente: para llevar a cabo esto, quitemos de en medio el oscuro «significa que», proveamos a la oración que reemplaza a «*p*» de una conectiva oracional apropiada, y suministremos a la descripción que reemplaza a «*o*» su propio predicado. El resultado plausible es

(V) *o* es *V* si y sólo si *p*.

Lo que exigimos de una teoría del significado para un lenguaje *L* es que sin apelar a ninguna noción semántica (adicional) coloque restricciones suficientes sobre el predicado «es *V*» para entrañar todas las oraciones obtenidas a partir del esquema *V* cuando «*o*» se reemplaza por una descripción estructural de una oración de *L* y «*p*» por esa oración.

Cualesquiera dos predicados que satisfagan esta condición tienen la misma extensión<sup>6</sup>, de modo que, si el metalenguaje es suficientemente rico, nada se interpone en el camino de expresar lo que yo llamo una teoría del significado en la forma de una definición explícita de un predicado «es *V*». Pero, ya esté definido explícitamente o caracterizado de modo recursivo, es claro que las oraciones a las que se aplica el predicado «es *V*» serán justamente las oraciones verdaderas de *L*, pues la condición que hemos colocado sobre las teorías del significado satisfactorias es en esencia la Convención *V* de Tarski que

<sup>6</sup> Suponiendo, desde luego, que la extensión de esos predicados se limita a las oraciones de *L*.

pone a prueba la adecuación de una definición semántica formal de verdad<sup>7</sup>.

El sendero hasta este punto ha sido tortuoso, pero la conclusión puede enunciarse de manera simple: una teoría del significado para un lenguaje *L* muestra «cómo los significados de las oraciones dependen de los significados de las palabras» si contiene una definición recursiva de verdad-en-*L*. Y, al menos hasta ahora, no tenemos ninguna otra idea de cómo producir el resultado deseado. Merece la pena subrayar que el concepto de verdad no desempeñaba ningún papel ostensible en el enunciado de nuestro problema original. Este problema, después de ciertos refinamientos, llevó al punto de vista de que una teoría adecuada del significado tiene que caracterizar a un predicado que cumpla ciertas condiciones. Constituyó un descubrimiento el que tal predicado se aplicase exactamente a las oraciones verdaderas. Espero que lo que estoy haciendo pueda describirse en parte como una defensa de la importancia filosófica del concepto semántico tarskiano de verdad. Pero mi defensa está solamente relacionada de manera distante, si es que lo está de algún modo, con la cuestión de si el concepto que Tarski ha mostrado cómo definir es la (o una) concepción filosóficamente interesante de verdad, o con la cuestión de si Tarski ha arrojado alguna luz sobre el uso ordinario de palabras como «verdadero» y «verdad». Es un hecho desafortunado el que el polvo proveniente de fútiles y confusas batallas sobre estas cuestiones haya impedido a aquéllos con un interés filosófico en el lenguaje —filósofos, lógicos, psicólogos y lingüistas por igual— el reconocer en el concepto semántico de verdad (bajo cualquier nombre) el sofisticado y poderoso fundamento de una teoría del significado competente.

No hay necesidad de suprimir, desde luego, la conexión obvia entre una definición de verdad del género que Tarski ha mostrado como construir, y el concepto de significado. Es ésta: la definición funciona dando las condiciones necesarias y suficientes para la verdad de toda oración, y dar las condiciones de verdad es una manera de dar el significado de una oración. Conocer el concepto semántico de verdad para un lenguaje es saber en qué consiste para una oración —para cualquier oración— ser verdadera, y esto equivale, en un buen sentido que podemos dar a la frase, a comprender el lenguaje. De todas formas, esto es mi excusa para un rasgo de la presente discusión que tiene propensión a escandalizar a perros viejos: mi uso desenfadado de la palabra «significado», pues lo que yo llamo teoría del significado ha resultado después de todo no hacer uso alguno de sig-

<sup>7</sup> Alfred Tarski, «The Concept of Truth in Formalized Languages», en *Logic, Semantics, Metamathematics*, Oxford, 1956, pp.152-278.

nificados, ya sean de oraciones o de palabras. Ciertamente, puesto que una definición de verdad de tipo tarskiano proporciona todo lo que hasta ahora hemos pedido a una teoría del significado, es claro que tal teoría cae confortablemente dentro de lo que Quine denomina la «teoría de la referencia» como algo que se distingue de lo que denomina la «teoría del significado». Con lo dicho hasta ahora basta ya para lo que llamo una teoría del significado, y basta también, quizás, en contra de mi llamarla así <sup>8</sup>.

Una teoría del significado (en mi ligeramente perverso sentido) es una teoría empírica, y su ambición es dar cuenta del funcionamiento del lenguaje natural. Al igual que cualquier teoría, puede ponerse a prueba comparando alguna de sus consecuencias con los hechos. En el caso presente esto es fácil, pues la teoría ha sido caracterizada como dando como resultado una plétora infinita de oraciones, dando cada una de ellas las condiciones de verdad de una oración; necesitamos solamente preguntar, en casos tomados como muestra, si lo que la teoría afirma que son las condiciones de verdad de una oración lo son realmente. Un caso de prueba típico podría incluir el decidir si la oración «La nieve es blanca» es verdadera si y sólo si la nieve es blanca. No todos los casos serán tan simples (por razones que se bosquejarán), pero es evidente que esta clase de prueba no invita a pensar que se trata de coser y cantar. Una concepción neta de lo que constituye una teoría en este dominio proporciona un contexto excitante para plantear profundas cuestiones sobre cuándo es correcta una teoría del lenguaje y cómo ha de probarse. Pero las dificultades son teóricas, no prácticas. En su aplicación, el problema es obtener una teoría que se aproxime al funcionamiento; cualquiera puede decir si es correcta <sup>9</sup>. Uno puede ver por qué esto es así. La teoría no revela nada nuevo sobre las condiciones bajo las cuales es verdadera una oración individual; no clarifica más esas condiciones en punto alguno de lo que lo hace la oración misma. El funcionamiento de la teoría consiste en relacionar las condiciones de verdad conocidas de cada oración con esos aspectos («palabras») de la oración que recurren en

<sup>8</sup> Pero Quine puede citarse en apoyo de mi usanza: «... en cuanto al significado ... puede decirse que una palabra está determinada hasta el punto en que está determinada la verdad o falsedad de sus contextos». «Truth by Convention», publicado por vez primera en 1936; ahora en *The Ways of Paradox*, Nueva York 1966, p. 82. Puesto que una definición de verdad determina el valor de verdad de toda oración del lenguaje objeto (relativo a una oración del metalenguaje), determina el significado de toda palabra y oración. Esto me parece que justifica el título Teoría del Significado.

<sup>9</sup> Para dar un solo ejemplo: es claramente un punto a favor de la teoría el que entrañe «“La nieve es blanca” es verdadera si y sólo si la nieve es blanca». pero diseñar una teoría que entrañe esto (y funcione para todas las oraciones relacionadas) no es trivial. No conozco ninguna teoría que tenga éxito con este mismo caso (el problema de los «términos-masa»).

otras oraciones, y a los que pueden asignarse idénticos papeles en otras oraciones. La potencia empírica de tal teoría depende del éxito en recuperar la estructura de una capacidad muy complicada —la capacidad de hablar y comprender un lenguaje—. Podemos decir con bastante facilidad cuándo los pronunciamientos particulares de la teoría se acomodan a nuestra comprensión del lenguaje; esto es consistente con una ligera idea del diseño de la maquinaria de nuestros logros lingüísticos.

Las observaciones del último párrafo se aplican sólo directamente al caso especial en el que se supone que el lenguaje para el que la verdad está siendo caracterizada es parte del lenguaje usado y comprendido por el caracterizador. Bajo estas circunstancias el constructor de una teoría se aprovechará cuando pueda como algo natural de la ventaja inherente a un metalenguaje con la garantía de una oración equivalente para cada oración del lenguaje objeto. Con todo, este hecho no debería persuadirnos de que pensásemos que es más correcta una teoría que entraña «La nieve es blanca» es verdadera si y sólo si la nieve es blanca» que una que entrañe en su lugar:

(O) «La nieve es blanca» es verdadera si y sólo si el césped es verde.

en el supuesto, desde luego, que estemos tan seguros de la verdad de (O) como lo estamos de la de su más célebre predecesora. Sin embargo (O) puede no fomentar la misma confianza de que una teoría que la entrañe merezca ser llamada una teoría del significado.

La amenazadora falta de vigor puede contrarrestarse como sigue. Lo grotesco de (O) no es en sí mismo nada en contra de una teoría de la cual es una consecuencia, en el supuesto de que la teoría dé los resultados correctos para toda oración (sobre la base de su estructura, no habiendo ninguna otra manera). No es fácil ver cómo (O) podría ser parte de tal empresa, pero, si lo fuese —esto es, si (O) se siguiese de una caracterización del predicado «es verdadero» que llevase al emparejamiento invariable de verdades con verdades y falsedades con falsedades—, entonces, pienso, no habría nada esencial a la idea de significado que quedara por capturar<sup>10</sup>.

Lo que aparece a la derecha del bicondicional en las oraciones de la forma «*o* es verdadera si y sólo si *p*» cuando tales oraciones son consecuencias de una teoría de la verdad desempeña su papel al determinar el significado de *o* no pretendiendo sinonimia, sino añá-

<sup>10</sup> Los críticos no han logrado a menudo darse cuenta de la estipulación esencial mencionada en este párrafo. El punto es que (O) no podría pertenecer a ninguna teoría razonablemente simple que diese también las condiciones de verdad correctas para «Eso es nieve», «Esto es blanco». (Véase la discusión de las expresiones indéxicas más adelante.) [Nota añadida en 1982.]

diendo una pincelada más al cuadro que, tomado como un todo, dice qué hay que conocer del significado de *o*; esta pincelada se añade en virtud del hecho de que la oración que reemplaza a «*p*» es verdadera si y sólo si *o* lo es.

Puede servir de ayuda el reflexionar que (*O*) es aceptable, si lo es, porque estamos seguros independientemente de la verdad de «La nieve es blanca», y «El césped es verde»; pero, en casos en los que no estamos seguros de la verdad de una oración, podemos tener confianza en una caracterización del predicado veritativo solamente si empareja esa oración con una respecto de la cual tenemos una buena razón para pensar que es equivalente. Sería una imprudencia por parte de alguien que tuviese algunas dudas sobre el color de la nieve o del césped aceptar una teoría que proporcionase (*O*), incluso si sus dudas eran de igual grado, a menos que pensase que el color de lo uno estaba enlazado con el color de lo otro <sup>11</sup>. La omnisciencia puede permitirse obviamente teorías del significado más bizarras que la ignorancia; pero, entonces, la omnisciencia tiene menos necesidad de comunicación.

Desde luego, tiene que ser posible para el hablante de un lenguaje construir una teoría del significado para el hablante de otro, aunque en este caso las comprobaciones empíricas de la corrección de la teoría ya no serán triviales. Como antes, la meta de la teoría será una correlación infinita de oraciones iguales en verdad. Pero esta vez no se tiene que suponer que el constructor de la teoría tiene intuición directa respecto de las equivalencias verosímiles entre su propia lengua y la ajena. Lo que tiene que hacer es averiguar, en la medida en que pueda, qué oraciones mantiene como verdaderas el extranjero en su propia lengua (o, mejor aún, hasta qué grado las mantiene como verdaderas). El lingüista intentará construir una caracterización de verdad-para-el-extranjero que proporcione, en la medida de lo posible, una correspondencia de oraciones mantenidas como verdaderas (o como falsas) por el extranjero con oraciones mantenidas como verdaderas (o como falsas) por el lingüista. Suponiendo que no se logra un ajuste perfecto, el residuo de oraciones mantenidas como verdaderas traducidas por oraciones mantenidas como falsas (y viceversa) es el margen de error (foráneo o doméstico). La caridad al interpre-

<sup>11</sup> Este párrafo es confuso. Lo que diría es que las oraciones de la teoría son generalizaciones empíricas sobre los hablantes, y de este modo tienen no sólo que ser verdaderas sino también legaliformes. Presumiblemente (*O*) no es una ley, puesto que no apoya contrafácticos apropiados. Es también importante que la evidencia para aceptar las condiciones de verdad (relativizadas por lo que respecta al tiempo y al hablante) para «Eso es nieve» se basa sobre la conexión causal entre el asentimiento de un hablante a la oración y la presentación demostrativa de nieve. Para una discusión adicional véase «Reply to Foster», en G. Evans y J. McDowell (eds.), Oxford University Press, 1966, pp. 33-41. [Nota añadida en 1982.]

lar las palabras y los pensamientos de otros es también inevitable en otra dirección: así como tenemos que maximizar el acuerdo, o corre-mos el riesgo de no conferir sentido a aquello sobre lo que el extran-jero está hablando, así también tenemos que maximizar la autocon-sistencia que le atribuimos, so pena de no comprenderle. No surge ningún único principio de caridad óptima; las constricciones, por lo tanto, no determinan ninguna teoría única. En una teoría de la tra-ducción radical (como Quine la llama) no hay cuestiones respecto de lo que el extranjero quiere decir que estén completamente desenma-ñaadas de las cuestiones sobre lo que él cree. No sabemos lo que al-guien quiere decir a menos que sepamos lo que él cree; no sabemos lo que alguien cree a menos que sepamos lo que él quiere decir. En la traducción radical somos capaces de romper este círculo, bien que sólo incompletamente, puesto que algunas veces podemos decir que una persona asiente a una oración que nosotros no entendemos <sup>12</sup>.

En las últimas páginas he estado preguntando cómo una teoría del significado que toma la forma de una definición de verdad puede ponerse a prueba empíricamente, y he ignorado alegremente la cues-tión prioritaria de si hay alguna posibilidad seria de que tal teoría pueda darse para un lenguaje natural. ¿Cuáles son las perspectivas para una teoría semántica formal de un lenguaje natural? Muy po-bres, de acuerdo con Tarski; y creo que la mayor parte de los lógicos, filósofos del lenguaje y lingüistas están de acuerdo <sup>13</sup>. Permítaseme que haga lo que pueda para disipar el pesimismo. Lo que pueda de una manera general y programática, desde luego; pues aquí la prue-ba del pastel estará ciertamente en la demostración de los teoremas correctos.

Tarski concluye la primera sección de su ensayo clásico sobre el concepto de verdad en los lenguajes formalizados con las observacio-nes siguientes, que él pone en cursiva:

... *La misma posibilidad de un uso consistente de la expresión*

<sup>12</sup> Este bosquejo de cómo puede someterse a comprobación una teoría del signifi-cado para una lengua ajena debe obviamente su inspiración a la explicación de Quine de la traducción radical en el capítulo II de *Word and Object*, Nueva York, 1960. [Versión castellana: *Palabra y objeto*, Labor, Barcelona 1968.] Al sugerir que una teoría aceptable de la traducción radical toma la forma de una caracterización recursiva de la verdad, voy más allá de cualquier cosa explícita en Quine. Hacia el final de este artículo, en la discusión de los demostrativos, aparecerá otro marcado punto de acuerdo.

<sup>13</sup> Hasta donde estoy enterado, ha habido muy poca discusión sobre si puede darse una definición formal de verdad para un lenguaje natural. Pero, en una vena más general, diversas personas han defendido insistentemente que los conceptos de la semántica formal se aplican al lenguaje natural. Véanse, por ejemplo, las contribucio-nes de Yehoshua Bar-Hillel y Evert Beth a *The philosophy of Rudolph Carnap* (edita-do por Paul A. Schilpp), La Salle, Ill., 1963 y «Logical Syntax and Semantics» de Bar-Hillel, en *Language*, 30, pp. 230-237.



*«oración verdadera» que esté en armonía con las leyes de la lógica y el espíritu del lenguaje diario parece ser muy cuestionable, y consecuentemente la misma duda afecta a la posibilidad de construir una definición correcta de esta expresión*<sup>14</sup>.

Más adelante, en el mismo ensayo, vuelve al tema:

... el concepto de verdad (así como otros conceptos semánticos) cuando se aplica al lenguaje coloquial en conjunción con las leyes normales de la lógica lleva inevitablemente a confusiones y contradicciones. Quienquiera que desee, a pesar de todas las dificultades, dedicarse a la semántica del lenguaje coloquial con la ayuda de métodos exactos se verá impulsado en primer lugar a emprender la desagradecida tarea de reformar este lenguaje. Encontrará que es necesario definir su estructura, superar la ambigüedad de los términos que ocurren en él, y finalmente dividir el lenguaje en series de lenguajes de mayor y mayor alcance, cada uno de los cuales está en la misma relación con el siguiente que en la que está un lenguaje formalizado con su metalenguaje. Puede, sin embargo, dudarse, si el lenguaje de la vida diaria preservaría, después de ser «racionalizado» de esta manera, su naturalidad y si no asumiría más bien los rasgos característicos de los lenguajes formalizados<sup>15</sup>.

Surgen aquí dos temas: que el carácter universal de los lenguajes naturales lleva a contradicción (las paradojas semánticas), y que los lenguajes naturales son demasiado confusos y amorfos para permitir la aplicación directa de métodos formales. El primer punto merece una respuesta seria, y me gustaría tenerla. Ya que esto es así, diré solamente por qué pienso que estamos justificados en proseguir sin haber desinfectado esta fuente particular de ansiedad conceptual. Las paradojas semánticas surgen cuando el rango de los cuantificadores en el lenguaje objeto es demasiado generoso en ciertas maneras. Pero no está realmente claro cuán injusto se habría de ser respecto al urdu o al dórico para contemplar el rango de sus cuantificadores como insuficiente para proporcionar una definición explícita de «verdadero-en-urdu» o «verdadero-en-dórico». O para plantear el asunto de otra manera (si no más seriamente): puede haber siempre algo en la naturaleza del caso que captamos al comprender el lenguaje de otro (el concepto de verdad) que no podemos comunicarle a él. En cualquier caso, la mayor parte de los problemas de interés filosófico general surgen dentro de un fragmento del lenguaje natural relevan-

<sup>14</sup> Tarski, *ibíd.*, p. 165.

<sup>15</sup> *Ibíd.*, p. 267.

to que puede concebirse como conteniendo muy poca teoría de conjuntos. Desde luego estos comentarios no hacen frente a la afirmación de que los lenguajes naturales son universales. Pero me parece que esta afirmación, ahora que sabemos que tal universalidad lleva a paradojas, es sospechosa.

El segundo punto de Tarski es que tendríamos que reformar un lenguaje natural hasta el punto de dejarlo irreconocible antes de que pudiésemos aplicarle métodos semánticos formales. Si esto es verdad, es fatal para mi proyecto, pues la tarea de una teoría del significado tal como yo la concibo no es cambiar, mejorar o reformar un lenguaje, sino describirlo y entenderlo. Miremos al lado positivo. Tarski ha mostrado la manera de dar una teoría para lenguajes formales interpretados en varios géneros; selecciónese uno tan parecido al castellano como sea posible. Puesto que este nuevo lenguaje ha sido explicado en castellano y contiene gran parte de castellano, no sólo podemos sino que pienso que tenemos que contemplarlo como una parte del castellano para aquellos que lo entienden. Para este fragmento del castellano tenemos, *ex hypothesi*, una teoría de la clase requerida. No solamente eso, sino que al interpretar esta parte asociada del castellano en el antiguo castellano dimos necesariamente indicaciones que conectan antiguo y nuevo. Dondequiera que haya oraciones del antiguo castellano con las mismas condiciones de verdad que las oraciones de la parte asociada podemos extender la teoría de modo que las cubra. Mucho de lo que se pide consiste solamente en mecanizar hasta donde sea posible lo que ahora hacemos gracias a una habilidad cuando ponemos el castellano ordinario en una u otra notación canónica. El punto no es que la notación canónica sea mejor que la tosca expresión idiomática original, sino más bien que, si sabemos *para* qué expresión idiomática es canónica la notación canónica, tenemos una teoría tan buena para la expresión idiomática como para su compañera.

Los filósofos han estado ocupados desde hace tiempo en la ardua tarea de aplicar la teoría al lenguaje ordinario mediante el dispositivo de aparear oraciones vernaculares con oraciones para las que tienen una teoría. La impresionante contribución de Frege consistió en mostrar cómo «todo», «alguno», «cada uno», «ninguno» y pronombres asociados podrían domesticarse, en algunos de sus usos; por vez primera fue posible soñar en una semántica formal para una parte significativa de un lenguaje natural. Este sueño se convirtió en realidad de una manera palpable con la obra de Tarski. Sería una vergüenza pasar por alto el hecho de que como resultados de esos dos magníficos logros, el de Frege y el de Tarski, hemos ganado una profunda penetración en la estructura de nuestras lenguas maternas. Los filósofos de inclinación lógica han tendido a comenzar donde estaba la teoría y a desarrollar su trabajo hacia las complicaciones del len-

guaje natural. Los lingüistas contemporáneos, con una aspiración que no puede verse fácilmente que sea diferente, empiezan con el lenguaje ordinario y su trabajo avanza hacia una teoría general. Si cualquiera de las dos partes tiene éxito, ha de haber un encuentro. La obra reciente de Chomsky y otros está haciendo mucho para colocar dentro del alcance de una teoría semántica seria las complejidades de los lenguajes naturales. Para dar un ejemplo: supóngase que se tiene éxito al dar las condiciones de verdad para algún rango significativo de oraciones de la voz activa. Entonces, con algún procedimiento formal para transformar cada oración tal en una oración correspondiente en la voz pasiva, la teoría de la verdad podría extenderse de una manera obvia a este nuevo conjunto de oraciones <sup>16</sup>.

Un problema tocado de pasada por Tarski no tiene que ser resuelto, al menos en todas sus manifestaciones, para lograr éxito con la teoría: la existencia en los lenguajes naturales de «términos ambiguos». En la medida en que la ambigüedad no afecta a la forma gramatical, y puede traducirse, ambigüedad por ambigüedad, en el metalenguaje, una definición de verdad no nos dirá mentira alguna. El problema principal, para la semántica sistemática, con la frase «cree que» en castellano no es su vaguedad, ambigüedad o su inadecuación para ser incorporada en una ciencia seria: sea nuestro metalenguaje el castellano y, todos *esos* problemas serán traspasados sin pérdida o ganancia al metalenguaje. Pero el problema central de la gramática lógica de «cree que» continuará persiguiéndonos de manera persistente.

El ejemplo es adecuado para ilustrar otro punto, ciertamente relacionado, pues la discusión de las oraciones de creencia ha estado negativamente afectada por el fallo en observar una distinción fundamental entre tareas: descubrir la gramática lógica o forma de las ora-

<sup>16</sup> La *aproximación* que anticipadamente imagino entre la gramática transformativa y una teoría correcta del significado ha progresado mucho gracias a un cambio reciente en la concepción de la gramática transformativa descrito por Chomsky en el artículo al que se ha hecho referencia anteriormente (nota 5). Las estructuras generadas por la parte correspondiente a la estructura de frase de la gramática son aquellas adecuadas, como se ha reconocido desde hace algún tiempo, para interpretación semántica; pero este punto de vista es inconsistente con la idea, mantenida por Chomsky hasta hace muy poco, de que las operaciones recursivas se introducen solamente mediante las reglas de transformación. Chomsky cree ahora que las reglas de estructura de frase son recursivas. Puesto que los lenguajes a los que los métodos semánticos formales se aplican directa y naturalmente son aquéllos para los que es apropiada una gramática de estructura de frase (recursiva), es claro que la presente concepción de Chomsky de la relación entre las estructuras generadas por la parte correspondiente a la estructura de frase de la gramática, y las oraciones del lenguaje es muy parecida a la concepción que muchos lógicos y filósofos han tenido de la relación entre los lenguajes foralizados más ricos y el lenguaje ordinario. (En estas observaciones estoy en deuda con Bruce Vermazen.)

ciones (que es la provincia de la teoría del significado tal como la interpreto), y el análisis de las palabras o expresiones individuales (que son tratadas como primitivas por la teoría). Así, Carnap, en la primera edición de *Meaning and Necessity*, sugería que virtiésemos «Juan cree que la Tierra es redonda» como «Juan responde afirmativamente a "la Tierra es redonda" como una oración del castellano». Abandonó esto cuando Mates señaló que Juan podría responder afirmativamente a una oración y no a otra sin importar cuán próximas estuvieran en significado.<sup>17</sup> Pero hay aquí una confusión desde el comienzo. La estructura semántica de una oración de creencia, de acuerdo con esta idea de Carnap, es dada por un predicado de tres lugares con lugares reservados para expresiones que se refieren a una persona, a una oración y a un lenguaje. Es una clase de problema enteramente diferente intentar un análisis de este predicado de acuerdo, quizás, con líneas conductistas. No es el menor entre los méritos de la concepción de Tarski de una teoría de la verdad el que la pureza de método que exige de nosotros se siga de la formulación del problema mismo, no de la autoimpuesta moderación de algún puritanismo filosófico adventicio.

Pienso que son difíciles de exagerar las ventajas para la filosofía del lenguaje de tener presente esta distinción entre cuestiones de forma o gramática lógicas, y el análisis de los conceptos individuales. Otro ejemplo puede ayudar a poner de manifiesto este punto.

Si suponemos decididas las cuestiones de gramática lógica, oraciones como «Bardot es buena» no plantean problemas especiales para una definición de verdad. Las profundas diferencias entre términos descriptivos y evaluativos (emotivos, expresivos, etc.) no se muestran aquí. Incluso si mantenemos que hay algún sentido importante en el que las oraciones morales o evaluativas no tienen un valor de verdad (por ejemplo, porque no pueden ser verificadas), no deberíamos sobresaltarnos ante «"Bardot es buena" es verdadera si y sólo si Bardot es buena»; en una teoría de la verdad, esta consecuencia se seguiría con el resto, estando al tanto, como debe hacerse, de la localización semántica de tales oraciones en el lenguaje como un todo —de su relación con las generalizaciones, su papel en oraciones compuestas como «Bardot es buena y Bardot es tonta», y así sucesivamente—. Lo que es especial de las palabras evaluativas simplemente queda intocado: el misterio es transferido de la palabra «buena» en el lenguaje objeto a su traducción en el metalenguaje.

Pero «buena» tal como se presenta en «Bardot es una buena actriz» es otro asunto. El problema no es que la traducción de esta ora-

<sup>17</sup> B. Mates, «Sinonimity», en *Semantics and the Philosophy of language*, L. Linsky (ed.) University of Illinois Pres, Urbana, Illinois (1952).

ción no está en el metalenguaje —supongamos que está—. El problema es construir una definición de verdad tal que «“Bardot es una buena actriz” si y sólo si Bardot es una buena actriz» —y todas las oraciones iguales que ésta— sean consecuencias. Obviamente «buena actriz» no significa «buena y una actriz». Podríamos pensar en considerar «es una buena actriz» como un predicado inanalizado. Esto borraría toda conexión entre «es una buena actriz» y «es una buena madre», y no nos daría excusa alguna para pensar en «buena», en esos usos, como una palabra o elemento semántico. Pero, aún peor, nos impediría construir en absoluto una definición de verdad, pues no hay fin alguno a los predicados que habríamos de tratar como lógicamente simples (y, por consiguiente, acomodar en cláusulas separadas en la definición de satisfacción): «es un buen amigo de los perros», «es un buen conversador de veintiocho años», y así sucesivamente. El problema no es peculiar al caso: es el problema general de los adjetivos atributivos.

Es consistente con la actitud tomada aquí el considerar usualmente como un error estratégico el emprender un análisis filosófico de palabras y expresiones que no esté precedido por, o acompañado de alguna manera por, el intento de obtener directamente la gramática lógica. Pues ¿cómo podemos tener confianza alguna en nuestros análisis de palabras tales como «correcto», «debe», «puede» y «obligado», o las frases que usamos para hablar de acciones, eventos y causas, si no sabemos con qué partes del habla (lógicas, semánticas) hemos de tratar? Diría otro tanto sobre los estudios de la «lógica» de esas y otras palabras, y las oraciones que las contienen. Si el esfuerzo y el ingenio que se han gastado en el estudio de lógicas deónticas, lógicas modales, lógicas impertivas y erotéticas han sido o no en gran parte fútiles no puede saberse hasta que tengamos un análisis aceptable de las oraciones que tales sistemas pretenden tratar. Los filósofos y los lógicos hablan algunas veces como si fueran libres de elegir entre, digamos, el condicional veritativo-funcional y otros, o libres de introducir operadores oracionales no veritativo-funcionales tales como «Sea el caso que» o «Debe ser el caso que». Pero de hecho la decisión es crucial. Cuando nos separamos de las expresiones idiomáticas que podemos acomodar en una definición de verdad, nos deslizamos en (o creamos) un lenguaje para el que no tenemos ninguna explicación semántica coherente —esto es, ninguna explicación de cómo tal discurso puede integrarse en el lenguaje como un todo—.

Para volver a nuestro tema principal: hemos reconocido que una teoría del género propuesto deja todo el asunto de lo que significan exactamente las palabras individuales exactamente donde estaba. Incluso cuando el metalenguaje es diferente del lenguaje objeto, la teoría no ejerce presión alguna para mejorar, clarificar o analizar las palabras individuales, excepto cuando, por un accidente de vocabula-

rio, la traducción llana y simple falla. Del mismo modo que la sinonimia, como sucede entre expresiones, queda generalmente sin tratar, así también sucede con la sinonimia de oraciones y la analiticidad. Incluso oraciones tales como «Una yegua es un hembra de caballo» no llevan ninguna etiqueta especial a menos que nos plazca proporcionarle una. Una definición de verdad no distingue entre oraciones analíticas y otras, excepto para oraciones que deben su verdad a la presencia sola de las constantes que dan a la teoría su agarre a la estructura: la teoría entraña no solamente que esas oraciones son verdaderas, sino que permanecerán verdaderas bajo todas las refundiciones significativas de sus partes no lógicas. Así pues, habiendo dado a la noción de verdad lógica una aplicación limitada, las nociones relacionadas de equivalencia lógica y entrañamiento seguirán igual camino. Es difícil imaginar cómo una teoría del significado podría no lograr leer una lógica en su lenguaje objeto hasta este grado; y, hasta el punto en que lo hace, nuestras intuiciones de verdad lógica, equivalencia y entrañamiento podrían reclamarse para construir y poner a prueba la teoría.

Vuelvo ahora a una mosca más, y bien grande por cierto, atrapada en el panal: el hecho de que la misma oración pueda en un tiempo o en una boca ser verdadera y en otro tiempo o en otra boca ser falsa. Tanto los lógicos como aquellos que son críticos con los métodos formales parecen estar aquí ampliamente de acuerdo (aunque de ningún modo universalmente) en que la semántica formal y la lógica son incompetentes para tratar con los trastornos causados por los demostrativos. Los lógicos han reaccionado a menudo degradando el lenguaje natural e intentando mostrar cómo continuar sin demostrativos; sus críticos reaccionan degradando la lógica y la semántica formal. Nada de esto puede hacerme feliz: claramente los demostrativos no pueden eliminarse de un lenguaje natural sin pérdida o cambio radical; de este modo no hay otra opción que acomodar la teoría a ellos.

No resulta error lógico alguno si tratamos simplemente a los demostrativos como constantes<sup>18</sup>; ni tampoco surge problema alguno para dar una definición semántica de verdad. «Yo soy sabio» es verdadera si y sólo si yo soy sabio», con su gentil ignorancia del elemento demostrativo incorporado en «Yo» abandona la cadena de montaje a la par que «Sócrates es sabio» es verdadera si y sólo si Sócrates es sabio» con su gentil indiferencia hacia el elemento demostrativo incorporado en «es sabio» (el tiempo verbal).

<sup>18</sup> Quine tiene buenas cosas que decir sobre esto en *Methods of Logic*, Nueva York, 1950. [Versión castellana: *Los métodos de la lógica*, Ariel, Barcelona, 1962.] Véase § 8.

Lo que sufre en este tratamiento de los demostrativos no es la definición de un predicado veritativo, sino la plausibilidad de la afirmación de que lo que se ha definido es verdad. Pues esta afirmación es aceptable solamente si el hablante y las circunstancias de emisión de cada oración mencionada en la definición encajan con el hablante y las circunstancias de emisión de la definición de verdad misma. Podría también señalarse con justicia que parte del comprender los demostrativos consiste en conocer las reglas mediante las cuales ajustan su referencia a la circunstancia; la asimilación de los demostrativos a términos constantes oblitera este rasgo. Estas quejas pueden responderse aunque solamente, pienso, mediante una revisión de gran alcance de la teoría de la verdad. Sugeriré apenas cómo podría hacerse esto, pero todo lo que se necesita es la sugerencia escueta: la idea es técnicamente trivial y está completamente en la línea del trabajo que se viene haciendo sobre la lógica de los tiempos verbales.<sup>19</sup>

Podríamos considerar que la verdad es una propiedad no de oraciones, sino de emisiones, o de actos de habla, o de triplos ordenados de oraciones, tiempos y personas; pero es más simple considerar la verdad sólo como una relación entre una oración, una persona y un tiempo. Bajo tal tratamiento, la lógica ordinaria tal como ahora se interpreta se aplica de la manera usual, pero solamente a conjuntos de oraciones relativizadas al mismo hablante y tiempo; las relaciones lógicas adicionales entre oraciones proferidas en tiempos diferentes por hablantes diferentes pueden articularse mediante axiomas nuevos. Tal cosa no me preocupa. La teoría del significado experimenta un cambio sistemático aunque sin presentar dificultades: correspondiendo a cada expresión con un elemento demostrativo, tiene que haber en la teoría una frase que pone en relación las condiciones de verdad de las oraciones en las que la expresión ocurre con tiempos y hablantes cambiantes. Así pues, la teoría entrañará oraciones como la siguiente:

«Estoy cansado» es verdadera en tanto que (potencialmente) proferida por  $p$  en  $t$  si y sólo si  $p$  está cansado en  $t$ .

«Ese libro fue robado» es verdadera en tanto que (potencialmente) proferida por  $p$  en  $t$  si y sólo si el libro señalado por  $p$  en  $t$  es robado anteriormente a  $t$ .<sup>20</sup>

<sup>19</sup> Para una bibliografía puesta al día, y discusión, ver A.N. Prior, *Past, Present and Future*, Oxford 1967. [En la edición de 1982 esta nota ha sido cambiada por la que viene a continuación: <sup>19</sup> Esta afirmación ha resultado ser ingenuamente optimista. Para algún trabajo serio sobre el tema, véase S. Weinstein, «Truth and Demonstratives», *Noûs*, 8 (1974), pp. 179-84.]

<sup>20</sup> Hay algo más que una insinuación de este enfoque de los demostrativos y la

Claramente, este procedimiento no muestra cómo eliminar los demostrativos; no hay ninguna sugerencia de que «el libro señalado por el hablante» pueda sustituir a «ese libro» en todos los lugares *salva veritate*. El hecho de que los demostrativos son susceptibles de tratamiento formal debería incrementar en gran medida las esperanzas de una semántica seria del lenguaje natural, puesto que es probable que muchos problemas destacados, tales como el análisis de citas u oraciones sobre actitudes proposicionales, puedan resolverse si reconocemos una construcción demostrativa encubierta.

Ahora que hemos relativizado la verdad a tiempos y hablantes, resulta apropiado el echar una mirada hacia atrás, hacia el problema de comprobar empíricamente una teoría del significado para una lengua ajena. La esencia del método era, como se recordará, correlacionar oraciones mantenidas como verdaderas con oraciones mantenidas como verdaderas por medio de una definición de verdad, y dentro de los límites de error inteligible. Ahora el cuadro debe elaborarse para tomar en cuenta el hecho de que las oraciones son verdaderas, y se mantienen como verdaderas, relativas solamente a un hablante y a un tiempo. Las oraciones con demostrativos proporcionan obviamente un modo muy sensible de comprobar una teoría del significado, y constituyen el lazo más directo entre el lenguaje y los objetos macroscópicos recurrentes de interés y atención humanos<sup>21</sup>.

En este artículo he supuesto que los hablantes de un lenguaje pueden determinar efectivamente el significado o los significados de una expresión arbitraria (si ésta tiene significado), y que la tarea central de una teoría del significado es mostrar cómo esto es posible. He argumentado que una caracterización de un predicado veritativo describe el género requerido de estructura y proporciona un criterio claro y comprobable de una semántica adecuada para un lenguaje natural. Sin duda hay otras exigencias razonables que pueden imponerse a una teoría del significado. Pero una teoría que no haga más que definir verdad para un lenguaje se aproxima más a constituir una teoría del significado completa de lo que un análisis superficial podría sugerir; por ello, al menos, he estado abogando insistentemente.

Puesto que creo que no hay alternativa, he abrazado una visión optimista y programática de las posibilidades para una caracterización formal de un predicado veritativo para un lenguaje natural.

verdad en el artículo de Austin de 1950 «Truth», reimpresso en *Philosophical Papers*, Oxford, 1961. [Versión castellana: *Ensayos filosóficos*, Alianza, Madrid, 1989], ver pp. 89-90. Estoy en deuda con Daniel Bennett por señalarme la necesidad del parentético «potencialmente» en esas formulaciones.

<sup>21</sup> Estas observaciones se derivan claramente de la idea de Quine de que las «oraciones ocasionales» (aquellas con un elemento demostrativo) tienen que desempeñar un papel central en la construcción de un manual de traducción.



Pero hay que conceder que permanece una imponente lista de dificultades y enigmas. Para citar unos pocos: no conocemos la forma lógica de las oraciones contrafácticas o subjuntivas; ni de las oraciones sobre probabilidades o sobre relaciones causales; no tenemos ninguna buena idea de cuál es el papel lógico de los adverbios, ni el papel de los adjetivos atributivos; no tenemos teoría alguna para términos masa tales como «fuego», «agua», y «nieve», ni para oraciones sobre creencia, percepción o intención, ni para verbos de acción que impliquen propósito. Y, finalmente, están todas las oraciones que no parecen tener valores de verdad en absoluto: las imperativas, optativas, interrogativas, y toda una hueste más de ellas. Una teoría comprensiva del significado tiene que habérselas con éxito con cada uno de estos problemas <sup>22</sup>.

<sup>22</sup> Para las soluciones intentadas a algunos de esos problemas, véanse los ensayos «The Logical Form of Action Sentences» (1967), «Causal Relations» (1967), «The Individuation of Events» (1969), «Events as Particulars» (1970) y «Eternal vs. ephemeral Events» (1971), todos ellos recogidos en D. Davidson, *Actions and Events*, Oxford University Press, 1980. También los ensayos «Quotation» (1979), «On saying That» (1968) y «Mood and Performances» (1979), recogidos en D. Davidson, *Inquiries into Truth and Interpretation*, Oxford University Press, 1984 [versión castellana: *De la verdad y de la interpretación*, Gedisa, Barcelona, 1990]. Hay discusiones adicionales en los ensayos «True to the Facts» (1969), «Semantics for Natural languages» (1970), «Radical Interpretation» (1973) [recogido en este volumen pp. 354-369] y «Belief and the Basis of Meaning» (1974), todos ellos recogidos en *Inquiries into Truth and Interpretation*. Hay referencia a algunos progresos en la sección 1 de «Radical Interpretation» [pp. 354-369 de este volumen]. [Nota añadida en 1982.]

## SIGNIFICADO Y VERDAD \*

P.F. STRAWSON

Durante el último cuarto de siglo Oxford ha ocupado, o ha vuelto a ocupar, una posición que había detentado en el pasado, hace quizás seiscientos años: la de un gran centro filosófico del mundo occidental. Durante el mismo período mi predecesor en esta Cátedra, el profesor Gilbert Ryle, ha sido el centro de este centro. Debemos mucho a su visión, su iniciativa y su dedicación a modo de supervisor —un supervisor completamente no autocrático— del desarrollo y la organización, pero debemos mucho más a su fertilidad, su brillantez y su originalidad como filósofo.

Es característico de los filósofos el reflexionar sobre su propia actividad con el mismo temple con que reflexionan sobre los objetos de esa actividad; examinar filosóficamente la naturaleza, los propósitos y los métodos del examen filosófico. El profesor Ryle, cuando ha escrito en esta vena metafilosófica, ha presentado algunas veces el modelo de filósofo bajo una luz un tanto austera: como una persona cuyo papel es corregir una postura mental habitualmente lánguida, desenredar los embotellamientos en el tráfico de las ideas o prescribir los ejercicios correctos para nuestros calambres y confusiones intelectuales. El profesor Ryle ha hecho su contribución a este necesario trabajo de corrección. Pero, cuando contemplamos su producción filosófica como un todo, la impresión que produce no es de austeridad, sino de abundancia; de profusión de perspicacia, vivacidad de ilustración y prontitud de inventiva. Cada tópico recalcitrante es, a su vez, brillantemente iluminado por un método en el que el detalle, las imágenes, el contraste y la generalización se combinan poderosamente. Los tópicos cubren un extenso rango; muchos de ellos caen en las vastas regiones de la filosofía del significado y la filosofía de la mente; y si pudiera, tentativamente, expresar aquí un juicio de preferencia, seleccionaría su tratamiento del *pensar* —tratamiento sobre el cual él ya ha escrito mucho y en el que está todavía ocupado— como quizás lo más sutil y de mayor sensibilidad de todas sus exploraciones filosóficas.

---

\* Versión castellana de Luis Ml. Valdés.

En la obra del profesor Ryle, como en la de algunos otros pocos filósofos, el pensamiento y el estilo son uno: la acumulación de imágenes y epigramas, la antítesis aguda, las tensas y equilibradas oraciones no son adiciones decorativas a su argumentación, sino la forma misma de su pensamiento. Si se tuviera que nombrar una única cualidad como supremamente característica de ese pensamiento y de ese estilo, habrá de ser una característica que, por fuerza he nombrado ya dos veces: *brillantez*. Sus escritos forman una contribución brillante y perdurable no sólo a la filosofía, sino también —lo cual es de la mayor importancia— a las letras inglesas.

¿En qué consiste que algo tenga un *significado*, a la manera, o en el sentido, en que palabras, oraciones o señales tienen significado? ¿En qué consiste que una oración particular tenga el significado o los significados que tiene? ¿En qué consiste que una frase particular, o una palabra particular, tenga el significado o los significados que tiene? Son éstas, obviamente, preguntas conectadas. Cualquier explicación que demos del significado en general (en el sentido relevante) tiene que cuadrar con la explicación que demos de en qué consiste que las expresiones particulares tengan significados particulares; y hemos de reconocer, como dos verdades complementarias, primero, que el significado de una oración en general depende, de alguna manera sistemática, de los significados de las palabras que la componen y, segundo, que el que una palabra tenga un significado particular es un asunto concerniente al hecho de que dicha palabra haga una contribución particular sistemática a los significados de las oraciones en que aparece.

No voy a conformarme a intentar responder a esas preguntas tan obviamente conectadas. No es ésa una tarea para una conferencia; tampoco para un hombre. Quiero más bien discutir cierto conflicto, o conflicto aparente, más o menos débilmente discernible en los enfoques actuales de esas preguntas. Para darle una etiqueta, podríamos llamarlo el conflicto entre los teóricos de la intención comunicativa y los teóricos de la semántica formal. De acuerdo con los primeros, es imposible dar una explicación adecuada del concepto de significado sin hacer referencia a la posesión por parte de los hablantes de intenciones, de un cierto género complejo, dirigidas a un auditorio. Los significados particulares de palabras y oraciones son en gran parte, sin duda, un asunto de regla y convención; pero, la naturaleza general de tales reglas y convenciones sólo puede entenderse a la larga por referencia al concepto de intención comunicativa. El punto de vista opuesto, al menos en su aspecto negativo, es que esta doctrina simplemente elude el asunto de una manera errónea, o lo presenta de una manera errónea, o confunde lo contingente con lo esencial. Desde luego, podemos esperar una cierta regularidad de relación entre lo que la gente intenta comunicar al emitir ciertas oraciones y lo

que esas oraciones significan convencionalmente. Pero el sistema de reglas semánticas y sintácticas, cuyo dominio constituye el conocimiento de un lenguaje —las reglas que determinan el significado de las oraciones— no es en absoluto un sistema de reglas *para* comunicar. Las reglas pueden explotarse para ese propósito, pero esto queda al margen de su carácter esencial. Sería perfectamente posible que alguien comprendiese un lenguaje completamente —tuviese una competencia lingüística perfecta— sin tener tan siquiera el pensamiento implícito de la función de comunicación, con tal que, desde luego, el lenguaje en cuestión no contuviese palabras que se refiriesen explícitamente a esta función.

Una pugna sobre lo que parece ser una cuestión central en filosofía debería tener un cierto carácter homérico; y una pugna homérica exige dioses y héroes. Yo puedo al menos nombrar, aunque tentativamente, algunos capitanes vivientes y algunas benévolas sombras: de un lado, digamos, Grice, Austin y el último Wittgenstein, y del otro, Chomsky, Frege y el primer Wittgenstein.

En primer lugar, pues, lo que respecta a los teóricos de la intención comunicativa. La manera más simple, y más fácilmente inteligible, aunque no la única, de unirse a sus filas es presentar su teoría general del significado en dos estadios: en primer lugar, presentar y elucidar un concepto primitivo de *comunicación* (o intención comunicativa) en términos que no presupongan el concepto de *significado lingüístico*; mostrar a continuación que el último concepto puede ser, y ha de ser, explicado en términos del primero<sup>1</sup>. Para cualquier teórico que siga este sendero, el concepto fundamental de la teoría del significado es el de un hablante o, de manera general, un emisor que *quiere decir significativamente algo* por medio de una emisión dirigida a un auditorio en una ocasión particular. Una emisión es algo producido o ejecutado por un emisor; no necesita ser vocal; podría ser un gesto o un dibujo o el mover y colocar objetos de una manera determinada. Lo que un emisor quiere decir mediante su emisión se especifica incidentalmente al especificar la intención compleja con que el produce la emisión. El análisis del género de intención en cuestión es demasiado complejo para darlo aquí en detalle, de modo que me limitaré a una descripción incompleta. Un emisor podría tener, como una de sus intenciones al ejecutar su emisión, la de inducir a su auditorio a pensar que él, el emisor, cree alguna proposición, digamos la

<sup>1</sup> No la *única* manera, puesto que decir que un concepto  $\phi$  no puede ser elucidado adecuadamente sin referencia a un concepto  $\psi$  no es lo mismo que decir que es posible dar un análisis clásico de  $\phi$  en términos de  $\psi$ . Pero es la manera *más simple*, pues el método clásico de análisis es aquel en términos del cual, en nuestra tradición, pensamos más naturalmente.

proposición de que *p*; y podría intentar que esta intención fuese completamente abierta, que fuese claramente reconocida por el auditorio. O, por otra parte, podría tener la intención de inducir a su auditorio a pensar que él, el emisor, quiere que su auditorio realice alguna acción, digamos *a*; y podría intentar que esta intención suya fuese completamente abierta, fuese reconocida claramente por el auditorio. Entonces, dado que se cumplen otras determinadas condiciones respecto de la intención del emisor, puede decirse que el emisor, en el sentido relevante, quiere decir significativamente algo mediante su emisión: específicamente, que él quiere decir significativamente que *p* en el modo indicativo en el primer caso, y que él quiere decir significativamente, en el modo imperativo, que el auditorio ha de realizar la acción *a* en el segundo caso. Grice nos ha dado razones para pensar que, con suficiente cuidado, y con un refinamiento mucho mayor que el que yo he indicado, es posible exponer tal concepto de intención comunicativa o, como él lo llama, de significado del emisor, que se mantiene a prueba de objeciones y que no presupone la noción de significado lingüístico.

Digamos ahora unas palabras acerca de cómo se supone que procedería el análisis del significado lingüístico en términos del significado del emisor. Aquí tampoco voy a entrar en detalles. Los detalles serían muy complejos. Pero la idea fundamental es comparativamente simple. Estamos acostumbrados a pensar, y de manera razonable, en los significados lingüísticos en términos de reglas y convenciones semánticas y sintácticas. Y cuando consideramos la enorme elaboración de esas reglas y convenciones —su capacidad, como los lenguajes modernos subrayan, de generar un número infinito de oraciones en un lenguaje dado— podemos sentirnos infinitamente separados de la clase de situación de comunicación primitiva en la que pensamos naturalmente cuando intentamos comprender la noción de significado del emisor en términos que no presuponen claramente significado lingüístico. Pero las reglas o convenciones gobiernan las prácticas y actividades humanas que tienen algún propósito. Por tanto, deberíamos preguntar qué actividades que tienen algún propósito están gobernadas por esas convenciones. ¿Para qué sirven esas reglas? Y el pensamiento más simple del que he hablado y que está subyacente en el tipo de análisis sugerido es que esas reglas son, precisamente, reglas para comunicar, reglas mediante cuya observancia el emisor puede lograr su propósito, cumplir su intención comunicativa; y que éste es su carácter *esencial*. Esto es, que no se trata solamente de un hecho afortunado el que esas reglas nos permitan su uso para este propósito; más bien, la misma naturaleza de las reglas que nos ocupan sólo puede entenderse si se contemplan como reglas mediante las cuales puede lograrse este propósito.

Esta simple consideración puede parecer demasiado simple, y de

diversas maneras. Pues está claro que podemos comunicar, y de hecho lo hacemos, cosas muy complicadas mediante el uso del lenguaje; y si hemos de pensar en el lenguaje fundamentalmente como un sistema de reglas para facilitar el logro de nuestras intenciones de comunicación, y si el análisis no ha de ser circular, ¿no es cierto que debemos atribuirnos a nosotros mismos intenciones de comunicación extremadamente complicadas (o al menos, deseos), independientemente de que tengamos a nuestra disposición los medios lingüísticos para cumplir esos deseos? ¿Y no es esto absurdo? Pienso que sí lo es. Pero el programa de análisis no lo requiere. Todo lo que el análisis requiere es que podamos explicar la noción de convenciones de comunicación en términos de la noción de comunicación preconventional a un nivel más bien básico. Dado que podemos hacer esto, entonces hay más de una manera en la que podemos empezar a hacer avanzar nuestra posición sirviéndonos de nuestros cordones de bota lingüísticos, y parece que podemos explicar la noción de convenciones de comunicación en términos de la noción de comunicación preconventional a un nivel más bien básico.

Podemos, por ejemplo, contarnos a nosotros mismos una historietita de la variedad analítico-genética. Supongamos que un emisor tiene éxito en una comunicación preconventional ante un auditorio dado por medio de una emisión, digamos  $x$ . El emisor tiene una intención compleja, *vis-a-vis* al auditorio de la clase que cuenta como una intención comunicativa, y tiene éxito al cumplir esa intención emitiendo  $x$ . Supongamos que la intención primaria era tal que el emisor *quería decir significativamente* que  $p$  al emitir  $x$ ; y, ya que, por hipótesis, logró un éxito comunicativo, el emisor fue *comprendido*, consiguientemente, por su auditorio. Ahora bien, si se presenta más tarde al mismo emisor, en relación con el mismo auditorio, el mismo problema de comunicación, el hecho, conocido por ambos, de que el emisor quiso decir significativamente que  $p$  al emitir  $x$  anteriormente proporciona al emisor una razón para emitir de nuevo  $x$  y al auditorio una razón para interpretar la emisión de la misma manera que antes. (La razón que cada uno tiene es el conocimiento que el otro tiene el conocimiento que él tiene.) Así es fácil ver cómo la emisión de  $x$  llega a convertirse en algo establecido entre este emisor y este auditorio como un medio de querer decir significativamente que  $p$ . Puesto que ha funcionado, se convierte en algo establecido; y entonces funciona *puesto que* está establecido. Es fácil ver cómo podría contarse esta historietita de manera que incluyese no solamente un grupo de dos, sino uno más amplio. Así podemos tener un movimiento a partir de un emisor que preconventionalmente quiere decir significativamente que  $p$  por medio de una emisión de  $x$  hasta la emisión-tipo  $x$  que significa convencionalmente que  $p$  dentro de un grupo, y desde allí volver a los miembros emisores del grupo que quieren decir significati-

vamente que *p* mediante una instancia del tipo, pero ahora *p* de acuerdo con las convenciones.

Desde luego, esta explicación del significado convencional en términos del significado del emisor no es suficiente por sí misma. Pero sólo cubre el caso, o sólo cubre de manera obvia el caso, de emisiones-tipo sin estructura, esto es, de emisiones-tipo cuyo significado se deriva sistemáticamente de los significados de sus partes. Pero una de las características de las emisiones-tipo lingüísticas es que tienen estructura. El significado de una oración es una función sintáctica de los significados de sus partes y la disposición de estas últimas. Sin embargo, no existe en principio ninguna razón por la que una emisión preconventional no habría de tener una cierta complejidad, un género de complejidad que permitiese que un emisor, habiendo logrado un éxito en la comunicación, lograra otro repitiendo una parte de la emisión mientras varía la otra; lo que él quiere decir significativamente en la segunda ocasión tiene algo en común con, y algo que lo diferencia de, lo que él quiere decir significativamente en la primera ocasión. Y si el emisor logra un segundo éxito está abierto el camino para que lleve a establecerse un sistema rudimentario de emisiones-tipo, esto es, para que llegue a ser convencional dentro de un grupo.

Un sistema de convenciones puede modificarse para hacer frente a necesidades que difícilmente podemos imaginar que existan antes que el sistema existiese. Y esta modificación y enriquecimiento pueden crear, a su vez, la posibilidad de pensamientos tales que no podemos comprender lo que sería para alguien tenerlos, sin suponer que tal modificación y enriquecimiento han tenido lugar. De este modo podemos describir un género alternativo de desarrollo. Las intenciones de comunicación primitivas y los éxitos dan lugar al surgimiento de un sistema de significado convencional limitado, que hace posible su propio enriquecimiento y desarrollo que hacen posible, a su vez, la ampliación de pensamiento y las necesidades de comunicación hasta un punto en que otra vez se presiona sobre los recursos existentes del lenguaje que, a su vez, responde a tal presión... Y, desde luego, hay en esto un elemento de misterio, pero así es en cualquier caso en la creatividad social e intelectual humanas.

Todo lo anterior es un modo de bosquejo, lo más esquemático posible, de algunas características sobresalientes de una teoría del significado de la intención comunicativa, y de insinuación respecto de cómo podría hacerse frente a la obvia objeción de que ciertas intenciones comunicativas presuponen la existencia del lenguaje. Todo esto se ha dicho antes y con mucho mayor refinamiento. Pero esto servirá, espero, como base suficiente para la confrontación de los puntos de vista que deseo plantear.

Pasemos ahora, pues, al punto de vista opuesto, al menos aparentemente, que hasta ahora he caracterizado sólo en su aspecto negati-

Desde luego, aquellos que sostienen este punto de vista comparan una zona del campo con sus oponentes. Ambos están de acuerdo en que los significados de las oraciones de un lenguaje están en gran parte determinados por las reglas o convenciones semánticas y sintácticas de ese lenguaje. Ambos están de acuerdo en que los miembros de cualquier grupo o comunidad de personas que comparten el conocimiento de un lenguaje —que tienen una competencia lingüística común— disponen de un instrumento o medio más o menos poderoso de comunicación y, por tanto, de modificar mutuamente las creencias o actitudes o influenciar mutuamente las acciones. Ambos están de acuerdo en que esos medios se usan regularmente de una manera por completo convencional, que lo que la gente intenta comunicar mediante lo que dice está relacionado regularmente con los significados convencionales de las oraciones que emiten. Donde discrepan es en lo que respecta a las relaciones entre las reglas del lenguaje que determinan el significado, de un lado, y la función de comunicación, de otro: un bando insiste, y el otro (aparentemente) se niega a admitir, que la naturaleza general de esas reglas puede comprenderse sólo por referencia a esta función.

El rechazo saca a colación naturalmente una pregunta, a saber: ¿cuál es el carácter general de esas reglas que, en algún sentido, han de ser dominadas por cualquiera que hable y comprenda un lenguaje dado? La respuesta rechazada fundamenta su carácter general en la función social de la comunicación, por ejemplo, de creencias, deseos o instrucciones. Si se rechaza esta respuesta, debe ofrecerse otra. De este modo preguntamos otra vez: ¿cuál es el carácter de esas reglas que determinan el significado?

Me parece que hay solamente un tipo de respuesta que ha sido alguna vez seriamente avanzado o desarrollado, o que necesita ser considerado seriamente, que proporciona una alternativa posible a la tesis del teórico de la comunicación. Se trata de una respuesta que descansa en la noción de condiciones de verdad. El pensamiento de que el sentido de una oración está determinado por sus condiciones de verdad se encuentra en Frege y en el primer Wittgenstein, y lo encontramos de nuevo en muchos escritores posteriores. Tomo como ejemplo un reciente artículo del profesor Davidson. Davidson está directamente interesado en el punto de que una explicación adecuada de las reglas de significado para un lenguaje *L* mostrará cómo los significados de las oraciones dependen de los significados de las palabras en *L*; y una teoría del significado para *L* llevará a cabo esto —dice él— si contiene una definición recursiva de verdad-en-*L*. La «conexión obvia» —dice él— entre tal definición de verdad y el concepto de significado es ésta: «la definición funciona dando las condiciones necesarias y suficientes para la verdad de cada oración, y *dar las condiciones de verdad es un modo de dar el significado de una ora-*



ción. Conocer el concepto semántico de verdad para un lenguaje es saber lo que es para una oración —cualquier oración— ser verdadera, y esto equivale a, en un buen sentido que podemos dar a la frase *comprender el lenguaje*»<sup>2</sup>.

Davidson, en el artículo del que estoy citando, tiene un interés limitado. Pero este interés encuentra su lugar dentro de una idea más general; y la idea general, dicha lisa y llanamente, es que las reglas sintácticas y semánticas juntas determinan los significados de todas las oraciones de un lenguaje y hacen esto, precisamente, por medio de la determinación de sus condiciones de verdad.

Ahora bien, si hemos de llegar a la raíz del asunto, si hemos de aislar la cuestión crucial, me parece importante dejar de lado, al menos inicialmente, una clase de objeciones a la adecuación de tal concepción del significado. He dicho una clase de objeciones, pero se trata de una clase de objeciones que admite subdivisiones. Así pues, puede señalarse que existe algún género de oraciones —por ejemplo imperativas, optativas e interrogativas— para las que la noción de condiciones de verdad parece inapropiada en el sentido de que la emisión convencional de tales oraciones no tiene como resultado el decir nada verdadero o falso. O puede señalarse también que incluso las oraciones a las cuales la noción de condiciones de verdad puede parecer apropiada pueden contener expresiones que ciertamente dan lugar a una diferencia por lo que respecta a su significado convencional, pero no el tipo de diferencia que puede explicarse en términos de sus condiciones de verdad. Compárese la oración «Afortunadamente Sócrates está muerto» con la oración «Desafortunadamente Sócrates está muerto». Compárese una oración de la forma «*p* y *q*» con una oración de la forma «*p* pero *q*». Resulta claro que el significado de los miembros de cada par de oraciones es diferente; dista de estar claro que difieran sus condiciones de verdad. Y no hay solamente una o dos expresiones que den lugar a este problema, sino un gran número de tales expresiones.

Obviamente, tanto una teoría del significado general y comprensiva como una teoría semántica comprensiva para un lenguaje particular deben estar equipadas para habérselas con esos puntos. Con todo, pueden considerarse razonablemente como puntos periféricos. Porque es una verdad reconocida implícitamente por los mismos teóricos de la comunicación<sup>3</sup> el que en casi todas las cosas que contaría

<sup>2</sup> «Truth and Meaning», *Synthese* (1967), p. 310 [versión castellana: pp. 314-314 de este volumen]. Los subrayados son míos.

<sup>3</sup> Este reconocimiento está implícito probablemente, aunque no de manera muy clara, en el concepto austiniano de *significado locucionario* (cf. *How to do Things with Words*. Oxford, 1962 [versión castellana: *Palabras y acciones*, Paidós, Buenos Aires, 1969]; está ciertamente implícito en la distinción de Grice entre lo que los hablantes

mos entre las oraciones existe un núcleo central y substancial de significado que es explicable o en términos de condiciones de verdad o en términos de alguna noción relacionada derivable de manera completamente simple de la de condición de verdad, por ejemplo, la noción, como podríamos llamarla, de condición de obediencia en el caso de una oración imperativa o de condición de cumplimiento en el caso de una optativa. Si suponemos, por tanto, que puede darse una explicación de la noción de condición de verdad misma, una explicación que sea de hecho independiente de la referencia a la intención comunicativa, entonces podemos razonablemente pensar que la mayor parte de la tarea de una teoría general del significado se ha llevado a cabo sin tal referencia. Y corroborando esto, según la misma suposición, podemos pensar que la mayor parte de la teoría particular del significado de un lenguaje particular *L* puede también darse, libre de cualquier referencia, incluso implícita, de este tipo; pues puede darse estableciendo de manera sistemática las reglas semánticas y sintácticas que determinan las condiciones de verdad para las oraciones de *L*.

Desde luego, como ya se ha admitido, algo habrá de añadirse para completar nuestra teoría general y nuestras teorías particulares. Así, a una teoría particular habrá de añadirse una explicación de las transformaciones que producen oraciones con condiciones de obediencia o de cumplimiento a partir de oraciones con condiciones de verdad; y la teoría general tendrá que decir qué clase de cosa es en general, semánticamente hablando, una oración derivada de este tipo. Pero esto, aunque dé como resultado una abundante cosecha de oraciones, es en sí mismo una adición relativamente pequeña, tanto a la teoría particular como a la general. De nuevo, serán necesarias otras adiciones en conexión con las otras objeciones que he mencionado. Pero, movido a confianza por su hipotético éxito, el teórico puede pensar en habérselas con alguna de esas adiciones sin referencia esencial a la intención comunicativa; y, movido a generosidad por este hipotético éxito, puede darse el gusto de conceder derechos en alguna pequeña y externa porción del territorio *de facto* de la semántica teórica al teórico de la intención comunicativa, en lugar de confiar a este último a algún territorio menos apetitoso denominado pragmática teórica.

Espero que resulte claro ahora cuál es el punto central. No consis-

---

*dicen efectivamente*, en un sentido favorecido de «decir», y lo que implican (cf. «Utterer's Meaning, Sentence-Meaning and Word-Meaning», en *Foundations of Language*, 1968); y de nuevo en la distinción de Searle entre la *proposición* que se propone y el modo ilocucionario en que es propuesta (cf. *Speech Acts*, Cambridge, 1969 [versión castellana: *Actos de habla. Ensayo de filosofía del lenguaje*, Cátedra, Madrid, 1980]).

te en otra cosa que la aparentemente simple cuestión de si la noción de condiciones de verdad puede ella misma explicarse o comprenderse sin referencia a la función de comunicación. Se requiere una clarificación menor antes de que vuelva a examinar la cuestión directamente. He usado libremente la frase «las condiciones de verdad de las oraciones» y he hablado de esas condiciones de verdad como de terminadas por las reglas semánticas y sintácticas del lenguaje al que las oraciones pertenecen. En tal contexto nosotros entendemos de manera natural la palabra «oración» en el sentido de una «oración-tipo». (Entiendo por oración en el sentido de tipo el sentido en que hay precisamente una oración castellana que diga «Estoy sintiendo escalofríos», o precisamente una oración castellana que diga «Ella celebró ayer su decimosexto cumpleaños», donde una y la misma oración puede emitirse en incontables ocasiones diferentes, por personas diferentes y con diferentes referencias o aplicaciones.) Pero para muchas oraciones-tipo, tales como esta que acabamos de mencionar, la cuestión de si ellas, las *oraciones*, son verdaderas o falsas es una cuestión que no tiene aplicación natural: no es de las mismas oraciones-tipo invariantes de las que se dice naturalmente que son verdaderas o falsas, sino más bien de las cosas sistemáticamente variantes que las personas dicen, las proposiciones que expresan, cuando emiten esas oraciones en diferentes ocasiones particulares. Pero, si la noción de *valores* de verdad es en general inapropiada para las oraciones-tipo, ¿cómo puede ser apropiada la noción de *condiciones* de verdad? Pues, presumiblemente, las condiciones de verdad de algo son las condiciones bajo las cuales es verdadero.

La dificultad, sin embargo, se resuelve muy fácilmente. Todo lo que es necesario decir es que el enunciado de condiciones de verdad para muchas oraciones-tipo —quizás la mayor parte de las que se emiten efectivamente en la conversación ordinaria— ha de ser relativizado, y puede serlo, de una manera sistemática, a las condiciones contextuales de emisión. Un enunciado general de condiciones de verdad para tal oración será entonces, no un enunciado de condiciones bajo las cuales esa oración es una verdad, sino un enunciado general de un tipo de condiciones bajo las cuales diferentes emisiones particulares de ella darán por resultado diferentes verdades particulares. Y existen otras maneras más o menos equivalentes, aunque más bien menos naturales, de resolver la dificultad.

Vayamos ahora, al fin, al punto central. Para los teóricos de la semántica formal, como los he llamado, el peso total, o la mayor parte del peso, tanto de la teoría general del significado como de las teorías semánticas particulares reside en la noción de condiciones de verdad y, por consiguiente, en la noción de verdad. Acordamos dejar el asunto aquí. Pero no podemos estar todavía convencidos de que tenemos una comprensión general adecuada de la noción de significa-

do a menos que estemos convencidos de que tenemos una comprensión general adecuada de la noción de verdad.

Hay aquí una maniobra que bloquearía toda esperanza de lograr la comprensión adecuada; y, si no estoy equivocado, se trata de una maniobra que tiene un cierto atractivo para algunos teóricos de la semántica formal. Consiste en reaccionar ante la petición de una explicación general de la noción de verdad remitiéndonos a una concepción de la verdad-en-un-lenguaje-dado, *L*, tipo Tarski, una concepción que es elucidada precisamente mediante un enunciado recursivo de las reglas que determinan las condiciones de verdad para oraciones de *L*. Esto equivale a rechazar totalmente el encararse con la cuestión filosófica general. Una vez que estamos de acuerdo en el punto general de que los significados de las oraciones de un lenguaje están determinados, o determinados en gran medida, por las reglas que determinan las condiciones de verdad, planteamos a continuación la cuestión general de qué clase de cosa son las condiciones de verdad, o *de* qué son condiciones las condiciones de verdad; y se nos dice que el concepto de verdad para un lenguaje dado está definido por las reglas que determinan las condiciones de verdad para oraciones de ese lenguaje.

Evidentemente, no podemos quedar satisfechos con esto. Así volvemos a nuestra cuestión general sobre la verdad. E inmediatamente sentimos cierta turbación. Pues hemos llegado a pensar que hay muy poco que decir sobre la verdad *en general*. Pero veamos lo que podemos hacer con este poquito. Hay aquí una manera de decir algo incontrovertido y bastante general sobre la verdad. Alguien que hace un enunciado o aserción hace un enunciado verdadero si y sólo si las cosas son, al hacer el enunciado, como él enuncia que son. O de nuevo, alguien que expresa una suposición expresa una suposición verdadera si y sólo si las cosas son, al expresar la suposición, como él expresamente supone que son. Entretejamos ahora con tales observaciones tan inocuas como éstas los pensamientos en que estábamos de acuerdo sobre significado y condiciones de verdad. Tenemos entonces: primero, el significado de una oración está determinado por aquellas reglas que determinan cómo se enuncia que son las cosas por alguien que, al emitir la oración, hace un enunciado; o cómo se supone expresamente que son las cosas por alguien que, al emitir la oración, expresa una suposición. Y así, recordando que las reglas están relativizadas a condiciones contextuales, podemos parafrasear lo anterior como sigue: el significado de una oración está determinado por las reglas que determinan *qué* enunciado hace alguien que, al emitir la oración en condiciones dadas, hace un enunciado, o que determinan *qué* suposición expresa alguien que, al emitir la oración en condiciones dadas, expresa una suposición, y así sucesivamente.

Así nos vemos retrotraídos, por medio de la noción de verdad, a

la noción del *contenido* de actos de habla tales como enunciar, suponer expresamente, y así sucesivamente. Y aquí el teórico de la intención comunicativa ve su oportunidad. No hay esperanza, dice él, de elucidar la noción del contenido de tales actos de habla sin prestar alguna atención a las nociones de esos mismos actos de habla. Aunque bien, de todos los actos de habla en los cuales, de un modo u otro, puede proponerse algo verdadero o falso, resulta razonable considerar que el enunciado o aseveración tiene una posición especialmente central. (En ascuas tras certezas: valoramos primariamente la especulación puesto que valoramos la información.) Y no podemos eludir, mantiene el teórico, la noción de enunciar o aseverar excepto en términos de intención dirigida a un auditorio. Pues el caso fundamental de enunciar o aseverar en términos de los cuales deben comprenderse todas las variantes es el de emitir una oración con una cierta intención —una intención completamente abierta en el sentido requerido por el análisis del significado del emisor— que puede describirse incompletamente como la intención de hacer que un auditorio sepa, o la de inducirlo a pensar, que el hablante tiene una cierta creencia; como resultado de lo cual puede, o no puede, activarse o producirse en el auditorio esa misma creencia. Las reglas que determinan el significado convencional de la oración se asocian con las condiciones contextuales de su emisión para determinar lo que la creencia en cuestión es en tal caso primario y fundamental. Y al determinar lo que la creencia en cuestión es en tal caso, las reglas determinan qué enunciado se hace en ese caso. Determinar lo primero es determinar lo último. Pero esto es precisamente lo que nosotros buscábamos. Puesto que cuando partíamos del punto de acuerdo de que las reglas que determinan las condiciones de verdad determinan de este modo el significado, la conclusión a que fuimos llevados era precisamente que aquellas reglas determinaban qué enunciado hace alguien que, al emitir la oración, hace un enunciado. De esta manera, el punto de acuerdo, lejos de ser una alternativa a la teoría comunicativa del significado, nos lleva directamente a tal teoría del significado.

La conclusión puede parecer un poco demasiado rápida. Así pues, veamos si hay alguna manera de evitarla. La condición general para evitarla resulta clara. Consiste en que seamos capaces de dar una explicación de la noción de condiciones de verdad que no incluya referencia esencial a actos de habla comunicativos. La alternativa de rehusar el dar cualquier explicación en absoluto —de apoyarse sólo en la noción de condiciones de verdad— simplemente no nos está abierta, como ya he indicado, si estamos interesados en la elucidación filosófica de la noción de significado: nos dejaría con los conceptos de significado y verdad señalando el uno hacia el otro vacua e inutilmente. Tampoco serviría de ayuda, por más que podría ser

hablante en este punto, retroceder desde la noción de condiciones de verdad a la noción menos específica de correlación en general; decir, simplemente, que las reglas que determinan el significado de las oraciones lo hacen correlacionando las oraciones, contempladas como emitidas en ciertas condiciones contextuales, con ciertos estados de cosas posibles. Una razón por la que esto no ha de funcionar es que la noción de correlación en general es simplemente demasiado inespecífica. Hay muchos géneros de conducta (incluyendo la conducta verbal) —y podrían imaginarse muchos más géneros—, que están correlacionados mediante reglas con estados de cosas posibles sin que sea el caso que tal correlación les confiera el género de relación con esos estados de cosas posibles en los que estamos interesados.

Otra razón por la cual esto no funcionará es la siguiente: Considérese la oración «Estoy cansado». Las reglas que determinan su significado son en efecto tales que correlacionan la oración, considerada como emitida por un hablante particular en un tiempo particular, con el estado de cosas posible consistente en que el hablante está cansado en ese tiempo. Pero esta característica no es peculiar a esa oración o a los miembros de la clase de oraciones que tienen el mismo significado que ella. Pues considérese la oración «No estoy cansado». Las reglas que determinan su significado son también tales que correlacionan la oración, contemplada como emitida por un hablante determinado en un tiempo determinado, con el estado de cosas posible consistente en que el hablante está cansado en ese tiempo. Desde luego, los géneros de correlación son diferentes. Son tales, respectivamente, que a aquel que emitiese la primera oración se le comprendería normalmente como afirmando, y a aquel que emitiese la segunda oración se le comprendería normalmente como negando, que el estado de cosas en cuestión se daba; o, de nuevo, son tales que aquel que emite la primera oración cuando el estado de cosas en cuestión se da ha hecho un enunciado verdadero, y aquel que emite la segunda oración en esas circunstancias ha hecho un enunciado falso. Pero invocar esas diferencias sería precisamente renunciar a la idea de emplear sólo la noción no específica de correlación en general. No merece la pena elaborar más este punto. Pero fácilmente se verá que no sólo oraciones diferentes, incluso opuestas, en significado están correlacionadas, de una manera u otra, con el mismo estado de cosas posible, sino que una y la misma oración no ambigua está correlacionada, de una manera u otra, con muchos estados de cosas diferentes y en algunos casos mutuamente incompatibles. La oración «Estoy cansado» está correlacionada con el estado de cosas posible consistente en que el hablante está en un punto total de agotamiento, y también con el estado de cosas consistente en que está fresco como una rosa. La oración «Tengo más de cuarenta años» está correlacionada con cualquier estado de cosas posible respecto a la edad del hablante, y la de

«Los cisnes son blancos» con cualquier estado de cosas respecto del color de los cisnes.

La completamente inespecífica noción de correlación carece entonces de uso para el propósito que llevamos entre manos. Es necesario encontrar alguna manera de especificar una correlación particular en cada caso, a saber: la correlación de la oración con el estado de cosas posible cuyo darse sería necesario y suficiente para que se hubiera dicho algo *verdadero* al emitir la oración bajo cualesquiera condiciones contextuales que se contemplen. Así volvemos, una vez más, a la noción de condiciones de verdad y a la cuestión de si podemos dar una explicación de esta noción que no incluya referencias esenciales a actos de habla comunicativos, esto es, a la intención comunicativa.

En este punto solamente acierto a ver un recurso abierto, o aparentemente abierto, para el teórico del significado que aún mantenga que la noción de intención comunicativa no tiene un lugar esencial en el análisis del concepto de significado. Si no ha de tragar el anzuelo de su oponente, tiene que imitarle. Él ve ahora que no puede detenerse en la idea de verdad. Esta idea lleva directamente a la idea de *lo que es dicho*, del contenido de lo que es dicho, cuando se hacen las emisiones; y esto, a su vez, a la cuestión de lo que está siendo *hecho* cuando se hacen las emisiones. Pero ¿no puede el teórico recorrer algún trecho de este camino sin ir tan lejos como su oponente? ¿No sería posible *suprimir* la referencia a la intención comunicativa *preservando* una referencia a, digamos, expresión de creencia? ¿Y no será esto, dicho sea de paso, más realista en tanto que nosotros a menudo nos vocalizamos nuestros pensamientos a nosotros mismos, sin ninguna intención comunicativa?

La maniobra propuesta merece una descripción más completa. Es como sigue. En primer lugar, sigue al teórico de la intención comunicativa al responder al desafío de que proporcione una elucidación de la noción de condiciones de verdad invocando la noción de, por ejemplo, y centralmente, enunciado o aserción (aceptando el punto incontrovertido de que alguien hace un enunciado o aserción verdaderos cuando las cosas son como, al hacer esa aserción, él asevera que son). En segundo lugar, sigue de nuevo al teórico de la comunicación al responder al desafío de que proporcione una elucidación de la noción de aseverar haciendo una conexión con la noción de creencia (concediendo que hacer una aserción es, en el caso primario, dar expresión a una creencia; hacer una aserción verdadera es dar expresión a una creencia correcta; y una creencia es correcta cuando las cosas son como aquel que sostiene esa creencia, cree que son). Pero, en tercer lugar, se aparta del teórico de la comunicación en lo que respecta a la naturaleza de esta conexión entre aserción y creencia; esto es: niega que el análisis de la noción de aseverar inclu-

ya referencia esencial a una intención; por ejemplo, lograr que un auditorio piense que el que hace la aserción mantiene la creencia; niega que el análisis de la noción de aseverar incluya *algún* género de referencia a intención dirigida a un auditorio; mantiene, por el contrario, que es perfectamente satisfactorio aceptar aquí la noción de simplemente vocalizar o expresar una creencia. Entonces concluye que las reglas determinantes del significado para una oración del lenguaje son las reglas que determinan *qué* creencia es convencionalmente articulada por alguien que, en condiciones contextuales dadas, emite la oración. Como antes, determinar qué es esta creencia, es la misma cosa que determinar qué aserción se hace. Así se preservan todos los méritos de la teoría del oponente, mientras que la referencia a la comunicación es empujada hacia afuera.

Desde luego, este teórico y su oponente deben decir muchas más cosas. Pues las oraciones que pueden usarse para expresar creencias no tienen por qué ser usadas siempre de esta manera. Pero la observación ha de hacerse a ambas partes. Por esto podemos ignorarla por ahora.

Ahora bien, ¿servirá esto? Creo que no. Pero para ver que no ha de servir, es posible que tengamos que luchar duramente contra una determinada ilusión. Pues la noción de expresar una creencia puede parecerse perfectamente llana; y, por tanto, la noción de expresar una creencia de acuerdo con ciertas convenciones puede parecerse igualmente llana. Con todo, en tanto que la noción de expresar una creencia es la noción que necesitamos, puede tomar prestada toda su fuerza y aparente llaneza precisamente de la situación de comunicación de cuya dependencia se suponía que liberaba al análisis del significado. Podemos sentirnos tentados a argumentar como sigue. A menudo expresamos creencias con una intención dirigida a auditorio; intentamos que nuestro auditorio considere que tenemos la creencia que expresamos y quizás también que esa creencia sea activada o producida en el auditorio. Pero, entonces, ¿podría haber algo más sencillo que esto: que lo que hacemos con una intención dirigida a auditorio lo podemos hacer también sin ninguna intención semejante? Esto es, la intención dirigida a auditorio, cuando está presente, es algo sobreañadido a la actividad de expresar una creencia y no es de ningún modo esencial a ella, o al concepto de ella.

Ahora bien, ¡qué mezcla de verdad y falsedad, de perogrullada e ilusión tenemos aquí! Supóngase que reconsideráramos por un momento ese análisis del significado del emisor que fue someramente bosquejado al comienzo. El emisor produce algo —su emisión *x*— con una intención compleja dirigida a auditorio, que incluye, digamos, inducir al auditorio a pensar que él tiene una cierta creencia. No podemos separar o extraer del análisis un elemento que corresponda a su expresar una creencia sin ninguna intención tal, aunque podría-



mos producir efectivamente la descripción siguiente e imaginar su defensa: él actúa *como si* tuviese tal intención aunque, de hecho, no la tiene. Pero aquí la descripción depende de la descripción del caso en que él tiene tal intención.

Lo que estoy sugiriendo es que podemos sentirnos tentados, aquí como en otras partes, por un género de aritmética de conceptos ficticia. Dado el concepto de Expresión de Creencia Dirigida a Auditorio (ECDA), podemos en efecto pensar en Expresión de Creencia (EC) sin Dirección a Auditorio (DA), y encontrar casos de esto. Pero no se sigue que el concepto de ECDA es un género de compuesto lógico, integrado, a su vez, por los dos conceptos más simples DA y EC, y, por consiguiente, que EC es conceptualmente independiente de ECDA.

Estas observaciones no muestran, desde luego, que no hay un concepto independiente de expresión de creencia que haya de responder a las necesidades del teórico de la anticomunicación. Se trata solamente de observaciones dirigidas contra un argumento demasiado simple al efecto de que hay tal concepto.

Esto está muy claro. Si existe tal concepto esencialmente independiente de expresión de creencia que haya de responder a las necesidades del análisis de la noción de significado, no podemos detenernos justamente en la frase «expresar una creencia». Tenemos que ser capaces de dar alguna *explicación* de este concepto, de contarnos a nosotros mismos alguna historieta inteligible sobre él. Algunas veces podemos decir razonablemente de las acciones de un hombre o de su conducta que expresan una creencia cuando, por ejemplo, vemos que esas acciones están dirigidas hacia un fin o meta que es plausible adscribirle en tanto que es también plausible adscribirle esa creencia. Pero esta reflexión no nos lleva por sí misma muy lejos. Por una parte, según el presente programa, le está vedado hacer referencia al fin o meta de la comunicación como parte esencial de nuestra historieta. Por otra parte, la clase de conducta en la que estamos interesados ha de ser, o ha de ser capaz de ser, formalizada o convencionalizada de tal manera que pueda ser considerada como sujeta a, o realizada por, la observancia de reglas; y de reglas, además, que regulan la conducta precisamente en su aspecto como expresión de creencia. No valdrá el decir simplemente: podríamos suponer que una persona encuentra *alguna* satisfacción (inespecificada) o *algun* objeto (inespecificado) al realizar ciertas acciones (quizás vocales) en algunas ocasiones, estando relacionadas esas acciones sistemáticamente con el hecho de que esa persona tenga ciertas creencias. Pues supongamos que una persona tuviese una práctica de vocalizar de determinada manera siempre que ve que el sol sale y de otra manera, en parte similar, en parte diferente, siempre que ella ve que se pone. Entonces esta práctica estaría relacionada regularmente con

ciertas creencias, esto es, que el sol estaba saliendo o que se estaba poniendo. Pero esta descripción no nos da ninguna razón en absoluto para decir que cuando esa persona estaba entregándose a esa práctica estaba *expresando la creencia* de que el sol estaba saliendo o poniéndose, de acuerdo con una regla para hacerlo así. Realmente no tenemos una descripción suficiente para saber *qué* decir. Por lo que podemos contar, podríamos decir nosotros, esa persona tiene este ritual consistente en *saludar* la salida y la puesta del sol de esta manera. No sabemos qué necesidades tuyas satisface esto.

Sin embargo, supongamos —en atención al argumento— que podemos elaborar alguna concepción relevante de expresar una creencia que no presuponga nada que, según el presente programa, le está vedado presuponer; y que trazamos este concepto de expresar una creencia para dar cuenta, o proporcionar un análisis, según las líneas indicadas, de la noción de significado lingüístico. Resulta entonces una consecuencia interesante. Esto es, aparecerá como una verdad completamente contingente sobre el lenguaje el que las reglas o convenciones que determinan los significados de las oraciones del lenguaje son reglas o convenciones públicas o sociales. Esto será, por así decirlo, un hecho natural, un hecho de la naturaleza, de ninguna manera esencial al concepto de lenguaje, y que reclama una explicación natural a la que no se le debe permitir tocar o modificar ese concepto. No debe haber nada en el *concepto* que excluya la idea de que todo individuo podría tener su propio lenguaje que solamente él comprende. Pero entonces podría preguntarse: ¿Por qué habría de observar cada individuo sus propias o cualesquiera reglas? ¿Por qué no habría de expresar cualquier creencia que quisiese de cualquier modo que se le antojase cuando tuviese el impulso de expresarla? Hay al menos una respuesta a esta pregunta que le está vedado dar al teórico, aunque sólo sea en interés de su propio programa. El no puede decir: bien, una persona podría querer *registrar* sus creencias de modo que pudiese referirse a los registros más tarde y entonces encontraría conveniente el tener reglas para interpretar sus propios registros. Al teórico le está vedado el dar esta respuesta porque introduce, aunque de forma atenuada, el concepto de intención comunicativa: la persona primera se comunica con su yo posterior.

Podría haber una manera de acallar las dudas que surgen tan rápidamente a lo largo de esta senda. Consistiría en ofrecer explicaciones naturales posibles del supuesto hecho natural de que el lenguaje es público, de que las reglas lingüísticas son más o menos reglas socialmente comunes; explicaciones que evitasen con éxito cualquier sugerencia de que la conexión de las reglas públicas con la comunicación no era algo sino incidental y contingente. ¿Cómo podría funcionar tal explicación? Podríamos decir que era un punto de acuerdo el que la posesión de un lenguaje desarrolla la mente, que hay creencias que

no podrían expresarse sin un lenguaje para expresarlas, pensamientos que no se podrían considerar sin un sistema de expresiones gobernado por reglas para articularlos. Y es un hecho el que los seres humanos no adquirirían simplemente dominio de tal sistema a menos que estuvieran expuestos, cuando niños, a condicionamiento o entrenamiento por los miembros adultos de una comunidad. Sin que nos ocupemos de los orígenes remotos del lenguaje, podemos entonces suponer que los miembros adultos de una comunidad desean que sus sucesores tengan a su disposición este instrumento de desarrollo de la mente, y evidentemente el procedimiento total de entrenamiento se simplificará si todos ellos enseñan el mismo, el lenguaje común. Podemos suponer razonablemente que los aprendices, al principio, no se dan cuenta por completo de lo que habrán de hacer a la larga con el lenguaje; se trata para ellos, al principio, de aprender a hacer la cosa correcta más bien que de aprender a decir la cosa verdadera. Esto es, se trata de responder vocalmente a situaciones de una manera en la que alcanzarán recompensa o evitarán castigo, más bien que *expresar sus creencias*. Pero más tarde llegan a darse cuenta de que han dominado un sistema que les capacita para realizar esta (aun inexplicada) actividad siempre que lo deseen, y *entonces* están hablando un lenguaje.

Desde luego, tiene que admitirse que en el proceso son susceptibles también de adquirir la destreza *secundaria* de comunicar sus creencias. Pero esto es simplemente algo añadido, un beneficio extra y conceptualmente fuera de contrato que queda completamente al margen de la descripción de lo que es haber dominado las reglas de significado del lenguaje. Si, en efecto, se dirigen emisiones inequívocas, cuya función especial es expresión de creencia, a otro miembro de la comunidad, éste tenderá a considerar que se mantienen cualesquiera creencias que estén es cuestión y que efectivamente se intenta que él considere que esto es así; y este hecho puede dar lugar — en realidad ha de admitirse que da lugar — a un racimo completo de consecuencias sociales, y saca a la luz toda suerte de posibilidades de géneros de comunicación lingüística distintos del que se basa en expresión de creencia. Por ello, como ya se ha reconocido, podemos tener que permitir a la larga alguna referencia esencial a la intención comunicativa en porciones externas de nuestra teoría semántica. Pero incurrimos en este riesgo solamente cuando vamos más allá del núcleo central de significado, determinado por las reglas que determinan las condiciones de verdad. Por lo que concierne al núcleo central, la función de comunicación permanece como secundaria, derivativa, conceptualmente inesencial.

Espero que esté claro que cualquier historieta de este tipo ha de ser demasiado perversa y arbitraria para satisfacer las exigencias de una teoría aceptable. Si es éste el modo en que ha de plantearse el

juego, entonces debe concedérsele al teórico de la comunicación el haber ganado.

Pero, finalmente, ¿ha de plantearse el juego de esta manera? Pienso, finalmente, que tiene que jugarse así. En efecto, es una cosa inocua y saludable el decir que conocer el significado de una oración es saber bajo qué condiciones el que la emite dice algo verdadero. Pero si deseamos una elucidación filosófica del concepto de significado, entonces el *dictum* representa no el final, sino el comienzo de nuestra tarea. Simplemente restringe, y cambia de lugar, nuestro problema, forzándonos a indagar qué es lo que contiene la frasecita «...dice algo verdadero». Desde luego hay muchas maneras en las que se puede decir algo que es de hecho verdadero, dar expresión, si se quiere, a una proposición, sin expresar, por ese medio, creencia en ello, sin aseverar esa proposición: por ejemplo, cuando las palabras en cuestión forman ciertos tipos de cláusulas subordinadas o coordinadas, cuando se está citando o actuando, y así sucesivamente. Pero cuando nos ponemos a intentar explicar en general qué es decir algo verdadero, expresar una proposición verdadera, es ineludible la referencia a creencia o a aseveración (y, por medio de esto, a creencia). Entonces podemos aventurar inocuamente: alguien dice algo verdadero si las cosas son como él dice que son. Pero éste «dice» tiene ya la fuerza de «asevera». O, para evitar el «dice» que equivale a «asevera», podemos aventurar inocuamente: alguien propone de una manera u otra una proposición verdadera si las cosas son como cualquiera que creyese lo que él propone, creería por este medio que son. Y aquí la referencia a creencia es explícita.

La referencia, directa o indirecta, a expresión de creencia es inseparable del análisis de decir algo verdadero (o falso). Y, como he intentado mostrar, no es realista hasta el punto de ser ininteligible —o, al menos, es de una perversidad extrema— el intentar liberar la noción de expresión lingüística de creencia de toda conexión esencial con el concepto de intención comunicativa.

Anteriormente insinué que el hábito de algunos filósofos de hablar como si «verdadero» fuese un predicado de oraciones-tipo era solamente una aberración menor, que podría bastante fácilmente acomodarse a los hechos. Y así se puede hacer. Pero no se trata de una simple pedantería el insistir en corregir la aberración. Pues si no andamos con cuidado es susceptible de inducirnos totalmente a error. Es susceptible, cuando indagamos sobre la naturaleza del significado, de hacernos olvidar *para* qué son las oraciones. Conectamos significado con verdad, y verdad, demasiado simplemente, con oraciones; y las oraciones pertenecen al lenguaje. Pero, como los teóricos, nosotros no sabemos nada del *lenguaje* humano, a menos que comprendamos el *habla* humana.

## INTERPRETACIÓN RADICAL \*

DONALD DAVIDSON

Kurt emite las palabras «Es regnet» y, bajo las condiciones correctas, sabemos que él ha dicho que está lloviendo. Una vez que hemos identificado su emisión como intencional y lingüística, somos capaces de continuar e interpretar sus palabras: podemos decir lo que sus palabras significaban en esa ocasión. ¿Qué podríamos saber que nos capacitase para hacer esto? ¿Cómo podríamos llegar a saberlo? La primera de esas preguntas no es la misma que la pregunta acerca de si sabemos *efectivamente* lo que nos capacita para interpretar las palabras de otros. Pues podría fácilmente haber algo que podríamos saber y no sabemos, cuyo conocimiento fuese suficiente para la interpretación, mientras que, por otra parte, no es absolutamente obvio que haya algo que de hecho sepamos y que desempeñe un papel esencial en la interpretación. La segunda pregunta, cómo podríamos llegar a tener un conocimiento que nos sirviese para producir interpretaciones, no es algo que, desde luego, concierna a la historia efectiva de la adquisición del lenguaje. Se trata, pues, de una pregunta doblemente hipotética: dada una teoría que hiciera posible la interpretación, ¿qué evidencia plausiblemente disponible a un intérprete potencial apoyaría la teoría hasta un grado razonable? En lo que sigue intentaré afinar esas preguntas y sugerir respuestas.

El problema de la interpretación es tanto doméstico como foráneo; se plantea entre hablantes del mismo lenguaje en la forma de la pregunta: ¿cómo se puede determinar que el lenguaje es el mismo? Los hablantes del mismo lenguaje pueden proceder de acuerdo con la suposición de que, para ellos, las mismas expresiones han de interpretarse de la misma manera, pero esto no indica lo que justifica la suposición. Toda comprensión del habla de otro implica interpretación radical. Pero nos ayudará a mantener que las suposiciones no pasen inadvertidas el concentrarnos en casos donde más claramente se requiere la interpretación: la interpretación en un idioma del habla en otro idioma <sup>1</sup>.

¿Qué conocimiento serviría para la interpretación? Una respuesta

\* Versión castellana de Luis Ml. Valdés Villanueva.

<sup>1</sup> Aquí, y a lo largo de este artículo, mi deuda con el trabajo de W.v.O. Quine será obvia. El término «interpretación radical» quiere sugerir una fuerte afinidad con la

la breve sería la siguiente: el conocimiento de lo que significa cada expresión significativa. En alemán, esas palabras que Kurt dijo significan que está lloviendo y Kurt estaba hablando alemán. Así, al emitir las palabras «Es regnet», Kurt dijo que estaba lloviendo. Esta réplica, como podría pensarse en principio, no vuelve a enunciar meramente el problema. Pues sugiere que al pasar de una descripción que no interpreta (su emitir las palabras «Es regnet») a una descripción interpretante (su decir que está lloviendo) tenemos que introducir un mecanismo de palabras y expresiones (que pueden o no simplificarse en emisiones efectivas), y esta sugerencia es importante. Pero la réplica no sirve de más ayuda, pues no dice lo que es saber lo que significa una expresión.

Hay ciertamente también una ligera indicación de que correspondientemente a cada expresión significativa hay una entidad, su significado. Se ha demostrado que esta idea, sino errónea, sirve de muy poca ayuda: en el mejor de los casos hipostatiza el problema.

El desencanto respecto a los significados en tanto que llevando a cabo una explicación viable de la comunicación o interpretación ayuda a explicar por qué algunos filósofos han intentado abandonar no solamente los significados, sino cualquier teoría seria en absoluto. Resulta tentador, cuando los conceptos que convocamos para intentar explicar la interpretación resultan ser más frustrantes que el *explanandum*, hacer la reflexión de que después de todo la comunicación verbal no consiste en nada más que perturbaciones elaboradas producidas en el aire que forman un eslabón causal entre las actividades no lingüísticas de los agentes humanos. Pero aunque los discursos interpretables no son nada más que (esto es, son idénticos con) acciones realizadas con variadas intenciones no lingüísticas (aconsejar, controlar, divertir, distraer, insultar), y esas acciones no son a su vez más que (son idénticas con) movimientos intencionales de los labios y la laringe, esta observación no nos hace avanzar nada hacia una explicación general inteligible de lo que podríamos saber que nos permitiese redescubrir emisiones no interpretadas como las emisiones correctamente interpretadas.

La apelación a los significados nos deja más varados aún que cuando empezábamos a partir de las actividades no lingüísticas que tenían que proporcionar la base evidencial para la interpretación; la actitud «nada más que» no proporciona ninguna clave de cómo está relacionada la evidencia con aquello para lo que seguramente es evidente.

---

\* «interpretación radical» de Quine (*Word and Object*, Cambridge, Mass., 1960 [versión castellana: Labor, Barcelona, 1968]). Sin embargo, afinidad no es identidad, e «interpretación» en lugar de «traducción» marca una de las diferencias: un mayor énfasis sobre lo explícitamente semántico.

Otras propuestas para llenar el vacío fracasan de varias maneras. Las teorías «causales» de Ogden y Richards y de Charles Morris intentaban analizar el significado de las oraciones, tomadas una por una, sobre la base de datos conductistas. Incluso si esas teorías hubiesen funcionado para las oraciones más simples (que claramente no lo hacen), no han tocado el problema de extender el método a oraciones de mayor complejidad y abstracción. Las teorías de otro género empiezan intentando conectar palabras más bien que oraciones con hechos no lingüísticos. Esto es prometedor puesto que las palabras son finitas en número mientras que las oraciones no lo son, y sin embargo cada oración no es más que una concatenación de palabras; esto ofrece la oportunidad de una teoría que interprete cada una de las infinitas oraciones usando solamente recursos finitos. Pero tales teorías fallan a la hora de alcanzar la evidencia, pues parece claro que las características semánticas de las palabras no pueden explicarse directamente sobre la base de fenómenos no lingüísticos. La razón es simple. Los fenómenos a los que tenemos que volvernos son los intenciones y actividades extralingüísticos al servicio de los cuales está el lenguaje, y las palabras están al servicio de ellos solamente en la medida en que las palabras están incorporadas (o siempre que suceda que lo están) en oraciones. Pero no hay oportunidad de dar una explicación fundacional de las palabras antes de darla de las oraciones.

Por razones completamente diferentes, la interpretación radical no puede esperar tomar como evidencia para el significado de una oración una explicación de las intenciones complejas y delicadamente discriminadas con las que se emite típicamente la oración. No es fácil ver cómo un enfoque tal podría habérselas con el rasgo estructural y recursivo del lenguaje que es esencial para explicar cómo pueden entenderse nuevas oraciones. Pero la dificultad central es que no podemos esperar añadir un sentido a la atribución de intenciones finamente discriminadas independientemente de la interpretación del habla. La razón es que no podemos plantear cuestiones necesarias, sino que interpretar las intenciones de un agente, sus creencias y sus palabras son partes de un proyecto único, ninguna de cuyas partes puede asumirse que esté completa hasta que el resto lo esté. Si esto es correcto, no podemos hacer de la panoplia completa de intenciones y creencias la base evidencial de una teoría de la interpretación radical.

Estamos ahora en posición de decir algo que sirva para hacer posible la interpretación. El intérprete tiene que ser capaz de entender cualquiera de las infinitas oraciones que el hablante podría emitir. Si hemos de enunciar explícitamente lo que el intérprete podría conocer de modo que lo capacitase para hacer esto, tenemos que ponerlo en forma finita<sup>2</sup>. Si ha de cumplirse este requisito, cualquier espe-

<sup>2</sup> En alguna época yo estuve convencido de que, a menos que una teoría tal ca-

anza de un método universal de interpretación debe abandonarse. Lo máximo que puede esperarse es explicar cómo un intérprete podría interpretar las oraciones de los hablantes de un único lenguaje (o un número finito de lenguajes): carece de sentido pedir una teoría que proporcionase una interpretación explícita para cualquier emisión en cualquier (posible) lenguaje.

No está aún claro, desde luego, en qué consiste para una teoría proporcionar una interpretación explícita de una emisión. La formulación del problema parece invitarnos a pensar en la teoría como la especificación de una función que toma las emisiones como argumentos teniendo las interpretaciones como valores. Pero entonces las interpretaciones no serían mejor que los significados y seguramente no mejor que entidades de algún género misterioso. Así pues, parece prudente describir lo que se quiere de la teoría sin referencia aparente a significados o interpretaciones: alguien que conoce la teoría puede interpretar las emisiones a las que la teoría se aplica.

El segundo requisito general que se impone a una teoría de la interpretación consiste en que pueda ser apoyada o verificada por evidencia plausible disponible a un intérprete. Puesto que la teoría es general —tiene que aplicarse a potencialmente infinitas oraciones— resulta natural pensar en la evidencia que está de su parte como instancias de interpretaciones particulares reconocidas como correctas. Y en este caso surge, desde luego, para el intérprete que trata con un lenguaje que él ya conoce. El hablante de un lenguaje no puede producir una teoría finita explícita para su propio lenguaje, pero puede poner a prueba una teoría propuesta puesto que puede decir si ésta proporciona interpretaciones correctas cuando se aplica a emisiones particulares.

En la interpretación radical, sin embargo, se supone que la teoría proporciona una comprensión de emisiones particulares que no se da de antemano, de modo que la última evidencia para la teoría no puede consistir en interpretaciones correctas de muestras. Para habérselas con el caso general, la evidencia tiene que ser de una clase que tendría que estar disponible a cualquiera que no sepa previamente cómo interpretar las emisiones para tratar con las cuales está diseñada la teoría: tiene que ser evidencia que pueda ser enunciada sin

---

caracterizada finitamente pudiera ser proporcionada para un lenguaje, el lenguaje no podría ser aprendido por una criatura con poderes finitos. (Ver Donald Davidson, «Theories of Meaning and Learnable Languages», en *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, Amsterdam, 1966, pp. 383-394 [versión castellana en: *De la verdad y la interpretación*, Gedisa, Barcelona, 1990].) Esto aún me parece igualmente correcto, pero George Kreisel me ha hecho comprender que la idea no es obvia.



uso esencial de conceptos lingüísticos tales como significado, interpretación, sinonimia y cosas parecidas.

Antes de decir qué género de teoría pienso que cumpliría estos propósitos, quiero discutir una última sugerencia alternativa a saber: que toda la teoría que se necesita es un método de traducción del lenguaje que se va a interpretar al lenguaje del intérprete. Tal teoría consistiría en el enunciado de un método efectivo para pasar de una oración arbitraria de la lengua ajena a una oración de un lenguaje familiar; de este modo satisfaría la exigencia de un método finalmente enunciado aplicable a cualquier oración<sup>3</sup>. Pero no pienso que un manual de traducción sea la mejor forma que puede tomar una teoría de la interpretación.

Cuando nuestro punto de mira es la interpretación, un método de traducción se ocupa de un tópico erróneo, una relación entre dos lenguajes, donde lo que se desea es una interpretación de uno (en otro, desde luego, pero esto no hace falta decirlo puesto que cualquier teoría es en algún lenguaje). No podemos sin ocasionar alguna confusión contar el lenguaje usado al enunciar la teoría como parte del tema de la teoría a menos que explícitamente lo hagamos así. En el caso general, una teoría de la traducción comporta tres lenguajes: el lenguaje objeto, el lenguaje sujeto, y el metalenguaje (los lenguajes a partir de que y en el cual tiene lugar la traducción, y el lenguaje de la teoría, que dice qué expresiones del lenguaje sujeto traducen qué expresiones del lenguaje objeto). Y, en este caso general, podemos saber qué oraciones del lenguaje sujeto traducen qué oraciones del lenguaje objeto sin saber qué significa ninguna de las oraciones de cualquiera de los dos lenguajes (en algún sentido esto permitiría, de todos modos, que alguien que entendiese la teoría interpretase las oraciones del lenguaje objeto). Si el lenguaje sujeto resulta ser idéntico con el lenguaje de la teoría, entonces alguien que entendiese la teoría podría sin duda usar el manual de traducción para interpretar emisiones ajenas; pero esto sucede porque pone en uso dos cosas que él sabe y que la teoría no enuncia: el hecho de que el lenguaje sujeto es el suyo propio, y su conocimiento de cómo interpretar emisiones de su propio lenguaje.

Resulta complicado intentar hacer explícita la suposición de que una oración mencionada pertenece al lenguaje de uno mismo. Po-

<sup>3</sup> La idea de un manual de traducción con constricciones empíricas apropiadas como un recurso para estudiar problemas en la filosofía del lenguaje es, desde luego, de Quine. La idea inspiró mucho de mi pensamiento sobre el presente tema, y mi propuesta está en importantes aspectos muy cercana a la de Quine. Puesto que Quine puede no haber intentado responder las cuestiones que yo he establecido, la pretensión de que el método de traducción no es adecuado como solución al problema de la interpretación radical puede no ser una crítica de ninguna doctrina de Quine.

intentamos probar, por ejemplo, de esta manera: «“Es regnet” en el lenguaje de Kurt se traduce como “Está lloviendo” en el mío»; pero la autorreferencia indéxica está fuera de lugar en una teoría que debiera funcionar para cualquier intérprete. Si decidimos aceptar esta dificultad, queda el hecho de que el método de traducción deja de manera tácita y fuera del alcance de la teoría lo que necesitamos saber de modo que nos permita interpretar nuestro propio lenguaje. Una teoría de la traducción debe de leer algún tipo de estructura en las oraciones, pero no hay ninguna razón para esperar que proporcione intuición alguna respecto a cómo los significados de las oraciones dependen de su estructura.

Una teoría satisfactoria para interpretar las emisiones de un lenguaje, incluido el nuestro, habrá de revelar estructuras semánticas significativas: la interpretación de las emisiones de oraciones complejas dependerán sistemáticamente de la interpretación, por ejemplo, de emisiones de oraciones más simples. Supongamos que hemos de añadir a una teoría de la traducción una teoría satisfactoria de la interpretación para nuestro propio lenguaje. Entonces tendríamos exactamente lo que queremos, pero en una forma innecesariamente abultada. El manual de traducción produciría mecánicamente, para cada oración del lenguaje a traducir, una oración del lenguaje del traductor; la teoría de la interpretación da a continuación la interpretación de esas oraciones familiares. Claramente la referencia al lenguaje de casa es superflua; es un intermediario del que no hay necesidad entre la interpretación y el idioma ajeno. Las únicas expresiones que una teoría de la interpretación tiene que mencionar son aquellas que pertenecen al lenguaje que ha de ser interpretado.

Una teoría de la interpretación para un lenguaje objeto puede ser contemplada entonces como el resultado de la fusión de una teoría de la interpretación para un lenguaje conocido que sea estructuralmente reveladora, y un sistema de traducción del lenguaje desconocido al conocido. La fusión convierte en ociosa cualquier referencia al lenguaje conocido; cuando esta referencia se elimina, lo que queda es una teoría de la interpretación estructuralmente reveladora para el lenguaje objeto —formulada, desde luego, en palabras familiares—. Tenemos tales teorías, sugiero, en las teorías de la verdad del género que Tarski mostró por vez primera cómo proporcionar<sup>4</sup>.

Lo que caracteriza una teoría de la verdad en el estilo de Tarski es que entraña, para toda oración  $o$  del lenguaje objeto, una oración de la forma:

$o$  es verdadera (en el lenguaje objeto) si y sólo si  $p$

<sup>4</sup> Alfred Tarski, «The Concept of Truth in Formalized Languages», en *Logic, Semantics, Metamathematics*. Oxford, 1956.

Se obtienen instancias de la forma (que llamaremos V-oraciones) reemplazando «o» por una descripción canónica de *o* y «p» por una traducción de *o*. La noción semántica importante no definida en la teoría es la de *satisfacción* que pone en relación las oraciones, abiertas o cerradas, con secuencias infinitas de objetos, que pueden considerarse que pertenecen al rango de las variables del lenguaje objeto. Los axiomas, en número finito, son de dos géneros: algunos dan las condiciones bajo las cuales una secuencia satisface una oración compleja sobre las bases de las condiciones de satisfacción de oraciones más simples; otras dan las condiciones bajo las cuales se satisfacen las oraciones más simples (abiertas). La verdad es definida para las oraciones cerradas en términos de la noción de satisfacción. Una teoría recursiva como ésta puede convertirse, como Tarski muestra, en una definición explícita de acuerdo totalmente con líneas familiares, en el supuesto de que el lenguaje de la teoría contenga suficiente teoría de conjuntos; pero no nos ocuparemos de este paso adicional.

Se presentan nuevas complejidades si los nombres propios y las expresiones funcionales son rasgos irreductibles del lenguaje objeto. Un asunto más intrincado es el que atañe a los dispositivos indéxicos. Tarski estaba interesado en los lenguajes formalizados que no contienen aspectos indéxicos o demostrativos. Por consiguiente, él podría tratar a las oraciones como vehículos de la verdad; la extensión de la teoría a emisiones es, en este caso, trivial. Pero los lenguajes naturales están indispensablemente repletos de rasgos indéxicos, como el tiempo verbal, y así sus oraciones pueden variar por lo que a la verdad respecta de acuerdo con el tiempo y el hablante. El remedio es caracterizar verdad para un lenguaje relativamente a un tiempo y a un hablante. La extensión a emisiones es de nuevo sencilla <sup>5</sup>.

Lo que sigue es una defensa de la afirmación de que una teoría de la verdad modificada para aplicarla a un lenguaje natural, puede usarse como una teoría de la interpretación. La defensa consistirá en intentos de responder a tres cuestiones:

1. ¿Es razonable pensar que puede darse una teoría de la verdad de la clase descrita para un lenguaje natural?
2. ¿Sería posible decir que tal teoría era correcta sobre la base de la evidencia plausiblemente disponible a un intérprete que no tenga ningún conocimiento anterior del lenguaje que va a ser interpretado?

<sup>5</sup> Para una discusión de cómo una teoría de la verdad puede manejar demostrativos, y cómo la Convención V debe ser modificada, ver Scott Weinstein, «Truth and Demonstratives», *Noûs*, 1974.

3. Si se supiese que la teoría era verdadera, ¿sería posible interpretar emisiones de los hablantes del lenguaje?

La primera cuestión se dirige a la suposición de que puede darse una teoría de la verdad para un lenguaje natural; la segunda y la tercera plantean si tal teoría satisficaría las exigencias adicionales que hemos impuesto a la teoría de la interpretación.

#### 1 ¿PUEDE DARSE UNA TEORÍA DE LA VERDAD PARA UN LENGUAJE NATURAL?

Nos ayudará a apreciar el problema el considerar brevemente el caso en el que se usa un fragmento significativo de un lenguaje (más uno o dos predicados semánticos) para enunciar su propia teoría de la verdad. De acuerdo con la Convención V de Tarski, es una prueba de la adecuación de la teoría el que ésta entrañe todas las V-oraciones. Esta prueba no puede aparentemente afrontarse sin asignar a las oraciones del lenguaje algo muy parecido a la forma cuantificacional estándar, y sin apelar, en la teoría, a una noción relacional de satisfacción<sup>6</sup>. Pero lo llamativo de las V-oraciones es que sea cual sea la maquinaria que tenga que operar para producirlas, y sean cuales sean las ruedas ontológicas que tengan que girar, al final una V-oración enuncia las condiciones de verdad de una oración usando recursos que no son más ricos que, puesto que son los mismos que, aquellos de la oración misma. A menos que la oración original mencione mundos posibles, entidades intensionales, propiedades, o proposiciones, el enunciado de sus condiciones de verdad no lo hace.

No hay un modo igualmente simple de proceder análogamente respecto de una lengua ajena sin apelar, como Tarski lo hace, a una noción no analizada de traducción. Pero lo que podemos hacer para nuestro propio lenguaje deberíamos poder hacerlo para otro; el problema, resultará, será el saber que estamos haciéndolo.

La restricción impuesta al exigir que una teoría satisfaga la Convención V parece ser considerable: no hay ningún método generalmente aceptado que se conozca en este momento para tratar, dentro de la restricción, con una hueste de problemas, por ejemplo, oraciones que atribuyen actitudes, modalidades, enunciados causales generales, contrafácticos, adjetivos atributivos, cuantificadores como «la mayor parte», y así sucesivamente. Por otro lado, hay lo que me parece que es un progreso bastante impresionante. Para mencionar al-

<sup>6</sup> Ver John Wallace, «On the Frame of Reference», *Synthese*, vol. 22 (1970), pp. 61-94.

gunos ejemplos, está la obra de Tyler Burge sobre nombres propios<sup>7</sup>, la de Gilbert Harman sobre «debe»<sup>8</sup>, la de John Wallace sobre términos masa y comparativos<sup>9</sup>, y está mi propia obra sobre atribuciones de actitudes y realizativos<sup>10</sup>, sobre adverbios, eventos y enunciados causales singulares<sup>11</sup>, y sobre el citar<sup>12</sup>.

Si estamos inclinados a ser pesimistas sobre lo que queda por hacer (¡o sobre algo de lo que se ha hecho!) deberíamos pensar en el magnífico logro de Frege sometiendo bajo control lo que Dummett llama «generalidad múltiple»<sup>13</sup>. Frege no tenía *in mente* una teoría de la verdad en el sentido de Tarski, pero es obvio que buscaba, y encontró, estructuras de un género para el que puede darse una teoría de la verdad.

El trabajo de aplicar en detalle una teoría de la verdad a un lenguaje natural se dividirá en la práctica con casi total certeza en dos estadios. En el primer estadio se caracterizará la verdad, no para la totalidad del lenguaje, sino para una parte cuidadosa y astutamente separada del lenguaje. Esta parte, aunque sin duda gramaticalmente pobre, contendrá infinitas oraciones que agotan el poder expresivo de todo el lenguaje. La segunda parte apareará cada una de las oraciones restantes con una o (en el caso de ambigüedad) más de una de las oraciones para las que la verdad ha sido caracterizada. Podemos pensar en las oraciones a las que se aplica el primer estadio de la teoría como dando la forma lógica, o estructura profunda, de todas las oraciones.

## 2. ¿PUEDE VERIFICARSE UNA TEORÍA DE LA VERDAD APELANDO A LA EVIDENCIA DISPONIBLE ANTES DE QUE HAYA COMENZADO LA INTERPRETACIÓN?

La Convención V dice que una teoría de la verdad es satisfactoria si genera una V-oración para cada oración del lenguaje objeto. Es entonces suficiente para demostrar que una teoría de la verdad es

<sup>7</sup> Tyler Burge, «Reference and Proper Names», *Journal of Philosophy*, vol. 70 (1973), pp. 425-439.

<sup>8</sup> Gilbert Harman, «Moral Relativism Defended», en prensa.

<sup>9</sup> John Wallace, «Positive, comparative, Superlative», *Journal of Philosophy*, vol. 69 (1972), pp. 773-782.

<sup>10</sup> Donald Davidson, «On Saying That», *Synthese*, vol. 19 (1968), pp. 130-146 [versión castellana en obra citada en nota 2].

<sup>11</sup> Donald Davidson, «Causal Relations», *Journal of Philosophy*, vol. 64 (1967) pp. 691-703.

<sup>12</sup> Donald Davidson, «Quotation», en *The Structure of Truth*, Oxford [versión castellana en obra citada en nota 2].

<sup>13</sup> Michael Dummett, *Frege*, Londres, 1973.

empíricamente correcta verificar que las V-oraciones son verdaderas (en la práctica, una muestra adecuada confirmará la teoría hasta un grado razonable). Las V-oraciones mencionan solamente las oraciones cerradas del lenguaje, de modo que la evidencia puede consistir enteramente en hechos sobre la conducta y actitudes de los hablantes en relación con las oraciones (sin duda por vía de las emisiones). Una teoría practicable tiene, desde luego, que tratar a las oraciones como concatenaciones de expresiones de longitud menor que la oracional, tiene que introducir nociones como las de satisfacción y referencia, y tiene que apelar a una ontología de secuencias y a los objetos ordenados por las secuencias. Todo este aparato se contempla propiamente como una construcción teórica más allá del alcance de la verificación directa. Ha llevado a cabo su misión en el supuesto de que sólo entrañe resultados comprobables en la forma de V-oraciones, y éstas no hagan mención de la maquinaria. Una teoría de la verdad reconcilia así la exigencia de una teoría que articule estructura gramatical con la exigencia de una teoría que pueda ser comprobada solamente por lo que dice sobre las oraciones.

En la obra de Tarski las V-oraciones se considera que son verdaderas porque se supone que la rama derecha del bicondicional es una traducción de la oración cuyas condiciones de verdad se están dando. Pero no podemos suponer de antemano que puede reconocerse la traducción correcta sin vaciar ya el objetivo de la interpretación radical; en las aplicaciones empíricas tenemos que abandonar la suposición. Lo que propongo es invertir la dirección de explicación: al suponer la traducción Tarski era capaz de definir verdad; la idea presente es considerar la verdad como básica y extraer una explicación de la traducción o interpretación. Las ventajas, desde el punto de vista de la interpretación radical, son obvias. La verdad es una propiedad de una pieza que se agrega, o deja de agregarse, a las emisiones, mientras que cada emisión tiene su propia interpretación; y la verdad es más propensa a conectar con actitudes bastante simples de los hablantes.

No hay ninguna dificultad en volver a expresar la Convención V sin apelar al concepto de traducción: una teoría aceptable de la verdad tiene que entrañar, para toda oración *o* del lenguaje objeto, una oración de la forma: *o* es verdadera si y sólo si *p*, donde «*p*» es reemplazada por alguna oración que es verdadera si y sólo si *o* lo es. Dada esta formulación la teoría es puesta a prueba por la evidencia de que las V-oraciones son simplemente verdaderas; hemos abandonado la idea de que tenemos también que decir si lo que reemplaza a «*p*» traduce *o*. Podría parecer que no hay oportunidad de que surja una teoría de la interpretación, si exigimos tan poco de las V-oraciones. Y desde luego sería así si considerásemos las V-oraciones aisladamente. Pero la esperanza es que imponiendo restricciones formales y

empíricas sobre la teoría como un todo, las V-oraciones individuales sirvan de hecho para proporcionar interpretaciones.

Tenemos aún que decir qué evidencia está disponible para un intérprete —evidencia, vemos ahora, de que las V-oraciones son verdaderas—. La evidencia no puede consistir en descripciones detalladas de las creencias e intenciones del hablante, puesto que las atribuciones de actitudes, al menos donde se requiere cierta sutileza, exigen una teoría que tiene que descansar sobre casi la misma evidencia que la interpretación. La interdependencia de creencia y significado es evidente de esta manera: un hablante mantiene que una oración es verdadera a causa de lo que la oración (en su lenguaje) significa, y a causa de lo que él cree. Sabiendo que mantiene que la oración es verdadera, y conociendo el significado, podemos inferir su creencia; si se nos da información suficiente sobre sus creencias, podríamos quizás inferir el significado. Pero la interpretación radical debe descansar sobre evidencia que no suponga conocimiento de significados o conocimiento detallado de creencias.

Un buen lugar para empezar es el de la actitud de mantener una oración como verdadera, de aceptarla como verdadera. Esto es, desde luego, una creencia, pero es una actitud de una sola pieza aplicable a todas las oraciones, y de este modo no se nos pide ser capaces de hacer distinciones finamente discriminadas entre creencias. Se trata de una actitud que un intérprete puede plausiblemente considerar ser capaz de identificar antes de que pueda interpretar, puesto que puede saber que una persona intenta expresar una verdad al emitir una oración sin tener idea alguna de *qué* verdad se trata. No se piense que la aserción sincera es la única razón para suponer que una persona mantiene que una oración es verdadera. Mentiras, órdenes, anécdotas, ironía, si son detectadas como actitudes, pueden revelar si un hablante mantiene que sus oraciones son verdaderas. No hay razón para descartar otras actitudes hacia oraciones, tales como desear verdadero, necesitar hacer verdadero, creer que uno va a hacer verdadero, y así sucesivamente, pero estoy inclinado a pensar que toda la evidencia de este género puede resumirse en términos de mantener que son verdaderas oraciones.

Supongamos, entonces, que la evidencia disponible es solamente que los hablantes del lenguaje a interpretar mantienen en ciertos tiempos y bajo circunstancias especificadas que varias oraciones son verdaderas. ¿Cómo puede usarse esta evidencia para apoyar una teoría de la verdad? Por una parte, tenemos V-oraciones de la forma:

(V) «Es regnet» es verdadera-en-alemán cuando es proferida por  $x$  en el tiempo  $t$  si y sólo si está lloviendo cerca de  $x$  en  $t$ .

Por otra parte, tenemos la evidencia, en la forma:

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

Deberíamos, pienso, considerar (E) como evidencia de que (V) es verdadera. Puesto que (V) es un condicional cuantificado universalmente, el primer paso sería reunir más evidencia para apoyar la afirmación de que:

(GE)  $(x) (t)$  (si  $x$ ) pertenece a la comunidad de habla alemana entonces  $x$  mantiene como verdadera «Es regnet» en  $t$  si y sólo si está lloviendo cerca de  $x$  en  $t$ .

La apelación a una comunidad de habla constituye un atajo pero no pide la cuestión: los hablantes pertenecen a la misma comunidad de habla si funcionan para ellos las mismas teorías de la interpretación.

La objeción obvia es que Kurt, o cualquier otro, puede estar equivocado respecto a si está lloviendo cerca de él. Y esto es, desde luego, una razón para no considerar (E) como evidencia conclusiva para (GE) o para (V); y una razón para no esperar que generalizaciones como (GE) sean más que generalmente verdaderas. El método es más bien el de conseguir un ajuste óptimo. Queremos una teoría que satisfaga las constricciones formales impuestas a una teoría de la verdad, y que maximice el acuerdo, en el sentido de hacer que Kurt (y otros) estén en lo cierto, hasta donde podamos determinarlo, tan a menudo como sea posible. El concepto de maximización no puede ser tomado aquí literalmente, puesto que las oraciones son infinitas en número, y en cualquier caso una vez que la teoría comienza a tomar forma tiene sentido aceptar el error inteligible y tener en cuenta la probabilidad relativa de diversos géneros de error.

El proceso de diseñar una teoría de la verdad para una lengua nativa desconocida podría proceder, a grandes rasgos, de la manera siguiente. En primer lugar buscamos el mejor modo de hacer encajar nuestra lógica con el nuevo lenguaje, hasta el punto requerido para obtener una teoría que satisfaga la convención V; esto puede significar leer en el lenguaje la estructura lógica de la teoría de la cuantificación de primer orden (más identidad), no considerando las constantes lógicas una por una, sino tratando esta considerable porción de lógica como una rejilla que ha de hacerse encajar con el lenguaje de una vez. La evidencia consiste aquí en clases de oraciones mantenidas siempre como verdaderas o mantenidas siempre como falsas por casi todo el mundo y casi siempre (verdades lógicas potenciales) y pa-



trones de inferencia. El primer paso identifica predicados, términos singulares, cuantificadores, conectivas e identidad; en teoría, establece los asuntos de la forma lógica. El segundo paso se concentra en oraciones con expresiones índicecas; aquellas oraciones mantenidas algunas veces como verdaderas y algunas veces como falsas de acuerdo con cambios descubribles en el mundo. Este paso en conjunción con el primero limita las posibilidades para interpretar predicados individuales. El último paso se ocupa de las oraciones restantes, aquellas respecto de las cuales no hay acuerdo uniforme, o aquellas cuyo valor de verdad mantenido no depende sistemáticamente de cambios en el entorno <sup>14</sup>.

Este método intenta resolver el problema de la interdependencia de creencia y significado manteniendo la creencia constante en la medida de lo posible mientras encuentra una solución al significado. Esto se lleva a cabo asignando condiciones de verdad a oraciones ajenas que hacen que los hablantes nativos estén en lo correcto cuando es plausiblemente posible, de acuerdo, desde luego, con nuestra propia visión de lo que es correcto. Lo que justifica el procedimiento es el hecho de que tanto el acuerdo como el desacuerdo son inteligibles solamente en contraste con un trasfondo de acuerdo substancial. Aplicado al lenguaje este principio reza: cuantas más oraciones conseguimos en aceptar o rechazar (sea o no por medio de la interpretación), mejor entendemos el resto, estemos o no de acuerdo sobre ellas.

El consejo metodológico de interpretar de una manera que optimice el acuerdo no debería concebirse como descansando en una suposición caritativa sobre la inteligencia humana que podría resultar ser falsa. Si no podemos encontrar un modo de interpretar las emisiones y otra conducta de una criatura en tanto que revelando un conjunto de creencias en gran parte consistentes y verdaderas de acuerdo con nuestros propios estándares, no tenemos razón alguna para tener en cuenta a esa criatura como racional, o como diciendo algo.

Querría insertar aquí una observación sobre la metodología de mi propuesta. En filosofía acostumbramos a usar definiciones, análisis, reducciones. Típicamente se pretende que estas cosas nos lleven de

<sup>14</sup> Los lectores que aprecian hasta qué punto esta explicación es paralela a la de Quine de la traducción radical en el capítulo segundo de *Palabra y objeto* observarán también las diferencias: la constricción semántica de mi método fuerza una estructura cuantificacional sobre el lenguaje a ser interpretado, lo cual probablemente no deja lugar para la indeterminación de la forma lógica; la noción de significado estimulativo no desempeña ningún papel en mi método, pero su lugar es ocupado por la referencia a los rasgos objetivos del mundo, que se alteran en conjunción con cambios de actitud hacia la verdad de las oraciones; el principio de caridad, que Quine subraya sólo en conexión con la identificación de las conectivas oracionales (puras), yo lo aplico de cabo a rabo.

conceptos mejor entendidos, o claros, o más básicos epistemológica u ontológicamente, a otros que queremos entender. El método que he sugerido no encaja con ninguna de esas categorías. He propuesto una relación más laxa entre los conceptos que hay que iluminar y los relativamente más básicos. En el centro está situada una teoría formal, una teoría de la verdad que impone una estructura compleja a oraciones que contienen las nociones primitivas de verdad y satisfacción. A esas nociones se les da aplicación por la forma de la teoría y la naturaleza de la evidencia. El resultado es un teoría parcialmente interpretada. La ventaja del método reside no en su apelación en estilo libre a la noción de apoyo evidencial sino en la idea de un poderosa teoría interpretada en el punto más ventajoso. Esto nos permite reconciliar la necesidad de una estructura semánticamente articulada con una teoría comprobable solamente en el nivel oracional. La ganancia más sutil es que esta muy ligera evidencia en apoyo de cada uno de los puntos potencialmente infinitos puede proporcionar ricos resultados, incluso con respecto a los puntos. Conociendo solamente las condiciones bajo las cuales los hablantes mantienen como verdaderas oraciones, podemos encontrarnos como resultado, dada una teoría satisfactoria, con una interpretación de cada oración. Queda por validar esta última afirmación. La teoría misma da en el mejor de los casos condiciones de verdad. Lo que necesitamos mostrar es que si tal teoría satisface las constricciones que hemos especificado, puede usarse para proporcionar interpretaciones.

### 3. SI SABEMOS QUE UNA TEORÍA DE LA VERDAD SATISFACE LOS CRITERIOS FORMALES Y EMPÍRICOS DESCRITOS, ¿PODEMOS INTERPRETAR LAS EMISIONES DEL LENGUAJE PARA EL QUE ES UNA TEORÍA?

Una teoría de la verdad entraña una V-oración para cada oración del lenguaje objeto, y una V-oración da condiciones de verdad. Resulta tentador, por lo tanto, decir simplemente que una V-oración «da el significado» de una oración. No, desde luego, nombrando o describiendo una entidad que es un significado, sino diciendo simplemente bajo qué condiciones es verdadera una emisión de una oración <sup>15</sup>.

<sup>15</sup> Esta idea, y otras rechazadas aquí, serán encontradas en varios de mis artículos: ver «Verdad y significado» [en este mismo volumen], *Synthese*, vol. 7 (1967), pp. 304-323; «Semantics for Natural Languages», en *Linguaggi nella Società e nella Tecnica*, Milán, 1970, pp. 177-188, y «True to the Facts», *Journal of Philosophy*, vol. 66 (1969), pp. 748-764 [versión castellana en obra citada en nota 2].

Pero si reflexionamos es claro que una V-oración no da el significado de la oración a la que atañe: las V-oraciones fijan el valor de verdad relativo a ciertas condiciones, pero no dicen que la oración del lenguaje objeto es verdadera *porque* las condiciones valen. Sin embargo, si los valores de verdad fuese todo lo que importase, la V-oración para «La nieve es blanca» podría también decir que es verdadera si y sólo si la hierba es verde o si  $2 + 2 = 4$  como dice que es verdadera si y sólo si la nieve es blanca. Podemos tener confianza, quizás, en que ninguna teoría satisfactoria de la verdad produzca tales V-oraciones anómalas, pero esta confianza no nos da licencia para tratar con especial consideración a las V-oraciones.

Un movimiento que podría parecer provechoso es afirmar que no es la V-oración sola, sino la demostración canónica de una V-oración la que nos permite interpretar la oración ajena. Una demostración canónica, dada una teoría de la verdad, es fácil de construir, moviéndose, por así decirlo, a través de una ristra de condicionales y estableciendo como requisito para unicidad solamente decisiones ocasionales para gobernar la precedencia izquierda y derecha. La demostración refleja la forma lógica que la teoría asigna a la oración, y de este modo podría pensarse que revela algo sobre el significado. Pero de hecho no sabríamos más que antes sobre cómo interpretar si todo lo que supiésemos fuese que una cierta secuencia de oraciones era la demostración, para alguna teoría verdadera, de una V-oración particular.

Una sugerencia final de acuerdo con esas líneas es que podemos interpretar una oración particular en el supuesto de que conozcamos una teoría de la verdad correcta que trata del lenguaje de la oración. Pues entonces no solamente conocemos la V-oración para la oración que ha de ser interpretada, sino que también «conocemos» las oraciones para todas las demás oraciones; y, desde luego, todas las demostraciones. Entonces veríamos el lugar de la oración en el lenguaje como un todo, conoceríamos el papel de cada parte significativa de la oración, y tendríamos conocimiento acerca de las conexiones lógicas entre esta oración y otras.

Si supiésemos que una V-oración satisfacía la convención V de Tarski, sabríamos que era verdadera, y podríamos usarla para interpretar una oración puesto que sabríamos que la rama derecha del bicondicional traducía la oración que había que interpretar. Nuestra presente dificultad surge del hecho de que en la interpretación radical no podemos suponer que una V-oración satisface el criterio de traducción. Lo que hemos estado pasando por alto es, sin embargo, que hemos proporcionado un criterio alternativo: este criterio es que la totalidad de las V-oraciones encajarían optimamente (en el sentido descrito más arriba) con la evidencia sobre las oraciones mantenidas como verdaderas por los hablantes nativos. La idea presente es

que lo que Tarski supuso de una vez para cada V-oración puede ser obtenido indirectamente por medio de una constricción holista. Si esta constricción es adecuada, cada V-oración proporcionará de hecho una interpretación aceptable.

Una V-oración para una teoría de la verdad puede, entonces, usarse para interpretar una oración en el supuesto de que conocemos también la teoría que la entraña, y que sabemos que es una teoría que cumple los criterios formales y empíricos. Pues, si las constricciones son adecuadas, el rango de teorías aceptables será tal que cualquiera de ellas proporciona alguna interpretación correcta para cada emisión potencial. Para ver cómo podría funcionar esto, aceptemos por un momento la hipótesis absurda de que las constricciones restringen las posibles teorías a una, y ésta implica la V-oración (V) discutida previamente. Entonces estamos justificados al usar esta V-oración para interpretar la emisión de Kurt de «Es regnet» como su decir que está lloviendo. No es probable, dada la naturaleza flexible de las restricciones, que todas las teorías aceptables vayan a ser idénticas. Cuando toda la evidencia esté en posesión nuestra quedarán, como Quine ha subrayado, los inestables equilibrios entre las creencias que atribuimos a un hablante y las interpretaciones que damos a sus palabras. Pero la indeterminación resultante no puede ser tan grande que ninguna teoría que pase las pruebas no sirva para proporcionar interpretaciones.

## ¿QUÉ ES UNA TEORÍA DEL SIGNIFICADO?

MICHAEL A. E. DUMMETT

De acuerdo con una concepción bien conocida, el mejor método de formular los problemas filosóficos en torno al concepto de significado y nociones emparentadas es preguntar qué forma debería tomar lo que se denomina «una teoría del significado» para un lenguaje en concreto; esto es, una especificación detallada de los significados de todas las palabras y de las operaciones formadoras de oraciones del lenguaje que arroje una especificación del significado de toda expresión y oración del lenguaje. No es que la construcción de una teoría del significado, en este sentido, para cualquier lenguaje se conciba como un proyecto práctico; pero se piensa que, tan pronto como podamos enunciar los principios generales de acuerdo con los que podría llevarse a cabo esa construcción, habremos llegado a una solución de los problemas concernientes al significado por los que los filósofos se sienten perplejos.

Comparto la creencia en que éste es el enfoque más fructífero a los problemas de este área de la filosofía, aunque no me siento capaz de ofrecer una demostración de que es así a alguien que lo negase: pero podemos ver algunas razones en su favor si tomamos como contraste ciertos casos diferentes. Por lo que sé, nadie ha sugerido jamás un enfoque paralelo de los problemas de la epistemología: nadie ha propuesto que el modo apropiado de ponerse a atacar los problemas filosóficos relacionados con el concepto de conocimiento sería considerar cómo se podría construir una teoría del conocimiento en el sentido de una especificación detallada de todo lo que puede decirse que conoce cualquier individuo o comunidad. La razón es, pienso yo, que nuestra captación del concepto de conocimiento es bastante más segura que nuestra captación del concepto de significado. Estamos en duda sobre lo que debe contar como conocimiento, estamos aún más en duda sobre cómo formular los principios a los que apelamos tácitamente para decidir si algo ha de contar o no como conocimiento; también tenemos cierta incertidumbre sobre el análisis semántico de una oración que atribuya conocimiento de algo a al-

---

\* Versión castellana de Alfonso García Suárez.

Ésta es la primera parte de una respuesta en dos partes a esta pregunta. La segunda apareció con el título «What is a Theory of Meaning? (II)», en G. Evans y J. McDowell (comps.), *Truth and Meaning: Essays in Semantics*, Clarendon Press, Oxford, 1976, pp. 67-137.

pero al menos estamos totalmente seguros de *cuáles* son las oraciones cuya forma lógica y cuyas condiciones de verdad buscamos analizar. Por contraste, aunque la mayoría de nosotros, incluido yo mismo, estaríamos de acuerdo en que el concepto de significado es fundamental e indispensable, ni siquiera tenemos claridad sobre la estructura superficial de los enunciados que involucran ese concepto.

¿Qué tipo de oración, del lenguaje natural, debe considerarse como la forma característica para la atribución de un significado particular a una cierta expresión o palabra dadas? No sólo no conocemos la respuesta: ni siquiera sabemos si se trata de la pregunta apropiada. Quizá sea imposible, en general, *enunciar* el significado de una expresión: quizá debamos, más bien, investigar por qué medios lingüísticos, o posiblemente incluso no lingüísticos, es posible *transmitir* el significado de una expresión sin enunciarlo explícitamente. O quizá incluso esto sea incorrecto: quizá la pregunta debiera ser no cómo expresamos que una expresión particular tiene un cierto significado, sino cómo debemos analizar las oraciones que involucran el concepto de significado de algún modo diferente. Es precisamente porque, en este área de la filosofía, sabemos aún menos de qué estamos hablando que en otras áreas, por lo que la propuesta de enfocar nuestros problemas considerando cómo podríamos intentar especificar los significados de las expresiones de un lenguaje entero no parece la pérdida de tiempo que parecería una propuesta análoga en epistemología.

Como es bien sabido, algunos, preeminentemente Quine, han preferido burlar esta dificultad investigando los principios que subyacen a la construcción, no de una teoría del significado para un lenguaje, sino de un manual de traducción del mismo a algún lenguaje conocido. La ventaja es que sabemos exactamente qué forma ha de tomar un manual de traducción, a saber, un conjunto efectivo de reglas que mapeen las oraciones del lenguaje traducido en oraciones del lenguaje en que se hace la traducción: podemos, por tanto, concentrarnos por entero en las cuestiones de cómo llegar a un sistema de traducciones tal y como se incorporan en ese manual y de qué condiciones han de satisfacerse para que ese sistema sea aceptable. La desventaja es que, aunque el interés de tal investigación tiene que residir en la luz que arroje sobre el concepto de significado, no podemos estar seguros de qué consecuencias para la noción de significado tengan los resultados de la investigación sobre la traducción, justamente porque se enuncian sin apelar directamente a esa noción. Captar el significado de una expresión es entender su papel en el lenguaje: una teoría completa del significado para un lenguaje es, por tanto, una teoría completa de cómo el lenguaje funciona como lenguaje. Nuestro interés por el significado, como concepto general, es, así, un interés por cómo funciona el lenguaje; una descripción directa del modo en que funciona el lenguaje —de todo lo que alguien tiene que

aprender cuando aprende el lenguaje— resolvería, en consecuencia, nuestras perplejidades de una forma en que no puede hacerlo una explicación indirecta, por medio de una traducción. Se dirá con toda justicia que el interés por la investigación de la traducción se aplica no a la traducción misma, sino a los criterios propuestos para juzgar la aceptabilidad de un esquema de traducción, y que éstos han de relacionarse con lo que puede observarse del funcionamiento del lenguaje a traducir. En realidad podría mantenerse plausiblemente que nada menor que una teoría completa del significado para un lenguaje —una descripción completa del modo en que funciona— podría ser una base adecuada para juzgar la corrección de un esquema de traducción propuesto. No intentaré pronunciarme sobre la validez de esta afirmación. Si es válida, entonces la ventaja aparente del enfoque vía traducción, en vez de preguntando sin más qué forma debe tomar una teoría del significado para el lenguaje, es totalmente ilusoria. Si es inválida —y ciertamente el procedimiento real del principal practicante del enfoque vía traducción sugiere que la considera inválida—, entonces se sigue que no hay ninguna inferencia inmediata que vaya de resultados sobre la traducción a conclusiones sobre el significado.

He dicho que es el oficio de una teoría del significado para un lenguaje dar cuenta de cómo funciona el lenguaje, esto es, de cómo se comunican sus hablantes por medio de él: aquí «comunicarse» no tiene más significación precisa que «hacer cuanto puede hacerse mediante la emisión de una o más oraciones del lenguaje». Y aquí repetiré lo que he mantenido en otros lugares, que una teoría del significado es una teoría de la comprensión; esto es, aquello de que una teoría del significado tiene que dar cuenta es aquello que alguien conoce cuando conoce el lenguaje, esto es, cuando conoce los significados de las expresiones y oraciones del lenguaje. Una pregunta sobre la forma que debe tomar una teoría del significado es si debe resultar en adscripciones directas de significado, esto es, en proposiciones de la forma «El significado de la palabra/oración *X* es...» o de la forma «La palabra/oración *X* significa...». Si la respuesta a esta pregunta es afirmativa, puede parecer que esa teoría del significado no tendrá necesidad de hacer explícitamente referencia a la noción de conocimiento: si la teoría nos permite decir que el significado de una palabra u oración dadas es esto o aquello, digamos *Q*, entonces, presumiblemente, querremos asimismo decir que alguien conoce el significado de esa palabra u oración si sabe que *Q* es lo que esa palabra u oración significa. Veremos después razones para dudar de esto; pero por el momento permítasenos suspender el juicio. Si la teoría del significado nos permite derivar esas adscripciones directas de significado, y si esas adscripciones directas de significado son tales que llevan de esta forma simple a una caracterización de lo que es conocer el sig-

nificado de cada palabra u oración del lenguaje, entonces, realmente, mi afirmación de que una teoría del significado tiene que ser una teoría de la comprensión no pretende tener un sentido tan fuerte que excluya una tal teoría meramente sobre la base de que no emplea la noción de conocimiento: sería adecuado aceptar una tal teoría como una teoría de la comprensión. Si, por otro lado, aunque la teoría del significado permite la derivación de adscripciones directas de significado, estas adscripciones están conformadas de manera que no permiten una caracterización inmediata de qué es lo que una persona conoce cuando conoce el significado de una palabra u oración, entonces, por hipótesis, la teoría es inadecuada para dar cuenta de un tipo extremadamente importante de contexto en que estamos dispuestos a usar la palabra «significado». Si, sin embargo, la teoría del significado no resulta en esas adscripciones directas de significado y si, además, no contiene dentro de sí misma ninguna descripción abierta de lo que alguien tiene que conocer a fin de conocer o captar el significado de cada expresión del lenguaje, sino que proporciona meramente una explicación de otros contextos en los que usamos la palabra «significado», tales como «X significa lo mismo que Y» o «X tiene significado», entonces me parece que será de nuevo inadecuada para la construcción a partir de ella de cualquier teoría de la comprensión. Esto es, si fuera posible dar una explicación de, por ejemplo, cuándo dos expresiones tienen el mismo significado, que no descansase abiertamente en una explicación de qué es conocer el significado de una expresión, entonces no sería posible derivar una explicación del conocimiento del significado a partir de ella. Hay, realmente, buenas razones para suponer que es imposible dar una explicación de la sinonimia excepto vía una explicación de la comprensión, puesto que es un requisito de la primera que quienquiera que conozca el significado de dos expresiones sinónimas tiene también que saber que son sinónimas: pero estoy diciendo meramente que, si una explicación tal de la sinonimia fuese posible, no habría ruta alguna desde ella a una explicación de la comprensión.

Cualquier teoría del significado que no fuese, o no produjese inmediatamente, una teoría de la comprensión no satisfaría el propósito que, filosóficamente, exigimos a una teoría del significado. Pues he argüido que a una teoría del significado se le exige que deje al descubierto el funcionamiento del lenguaje. Conocer un lenguaje es ser capaz de emplear un lenguaje; por tanto, una vez que tenemos una explicación explícita de aquello en que consiste el conocimiento de un lenguaje, tenemos por ello una explicación del funcionamiento de ese lenguaje; y nada menos que eso puede darnos lo que andamos buscando. A la inversa, también me parece que una vez que podamos decir en qué consiste el que alguien conozca un lenguaje, en el sentido de conocer los significados de todas las expresiones del len-



guaje, hemos resuelto esencialmente todo problema que pueda surgir acerca del significado. Por ejemplo, una vez que clarifiquemos lo que es conocer el significado de una expresión, las cuestiones acerca de si, en tal y cual caso, ha cambiado el significado de una palabra pueden resolverse preguntando si alguien que entendiese previamente la palabra tiene que adquirir nuevo conocimiento a fin de entenderla ahora.

Si una teoría del significado da cuenta del funcionamiento del lenguaje con el que está relacionada, entonces parece que ha de incorporar una explicación de todos los conceptos expresables en ese lenguaje, al menos por medio de expresiones unitarias. No es necesario que nos detengamos a investigar si, o en qué casos, de alguien que no posee los medios lingüísticos para expresar un concepto puede decirse que capta ese concepto: es suficiente reconocer que el caso prototípico de captación de un concepto es aquél en el que esa captación consiste en la comprensión de una cierta palabra o expresión, o ámbito de expresiones, en algún lenguaje. Por ello, si una teoría del significado es una teoría de la comprensión, como he afirmado, parecería seguirse que tal teoría del significado, al explicar lo que se ha de conocer a fin de conocer el significado de cada expresión del lenguaje, tiene que explicar simultáneamente qué es tener los conceptos expresables por medio de ese lenguaje.

La teoría del significado, por cierto, hará más que esto: es claro que no puede explicar meramente los conceptos expresables en el lenguaje, puesto que estos conceptos pueden ser captados por alguien que sea totalmente ignorante de ese lenguaje particular, pero que conozca otro lenguaje en que sean expresables. De ahí que la teoría del significado tenga también que asociar conceptos con palabras del lenguaje —que mostrar o enunciar qué conceptos son expresados por qué palabras—. Y una concepción alternativa sería que es sólo esta última tarea la que pertenece a la teoría del significado, que pedirle a la teoría del significado que sirva para explicarle nuevos conceptos a alguien que no los tiene ya es ponerle una carga demasiado pesada, y que todo lo que podemos exigir de tal teoría es que le dé la interpretación del lenguaje a alguien que ya tiene los conceptos exigidos. Llamemos a una teoría del significado que se propone cumplir solamente esta tarea restringida una teoría del significado «modesta», y a una que busca explicar realmente los conceptos expresados por los términos primitivos del lenguaje una teoría «vigorosa»\*. Una cuestión que deseo tratar de responder es si una teoría del significado modesta es sencillamente posible o si cualquier cosa que tenga el rango de una teoría del significado tiene que ser vigorosa.

Si se aceptase una bien conocida concepción, propuesta por Da

\* *Full-blooded.* (N. del T.)

además, de la forma que debe tomar una teoría del significado, creo que lo que ha de mantenerse es que una teoría del significado modesta es todo lo que tenemos derecho de pedir. Según esta concepción, el corazón de la teoría del significado será una teoría de la verdad, construida sobre el modelo de una definición de la verdad del tipo de la de Tarski (sin asumir, en general, que el lenguaje objeto sea un fragmento del metalenguaje): esa teoría de la verdad, sin embargo, carecerá del aparato requerido para convertirla en una definición explícita y no servirá para explicar el concepto de verdad en modo alguno, sino, tomándolo como ya conocido, para dar la interpretación del lenguaje objeto. La teoría de la verdad producirá una V-oración para cada oración del lenguaje objeto, es decir, o bien un bicondicional cuyo lado izquierdo es de la forma «La oración *O* es verdadera» o la cláusula universal de un bicondicional cuyo lado izquierdo es de la forma «Una emisión de la oración *O* por un hablante *X* en un tiempo *t* es verdadera». Sin embargo, al juzgar si las V-oraciones que produce la teoría son las correctas no se apela a la noción de traducción; por el contrario, hay constricciones que la teoría tiene que satisfacer, para ser aceptable, relativas a las oraciones tenidas por verdaderas por los hablantes del lenguaje (dándose por supuesto que podemos construir criterios adecuados de si un hablante mantiene que una oración dada es verdadera): en primer lugar que, en conjunto, las V-oraciones derivables en la teoría de la verdad enuncien, en sus lados derechos, las condiciones bajo la que de hecho los hablantes tienen por verdaderas las oraciones nombradas en sus lados izquierdos.

Los axiomas de la teoría de la verdad, cuando ésta forma parte de la teoría del significado para un lenguaje bajo esa concepción, enunciarán las denotaciones de los nombres propios del lenguaje, darán las condiciones para la satisfacción de los predicados primitivos, etc. Si un predicado primitivo del lenguaje expresa un cierto concepto, parecería totalmente fuera de lugar afirmar que una teoría del significado de este tipo, o, en particular, el axioma de la teoría de la verdad que gobierne ese predicado, proporcionen explicación alguna de ese concepto. Por el contrario, la teoría le sería inteligible sólo a alguien que ya captase el concepto. Una teoría davidsoniana del significado es una teoría modesta.

Ya he observado que un manual de traducción ha de contrastarse con una teoría del significado y no puede pretender serlo. Una teoría del significado describe directamente el modo en que funciona el lenguaje, un manual de traducción proyecta meramente ese lenguaje sobre otro cuyo funcionamiento ha de darse ya por conocido para que la traducción sea de uso práctico. Ese punto ha sido igualmente subrayado por Davidson, que lo ha expresado diciendo que el manual de traducción nos dice sólo que ciertas expresiones de un lenguaje significan lo mismo que ciertas expresiones del otro, sin decir-

nos qué significan, específicamente, las expresiones de ambos lenguajes; sería, dice él, posible en principio saber, de cada oración de un lenguaje dado, que significa lo mismo que alguna oración particular de otro lenguaje, sin saber en absoluto qué significado tiene ninguna de estas oraciones. Esta objeción a considerar que un manual de traducción constituya en sí mismo una teoría del significado es evidentemente justa: pero podemos preguntarnos por qué se pone un énfasis tan fuerte en la distinción entre un manual de traducción y una teoría del significado cuando de la teoría del significado se exige que sea, no vigorosa, sino sólo modesta. Un manual de traducción lleva a una comprensión del lenguaje traducido sólo vía una comprensión del lenguaje al que se hace la traducción, una comprensión que aquél no suministra en sí mismo; de ahí que podamos decir que no exhibe directamente en qué consiste una comprensión del lenguaje traducido. Pero una teoría del significado modesta, asimismo, lleva a una comprensión del lenguaje objeto sólo vía una captación de los conceptos expresados por sus expresiones primitivas, que ella no explica en sí misma; parece, por tanto, que deberíamos decir simularmente que una tal teoría del significado no exhibe plenamente en qué consiste una comprensión del lenguaje objeto. Esto es así especialmente porque nuestro mejor modelo —y, en muchos casos, nuestro único modelo— de la captación de un concepto lo proporciona el dominio de una cierta expresión o ámbito de expresiones en algún lenguaje. Así, un manual de traducción presupone un dominio de algún otro lenguaje —aquél al que se hace la traducción— para derivar de él una comprensión del lenguaje traducido; pero una teoría del significado modesta presupone un dominio de *algún*, aunque inespecificado, lenguaje para derivar de ella una comprensión del lenguaje objeto. El contraste significativo, sin embargo, parecería que no es entre una teoría que (como un manual de traducción) hace una presuposición específica y una que (como una teoría del significado modesta) hace una presuposición tan fuerte, aunque menos específica, sino entre teorías que (como esas dos) descansan en presuposiciones extrínsecas y las que (como las teorías del significado vigorosas) no involucran en absoluto ninguna presuposición así.

Volvamos a la pregunta: ¿debe resultar en adscripciones directas de significado una teoría del significado? Una teoría del significado debe decirnos, naturalmente, para cada expresión del lenguaje lo que significa: pero sería superficial concluir de aquí que por tanto ha de ser posible derivar de la teoría enunciados que comiencen «La expresión *X* significa...». Para dar un ejemplo trivial, una teoría exitosa de un crimen, digamos de un asesinato, debe decirnos la identidad del asesino: eso no entraña que debamos poder derivar de la teoría un enunciado que comience «La identidad del asesino es...»; en realidad (donde «es» es el signo de identidad) no hay enunciados bien

formados que comiencen de ese modo. Como ejemplo más serio, podemos advertir que «química» no es él mismo un concepto de la teoría química. Exigimos ciertamente de la teoría química que nos capacite para decir qué propiedades de una sustancia son propiedades químicas, qué interacciones son químicas, etc.; asimismo puede exigirse de una teoría del significado que nos capacite para decir qué propiedades de una expresión son semánticas, esto es, dependientes de y sólo de su significado: pero no podemos exigir que «significado» sea él mismo un concepto de la teoría del significado, al menos si esto se interpreta entrañando que nos capacite para caracterizar las propiedades semánticas de una expresión por medio de un enunciado que comience «El significado de la expresión es...» o «La expresión significa...».

Para expresiones menores que oraciones, y en particular para conectivas, preposiciones, etc., hay cierta dificultad en construir incluso una forma gramaticalmente correcta para una adscripción directa de significado (donde, naturalmente, no queramos emplear como complemento directo del verbo «significa» un término que denote una expresión, de manera que «significa» se volviera reemplazable por «significa lo mismo que»). Sin embargo, no es mi propósito investigar cómo, o incluso si, esas dificultades pueden resolverse: podemos limitar nuestra atención al caso de las oraciones, para el que no surge la dificultad. El propio Davidson admite que, a partir de una teoría del significado del tipo que él auspicia, será derivable una adscripción directa de significado, al menos para oraciones. Dada una V-oración derivable de una teoría de la verdad que satisfaga las constricciones requeridas, por ejemplo la oración «“La terra si muove” es verdadera si y sólo si la tierra se mueve», podemos convertirla legítimamente en lo que podemos llamar una S-oración, en este caso «“La terra si muove” significa que la tierra se mueve». Pues bien, anteriormente consideramos la cuestión de si una teoría del significado que no hiciese ninguna alusión abierta al conocimiento podría no obstante capacitarnos para derivar de ella una explicación, para cada expresión, de en qué consiste el conocimiento del significado de esa expresión; y, en particular, dimos por argumentado que, si la teoría permite la derivación, para cada expresión, de una adscripción directa de significado, entonces también tiene que proporcionarnos una explicación de qué es conocer el significado de una expresión dada, concretamente que esto sería conocer lo que es enunciado por la adscripción directa del significado de la expresión. Pero, ahora, si se nos pregunta si la S-oración «“La terra si muove” significa que la tierra se mueve» expresa lo que alguien tiene que saber a fin de saber lo que significa la oración italiana «La terra si muove», difícilmente podemos hacer otra cosa que responder afirmativamente: saber que «La terra si muove» significa que la tierra se mueve *es* jus-

tamente saber lo que significa «La terra si muove», pues eso es precisamente lo que significa. Si, por otro lado, se nos pregunta si una explicación adecuada de en qué consiste un conocimiento del significado de «La terra si muove» lo proporciona el decir que se debe conocer lo que es enunciado por la S-oración relevante, entonces, igualmente, nos vemos impelidos a responder negativamente: pues la S-oración, tomada en sí misma, es, aunque en modo alguno no informativa, sí señaladamente no explicativa. Si estas reacciones son correctas, se sigue que el hecho de que una teoría del significado resulte en adscripciones directas de significado no es en sí mismo base suficiente para afirmar que proporcione una explicación adecuada de en qué consiste el conocimiento del significado.

Uno de nuestros problemas aún sin resolver era descubrir qué ventaja podría tener una teoría del significado modesta sobre un mero manual de traducción. Un manual de traducción nos informará, por ejemplo, de que «La terra si muove» significa lo mismo que «La tierra se mueve»: pero se dijo que su inadecuación residía en el hecho de que alguien podría saber que las dos oraciones son sinónimas sin saber qué significa ninguna de ellas. A fin de derivar, del conocimiento de que las dos oraciones son sinónimas, el conocimiento de lo que significa la oración italiana, lo que alguien tiene que saber adicionalmente es, obviamente, lo que significa la oración española. Obviamente por igual, lo que, además de saber que las dos oraciones son sinónimas, ha de saberse a fin de saber que la oración italiana significa que la tierra se mueve, es que es eso lo que significa la oración española. Se sigue que si se nos dijera que el conocimiento del significado de la oración italiana consiste en saber que significa que la tierra se mueve, tenemos también que mantener que saber lo que significa la oración española «La tierra se mueve» consiste en saber que *ella* significa que la tierra se mueve. Una S-oración como «“La tierra se mueve” significa que la tierra se mueve», para un lenguaje objeto que es parte del metalenguaje, parece totalmente no explicativa porque, esta vez, es totalmente no informativa; aunque aún parece imposible negar que alguien sabe lo que significa «La tierra se mueve» justamente en el caso en que sabe que significa que la tierra se mueve.

En este contexto, es importante observar una distinción que, en muchos contextos, puede pasarse por alto: la que hay entre saber, de una oración, que es verdadera y conocer la proposición expresada por la oración. Al usar la frase «conocer la proposición expresada por la oración» no pretendo reconocimiento alguno de las proposiciones en cuanto entidades, compromiso alguno con una ontología de proposiciones: empleo la frase simplemente como un medio conveniente de expresar la generalización de la distinción entre, por ejemplo, decir de alguien que él sabe que la oración «19 es primo» es verdadera

ni y decir de alguien que él sabe que 19 es primo. La razón por la que la S-oración «“La tierra se mueve” significa que la tierra se mueve» parece totalmente no informativa es que no cabría la posibilidad de mantener que el conocimiento del significado de «La tierra se mueve» consistiese en el conocimiento de que esa S-oración es verdadera; pues cualquiera que haya captado los principios más simples que gobiernan el uso del verbo «significar» y que sepa que «La tierra se mueve» es una oración española, tiene que saber que esa oración es verdadera, aun cuando pueda no saber qué significa en particular «La tierra se mueve». El caso es análogo al ejemplo de Kripke de la oración «Los caballos se llaman “caballos”». Kripke dice que cualquiera que conozca el uso de «se llama» en español tiene que saber que la oración expresa una verdad, independientemente de que sepa lo que son los caballos: claramente, todo lo que necesita saber es que «caballo» es un término general significativo en español y, claramente por igual, el sentido relevante de «saber lo que son los caballos» es «aquél en que es sinónima de «saber lo que significa “caballo”». Kripke admite, sin embargo, que alguien que no sepa lo que son los caballos no sabrá *qué* verdad expresa «Los caballos se llaman “caballos”». Parece razonable suponer que, con esta concesión, Kripke pretende negar que podamos decir de esa persona que ella sabía que los caballos se llaman «caballos», aunque no es explícito sobre esto: esto es, en mi terminología, esa persona puede saber que la oración «Los caballos se llaman “caballos”» es verdadera sin conocer la proposición expresada por esa oración.

Pudiera objetarse que alguien que sepa que una oración es verdadera tiene también que conocer la proposición expresada por la oración, sobre la base de que si sabe bastante sobre el significado de la palabra «verdadera» para atribuirle el conocimiento de que la oración es verdadera, tiene que conocer la conexión que hay entre conocer algo y saber que es verdadero (y entre creer algo y creer que es verdadero, etc.); una conexión que es exhibida por las V-oraciones. Por ejemplo, tiene que saber que «Los caballos se llaman “caballos”» es verdadera si y sólo si los caballos se llaman «caballos»: de ahí que, puesto que por supuesto él sabe que «Los caballos se llaman “caballos”» es verdadera, si es capaz de realizar una inferencia simple, también será capaz de saber que los caballos se llaman «caballos». Pero esta objeción deriva su plausibilidad de ignorar en su premisa la distinción cuya falta de substancia se proponía demostrar, a saber; la que hay entre saber que una oración es verdadera y conocer la proposición que expresa. A alguien que no sepa lo que significa «caballo», pero que sepa que es un término general significativo, podemos atribuirle justificadamente el conocimiento de que la V-oración «“Los caballos se llaman ‘caballos’” es verdadera si y sólo si los caballos se llaman “caballos”» es verdadera: pero suponer, como requiere

el argumento, que él sepa que «Los caballos se llaman “caballos” es verdadera si y sólo si los caballos se llaman «caballos» es pedir la cuestión.

Decir, de alguien que no conoce lo que significa «La tierra se mueve», que él no sabe que «La tierra se mueve» significa que la tierra se mueve, sino sólo sabe que la S-oración es verdadera, no es en absoluto decir que él no esté preparado para emitir esa S-oración asertóricamente, sino sólo la oración «La oración “La tierra se mueve” significa que la tierra se mueve” es verdadera». Ni siquiera es decir que él no pudiera ofrecer excelentes fundamentos para la anterior emisión: por el contrario, puede ofrecer fundamentos enteramente concluyentes, concretamente la apelación al uso de «significa» en español. Pero hemos aprendido de la paradoja de Gettier que no toda justificación cabal de una creencia verdadera es suficiente para autorizar al que sostiene la creencia a pretender conocimiento; la justificación ha de relacionarse apropiadamente con lo que hace verdadera a la creencia. Una justificación de una emisión de la S-oración que fundamentaría la adscripción al hablante de un conocimiento de la proposición expresada por esa S-oración tendría que ser una que dependiese del significado específico de la oración de la que tratase la S-oración, en nuestro caso la oración «La tierra se mueve», aun cuando, en circunstancias ordinarias, nadie pensara en justificar esa emisión de manera tan complicada.

Todo esto muestra que estábamos enteramente en lo correcto en nuestra primera inclinación, al considerar como una condición necesaria y suficiente para que alguien sepa lo que significa «La tierra se mueve» que él sepa que significa que la tierra se mueve, esto es, que él conozca la proposición expresada por la correspondiente S-oración. Pero muestra igualmente que estábamos también en lo correcto al considerar que la S-oración es totalmente no explicativa de qué es conocer el significado de la oración «La tierra se mueve». El modo más simple que tenemos de enunciar su carácter no explicativo es observar que hasta aquí no tenemos ninguna caracterización independiente de qué más tiene que conocer alguien que sepa que la S-oración es verdadera a fin de conocer la proposición que expresa: salvo que él tiene que saber lo que significa «La tierra se mueve»: el conocimiento de esa proposición no puede desempeñar, por tanto, ningún papel en una explicación de aquello en lo que consiste la comprensión de esa oración. Y, si una S-oración para la cual el metalenguaje contiene el lenguaje objeto es no explicativa, entonces una S-oración para un lenguaje objeto disyunto del metalenguaje es igualmente no explicativa. En el último caso, la S-oración no proporciona en realidad información alguna: pero el conocimiento de la verdad de una tal S-oración (en cuanto opuesto a un conocimiento de la proposición que expresa) no requiere la posesión de ninguna infor-

mación que no esté también contenida en la correspondiente oración de un manual de traducción.

Las consideraciones sobre la conexión entre conocimiento y justificación que vimos que subyacían a la distinción entre conocer la verdad de una oración y conocer la proposición expresada por ella pueden generalizarse a casos en los que no es precisamente esta distinción la que está en cuestión. La expresión «sabe que» se usa, por cierto, frecuentemente en el discurso cotidiano y en contextos filosóficos en los que no se centra la atención sobre el concepto de conocimiento, meramente como sinónima de «se apercibe de que». Cuando «conocimiento» se usa en un sentido más estricto, sin embargo, el conocimiento de un hecho trasciende la mera apercepción del mismo en que involucra que su apercepción se alcanzó de alguna manera canónica, esto es, que se *derivó* de alguna forma especial. Si, entonces, intentamos explicar en qué consiste una cierta capacidad diciendo que consiste en tener un cierto conocimiento, y si la plausibilidad de esta explicación depende de tomar «conocimiento» en el sentido estricto, y no como mera apercepción, la pretendida representación de la capacidad seguirá siendo inadecuada mientras se quede en enunciar simplemente el *objeto* de conocimiento —lo que ha de ser conocido, en el sentido estricto de «conocer», para que alguien tenga esa capacidad—. Para dar una explicación adecuada de la capacidad en cuestión, la descripción tiene que hacer algo más que especificar simplemente el hecho que tiene que ser conocido: tiene que indicar cómo tiene, en particular, que alcanzarse la apercepción de ese hecho, esto es, qué proceso de derivación se requiere para que cuente como *conocimiento*, en el sentido estricto.

Puede objetarse que nadie ha supuesto jamás que pueda darse una explicación adecuada del significado, o de la comprensión, de una oración aludiendo meramente a la S-oración con ella relacionada. En los términos en que acabo de discutir el asunto, toda la gracia de la teoría del significado está en exhibir los medios canónicos por los que ha de derivarse la S-oración: sólo de alguien que fue capaz de derivarla así se podría decir que la conoce, en el sentido estricto, o, como lo expresé anteriormente, que conoce la proposición que expresa. Tal objeción es enteramente justa: mi propósito al discutir las S-oraciones tan largamente no era refutar una tesis que nadie ha mantenido, sino analizar las razones intuitivas que todos compartimos para rechazar esa tesis, a fin de sacar a la luz algunos puntos generales que podemos aplicar en otro lugar.

A fin de ver, pues, en qué consiste, según una explicación davidsoniana, el conocimiento del significado de una oración, tenemos que mirar al modo en que la S-oración con ella relacionada se deriva en la teoría del significado. La S-oración se obtiene, como hicimos notar, reemplazando «es verdadera si y sólo si» en la correspondiente



V-oración por «significa que»: y la V-oración, a su vez, se deriva de los axiomas de la teoría de la verdad que gobiernan las palabras constituyentes de la oración y de los que gobiernan los métodos de formación oracional ejemplificados por ella. Naturalmente, esto está enteramente de acuerdo con nuestra convicción intuitiva de que un hablante deriva su comprensión de una oración de su comprensión de las palabras que la componen y del modo en que se ensamblan. Lo que desempeña el papel, dentro de una teoría del significado del tipo de la de Davidson, de una captación de los significados de las palabras es el conocimiento de los axiomas que gobiernan esas palabras. En nuestro ejemplo, esos axiomas pueden formularse como «“La tierra” denota la tierra» y «Es verdadero decir “Se mueve” si y sólo si esa cosa se mueve». (Esta última formulación del axioma que gobierna «se mueve» evita la apelación al dispositivo técnico de la satisfacción por una secuencia infinita y es sólo una indicación aproximada de lo que se pretende: pero, si pretendemos una representación sería de lo que es conocido por cualquiera capaz de hablar español, no podemos atribuirle literalmente una comprensión de ese dispositivo técnico.)

No es suficiente, para que alguien sepa lo que significa la oración «La tierra se mueve», que sepa que la S-oración con ella relacionada es verdadera; debe conocer la proposición expresada por esa S-oración. Y la forma natural de caracterizar lo que tiene que conocer adicionalmente alguien que conoce la verdad de la S-oración, a fin de conocer la proposición que expresa, es éste: los significados de las palabras componentes. Ahora bien, si explicamos que la comprensión de las palabras componentes consiste en el conocimiento de los axiomas de la teoría de la verdad que gobiernan esas palabras, surge la misma pregunta: ¿es suficiente que él sepa que esos axiomas son verdaderos o debe conocer las proposiciones que expresan? La objeción que hay a exigir solamente que sepa que los axiomas son verdaderos es paralela a la que admitimos en el caso de la S-oración: cualquiera que conozca el uso de «denota» y que sepa que «la tierra» es un término singular en español tiene que saber que la oración «“La tierra” denota la tierra» es verdadera, aunque no sepa qué significa específicamente, «la tierra» o qué denota.

A esto podría objetarse, sin embargo, sobre la base de que, si cambiásemos el ejemplo de «La tierra se mueve» por el de «Homero era ciego», resultaría que, a fin de saber que «“Homero” denota a Homero» es verdadera, no basta con saber que «Homero» es un nombre propio: se debe saber también que no es un nombre vacío. Esta objeción está fuera de lugar porque, para cualquier lenguaje en el que esté abierta la posibilidad de que «Homero» sea un nombre vacío, el axioma relevante de la teoría de la verdad no adoptará la forma simple «“Homero” denota a Homero»; al menos no lo hará si

la vacuidad del nombre priva de la verdad a la oración «“Homero” denota a Homero». Sólo en una teoría de la verdad para un lenguaje de tipo fregeano, en la que todos los términos singulares se conciben con denotación garantizada, adoptará esa forma simple el axioma que gobierna cada nombre propio. Para lenguajes de otros tipos, el axioma que gobierna un nombre como «Homero» tendrá que adoptar una forma diferente. Por ejemplo, para cualquier lenguaje en el que se considerase al predicado «...es Homero» verdadero del referente de «Homero», si hay alguno, y falso de todos y de todo lo demás, el axioma podría adoptar la forma «Para todo  $x$ , “Homero” denota a  $x$  si y sólo si  $x$  es Homero». Ahora bien, si el lenguaje fuera russelliano, de modo que la presencia de un nombre vacuo en una oración atómica convirtiera a esa oración en falsa, ulteriores axiomas apropiados arrojarían la V-oración «“Homero era ciego” es verdadera si y sólo si Homero era ciego». Por otro lado, si el lenguaje fuese tal que la presencia de un nombre vacuo en una oración, salvo cuando siguiese al signo de identidad, convirtiera a esa oración en ni verdadera ni falsa, entonces no querríamos que esa V-oración fuese derivable, ya que, si «Homero» fuese un nombre vacuo, el lado izquierdo sería falso mientras que el lado derecho no era falso. Querriamos, en cambio, la V-oración no estándar «“Homero era ciego” es verdadera si y sólo si, para algún  $x$ ,  $x$  es Homero y  $x$  era ciego». De ahí que la demanda de que, a fin de poder derivar la V-oración relativa a «Homero era ciego», tendríamos primero que saber si «Homero» es o no es un nombre vacuo, sea enteramente injustificada.

Esto sólo podría negarse si se mantuviera que para conocer el significado de «Homero» se debe saber si hubo o no de hecho un hombre tal: pues la teoría de la verdad es una parte de la teoría del significado para el lenguaje e incorporará todo y sólo lo que se requiere para la comprensión del lenguaje. Pero, claramente, a fin de conocer el uso del nombre «Homero» en nuestro lenguaje, no es necesario saber si tiene o no tiene una denotación: lo más que pudiera exigirse es que se supiera si es o no es sabido que el nombre tenga una denotación. Esto es, podría mantenerse que, para un nombre del que se sabe que tiene denotación, este conocimiento entra en la comprensión del nombre: si es así, entonces, para un nombre como ese, «Londres» pongamos por caso, el axioma que lo gobierna adoptará la forma simple «“Londres” denota Londres». El conocimiento de si «Homero» tiene denotación o no, por otro lado, no puede ser parte de lo que está involucrado en el uso de ese nombre, por la razón obvia de que ese conocimiento no es poseído por los hablantes del lenguaje.

Si se supone que cualquiera que intente una investigación en serio sobre si hay un lugar tal como Londres mostraría con ello que no tenía dominio del uso aceptado del nombre «Londres», entonces

será cierto que alguien que sepa de la palabra «Londres» solamente que es un nombre propio no puede siquiera reconocer que el axioma que lo gobierna es verdadero: debe saber también que es un nombre del que estamos seguros que no es vacío. Pero, claramente, podríamos estar informados de este hecho, y concluir por ello la verdad de la oración «“Londres” denota Londres», sin saber qué significa exactamente «Londres»; y así hemos aún de concluir que el conocimiento de la verdad del axioma es insuficiente para la comprensión del nombre. Sería incorrecto argüir contra esto que estar meramente informado de que «Londres» es un nombre del que se sabe con certeza que no es vacío no es *conocer* ese hecho, que conocerlo, en el sentido estricto, involucra saber específicamente cómo se usa el nombre «Londres». Si tal argumento fuese correcto, la objeción de Davidson a considerar como una teoría del significado a un manual de traducción, que se podría saber, por ejemplo, que «la terra» significa lo mismo que «la tierra» sin saber lo que significa ninguna de ellas, sería incorrecta: pues podríamos argüir, del mismo modo, que, aunque alguien pudiera estar informado de su sinonimia, no podría, en sentido estricto, *conocerlo* sin conocer lo que significan ambas palabras. La objeción burlaría malamente el principio metodológico que hemos adoptado, a saber: no aceptar como parte de una explicación un requisito de que alguien sepa algo, donde «conocimiento» se toma en el sentido estricto, como algo que trasciende la mera apercepción, sin dar cuenta de qué constituiría ese conocimiento.

Es esencial observar este principio para evitar explicaciones nulas o circulares. Supongamos que fuese verdad —por dudoso que a mí me parezca— que no se pudiera saber, en el sentido estricto, que un nombre denota un objeto bien conocido aún existente, sin conocer el uso preciso del nombre. Esto debe ser porque, para que cuente como *conocimiento*, la apercepción del hecho debe derivarse de una manera particular. Uno de los méritos de una teoría del significado que representa el dominio de un lenguaje como el conocimiento, no de proposiciones aisladas, sino de proposiciones deductivamente conectadas, es que muestra debido reconocimiento del hecho indudable de que en la comprensión de una oración va involucrado un proceso de derivación de algún tipo. Donde una tal teoría no hace apelación a ningún proceso de derivación es, naturalmente, en el reconocimiento de la verdad de los axiomas. La insistencia en que tal reconocimiento equivale a conocimiento en el sentido estricto, haría, sin embargo, apelación tácita a un proceso por el que se derivó su verdad, un proceso que la teoría no lograría hacer explícito. Sería, por ejemplo, simplemente circular decir que la comprensión del nombre «Londres» consiste en el conocimiento, en el sentido estricto, de la verdad de la oración «“Londres” denota Londres» y luego continuar diciendo que una condición para tener ese conocimiento es la captación del

uso preciso del nombre: lo que estábamos buscando era una caracterización de qué constituye una captación del uso de ese nombre.

No hay, pues, posibilidad ninguna de mantener que la comprensión de las palabras componentes de una oración consiste sólo en una apercepción de la verdad de los axiomas que las gobiernan: tendríamos que conocer las proposiciones expresadas por esos axiomas. La teoría del significado debe ser capaz, por tanto, de explicar lo que diferencia un conocimiento de las proposiciones expresadas por esos axiomas de una mera apercepción de su verdad. Ahora bien, el propio Davidson ha reconocido plenamente la obligación que pesa sobre la teoría del significado de producir una teoría de la comprensión: ha sido totalmente explícito sobre en qué consiste, en su opinión, la comprensión de una oración, concretamente en el conocimiento tanto de la V-oración relevante como del hecho de que esa V-oración se derivó de una teoría de la verdad para el lenguaje que satisface las constricciones impuestas sobre una teoría tal para que sea aceptable. La condición análoga, para la comprensión de una palabra, sería presumiblemente el conocimiento del axioma de una teoría de la verdad que satisface aquellas constricciones. Esta vez la sugerencia es, por tanto, que podemos representar el conocimiento de las proposiciones expresadas por las oraciones que sirven de axiomas como consistente en una apercepción de su verdad suplementada por cierto trasfondo de conocimiento acerca de esas oraciones.

Me parece que se necesita sólo una pequeñísima consideración para reconocer que esta apelación a un trasfondo de información no puede suministrar lo que necesitamos. Si alguien no sabe lo que significa «la tierra», aprenderá algo de que se le diga que la oración «“La tierra” denota la tierra» es verdadera, supuesto que entienda el verbo «denota»: concretamente, aprenderá que «la tierra» es un término singular y no es vacío. Ahora bien, si pide que se le diga el significado específico del término, en nada le ayudará el que se le diga que la oración en cuestión es un axioma en una teoría de la verdad para el español que satisface ciertas constricciones particulares. Obviamente, lo que le dice qué denota específicamente «Londres» es la oración misma «“Londres” denota Londres» y, en particular, el complemento del verbo «denota» en esa oración, y no ninguna información extrínseca *sobre* esa oración. Lo que se atribuye a quien sabe español no es meramente la apercepción de que esa oración (y otras como ella) es verdadera, sino esa apercepción junto con una comprensión de la oración; en otras palabras, un conocimiento de la proposición expresada por la oración. Naturalmente, cuando consideramos el caso degenerado en que el metalenguaje es una extensión del lenguaje objeto, el requisito de que se entienda el metalenguaje se vuelve circular; a fin de derivar del axioma un conocimiento de lo que denota «Londres», se tendría ya que entender el nombre «Lon-

dres». Pero no hay exigencia ninguna de que la teoría de la verdad se exprese en una extensión del lenguaje objeto: si el axioma reza así «“Londres” denota Londra», entonces lo que se necesitaría para aprender la denotación de «Londres» sería una comprensión del término «Londra», y no habría circularidad.

Esto es bastante razonable en sí mismo, pero no nos ayuda a entender qué diferencia significativa hay entre una teoría del significado modesta de este tipo y un manual de traducción. Se ve ahora claramente que a cualquiera capaz de usar la teoría de la verdad para obtener una interpretación del lenguaje objeto hemos de adscribirle la posesión de una comprensión previa del metalenguaje. Esto se ve incluso más claramente cuando le atribuimos una apercepción de que la teoría de la verdad satisface las constricciones requeridas, ya que esas constricciones aluden a las condiciones *enunciadas* en los lados derechos de las V-oraciones, una noción que no podemos explicar en términos de la teoría formal, sino que presupone una interpretación de la misma. De ahí que una teoría del significado de este tipo exhiba meramente qué es llegar a una interpretación de un lenguaje vía una comprensión de otro, que es justamente lo que hace un manual de traducción: no explica qué es tener un dominio de un lenguaje, nuestra lengua materna pongamos por caso, independientemente de un conocimiento de cualquier otro.

Esta conclusión sólo podría evitarse si a un hablante del lenguaje objeto le pudiéramos adscribir el conocimiento de las proposiciones expresadas por las oraciones de la teoría de la verdad, independientemente de cualquier lenguaje en que pudieran expresarse esas proposiciones. Si es ésa la intención de una tal teoría del significado, parece profundamente insatisfactoria, ya que carecemos de modelo, y la teoría no nos proporciona ninguno, de en qué pudiera consistir una aprehensión de esas proposiciones excepto en la capacidad de enunciarlas lingüísticamente.

Puede replicarse que la aprehensión de esas proposiciones no puede explicarse pieza por pieza, para cada oración de la teoría de la verdad tomada separadamente, sino que el conocimiento de la teoría de la verdad como un todo resulta, precisamente, en la capacidad de hablar y entender el lenguaje objeto, de manera que no hay ninguna laguna. Lo que se nos está dando es un modelo teórico de una capacidad práctica, la capacidad de usar el lenguaje. Puesto que es un modelo teórico, la representación es en términos del conocimiento de un sistema deductivamente conectado de proposiciones; y, puesto que sólo podemos expresar las proposiciones con oraciones, el modelo tiene que ser descrito en términos de un sistema deductivamente conectado de oraciones. No se pretende presumir que un hablante del lenguaje objeto tenga de hecho una comprensión previa del lenguaje en el que se construyen esas oraciones —es por eso por lo que

es inocuo construirlas en un lenguaje que sea de hecho una extensión del lenguaje objeto—; pero, igualmente, no queda incumplida ninguna obligación de decir en qué consiste la captación de las proposiciones expresadas por la teoría: consiste en esa capacidad práctica de la que estamos dando un modelo teórico.

Es justamente aquí donde sale a la vista la conexión entre una teoría del significado que procede vía una teoría de la verdad y una visión holista del lenguaje, una conexión enigmática a primera vista. Una semántica que resulta en un enunciado de condiciones de verdad para cada oración, derivadas de una pluralidad finita de axiomas, cada uno de los cuales gobierna una única palabra o construcción parece al principio una realización de una concepción atomista del lenguaje bajo la cual cada palabra tiene un significado individual y cada oración un contenido individual: la V-oración para una oración dada del lenguaje se deriva de solamente aquellos axiomas que gobiernan las palabras y construcciones que aparecen en esa oración. Pero la conexión entre esa concepción y la visión holista del lenguaje reside en el hecho de que nada se especifica acerca de en qué consiste el conocimiento de las proposiciones expresadas por los axiomas, o por las V-oraciones: las únicas constricciones a la teoría son globales, relativas al lenguaje como un todo. Según eso, no puede haber respuesta a la pregunta de qué constituye la comprensión por parte de un hablante de cualquier palabra u oración: sólo se puede decir que el conocimiento de la teoría de la verdad entera resulta en una capacidad de hablar el lenguaje y, en particular, en una propensión a reconocer que sus oraciones son verdaderas bajo condiciones correspondientes, en conjunto, a las enunciadas por las V-oraciones.

Así pues, la apelación al conocimiento de que la teoría de la verdad satisface las constricciones externas no sirve para explicar la comprensión por parte del hablante de cualquier palabra u oración individuales, para salvar la hendidura que hay entre su conocimiento de la verdad de un axioma o teorema de la teoría de la verdad y su conocimiento de la proposición por él expresada: meramente media entre su conocimiento de la teoría como un todo y su dominio del lenguaje completo. Pues bien, el encanto de una teoría del significado de este tipo es que parece refutar la sospecha de que una visión holista del lenguaje tiene que ser antisistemática; puesto que hablar un lenguaje es tener la capacidad de emitir sus oraciones de acuerdo con su significación convencional, parece no haber esperanza de ninguna descripción sistemática del uso de un lenguaje completo que no produzca una descripción de la significación de las emisiones individuales. Una teoría del significado davidsoniana, por otro lado, combina el principio básico del holismo con lo que pretende ser una descripción del modo en que el significado de cada oración individual se determina a partir de los significados de sus palabras constituyentes.

Esta apariencia es, sin embargo, una ilusión. La articulación de la teoría de la verdad no se toma como correspondiente a articulación ninguna de la capacidad práctica cuya posesión es la manifestación de aquel conocimiento del cual la teoría se presenta como modelo teórico. El conocimiento que un hablante tiene del significado de una oración individual se representa como consistente en su captación de una parte de una teoría deductiva, y esto se conecta con sus emisiones efectivas sólo por el hecho de que se supone que una captación de la teoría completa resulta, de alguna manera de la que no se da explicación ninguna, en su dominio del lenguaje en su totalidad; pero no se proporciona ninguna manera, incluso en principio, de segmentar su capacidad de usar el lenguaje como un todo en distintas capacidades componentes que manifiesten su comprensión de palabras, oraciones o tipos de oración individuales. Para efectuar cualquier segmentación así, sería necesario dar una descripción detallada de la capacidad práctica en que consiste la comprensión de una palabra u oración particulares, mientras que, según la visión holista, no sólo el dominio que un hablante tiene de su lenguaje no puede segmentarse así, sino que no puede darse en absoluto ninguna descripción detallada de en qué consiste. De ahí que la articulación de la teoría no desempeñe ningún papel auténtico en la descripción de lo que constituye el dominio de su lenguaje por parte de un hablante.

Contra esto puede objetarse que la teoría de la verdad sí que nos dice algo sobre el uso de cada oración individual: pues enuncia condiciones bajo las que un hablante probablemente mantendrá que es verdadera. Ahora bien, ciertamente sucede que una teoría del significado basada en una teoría de la verdad reflejaría una visión molecular, en vez de holista, del lenguaje si pudiéramos considerar que los lados derechos de las V-oraciones enuncian condiciones bajo las que los hablantes del lenguaje mantendrían invariablemente la verdad de las oraciones nombradas en los lados izquierdos. Esta no es una manera posible de interpretar la teoría, por dos razones. Primero, para cualquier lenguaje natural, las condiciones enunciadas en los lados derechos de las V-oraciones no serán, en general, condiciones que seamos capaces de reconocer que se dan siempre que se den. Una teoría molecular del significado basada en la noción de condiciones de verdad tiene que atribuirle a quien entiende una oración el conocimiento de las condiciones que deben darse para que sea verdadera, no una capacidad de reconocer a esa oración como verdadera justamente en caso de que esa condición se dé. En segundo lugar, tal explicación no dejaría lugar para errores. A fin de dejarles lugar, hemos de afirmar que una teoría de la verdad aceptable dará el ajuste *mejor posible* entre las condiciones para la verdad de una oración y las condiciones bajo las que se la tiene por verdadera, pero no un ajuste *perfecto*: se sigue que la comprensión de una oración por parte

de un hablante no puede juzgarse excepto en relación a su empleo del lenguaje completo. (En realidad, es un tanto dudoso el que el dominio del lenguaje por parte de un hablante individual pueda siquiera juzgarse. Si hemos identificado una comunidad lingüística desde el exterior, entonces una teoría davidsoniana del significado nos dará una guía bastante buena, aunque necesariamente imperfecta, de qué oraciones tienen por verdaderas sus hablantes. Habrá divergencias por parte de la comunidad completa —casos en los que diremos, sobre la base de la teoría del significado, que la comunidad comparte una creencia errónea—. Habrá también desacuerdos entre hablantes individuales. ¿Cómo vamos a distinguir entre un desacuerdo como el que puede darse entre dos hablantes que aceptan ambos tácitamente la misma teoría del significado para su lenguaje común y uno que refleja interpretaciones divergentes de ese lenguaje? Presumiblemente, si un miembro de la comunidad lingüística mantiene una teoría divergente de la verdad para el lenguaje, tenderá a divergir más de la mayoría en sus juicios de lo que lo hacen la mayor parte de los hablantes. Pero, puesto que ningún conjunto finito de esas divergencias revelará, en sí mismo, su apoyatura en una teoría no estándar de la verdad, es difícil ver cómo ya sea él o los demás hablantes o nosotros como espectadores podríamos detectarlo jamás, o cómo, una vez descubierto, podría corregirse. La dificultad surge precisamente porque no hay modo de determinar, dentro de tal teoría, el contenido individual del que cualquier hablante dota a una oración.)

Davidson hace de la necesidad virtud y, como ustedes oyeron en su brillantemente clara exposición en la primera conferencia de esta serie \*, usa la hendidura que hay entre la condición de verdad de una oración y la condición bajo la que es tenida por verdadera para explicar la génesis del concepto de creencia. Esto es, sin embargo, renegar de lo que estamos autorizados a esperar de una teoría del significado: una tal teoría debe poder distinguir entre desacuerdos provenientes de una diferencia de interpretaciones y desacuerdos de sustancia (desacuerdos acerca de los hechos); debe poder explicar cómo es posible que se dé un desacuerdo sobre el valor de verdad de las oraciones aun cuando hay un acuerdo sobre su significado. Quine nos ha enseñado, ciertamente, a mirar con sospecha esta distinción; e innegablemente ocurre que los significados de las expresiones del lenguaje natural son frecuentemente borrosos y que, en consecuencia, la distinción se empaña. Es igualmente cierto que, como

---

\* Se refiere a las Conferencias del Wolfson College de 1974. Esa primera conferencia de Davidson lleva el título «Thought and Talk» y está recogida, al igual que la conferencia de Dummett que aquí traducimos, en S. Guttenplan (ed.), *Mind and Language*, Clarendon, Oxford, 1975. (N. del T.)



Davidson ha señalado en su conferencia, no debemos asumir a la ligera que todo desacuerdo sobre valor de verdad, por ejemplo de la oración «La tierra es redonda», deba considerarse como un desacuerdo de substancia y no de interpretación. Pero una teoría del significado que niegue en principio la viabilidad de la distinción corre el riesgo de volverse solipsista. Un desacuerdo entre hablantes individuales del mismo lenguaje al mismo tiempo o bien no puede ser explicado en absoluto o habría de ser explicado atribuyéndoles teorías divergentes de la verdad para el lenguaje; y lo mismo se aplica a un cambio de opinión por parte de un solo individuo. Si se toma el último camino, perdemos la concepción de la comunidad lingüística: un lenguaje, considerado en cuanto determinado por una teoría del significado, se convierte en algo hablado por un único individuo en un cierto período.

El hecho objetivo obvio es que los juicios que hacemos no están directamente correlacionados con los estados de cosas que los vuelven verdaderos o falsos. Aun cuando la teoría del significado correcta para nuestro lenguaje representase nuestra captación del significado de cada oración como consistente en nuestro conocimiento de la condición que ha de darse para que sea verdadera, no llegamos, en general, a nuestra evaluación de la verdad de una oración por reconocimiento directo de que la condición apropiada se da, puesto que, en la mayor parte, esa condición no es una que seamos capaces de reconocer así. ¿Habríamos, por tanto, de decir que una teoría adecuada del significado tiene que poder dar cuenta no meramente de qué determina la corrección o incorrección de nuestros juicios, sino también de cómo llegamos a ellos, puesto que esto también depende de los significados que asignemos a las oraciones cuyo valor de verdad estamos juzgando; y que esta explicación tiene que poder mostrar cómo, en el proceso, somos capaces de extraviarnos, aun cuando compartamos con otros hablantes una interpretación común de la oración? El que digamos o no esto es en parte una cuestión de gusto, de cuánto deseemos contar como perteneciente a una teoría del significado; tal explicación pertenece ciertamente a una descripción completa del funcionamiento del lenguaje. Si una teoría del significado, basada en una visión molecular del lenguaje, permite dar un contenido claro a la asociación de un cierto significado con una oración por parte de un individuo, un significado que determine cuándo puede juzgarse correctamente que esa oración es verdadera, entonces tenemos también un criterio claro de cuándo un juicio representa un error de hecho; si entonces optamos por decretar que una explicación del proceso que lleva a tales errores no pertenece a la teoría del significado, sólo se trata de una disputa sobre demarcación. Pero una teoría del significado basada en una visión holista, que carece de criterio acerca de la asociación de un significado específico con cual-

quier oración por parte de un hablante, salvo su inclinación a tenerla por verdadera o falsa, y que por tanto no se propone dar cuenta de su comprensión de esa oración, sino sólo del lenguaje entero, no puede dar ningún contenido a la noción de un error, noción que invoca solamente para dar cuenta de la falta de ajuste entre la teoría de la verdad y los juicios realmente hechos por los hablantes. Sería absurdo esperar que una teoría del significado adscriba a toda expresión un significado completamente preciso; estoy defendiendo, sin embargo, que se requiere que se deje un lugar para la distinción entre un desacuerdo de substancia y un desacuerdo sobre significado, una distinción que, después de todo, no fue inventada por teóricos descaminados, sino que se emplea realmente dentro de nuestro lenguaje. Cualquier teoría que asocie oraciones meramente con condiciones de verdad, sin intentar explicación ninguna de los medios por los que reconocemos o juzgamos que esas condiciones de verdad se cumplen, ni proporcionar medio ninguno de determinar que un hablante individual, o incluso la comunidad completa, asocia una condición de verdad particular con una oración particular, salvo una tosca concordancia entre las condiciones de verdad de todas las oraciones bajo una teoría dada y los juicios que se hacen acerca de ella es incapaz de ofrecer ningún lugar para esa distinción.

Ahora bien, podría replicarse que estoy totalmente equivocado al negar que Davidson pueda representar la captación del significado de una oración particular por parte de un hablante individual: en la conferencia que inauguró esta serie, él afirmó que la comprensión de una oración por parte de un individuo consiste en que de la V-oración relevante sepa que es derivable de alguna teoría de la verdad para el lenguaje que satisface las constricciones requeridas, sin que tenga realmente que conocer esa teoría de la verdad. Pero ¿cómo juzgar que un individuo sabe eso? ¿Qué ha de hacer él, en realidad, con la información si la tiene? Podría afirmarse que él manifestará su conocimiento juzgando que esa oración es verdadera sólo en caso de que se dé la condición enunciada en la V-oración. Pero ¿por qué habría de hacer eso? Bueno, puede decirse, él sabe que la teoría de la verdad que produce esa V-oración logra el mejor ajuste con los juicios hechos por otros hablantes y desea maximizar la concordancia de sus juicios con los de ellos. Es verdad, por hipótesis, que esta teoría de la verdad logrará el mejor ajuste posible *para una teoría de la verdad*: pero puesto que no logrará un ajuste perfecto, mejor haría él, al maximizar la concordancia, no guiándose exclusivamente por ninguna teoría de la verdad. ¿Cómo puede saber que no lograría una concordancia mejor desatendiendo la teoría de la verdad en este caso? Después de todo, no puede ocurrir que los demás hablantes sigan todos la política de juzgar los valores de verdad de las oraciones sólo de acuerdo con la teoría dada de la verdad, de lo con-

trario el ajuste *sería* perfecto: ¿así que por qué habría de hacerlo el? A esto sólo puede responderse que los demás hablantes sí que tratan de seguir esa política, pero cometen errores al hacerlo. Hemos ahora vuelto circularmente una vez más a la pregunta: ¿qué es un error? Al atribuir a los hablantes una política de conformar sus juicios a una teoría de la verdad, les hemos adscrito subrepticamente una capacidad de juzgar si se cumplen las condiciones de verdad de las oraciones —juicios que no siempre serán correctos; pero no hemos dado ningún contenido a la noción de un tal juicio en cuanto distinto de un juicio con respecto al valor de verdad de una oración—.

El que habríamos de apelar a la noción de error a fin de explicar la falta de ajuste entre una teoría de la verdad y los juicios reales hechos por los hablantes del lenguaje suena plausible sólo porque ya encontramos inteligible la noción de un tal error: estamos lo bastante familiarizados con la idea de que alguien puede asignar un significado determinado a una oración y con todo juzgar erróneamente que es verdadera. Pero una teoría que no ofrezca ninguna explicación de cómo ocurren esos errores no tiene derecho ninguno a apelar a esta noción. Podemos verlo claramente si consideramos cualquier teoría que no tenga el lenguaje como su objeto, por ejemplo una teoría física. No sería tolerable, digamos, afirmar que una teoría de los movimientos de los planetas es la que logra el mejor ajuste posible con sus movimientos observados, debiéndose cualquier discrepancia a errores por parte de los planetas. Si todo lo que tuviéramos para proceder, en la construcción de una teoría del significado, fuesen juicios de los hablantes con respecto a la verdad o falsedad de oraciones y las condiciones que prevalecen cuando se hiciesen esos juicios, entonces estaríamos autorizados a exigir, de cualquier teoría que se nos pidiera que aceptemos, que el ajuste sea perfecto, salvo por pequeñas discrepancias atribuibles a errores de observación. Afortunadamente, esto no es todo lo que tenemos para proceder.

El resultado de nuestra discusión es, por tanto, éste. Si una teoría del significado de este tipo se toma literalmente, en cuanto relacionada con una teoría de la verdad construida en oraciones reales, no tiene ventaja alguna sobre un manual de traducción, ya que ha de presuponer una comprensión del metalenguaje. Si, por otro lado, se la interpreta atribuyéndole a un hablante un conocimiento no verbalizado de las proposiciones expresadas por las oraciones de la teoría, su fuerza explicativa se evapora, ya que no proporciona medios con los que podamos explicar la adscripción a un hablante individual de las diversas proposiciones diferentes y de su interconexión deductiva. Es decir: una teoría del significado modesta o bien no lleva a cabo más que un manual de traducción, y por ello no logra explicar lo que alguien, en general, sabe cuando conoce un lenguaje; o ha de interpretarse holistamente, en cuyo caso su pretensión de dar una explica-

ción sistemática del dominio de un lenguaje es espúrea, ya que una visión holista del lenguaje impide la posibilidad de cualquier explicación así.

Hemos hecho notar que una teoría del significado, si representa la comprensión de una expresión como consistente en la posesión de un cierto conocimiento, no puede contentarse con especificar el objeto de este conocimiento e insistir en que «conocimiento» ha de tomarse en un sentido estricto; tiene también que exhibir el modo en que ese conocimiento tiene que ser derivado a fin de contar como conocimiento. Pero nuestras consideraciones más recientes se relacionaban con un punto diferente. En muchos contextos, podemos tomar como no problemática la adscripción a alguien de la apercepción de algún hecho, ya que podemos atribuirle una comprensión del lenguaje y la manifestación de su apercepción consistirá primariamente en su capacidad de enunciar el hecho o en su propensión a asentir a un enunciado del mismo. Pero, donde nos ocupamos de una representación de alguna capacidad práctica en términos de conocimiento proposicional y, en particular, donde esa capacidad práctica es precisamente el dominio de un lenguaje, estamos bajo la obligación, para que nuestra explicación sea satisfactoria, no sólo de especificar lo que alguien tiene que saber para que él tenga esa capacidad, sino también en qué consiste el que él tenga ese conocimiento, esto es, qué consideramos que constituye una manifestación del conocimiento de esas proposiciones; si no logramos hacerlo, la conexión entre la representación teórica y la capacidad práctica que pretende representar no quedará hecha. No estoy haciendo una objeción a la idea de una representación teórica de una capacidad práctica como tal, ni ciertamente a la representación del dominio del lenguaje por medio de una teoría deductiva: estoy diciendo solamente que esa representación está vacía de poder explicativo a no ser que la captación de las proposiciones individuales de la teoría se explique en términos de una capacidad práctica específica del hablante. No sé si esto es posible; no sé si el holismo es una concepción del lenguaje incorrecta. Pero estoy afirmando que la aceptación del holismo habría de conducir a la conclusión de que es imposible cualquier teoría del significado sistemática, y que el intento de resistirse a esta conclusión sólo puede conducir a la construcción de pseudoteorías; mi propia preferencia es, por tanto, asumir como principio metodológico que el holismo es falso.

La siguiente cuestión que surge naturalmente es si una teoría del significado vigorosa podría darse en términos de la noción de las condiciones de verdad de una oración: los aliviará escuchar que les ahorraré la extensa discusión que exigiría una respuesta. Pero estamos en posición de tratar brevemente de otra cuestión acerca de la forma que debe tomar una teoría del significado, a saber: si, en terminolo-

gía que tomo prestada del Sr. John McDowell, debería ser rica o austera. Si la teoría del significado se da en términos de condiciones de verdad, entonces, en lo concerniente a un nombre propio, una teoría rica le atribuirá a un hablante que entiende el nombre un conocimiento de la condición que cualquier objeto ha de satisfacer para ser el portador del nombre, mientras que una teoría austera lo representará simplemente como sabiendo, del objeto por el que de hecho está el nombre, que es el portador. Para este caso al menos, donde la teoría se construye en términos de condiciones de verdad, la distinción parece coincidir con la que hay entre una teoría vigorosa y una modesta, aunque se formule de modo diferente. Para un tipo de teoría más verificacionista, una teoría austera le atribuirá a cualquiera que entienda un nombre la capacidad de reconocer su portador cuando se lo encuentra, mientras que una teoría rica, en cambio, lo representará preparado para reconocer cuanto se considere que establece, para cualquier objeto dado, que es el portador. En favor de esta teoría rica pudiera decirse: «No reconocemos *simplemente* los objetos: los reconocemos *por* algún rasgo.» En nombre de la teoría austera pudiera responderse que *cómo* reconozcamos un objeto es un asunto psicológico, irrelevante para una teoría del significado, y que, en cualquier caso, no tiene por qué haber medios por los que los reconocamos; nadie podría, por ejemplo, explicar demasiado los medios por los que reconocemos que el predicado «... es rojo» se aplica a algo. Así, supongamos que nos encontramos con ciertas criaturas racionales pero no humanas que tienen un lenguaje que contiene lo que parecen ser nombres de ríos: aunque identifican ríos con estos nombres con notable precisión, no podemos descubrir los medios por los que hacen esas identificaciones ni pueden ellos dar ninguna explicación de ello. Continúa sucediendo sin embargo que si una de estas criaturas ha identificado dos tramos distintos de agua con el mismo nombre de río y se demuestra posteriormente, al trazar sus cursos, que no hay ningún flujo de agua que vaya del uno al otro, entonces tiene que retirar una u otra identificación; al menos, si estas criaturas no reconocen esa necesidad, sus palabras no pueden considerarse nombres *de ríos*. Las llamadas teorías de la referencia son teorías acerca de qué consideraríamos, en casos problemáticos, que establece qué objeto es el portador de un nombre propio dado, en caso de que lo sea alguno, y por tanto debieran llamarse más precisamente teorías del sentido para los nombres propios: el hecho de que sean tan discutibles muestra cuán inexplicita es nuestra captación de nuestro propio uso de los nombres propios. Pero si nuestras criaturas imaginarias usan los nombres de manera tal que, en un caso de desacuerdo, no aceptarían para decidir la cuestión de qué objeto es el portador cuanto aceptaríamos de hecho nosotros, entonces ellos no entienden esos nombres de la manera en que nosotros entendemos

los nuestros. Tales ejemplos ponen agudamente de relieve el mérito de la idea de que lo que determina el significado de una palabra no es tanto lo que en la práctica provoca su aplicación como aquello sobre lo que se está de acuerdo en que establece conclusivamente su aplicación correcta en casos de disputa: argüir que, en las ocasiones ordinarias de uso, no necesitamos basarnos en ningún principio que nos guíe en su aplicación, es no apreciar el quid de esta idea familiar.

Concluyo, por tanto, que una teoría del significado, para ser posible en cualquier forma, debe concordar con una concepción del lenguaje atomista, o al menos molecular, y no con una holista; que debe ser vigorosa, no modesta, y rica, no austera. No necesita resultar en adscripciones directas de significado; pero debe dar cuenta explícitamente, no sólo de lo que debe conocer cualquiera a fin de conocer el significado de cualquier expresión dada, sino también de lo que constituye tener ese conocimiento. Como advertí, el siguiente paso sería preguntar si esa teoría del significado habría de basarse en la noción de condiciones de verdad o en alguna otra noción. Cuando comencé a componer esta conferencia, tenía la absurda idea de que tendría tiempo para continuar discutiendo no sólo eso, sino también la cuestión planteada por el Profesor Strawson en su extremadamente interesante conferencia inaugural, concerniente a la relación entre teorías del significado como las hemos estado discutiendo y la explicación del significado dada por Grice, y concluir entonces examinando la noción de acto lingüístico y la relación que hay entre esos actos y sus interiorizaciones, por ejemplo entre aserción y juicio. Sólo tratando estos temas se podría pretender haber respondido a la pregunta que he tomado como mi título: pero, se alegrarán de saberlo, pensé que es mejor no tratar de completar la respuesta ahora.

## APÉNDICE

Es útil ver la teoría davidsoniana del significado contrastándola con la teoría del sentido y la referencia de Frege. Frege tuvo dos tipos de argumento en favor de la necesidad de una noción de sentido junto con la noción de referencia. El primero se relaciona con el conocimiento de un lenguaje por parte de un hablante y consiste, fundamentalmente, en la observación de que es ininteligible atribuirle a cualquiera un conocimiento cuya explicación *total* es que él conoce el referente de una expresión dada: si alguien sabe cuál es el referente de una expresión, ese referente le tiene que ser dado de algún modo particular, y el modo en que es dado constituye el sentido que él da a la expresión. Atribuir a alguien un conocimiento de, pongamos por caso, la referencia del nombre «Oxford» es decir de él que sabe, de la ciudad de Oxford, que es el referente de ese nombre. Decir de él que

conoce la referencia del nombre sin darle ningún sentido particular es decir que la explicación *completa* de su posesión de este conocimiento es dada diciendo que él sabe, de la ciudad, que es el referente del nombre; y esto equivale a decir que este conocimiento no puede ser caracterizado ulteriormente diciendo de él algo de la forma «El sabe que la ciudad que... es el referente de "Oxford"». Asimismo, atribuir a alguien un conocimiento de la referencia (extensión) de, pongamos por caso, el predicado «*x* es dúctil» es decir de él que sabe, de las cosas que son flexibles, que el predicado es verdadero de ellas: mientras que decir que él conoce la referencia del predicado sin darle ningún sentido particular es decir que esta atribución constituye una explicación *completa* de este conocimiento particular; y esto equivale a negar que este conocimiento pueda ser caracterizado ulteriormente diciendo de él cualquier cosa de la forma «Él sabe que "x es dúctil" es verdadero de cualquier objeto que...».

Es decir, una atribución a alguien de un conocimiento de la referencia de una expresión ha de entenderse como un enunciado de la forma [*X* sabe, de *a*, que es *F*] o de la forma [*X* sabe, que los *Gs*, que son *F*], esto es, un enunciado en que el sujeto de la cláusula-«que» está, en un contexto transparente, fuera de la cláusula-«que»; llamemos a tal enunciado una «atribución de conocimiento acerca de un objeto u objetos». Y la aserción de que alguien conoce la referencia de una expresión sin darle ningún sentido particular equivale a atribuirle conocimiento acerca de un objeto, u objetos, negando a la vez que haya ninguna caracterización ulterior de ese conocimiento por medio de un enunciado de la forma [*X* sabe que *b* es *F*] o [*X* sabe que los *Gs* son *F*], esto es, un enunciado en que el sujeto de la cláusula-«que» aparezca dentro de ella y, por tanto, en un contexto opaco, llamemos a tal enunciado una «atribución de conocimiento proposicional». Pero, de acuerdo con el argumento fregeano, una atribución de conocimiento acerca de un objeto u objetos es ininteligible si va acompañada de la afirmación de que no es posible ninguna caracterización ulterior, en términos de conocimiento proposicional. Pues, según esta concepción, el conocimiento proposicional es básico: siempre que es correcta una atribución de conocimiento acerca de un objeto u objetos, tiene que haber alguna atribución correcta de conocimiento proposicional de la que se siga. De ahí que no pueda haber una cosa tal como un conocimiento *neto* de la referencia de una expresión, esto es, un conocimiento de la referencia no mediado por sentido alguno que se dé a la expresión.

Debiera advertirse que este argumento, tal como es formulado, *no* entraña la llamada «teoría descriptiva de los nombres» que sus oponentes adscriben tendenciosamente a Frege. La «teoría causal de los nombres», por ejemplo, ofrece una descripción de la condición que un objeto ha de satisfacer para ser el portador del nombre. El de

acuerdo crucial entre la teoría causal y la teoría descriptiva no es sobre si existe alguna condición tal, sino sobre si es posible enunciarla sin referencia esencial al nombre mismo. (La atribución de la teoría descriptiva a Frege es tendenciosa porque no hay ningún argumento avanzado por él que pretenda mostrar que eso sea siempre posible.)

A pesar de todo lo que esta línea de argumentación puede mostrar, el sentido que cada hablante da a una expresión pudiera ser diferente, aunque cada uno tiene que darle *algún* sentido. La segunda línea de argumentación de Frege concierne a la contribución que hace una oración que entendemos a nuestro conocimiento no lingüístico, en el momento en que por primera vez la reconocemos como verdadera. Este argumento es sumamente familiar con aplicación a los enunciados de identidad: si, para entender un nombre propio, un hablante tiene que saber, del referente, que es el referente, entonces es incomprendible cómo un enunciado de identidad verdadero [ $a = b$ ] puede transmitirle nuevo conocimiento, puesto que él tiene ya que saber, del objeto que es el referente de los dos nombres, que es el referente de cada uno de ellos. En realidad, el argumento funciona igualmente bien para cualquier enunciado atómico: bajo el supuesto anterior sobre los nombres y bajo el supuesto correspondiente de que, para entender un predicado, un hablante tiene que saber, de cada objeto del que es verdadero el predicado, que el predicado es verdadero de él, es igualmente incomprendible cómo un enunciado verdadero formado insertando un nombre en el lugar argumental de un predicado puede transmitirle nueva información. Si suponemos que una explicación del uso del lenguaje en la comunicación exige que cada oración posea un contenido cognitivo común para todos los hablantes, entonces este argumento proporciona un fundamento para adscribirle a cada expresión un sentido constante de hablante a hablante.

La conclusión del primer argumento es, en efecto, que necesitamos adscribirle a un hablante *más* que sólo un conocimiento neto de la referencia de cada expresión, mientras que la conclusión del segundo argumento es que para que las oraciones sean informativas, no podemos, en general, atribuirles a los hablantes *tanto como* un conocimiento de la referencia de las expresiones. No hay tensión real aquí. Si exigimos meramente, para decir de alguien que él sabe de un objeto  $x$  que es el referente del nombre  $N$ , que haya algún término  $t$  que esté por  $x$  y sea tal que sea verdadero decir de esa persona [Él sabe que  $t$  es el referente de  $N$ ], entonces *no* se sigue que alguien que sepa, de un cierto objeto, tanto que es el referente de un nombre y que es el referente de otro, que él vaya a saber que los nombres tienen la misma referencia; por el contrario, tenemos aquí en forma esquemática precisamente la explicación en términos de sentido que



Frege propone como la solución del problema. La suposición que el segundo argumento trata de reducir al absurdo es, más bien, la de que la comprensión de una expresión consista en un conocimiento *neto* de la referencia. Lo que añade al primer argumento es sólo un fundamento para pensar que el sentido tiene que ser común a diferentes hablantes.

Ahora bien, a primera vista la teoría de Davidson es una teoría que lo explica todo en términos de la referencia, sin traer a colación el sentido; pero esta primera apariencia es totalmente desorientadora. La atribución por parte de Davidson a un hablante de un conocimiento (implícito) de la proposición expresada por el axioma que gobierna el nombre «Oxford» no consiste en mantener que ese hablante sepa, de la ciudad de Oxford, que el nombre «Oxford» la denota, sino, más bien, que sepa que «Oxford» denota la ciudad de Oxford. Así Davidson no está ciertamente atribuyéndole a cada hablante un conocimiento *neto* de la referencia de cada expresión que entienda, en el sentido en el que los argumentos de Frege hablan contra esa atribución. (En la conferencia, sí que interpreté que la noción de una teoría austera del significado de McDowell involucra un conocimiento neto de la referencia. Probablemente esto haya sido malinterpretar la intención de McDowell.)

Surge realmente la cuestión de exactamente *qué* conocimiento le atribuimos a un hablante cuando lo representamos como sabiendo que «Oxford» denota Oxford, dado que queremos atribuirle más que sólo el conocimiento trivial de la verdad de la oración ««Oxford» denota Oxford»: y aquí nos sentimos inclinados a decir que la teoría de Davidson es modesta porque, aunque no choca con los argumentos de Frege adscribiéndoles a los hablantes un conocimiento neto de la referencia, sino que admite que les den sentidos particulares a las expresiones, no intenta explicar qué son esos sentidos. Ésta es esencialmente la línea que seguí en la conferencia cuando criticaba la noción de una teoría del significado modesta; aunque, cuando llegué a considerar el holismo de Davidson, me incliné a considerar que entraña la imposibilidad de suministrar cualquier explicación del sentido.

Posteriores reflexiones me han sugerido, sin embargo, que puede que ésta no haya sido la línea correcta a adoptar. ¿Qué es una teoría del significado modesta? ¿Es una que deja lugar para una explicación de los sentidos que los hablantes dan a sus palabras (los conceptos que asocian con ellas), pero que no proporciona ella misma esa explicación? ¿O es una teoría que niega en principio la posibilidad de dar cualquier explicación así? Si consideramos la teoría de Davidson modesta en el primer sentido, entonces sigue abierta la posibilidad de rellenarla con una explicación de los sentidos específicos que los hablantes dan a las palabras del lenguaje y convertir así la teoría en una vigorosa, atomista: pero, en ese caso, ¿qué pasa con el aspecto holis-

ta de la teoría? El holismo que quedara haría relación sólo a la descripción del modo en que, a partir de la observación de la conducta lingüística y de otro tipo de los hablantes, podría llegarse a una teoría del significado para un lenguaje que, inicialmente, no se conociera: al planear una teoría del significado, tendríamos que ajustar la teoría a toda la evidencia proporcionada por los juicios de los hablantes en cuanto a la verdad y falsedad de sus oraciones. El holismo con respecto a la evidencia para una teoría del significado es, sin embargo, una cosa enteramente diferente de la visión holista del lenguaje de la que hablé en la conferencia. La última concierne a la teoría del significado misma, no al modo en que un no hablante pudiera llegar a ella; especialmente, hace relación a la explicación que se dé del modo en que una captación implícita de la teoría del significado que le es atribuida al hablante resulta en su empleo del lenguaje y, por tanto, en el contenido de esa teoría. El holismo con respecto meramente a cómo se podría llegar, partiendo de cero, a una teoría del significado para un lenguaje no tiene en sí mismo, por otro lado, esas implicaciones y es, por lo que yo puedo ver, inobjetable y casi banal. Es cierto que Davidson pretende que su holismo sea una doctrina con más mordiente que éste.

Davidson pudiera subscribir una doctrina tendenciosa del holismo, aun cuando su concepción de una teoría del significado fuera, en sí misma, neutral respecto de una visión del lenguaje holista, una molecular y una atomista; pero es improbable que no haya una conexión más orgánica entre los diferentes rasgos de su filosofía del lenguaje. Si, por otro lado, consideramos su teoría del significado modesta en el segundo de los sentidos indicados más arriba, se vuelve difícil ver cómo difiere de una teoría que repudie enteramente la noción de sentido y les adscriba a los hablantes un conocimiento neto de las referencias de sus palabras. La conclusión a la que me veo arrastrado es que, después de todo, es un error ver una teoría davidsoniana del significado como una teoría modesta en cualquier sentido. Veamos cómo puede ser esto.

Hay muchos tipos diferentes de consideración que se han aducido en favor del holismo lingüístico: el más relevante para nuestro propósito es el que generaliza las observaciones de Wittgenstein sobre el nombre «Moisés». La tesis de Wittgenstein es que hay una cantidad de cosas que creemos ordinariamente que son verdaderas de Moisés —que fue criado en un palacio real, que sacó a su pueblo de la esclavitud, que les dio la Ley, etc., etc.— Ninguna de ellas tiene que continuar manteniéndose verdadera so pena de que perdamos el uso del nombre «Moisés»: a condición de que continuemos creyendo que hubo sólo un hombre de quien un gran número de esas cosas son verdaderas, podemos rechazar el resto. De ahí que pueda admitirse que demos más peso a algunas cosas que creemos de Moisés que a otras;

por lo que respecta a la determinación del portador del nombre, podemos no darles ningún peso a algunas de ellas. Wittgenstein trató sólo del caso en que nos ocupamos de determinar el referente de un único nombre; pero es claro que podemos adaptarlo a aquél en que nos ocupamos de la determinación simultánea de los referentes de dos nombres, «Moisés» y «Aarón» pongamos por caso. Hay una cantidad de oraciones que contienen un nombre o el otro y que consideramos verdaderas, algunas de las cuales, como «Moisés y Aarón era hermanos», contienen ambos nombres. Podemos ahora hacer una estipulación del siguiente tenor. Si existe un único par de individuos  $m$  y  $a$ , tales que, cuando se los toma como los referentes respectivos de «Moisés» y «Aarón», una mayoría (ponderada) de las oraciones que contienen «Moisés» resultan verdaderas y también una mayoría (ponderada) de las oraciones que contienen «Aarón» resultan verdaderas, entonces esos individuos son los referentes reales de los nombres. Si no hay ningún par así, o hay más de un par así, pero hay un único individuo  $m$  tal que, cuando  $m$  se toma como el referente de «Moisés» y todas las oraciones que contienen «Aarón» se toman como falsas, una mayoría (ponderada) de las oraciones que contienen «Moisés» resultan verdaderas, entonces  $m$  es el referente real de «Moisés», mientras que «Aarón» carece de referente; y correspondientemente para el caso en que «Aarón» tiene referente pero «Moisés» es vacuo. Si ninguno de estos casos se da, ambos nombres carecen de referente.

No estoy abogando por esa doctrina; pero es fácilmente inteligible y tiene una obvia plausibilidad. Según esto, el sentido de un nombre propio es tal que hemos dispuesto de antemano que cualquiera de las cosas que consideramos parcialmente determinante del referente puede resultar falsa sin que el nombre quede privado de referencia. Esto no significa, naturalmente, que, cuando repudiamos por falso algo que anteriormente habíamos considerado verdadero y determinante en parte de la referencia, no sufra ninguna alteración el sentido del nombre; por el contrario, la sufre, porque ya no contamos al enunciado rechazado entre aquéllos una mayoría de los cuales tienen que ser verdaderos del portador del nombre.

La plausibilidad de la explicación de Wittgenstein no se limita a los nombres propios personales; es natural aplicarla a palabras de otras clases, por ejemplo a términos de masa. Llegamos a una forma de holismo si generalizamos simultáneamente la tesis a todas las palabras del lenguaje, incluyendo predicados, con la sola excepción de las constantes lógicas y tal vez las preposiciones y otras por el estilo. Supongamos que tenemos alguna gran clase (V) de oraciones consideradas verdaderas y conjuntamente determinantes de las referencias de nuestras palabras (nombres y predicados). Supongamos también, como una gran simplificación, que se nos da un determinado

universo de objetos sobre el que puedan tomarse como definidos los predicados y dentro del cual caigan las denotaciones de los nombres. Consideremos ahora todas las *asignaciones totales* posibles de referencias a los nombres y predicados del lenguaje: cada una de esas asignaciones totales constituirá una interpretación del lenguaje, relativa al universo dado, en el sentido de la semántica estándar de un lenguaje clásico de primer orden, salvo que una asignación total puede admitir que uno o más nombres no tengan referentes; asignará referentes a los demás nombres y extensiones a los predicados. Cualquier asignación total determinará valores de verdad para las oraciones atómicas del lenguaje y la evaluación se extenderá a todas las oraciones vía los axiomas de la teoría de la verdad que gobiernan los operadores formadores de oraciones. Podemos ahora especificar que los referentes reales de los nombres y las extensiones reales de los predicados sean aquellos que tienen bajo la asignación total *preferida* o *correcta*, explicándose a su vez esta última noción de alguna manera apropiada en términos de la clase V. La explicación más simple, y la que más probablemente favorecería un holista, sería decir que la asignación total preferida es la única (si es que la hay) que hace aparecer verdaderas a un número máximo de oraciones de V<sup>1</sup>.

Si ahora interpretamos que la teoría del significado de Davidson incorpora una explicación holista, siguiendo estas líneas, de cómo se determinan las referencias de las palabras primitivas no lógicas del lenguaje, ya no podemos considerar que carezca de una explicación de la captación de los sentidos de esas palabras por parte de un ha-

<sup>1</sup> Una explicación más fiel al modelo original wittgensteiniano tendría que ser un tanto complicada. Podríamos decir que una asignación total es *admisible* si, para cada palabra a la que se asigna un referente, hace verdaderas una mayoría de las oraciones de V que contienen esa palabra, y podríamos llamar el *grado* de una asignación el número de nombres a los que asigna un referente; podría entonces estipularse que la asignación total preferida es aquella única asignación admisible, si es que la hay, que es de grado máximo entre las asignaciones admisibles. La complejidad de esta formulación parece inevitable si ha de seguirse el patrón establecido por el caso de sólo dos nombres propios interconectados, como «Moisés» y «Aaron». Pues consideremos un caso en el que tenemos dos nombres propios así, «a» y «b», y sólo cinco oraciones que los contienen y que tenemos por verdaderas, «Fa», «Ga», «Rab», «Hb» y «Kb»; doy por fijadas las extensiones de los predicados. Supongamos que hay sólo cuatro individuos, *i*, *j*, *m* y *n*, que son candidatos a referentes de esos nombres, que *i* y *m* están en la extensión de «F», solamente *m* en la extensión de «G», *j* y *n* en la extensión de «H», y solamente *n* en la extensión de «K», mientras que el par  $\langle i, j \rangle$  es el único que está en la relación denotada por «R». Entonces, si asignamos *i* a «a» y *j* a «b», dos de las tres oraciones que contienen «a» se vuelven verdaderas, y dos de las tres que contienen «b» se vuelven verdaderas; pero exactamente el mismo resultado se obtiene asignando *m* a «a» y *n* a «b». Querríamos decir, en este caso, supongo, que la indeterminación privaba a los nombres de referencia; no habría fundamento alguno para estipular que sólo uno carecía de referencia, puesto que no habría base para decidir cuál de ellos.

blante: por el contrario, lo que un hablante sabe implícitamente es que la referencia se determina de esta manera holista. Ese conocimiento entra en el conocimiento por parte del hablante de las proposiciones expresadas por los axiomas de la teoría de la verdad. Por ejemplo, lo que un hablante sabe cuando sabe que «Oxford» denota Oxford es, según esta explicación, que «Oxford» denota el objeto al que se le asigna al nombre «Oxford» bajo la asignación total preferida a los nombres y predicados del español; lo que sabe cuando sabe que « $x$  es dúctil» es verdadero de un objeto si y sólo si ese objeto es flexible es que « $x$  es flexible» es verdadero de un objeto si y sólo si ese objeto pertenece al conjunto de objetos que se asigna como extensión de « $x$  es dúctil» bajo la asignación total preferida; lo que sabe cuando sabe que «La tierra se mueve» es verdadera si y sólo si la tierra se mueve es que «La tierra se mueve» es verdadera si y sólo si el objeto que se asigna a «la tierra» bajo la asignación total preferida es miembro del conjunto que se asigna al predicado « $x$  se mueve» bajo esa asignación.

Vista de este modo, una teoría davidsoniana del significado aparece como inerradicablemente holista, pero ya no como una teoría modesta en ningún sentido: considerada así, no es dañada por quejas, como las que hice en la conferencia, de que la teoría no da ninguna explicación de en qué consiste el conocimiento que se les atribuye a los hablantes del lenguaje. Yo aún defendería que la concepción entera de una teoría del significado modesta es ilegítima; pero pienso que la impresión que tuvimos, no sólo yo, sino, creo yo, una serie de partidarios de Davidson, de que una teoría del significado de este tipo ha de interpretarse como modesta, ha de rechazarse. Una gran parte de la razón para interpretarla así reside en el hecho de que Davidson siempre ha representado la colección de datos sobre los juicios realmente hechos por los hablantes en cuanto a la verdad y falsedad de las oraciones como estando en una relación de *evidencia* respecto de la teoría de la verdad resultante; mientras que, según la concepción holista del sentido que esbozé más arriba, no proporcionan apoyo externo a la teoría, sino que son integrantes de ella. Pues consideremos el modelo del que partimos, la explicación que Wittgenstein da del nombre «Moisés». De alguien que no tenga idea de *qué* oraciones que contengan el nombre «Moisés» son tenidas generalmente por verdaderas, pero que sepa meramente que el nombre denota a aquel individuo único, si lo hay, de quien una mayoría de esas oraciones, sean las que sean, son verdaderas, Wittgenstein no diría que capta el uso del nombre «Moisés»: tiene meramente una explicación esquemática de la forma que ha de adoptar una especificación de su uso —o del de cualquier otro nombre—. A fin de conocer el uso específico del nombre «Moisés», debe saber qué oraciones particulares que involucren el nombre son tenidas generalmente por

verdaderas. Admitidamente, los hablantes individuales explotan frecuentemente la existencia de un uso establecido para un nombre u otra palabra, ateniéndose a los medios establecidos para determinar la aplicación de la palabra sin tener ellos mismos un dominio completo de la misma; esto a menudo se aplica llamativamente a los nombres de lugares. Esto es una consecuencia del hecho de que un lenguaje es un fenómeno social en vez de una familia de dialectos semejantes y no afecta al punto fundamental. Para poder usar un nombre, u otra palabra, de una manera que no sea como un aparato registrador, un hablante ha de saber algo específico sobre el modo en que se determina su referencia, aunque no sepa todo lo que sea relevante; y el hecho de que hay una aplicación socialmente establecida a la que él se atiene depende de que haya un medio de descubrir lo que gobierna esa aplicación.

De la misma manera, según una teoría holista, no puede decirse que un hombre conozca el axioma que gobierna «la tierra», esto es, que conozca que «la tierra» denota la tierra, si sabe meramente que la expresión denota el objeto que se asigna a «la tierra» bajo la asignación total a las expresiones primitivas del español que hace aparecer verdaderas al máximo número de oraciones tenidas generalmente por verdaderas por los hablantes de español, sean cuales fueren esas oraciones. Al saber eso, sólo conoce el esquema general de acuerdo con el cual ha de darse la explicación particular de cualquier término singular en cualquier lenguaje y no sabe, adicionalmente, más que que «la tierra» es un término singular del español; podría tener ese conocimiento sin saber ninguna otra cosa más sobre la lengua española y, en tal caso, difícilmente podría decirse que conoce lo que significa «la tierra» ni, por tanto, la proposición expresada por «“La tierra” denota la tierra». Para conocer el significado específico de «la tierra», para conocer la proposición expresada por ese axioma, tiene que saber qué oraciones particulares componen la clase V, por relación a la cual se determina qué asignación total es la preferida. (El holismo surge del hecho de que lo requerido para una captación de los sentidos de todos los nombres y predicados del lenguaje dado es el mismo conocimiento específico.) Así, lo que Davidson llama la «evidencia» en favor de la teoría de la verdad es de hecho interna a ella. La teoría no es algo que basemos en la «evidencia», pero que pueda entenderse sin saber cuál puede ser la evidencia en su favor: no podemos captar ni transmitir el contenido de la teoría sin mención explícita, con detalle, de las oraciones que determinan conjuntamente las referencias de nuestras palabras; pues sin esa mención no podemos saber cuáles son las referencias que la teoría de la verdad asevera que tienen esas palabras.

Mi interés primario, en la conferencia, era llegar a ciertos principios básicos que regulan la construcción de una teoría viable del sig-

nificado; y la mayoría de estas conclusiones se mantienen en pie, aun cuando me equivocase al interpretar la concepción davidsoniana de una teoría del significado como una modesta. Una importante conclusión requiere, no obstante, reconsideración, concretamente que la adopción de una visión holista del lenguaje haga imposible la construcción de una teoría sistemática del significado. Eso depende ahora de si una teoría davidsoniana, interpretada a la manera holista anteriormente esbozada, es sencillamente creíble o no. Vimos que enunciar los principios que subyacen la determinación simultánea de las referencias de dos nombres propios de una manera wittgensteiniana era bastante complejo: pero en ese contexto se estaban dando por ya conocidos los sentidos de las demás palabras que aparecen en las diversas oraciones que contienen esos nombres; y, a causa de la fijeza supuesta de las aplicaciones de los términos generales, los resultados de una investigación en cuanto a los referentes de los nombres propios, y, en consecuencia, los valores de verdad de las oraciones que los contienen, podrían concebirse como susceptibles de ser *enunciados* por el uso de términos generales. Pero cuando tratamos de tomar seriamente la idea de que las referencias de todos los nombres y predicados del lenguaje sean conjunta y simultáneamente determinadas, queda claro que con ello le estamos atribuyendo a un hablante una tarea que está totalmente más allá de las capacidades humanas. En esa determinación simultánea, no hay razón por la que la referencia de cualquier palabra resultase ser tal que hace aparecer como verdaderas al máximo número de oraciones de *V* que contienen esa palabra; pero, aunque así fuera, sería de poca ayuda para guiar al hablante el pensar que el referente de un nombre es aquel individuo del que fuesen verdaderos la mayoría de los predicados extraídos de esas oraciones. Esto lo guiaría poco porque no podría dar por ya sentido en qué consiste el que cualquiera de esos predicados sea verdadero de cualquier individuo particular: por el contrario, eso sería necesario determinarlo al mismo tiempo, vía la determinación de las extensiones de los predicados primitivos que aparecen en esas oraciones y, últimamente, de todos los del lenguaje. Por la misma razón, el resultado del proceso de determinar la referencia de cualquier palabra nunca podría formularse verbalmente, excepto quizá cuando el referente fuese un posible objeto de ostensión, ya que las palabras que pudieran usarse para formularlo no podrían tomarse como si tuvieran una aplicación dada con antelación a la determinación de la referencia de la palabra en cuestión. Admitidamente, mientras que una demostración concluyente de la verdad de cualquier oración requeriría que se llevase a cabo efectivamente la tarea de descubrir los referentes de sus palabras constituyentes bajo la asignación total preferida, el hacer un solo juicio en cuanto al valor de verdad no necesita esperar por esa tarea, del mismo modo que un juicio sobre Moisés no

tiene que esperar por una decisión definitiva en cuanto a cuáles de las cosas que normalmente creemos de él son verdaderas; se podría en realidad inferir de la teoría holista que nunca podría proporcionarse una demostración concluyente de la verdad. Queda en pie el hecho de que, así como en la teoría de Wittgenstein, para conocer el contenido de cualquier oración que contenga el nombre «Moisés», se ha de conocer cómo se determina el referente de ese nombre y también las cosas particulares que creemos acerca de Moisés, así también en la teoría holista, para captar el contenido de cualquier oración individual, se ha de conocer la composición de la entera totalidad V y también se ha de tener la concepción de una determinación simultánea de las referencias de nuestras palabras en relación a ella.

La dificultad de hacer plausible la explicación holista se vuelve más manifiesta cuando indagamos en la composición de la totalidad de base V. Sería de alguna manera contrario al espíritu del holismo admitir que exista una clase especial de oraciones privilegiadas, entre todas las consideradas generalmente verdaderas, que pudiéramos llamar «cuasi-analíticas»: oraciones que no son individualmente inmutables a revisión (aunque el rechazo de cualquiera de ellas efectúe un cambio en los sentidos de nuestras palabras), pero que desempeñan un papel especial en la determinación de las referencias de nuestras palabras, papel que no desempeñan otras oraciones reconocidas como verdaderas. El holista sin embargo se enfrenta a una elección, en la formulación exacta de su doctrina, sobre si ha de admitir o no desacuerdos entre hablantes. Si no lo hace, entonces tiene que considerar que V comprende sólo aquellas oraciones que todos los hablantes aceptan como verdaderas o, al menos, que muchos aceptan como verdaderas y que ninguno rechaza como falsas, y, por tanto, tiene que considerar que incluye sólo oraciones que no tienen ningún rasgo indético significativo. Pero en este caso se hace implausible que V sea adecuada para determinar la aplicación de muchos predicados, por ejemplo de «... es dúctil»: aunque la mayoría de los hispanoablantes estuviesen de acuerdo en cualquier aplicación particular de ese predicado, hay demasiadas pocas oraciones reales que contengan esa palabra, y cuya verdad reconozcan la mayoría de los hablantes, para determinar su extensión. A la vista de esta dificultad, es más probable que el holista haga la otra elección y considere que la totalidad de base V consiste no en oraciones, sino más bien en juicios individuales de valor de verdad hechos por hablantes particulares. En este caso, V contendrá no sólo juicios divergentes sobre oraciones no indéticas, sino también juicios relativos a oraciones con elementos indéticos o, más precisamente, a enunciados (donde un enunciado es considerado como un triple de una oración, un hablante y un tiempo). Esta elección, sin embargo, conlleva una diferente implausibilidad: donde V se considera como la totalidad de todos los juicios



efectivamente hechos por hablantes del lenguaje, ningún hablante se acercará lo más mínimo a tener una captación de la teoría del significado correcta para ese lenguaje, ya que la vasta mayoría de esos juicios le será desconocida.

A fin de escapar a este absurdo, el holista está sujeto a una fuerte tentación a comprimir la noción de lenguaje hasta convertirla en la de idiolecto; a cada hablante individual se lo concibe ahora como poseedor de su personal teoría de la verdad para el lenguaje que habla, una teoría que incorpora, en su totalidad de base V, todos los juicios que él hace personalmente, pero ninguno de los demás hablantes, ya que son irrelevantes para su idiolecto. Tal concepción invierte la verdadera relación que hay entre la noción de idiolecto y la de lenguaje en el sentido cotidiano de «lenguaje». Un lenguaje, en el sentido cotidiano, es algo esencialmente social, una práctica en la que toman parte muchas personas; y es esta noción, en vez de la de idiolecto, la que debe tomarse como primaria. No podemos, ciertamente, prescindir de la noción de idiolecto, que representa la comprensión siempre parcial y a menudo en parte incorrecta, que un individuo tiene de su lenguaje; pero necesita ser explicada en términos de la noción de un lenguaje compartido, y no a la inversa. Una entre muchas razones para sostener esto es el fenómeno llamado por Putnam la «división lingüística del trabajo»; pero es innecesario proseguir aquí este punto con detalle, ya que el cambio de un lenguaje común a un idiolecto no salva al holista de su dificultad.

Si el dominio que un hablante tiene de su lenguaje consiste en una captación implícita de una teoría del significado para ese lenguaje, entonces, si la teoría es holista, él tiene que apercebirse de los juicios que comprenden la totalidad de base. Aun cuando el lenguaje sea su propio idiolecto personal, esa totalidad no puede, por tanto, contener una multitud de juicios casuales que él ha hecho pero ha olvidado posteriormente; sólo puede contener, en cualquier momento dado, aquellos juicios que puedan obtenerse de él en ese momento. Esto hace aún altamente improbable que la totalidad pueda ser suficientemente extensa para determinar las referencias de todas las palabras de su lenguaje.

De ciertas palabras es perfectamente razonable mantener la tesis de que la referencia de cada una de ellas es determinada por el requisito de que una o más oraciones que la contenga resulte verdadera. Siempre que pueda mantenerse que hay un modo esencialmente único de definir una palabra, este hecho puede expresarse aplicando esa tesis a una única oración que incorpore la definición; y la tesis puede aplicarse a cualquier palabra que deba, o incluso pueda, presentarse por medio de una explicación verbal, equivalga o no esa explicación a una definición real. La teoría descriptiva de los nombres propios deriva la considerable plausibilidad que posee precisamente

del hecho de que los nombres propios pueden ser, y a menudo son, presentados a alguien que no los conoce por medio de una explicación verbal; y este hecho subyace también a la descripción que Wittgenstein hace del nombre «Moisés», la cual, como Kripke ha observado, es una modificación de la teoría descriptiva. La modificación tiene dos rasgos: primero, deja lugar para el hecho de que hay habitualmente más de un modo legítimo de presentar un nombre propio y de que esos diferentes modos, tomados conjuntamente, suministran más de lo necesario para determinar su referencia; y, en segundo lugar, previene de antemano la resolución de cualquier conflicto que pueda surgir entre los medios alternativos de fijar el referente. Esta explicación puede, también, representarse mediante la tesis de que la referencia de un nombre tal es determinada por el requisito de que una mayoría ponderada de las oraciones que podrían usarse para presentarlo han de volverse verdaderas. Entre los términos generales, algunos se comportan a este respecto como nombres propios, mientras que, para algunos otros, no hay criterios múltiples de aplicación que puedan entrar en conflicto, sino, por el contrario, esencialmente sólo un modo correcto de explicarlos. Otros, a su vez, ocupan una posición intermedia: su explicación es compleja, en el sentido de que se podría representar su extensión como determinada por el requisito de que una serie de oraciones diferentes resultasen verdaderas; pero el conflicto que sería provocado por el descubrimiento de que es imposible mantener todas esas oraciones, hasta aquí consideradas como constitutivas de su significado, sería mucho más grave que en el caso de un nombre como «Moisés» y la medida que adoptaríamos para resolver ese conflicto no es prevista de antemano.

Indudablemente, es una falacia suponer que se pueda siempre hacer equivaler simplemente el sentido de lo que se dice como explicación de una palabra con el sentido de esa palabra; y, sea lo que fuere lo que se piense en detalle de las concepciones de los nombres propios de Kripke, sirven para subrayar la falacia: en la medida en que haya una diferencia generalmente entendida en el empleo de una descripción definida y de un nombre propio, el oyente hará una tácita admisión de esa diferencia cuando se le presente un nombre propio por medio de una descripción definida. Esa concesión no invalida, sin embargo, la idea de que el medio que empleemos para transmitirle a alguien el sentido de una palabra que no entendía previamente exhibe el sentido que lleva en el lenguaje, donde captar el sentido de una palabra se hace equivaler a entender su uso aceptado. Si, por ejemplo, existe un medio establecido de fijar la referencia de un nombre, será necesariamente parte integrante del sentido de ese nombre.

La tesis de que la referencia es determinada por el requisito de

que todas, o la mayoría de, las oraciones de cierto conjunto resulten verdaderas puede, así pues, retenerse con respecto a un gran número de palabras. Pierde su plausibilidad, sin embargo, cuando el holista la generaliza aplicándola a todas las palabras del lenguaje simultáneamente. Esto sucede porque era, al principio, un modo particular de representar el sentido de una palabra que es posible introducir por medio de una explicación verbal: su plausibilidad se extiende por tanto mientras se aplique sólo a palabras que puedan ser introducidas así, y se mantiene sólo en relación a aquellas oraciones que podrían usarse legítimamente para dar esa explicación. Si la captación del sentido de una palabra por parte de un hablante se ha de representar como consistente en su conocimiento de que su referencia es determinada por un conjunto de oraciones que contienen esa palabra, entonces esas oraciones tienen que ser oraciones que pudieran de hecho obtenerse de él para explicar la palabra; y, si estamos considerando a la palabra parte de un lenguaje común, entonces tienen que ser oraciones aceptadas generalmente como verdaderas y también como determinantes del sentido de la palabra, esto es, como oraciones que sería legítimo citar para explicar su sentido. El holista se equivoca por tanto al incluir, en su totalidad de base V, juicios particulares de hablantes individuales, o juicios que un hablante individual no recuerde haber hecho o a los que no aludiría para explicarle una palabra a alguien que no la entendiese. Se sigue que la tesis de que las referencias de nuestras palabras son determinadas por el requisito de que ciertas oraciones sean verdaderas no puede generalizarse, como el holista desea generalizarla, para dar cuenta de cómo se fijan las referencias de *todas* las palabras del lenguaje: hay muchas palabras en el lenguaje que no son, y no pueden ser, introducidas por medio de explicaciones puramente verbales y a éstas la tesis sencillamente no se aplica. Nuestro lenguaje es una estructura de muchos pisos y la posibilidad de introducir nuevas expresiones —en el lenguaje o en el vocabulario de un hablante particular— por medio de explicaciones lingüísticas depende de que hayamos construido previamente los pisos inferiores con medios diferentes; notoriamente el holismo muestra su mayor debilidad en la explicación que da de la adquisición progresiva del lenguaje. Pero a una teoría del significado correcta se le exige que dé una explicación de lo que es dominar un lenguaje *en general*: un modelo que dé una representación solamente de cómo, mediante el uso de una parte fundamental del lenguaje, se puede llegar a captar los sentidos de expresiones de niveles superiores es un mal modelo general a emplear en la construcción de esa teoría.

Como ya hemos visto, los juicios hechos por los hablantes individuales juegan un papel dual en la explicación de Davidson: por un lado, forman la evidencia que pudiera usar alguien sin conocimiento

previo del lenguaje que deseara construir para él una teoría del significado; por otro lado, se convierten en un ingrediente de la propia teoría, al componer la totalidad  $V$  que determina las referencias de las palabras. En el primer papel, no puede plantearse ninguna objeción contra la apelación que se hace a ellos: si estamos tratando de descubrir, a partir de la observación de la conducta lingüística de alguien, el sentido que él da a una cierta palabra, prestaremos naturalmente atención a todos los juicios de valor de verdad que haga con respecto a las oraciones que contienen esa palabra, ya que esos juicios exhiben obviamente la propensión que él tiene a emplear la palabra de una cierta manera. Pero la idea de que entonces podemos obtener, por referencia a la totalidad de todos los juicios hechos por los hablantes, una sola representación uniforme de la manera en que se determinan los portadores de todos los nombres, y las extensiones de todos los predicados del lenguaje, pasa por alto la diversidad de los muchos tipos de expresión que contiene nuestro lenguaje y las gradaciones de nivel al que están. Esto puede parecer una afirmación dura, en vista del hecho de que fue Quine, el principal exponente contemporáneo del holismo lingüístico, quien avanzó la célebre imagen del lenguaje como una estructura articulada cuyas oraciones están a diferentes profundidades respecto de la periferia; pero el caso es que esa imagen en modo alguno representa una visión esencialmente holista del lenguaje y, en realidad, concuerda bastante mal con esa visión. Para el holismo, el lenguaje no es una estructura de muchos pisos, sino más bien un vasto complejo de un solo piso; sus dificultades al dar cuenta de nuestra adquisición pieza por pieza del lenguaje brotan del hecho de que no puede darle ningún sentido a la idea de conocer parte de un lenguaje. Como en el caso actual, las intuiciones que proporcionan los puntos de arranque para los argumentos conducentes a una visión holista son perfectamente genuinas; el holismo surge de sucumbir a la tentación de generalizarlas más allá de su ámbito de aplicación, para llegar a una sola fórmula que cubra todo caso.

## V. ACTOS DE HABLA

V. ACTOS DE HABLA

El tipo de oración por el que tradicionalmente se han interesado los filósofos ha sido el declarativo, el de las oraciones mediante las que típicamente, pero no necesariamente, se hacen enunciados. Con ello su interés ha estado dirigido hacia las oraciones que *figuran* o *representan* hechos y son por ello calificables de verdaderas o falsas. Ya desde Aristóteles, el estudio de las oraciones que no poseían esta característica fue transferido a la *Retórica* o a la *Poética*, y la fundación de la moderna filosofía del lenguaje no llevó aparejado cambio significativo alguno. Tanto Frege, como Russell, como el autor del *Tractatus* o los positivistas lógicos compartían, por encima de sus diferencias, el punto de vista de acuerdo con el cual el objeto primario del lenguaje es representar y comunicar información fáctica, esto es: la parte del lenguaje que contaba era la *cognitiva*, que es independiente de las intenciones, deseos o creencias que los hablantes tengan: de este modo el significado se entendía estrictamente en términos de condiciones de verdad. Pero esto es una concepción cuando menos parcial de aquello en lo que consiste un lenguaje; hablar no es *solamente* emitir oraciones para comunicar información fáctica. Ciertamente los filósofos sabían esto desde hacía mucho tiempo, pero no prestaron la atención debida al componente de *acción* que el lenguaje conllevaba. Es más: la tendencia dominante fue asimilar todos los casos de discurso significativo al modelo enunciativo. Así, a la pregunta «¿Cómo puede una promesa tener un significado tal que cree una obligación?» se solía responder diciendo: «Porque mi emisión de “Prometo tal-y-tal” describe un acto mental mío: mi firme resolución de hacer tal-y-tal: ése es su significado.» Este tipo de respuesta es una instancia de lo que Austin llamará «falacia descriptiva».

Esta concepción entró definitivamente en crisis al intentar llevarla hasta sus últimas consecuencias. Durante los años treinta el positivismo lógico extrajo como consecuencia, aplicando el principio de verificación, que la mayor parte del discurso ético, estético, filosófico e incluso ordinario carecía, estrictamente hablando, de sentido. Y esto que, debe reconocerse con Austin, tuvo un efecto temporalmente saludable en el planeta filosófico, era más de lo que sensatamente podía aceptarse e incluso sirvió en cierto sentido de aguijón al propio Wittgenstein para revisar su doctrina del *Tractatus* y abrazar en las *Investigaciones* la tesis de que había que «romper con la idea de que el lenguaje funciona siempre de *una* manera, que tiene siempre el mismo objeto: transmitir pensamientos». Por esa misma época J. L. Austin trabajaba, sin al parecer ninguna conexión con Wittgenstein, sobre lo que sería su teoría de los actos de habla. El artículo «Emisio-

nes realizativas» es una de las primeras exposiciones de ella: en él se presenta la célebre distinción entre emisiones realizativas y emisiones constatativas y se argumenta que las primeras son, en contra de lo que podría pensar algún positivista lógico, perfectamente significativas sin que tengan valores de verdad. John R. Searle es el filósofo que más ha contribuido a sistematizar y divulgar (particularmente entre los lingüistas) la teoría de los actos de habla. «¿Qué es un acto de habla?» presenta las tesis de Austin (que, se diría, están originalmente formuladas en términos de «aires de familia») en un armazón más rígido y sistemático: en términos de condiciones necesarias y suficientes. Tales condiciones no son, sin embargo, suficientes para realizar una clasificación exhaustiva de los actos de habla, tarea que con éxito dudoso (¿qué ventajas, podría preguntarse, reportaría tal clasificación?), Searle emprende en «Una taxonomía de los actos de habla».



## EMISIONES REALIZATIVAS \*

J. L. AUSTIN

### I

Tienen ustedes más que derecho a no saber lo que significa la palabra «realizativo». Es una palabra nueva y una palabra fea, y acaso no signifique nada demasiado. Pero en cualquier caso hay algo en su favor, que no es una palabra profunda. Recuerdo que una vez, cuando yo había hablado de este tema, alguien dijo después: «Sabes, no tengo la menor idea de lo que quiere decir, a menos que pudiera ser que simplemente quiera decir lo que dice.» Pues bien, esto es lo que me gustaría querer decir.

Consideremos primero cómo surge este asunto. No tenemos que retroceder muy lejos en la historia de la filosofía para encontrar filósofos dando por sentado como algo más o menos natural que la única ocupación, la única ocupación interesante, de cualquier emisión —es decir, de cualquier cosa que decimos— es ser verdadera o al menos falsa. Naturalmente, siempre han sabido que hay otros tipos de cosas que decimos —cosas como imperativos, las expresiones de deseos, y exclamaciones— algunas de las cuales han incluso sido clasificadas por los gramáticos, aunque tal vez no era demasiado fácil decir siempre cuál era cuál. Pero con todo, los filósofos han dado por sentado

\* Versión castellana de Alfonso García Suárez.

Traduzco *utterance* como «emisión» y *performative* como «realizativo». La alternativa de traducir la primera como «expresión» —tal como hacen G. R. Carrió y E. A. Rabossi en *Palabras y acciones*, Paidós, Buenos Aires, 1971 (traducción de las conferencias de Austin *How to do Things with Words*)— no me parece aceptable porque prefiero reservar esta palabra para *phrase*: segmento lingüístico que no llega a oración: *i. e.*, lo que los gramáticos europeos llaman sintagma. Tampoco «locución» parece adecuada porque Austin la utiliza, en *Palabras y acciones*, para referirse al contenido de un acto locucionario. «Emisión» goza además de la ventaja —de la que carece «locución»— de tener asociado el verbo *emitir*, correspondiente al inglés *to utter*. En cambio la traducción de Carrió y Rabossi de *performative* por «realizativo» me parece perfecta y preferible a la alternativa «ejecutivo». La razón es que «realizativo» es un neologismo sobre la base de palabras ya en uso: «realizar» y «realización» —del mismo modo que *performative* es, como dice Austin, una nueva y fea palabra sobre la base de *perform* y *performance*—. «Ejecutivo» equivaldría, en cambio, a *executive* (cfr. «poder ejecutivo»). (N. del T.)

que las únicas cosas en las que están interesados son las emisiones que registran hechos o que describen situaciones con verdad o con falsedad. En los tiempos recientes este tipo de postura ha sido puesto en duda —en dos etapas, creo yo—. Lo primero de todo la gente comenzó a decir: «Está bien, si estas cosas son verdaderas o falsas debe ser posible decidir qué son, y si no podemos decidir qué son no son nada buenas sino que son, en resumen, sinsentidos.» Y esta nueva postura hizo muchísimo bien; una gran cantidad de cosas que probablemente son sinsentidos se descubrieron como tales. Sin embargo, no creo que sea cierto que se hayan clasificado adecuadamente todos los tipos de sinsentido, y tal vez algunas cosas que han sido rechazadas por sinsentidos no lo sean realmente; pero con todo este movimiento, el movimiento verificacionista, fue, a su manera, excelente.

Con todo, llegamos entonces a la segunda etapa. Después de todo, ponemos unos límites a la cantidad de sinsentido que decimos, o al menos a la cantidad de sinsentido que estamos dispuestos a admitir que decimos; y así la gente comenzó a preguntarse si después de todo algunas de las cosas que, tratadas como enunciados, estaban en peligro de ser rechazadas como sinsentidos fueron después de todo realmente propuestas como enunciados. ¿No podrían tal vez no pretender registrar hechos sino influir en la gente de esta o aquella manera, o dar rienda suelta a las emociones de esta o aquella manera? ( ) tal vez en cualquier caso algunos elementos de estas emisiones realizaban esas funciones, o, por ejemplo, llamaban la atención de alguna forma (sin registrarlo efectivamente) hacia algún rasgo importante de las circunstancias en que la emisión se hacía. Sobre estas líneas la gente ha adoptado ahora un nuevo eslogan, el eslogan de los «diferentes usos del lenguaje». La vieja postura, la vieja postura enunciativa, es incluso llamada a veces una falacia, la falacia descriptiva.

Ciertamente hay una gran cantidad de usos del lenguaje. Es más bien una pena el que la gente tienda a invocar un nuevo uso del lenguaje siempre que se sienten inclinados a hacerlo, para que les ayude a salir de este, de aquel o del otro bien conocido enredo filosófico, necesitamos más de un entramado en el que discutir estos usos del lenguaje; y también creo que no debiéramos desesperarnos tan fácilmente y hablar, como tiende a hacer la gente, de los *infinitos* usos del lenguaje. Los filósofos hacen esto cuando han enumerado tantos como, digamos, diecisiete; pero incluso si hubiese unos diez mil usos del lenguaje, seguro que podríamos enumerarlos todos con tiempo. Esto, después de todo, no es mayor que el número de especies de escarabajo que los entomólogos se han tomado la molestia de enumerar. Pero sean cuales fueren los defectos de cualquiera de ambos movimientos —el movimiento «verificacionista» o el movimiento del «uso del lenguaje»—, en cualquier caso han dado lugar, nadie podría negarlo, a una gran revolución en filosofía y, dirían muchos, la más

saludable en su historia. (Una pretensión, si se paran a pensarlo, no muy inmodesta.)

Pues bien, es una de esas suertes de uso del lenguaje la que quiero examinar aquí. Quiero discutir un tipo de emisión que parece un enunciado y supongo que gramaticalmente sería clasificada como un enunciado que no es carente de sentido, y sin embargo no es verdadera o falsa. Estas no van a ser emisiones que contienen verbos curiosos como «pudo» o «podría», o palabras curiosas como «bueno», que muchos filósofos consideran hoy en día sencillamente como señales peligrosas. Serán emisiones perfectamente claras, con verbos corrientes en primera persona del singular del presente de indicativo de la voz activa, y no obstante veremos de inmediato que no tienen la posibilidad de ser verdaderas o falsas. Más aún, si una persona hace una emisión de este tipo, diríamos que está *haciendo* algo en vez de meramente *diciendo* algo. Esto puede sonar un poco extraño, pero los ejemplos que daré de hecho no son extraños en absoluto, y puede que incluso parezcan decididamente grises. He aquí tres o cuatro. Supongamos, por ejemplo, que en el transcurso de una ceremonia nupcial digo, como la gente hace, «Sí quiero» \* —(sc., tomar a esta mujer por mi esposa legalmente desposada). O también, supongamos que le piso a usted en el pie y digo «Le pido disculpas» \*\*. O también, supongamos que tengo la botella de champán en la mano y digo «Bautizo este barco el *Queen Elizabeth*». O supongamos que digo «Te apuesto cinco duros que lloverá mañana». En todos estos casos sería absurdo considerar la cosa que digo como un registro de la realización de la acción que indudablemente se hace —la acción de apostar, o bautizar, o disculparse—. Diríamos más bien que, al decir lo que digo, realizo efectivamente esa acción. Cuando digo «Bautizo este barco el *Queen Elizabeth*» no describo la ceremonia de bautizo, realizo efectivamente el bautizo; y cuando digo «Sí quiero» (sc., tomar esta mujer como mi esposa legalmente desposada), no estoy informando de un matrimonio, estoy satisfaciéndolo.

Pues bien, estos tipos de emisiones son las que llamamos emisiones *realizativas*. Esta es una palabra un poco fea, y una palabra nueva, pero parece que no hay ya en existencia ninguna palabra que haga su oficio. La aproximación más cercana que se me ocurre es la palabra «operativo», tal como la usan los abogados. Los abogados cuando hablan de instrumentos legales distinguen entre el preámbulo, que recita las circunstancias en que se efectúa una transacción, y por otro lado la parte operativa —la parte que realiza efectivamente el acto legal que el instrumento se propone realizar—. De manera

\* «I do.» (N. del T.)

\*\* «I apologize.» (N. del T.)

que la palabra «operativo» está muy cerca de lo que queremos. «Lego mi reloj a mi hermano» sería una cláusula operativa y es una emisión realizativa. Sin embargo, la palabra «operativo» tiene otros usos, y parece preferible tener una palabra especialmente asignada para el uso que queremos.

Ahora bien, en este punto alguien podría objetar, quizá incluso con cierta alarma, que parezco estar sugiriendo que casarse es simplemente decir unas cuantas palabras, que justamente el decir unas cuantas palabras es casarse. Bien, esto ciertamente no es el caso. Las palabras tienen que decirse en las circunstancias apropiadas, y ésta es una cuestión que volverá a surgir después. Pero lo que no debemos suponer es que lo que se necesita en tales casos además de decir las palabras es la realización de un acto espiritual interno, del cual las palabras serán entonces un registro. Es muy fácil deslizarse hacia esta opinión al menos en casos difíciles, portentosos, aunque quizá no sea tan fácil en casos simples como el de pedir disculpas. En el caso de prometer —por ejemplo, «Prometo estar allí mañana»— es muy fácil pensar que la emisión es simplemente el signo externo y visible (es decir, verbal) de la realización de un acto espiritual interior de prometer, y esta opinión ciertamente ha sido expresada en muchos lugares clásicos. Tenemos el caso del Hipólito de Eurípides que dijo «Mi lengua lo juró, pero mi corazón no» —quizá debiera ser «mente» o «espíritu» en vez de «corazón», pero en cualquier caso una especie de artista de candilejas—. Ahora bien, es claro con este tipo de ejemplo que, si nos deslizamos hacia la creencia de que esas emisiones son registros, verdaderos o falsos, de la realización de actos espirituales e interiores, abrimos una fisura a perjuros y estafadores y bigamos, etc., de manera que tiene desventajas el ser excesivamente solemne de esta forma. Tal vez sea mejor aferrarnos al viejo dicho de que la palabra empeña.

Sin embargo, aunque estas emisiones no registran ellas mismas hechos y no son ellas mismas verdaderas o falsas, el decir estas cosas muy a menudo *implica* que determinadas cosas son verdaderas y no falsas, en algún sentido al menos de la palabra un tanto enredosa «implicar». Por ejemplo, cuando digo «Tomo a esta mujer como mi esposa legalmente desposada», o alguna otra fórmula de la ceremonia matrimonial, implico que no estoy ya casado, con esposa viva, cuerda, no divorciada, y demás cosas. No obstante, es muy importante darse cuenta de que implicar que esto o lo otro es verdadero, no es en absoluto lo mismo que decir algo que es ello mismo verdadero.

Estas emisiones realizativas no son, pues, verdaderas o falsas. Pero sufren de ciertas incapacidades propias. Pueden fracasar de maneras especiales y esto es lo que quiero considerar ahora. Las diversas maneras en que una emisión realizativa puede ser insatisfactoria

los llamamos, por darles un nombre, infortunios \*; y un infortunio surge —es decir, la emisión es desafortunada— si se rompen determinadas reglas, reglas transparentemente simples. Mencionaré algunas de estas reglas y daré entonces ejemplos de algunas infracciones.

Ante todo, es obvio que debe efectivamente existir el procedimiento convencional que nos estamos proponiendo emplear con nuestra emisión. En los ejemplos dados aquí este procedimiento será verbal, un procedimiento verbal para casarse o donar o lo que fuere; pero debe tenerse en cuenta que hay muchos procedimientos no verbales por los que realizar exactamente los mismos actos que realizamos por estos medios verbales. Vale la pena recordar también que una gran cantidad de las cosas que hacemos son al menos en parte de este género convencional. Los filósofos al menos tienden demasiado a dar por sentado que una acción es siempre en último extremo el llevar a cabo un movimiento físico, mientras que es usualmente, al menos en parte, una cuestión de convención.

La primera regla es, pues, que la convención invocada debe existir y ser aceptada. Y la segunda regla, también muy obvia, es que las circunstancias en que nos proponemos invocar este procedimiento deben ser apropiadas para su invocación. Si esto no se observa, entonces el acto que nos proponemos realizar no saldría —será, podríamos decir, un fallo \*\*. Esto también ocurrirá si, por ejemplo, no llevamos a cabo el procedimiento —sea lo que fuera— correcta y completamente, sin ningún defecto \*\*\* y sin ninguna obstrucción \*\*\*\*. Si alguna de estas reglas no se observa, decimos que el que nos proponíamos realizar es nulo, sin efecto. Si, por ejemplo, el pretendido acto era un acto de casarse, entonces diríamos que «tomamos parte en una formalidad» de matrimonio, pero que no logramos efectivamente casarnos.

He aquí algunos ejemplos de este tipo de fallo. Supongamos que, viviendo en un país como el nuestro, deseamos divorciarnos de nuestra esposa. Podemos intentar colocarla directamente frente a nosotros en la sala y decir, en voz lo bastante alta para que lo oigan todos, «Me divorcio de ti». Ahora bien, este procedimiento no es aceptado. No hemos logrado con él divorciarnos de nuestra esposa, al menos en este país y otros como él. Este es un caso en que la convención, diríamos, no existe o no es aceptada. Por otra parte, supongamos que, escogiendo compañeros en una fiesta infantil, digo «Escojo a Jorge». Pero Jorge se sonroja y dice «No juego». En este caso obviamente,

\* *infelicitities*. (N. del T.)

\*\* *misfire*. (N. del T.)

\*\*\* *flaw*. (N. del T.)

\*\*\*\* *bûche*. (N. del T.)

por una razón u otra, no escogí a Jorge —ya sea porque no existe la convención de que puedes escoger gente que no juega, o porque Jorge en esas circunstancias es un objeto inapropiado del procedimiento de escoger—. O consideremos el caso en que digo «Te nombro cónsul», y resulta que ya has sido nombrado —o quizá pueda incluso trascender que eres un caballo—; aquí de nuevo tenemos el infortunio de circunstancias inapropiadas, objetos inapropiados, o cosas por el estilo. Ejemplos de defectos y obstrucciones tal vez apenas sean necesarios —una parte en la ceremonia de matrimonio dice «Sí querré», la otra dice «No querré»; yo digo «Apuesto cinco duros», pero nadie dice «Hecho», nadie acepta la oferta—. En todos estos casos y otros similares, el acto que nos proponemos realizar, o que nos ponemos a realizar, no es logrado.

Pero hay otra manera un tanto diferente en que este tipo de emisión puede ir mal. Una buena cantidad de estos procedimientos verbales están diseñados para ser usados por personas que sostienen determinadas creencias o tienen determinados sentimientos o intenciones. Y si usted usa una de estas fórmulas cuando no tiene los pensamientos o sentimientos o intenciones requeridos entonces hay un abuso del procedimiento, hay una insinceridad. Tomemos, por ejemplo, la expresión «Te felicito». Esta expresión está diseñada para ser usada por personas que se alegran de que la persona aludida haya logrado una determinada hazaña, que creen que ella fue personalmente responsable del éxito, etc. Si digo «Te felicito» cuando no me alegro o cuando no creo que el mérito fuese tuyo, entonces hay una insinceridad. Asimismo si digo que prometo hacer algo, sin tener la menor intención de hacerlo o sin creerlo factible. En estos casos hay algo que va mal ciertamente, pero no es igual que un fallo. No diríamos que yo no prometí de hecho, sino más bien que prometí pero prometí insinceramente; te felicité pero las felicitaciones fueron huecas. Y puede haber un infortunio de una especie un tanto semejante cuando la emisión realizativa compromete al hablante a conducta futura de determinado tipo y luego en el futuro él no se comporta de hecho de la manera esperada. Esto es muy obvio, naturalmente, si prometo hacer algo y luego rompo mi promesa, pero hay muchos tipos de compromiso de una forma bastante menos tangible que la del caso de prometer. Por ejemplo, yo puedo decir «Te doy la bienvenida», dándote por bienvenido a mi casa o a donde fuere, pero luego empiezo a tratarte como si fueses extremadamente mal recibido. En este caso se ha abusado del procedimiento de decir «Te doy la bienvenida» de una manera un tanto diferente de la de simple insinceridad.

Podríamos preguntarnos ahora si esta lista de infortunios es completa, si los tipos de infortunio son mutuamente exclusivos, etc. Pues bien, no es completa, y no son mutuamente exclusivos; nunca lo son.

Supongamos que estás a punto de bautizar el barco, has sido nombrado para bautizarlo, y estás a punto de estallar la botella contra el casco; pero en ese mismo instante un tipo burdo aparece, te arrebató la botella de la mano, la rompe contra el casco, exclama «Bautizo este barco el *Generalísimo Stalin*, y luego retira las cuñas con buen compás. Pues bien, estamos de acuerdo naturalmente en varias cosas. Estamos de acuerdo en que el barco no se llama ahora el *Generalísimo Stalin*, y estamos de acuerdo en que se trata de una infernal vergüenza, etc., etc. Pero puede que no estemos de acuerdo en cuanto a cómo clasificar el infortunio concreto de este caso. Podríamos decir que aquí tenemos el caso de un procedimiento perfectamente legítimo y admitido que, sin embargo, ha sido invocado en circunstancias incorrectas, concretamente por la persona incorrecta, este tipo burdo en vez de la persona designada para hacerlo. Pero por otro lado podríamos verlo de manera diferente y decir que éste es un caso en que el procedimiento en su totalidad no se ha llevado a cabo correctamente, porque parte del procedimiento para bautizar un barco es que tú hayas sido designado como la persona que debe hacer el bautizo y esto es lo que este compadre no fue. Así pues, la forma en que debemos clasificar los infortunios en casos diferentes quizá sea un asunto bastante difícil, e incluso puede que en último extremo sea un tanto arbitrario. Pero por cierto los abogados, que tienen que vérselas muchísimo con este género de cosas, han inventado todos los tipos de términos técnicos y han formulado numerosas reglas sobre diferentes tipos de casos, que les permiten clasificar bastante rápidamente lo que en particular anda mal en cualquier caso dado.

En cuanto a si esta lista es completa, ciertamente no lo es. Una forma ulterior por la que las cosas pueden ir mal es, por ejemplo, por lo que en general puede llamarse malentendido. Puede que no oigas lo que digo, o puede que entiendas que me refiero a algo distinto de aquello a lo que pretendí referirme, etc. Y aparte de ulteriores adiciones que pudiéramos hacer a la lista, está la prevalente consideración general de que, en cuanto que estamos realizando un acto cuando proferimos estas emisiones realizativas, puede por cierto que lo estemos haciendo bajo coerción o en otras circunstancias que no nos hagan enteramente responsables de hacer lo que estamos haciendo. Esto sería ciertamente una infelicidad de una cierta especie —cualquier especie de no-responsabilidad podría llamarse una infelicidad; pero naturalmente es un tipo de cosa completamente diferente de aquellas de las que hemos estado hablando. Y podría mencionar que, de nuevo de manera muy diferente, podríamos proferir cualquiera de estas emisiones, tal como podemos proferir una emisión de cualquier tipo, en el transcurso, por ejemplo, de representar un papel o de hacer un chiste o escribir un poema—, en cuyo caso na-

turalmente no sería presentado seriamente y no podremos decir que realizamos seriamente el acto en cuestión. Si el poeta dice «Ve y recoge una estrella cadente» o lo que fuere, no profiere seriamente un orden. Consideraciones de este género se aplican absolutamente a cualquier emisión, no solamente a las realizativas.

Esto, pues, tal vez sea bastante para seguir con ello. Hemos discutido la emisión realizativa y sus infortunios. Ello nos equipa, podemos suponer, con dos nuevas y relucientes herramientas para resquebrajar acaso la cuna de la realidad. También nos equipa —siempre lo hace— con dos nuevas y relucientes zapatitas bajo nuestros pies metafísicos. La cuestión está en cómo las usamos.

## II

Hasta aquí hemos estado haciendo firmes progresos, sintiendo deslizarse bajo nuestros pies el firme piso del prejuicio, lo cual es siempre bastante regocijante, ¿pero qué pasa ahora? Ustedes estarán esperando el momento en que nos embarranquemos, el momento en que nos retractemos de todo, y con bastante seguridad que llegará pero tomará tiempo. Ante todo formulémosnos una pregunta bastante sencilla. ¿Cómo podemos estar seguros de, cómo podemos distinguir, si una emisión cualquiera ha de ser clasificada como realizativa o no? Seguramente, sentimos, debemos poder hacerlo. Y obviamente nos gustaría muchísimo poder decir que hay un criterio gramatical para ello, algún medio gramatical de decidir si una emisión es realizativa. Todos los ejemplos que he dado hasta aquí tienen de hecho la misma forma gramatical; todos ellos comienzan con el verbo en primera persona del singular del presente de indicativo de la voz activa —no precisamente cualquier tipo de verbo por cierto, pero con todo todos ellos son de hecho de esa forma—. Además, en el caso de los verbos que he usado hay una asimetría típica entre el uso de esta persona y tiempo del verbo y el uso del mismo verbo en otras personas y otros tiempos, y esta asimetría es una clave ciertamente importante.

Por ejemplo, cuando decimos «Yo prometo que...», el caso es muy diferente de cuando decimos «Él promete que...», o en tiempo pasado «Yo prometí que...». Pues cuando decimos «Yo prometo que...» realizamos un acto de prometer —hacemos una promesa—. Lo que *no* hacemos es informar de que alguien realiza un acto de prometer —en particular, no informamos de que alguien usa la expresión «Yo prometo»—. Efectivamente la usamos y hacemos la promesa. Pero si digo «Él promete», o en tiempo pasado «Yo prometí», informo precisamente de un acto de prometer, esto es, de un acto de usar la fórmula «Yo prometo» —informo de un acto presente de pro-



meter por su parte, o de un acto pasado por mi parte—. Hay así una clara diferencia entre la primera persona del singular del presente de indicativo de la activa, y las demás personas y tiempos. Esto es puesto de manifiesto con el típico caso del pequeño Guillermito, cuyo tío dice que le dará media corona si promete no fumar nunca hasta que tenga 55 años. El ansioso padre del pequeño Guillermito dice «Naturalmente que lo promete, ¿verdad, Guillermito?», dándole un codazo, y el pequeño Guillermito ni siquiera dice esta boca es mía. La gracia reside aquí en que él debe hacer la promesa por sí mismo diciendo «Lo prometo», y su padre va demasiado rápido al decir que promete.

Esto, pues, es un ejemplo de una prueba de si una emisión es realizativa o no, pero no debemos suponer que toda emisión realizativa haya de adoptar esta forma estándar. Hay al menos otra forma estándar, igual de común que ésta, en que el verbo está en la voz pasiva (o reflexiva) \* y en la segunda o tercera persona, no en la primera. El tipo de caso al que me refiero es el de un aviso que reza «Se advierte a los pasajeros de que crucen las vías por el puente solamente», o de un documento que dice «Por la presente está usted autorizado» a hacer tal y cual. Éstos son indudablemente realizativos, y de hecho a menudo se requiere una firma a fin de mostrar quién es el que está haciendo el acto de advertir, o autorizar, o lo que fuere. Muy típico de este tipo de realizativo —especialmente susceptible de figurar en documentos escritos naturalmente— es que la expresión «por la presente» \*\* o bien figura de hecho o podría con naturalidad insertarse.

Desgraciadamente, sin embargo, todavía no nos es posible sugerir que toda emisión que vaya a ser clasificada como un realizativo haya de adoptar una u otra de estas dos, por llamarlas así, formas estándares. Después de todo sería una emisión realizativa muy típica decir «Te ordeno que cierres la puerta». Satisface todos los criterios. Realiza el acto de ordenarte cerrar la puerta, y no es verdadera o falsa. Pero en las circunstancias apropiadas seguramente que pudimos realizar exactamente el mismo acto diciendo sencillamente «Cierra la puerta», en imperativo. O también, supongamos que alguien coloca un aviso «Este toro es peligroso», o simplemente «Toro peligroso», o simplemente «Toro». ¿Difiere esto necesariamente de colocar un aviso, apropiadamente firmado, diciendo «Por la presente se advierte que este toro es peligroso»? Parece que el simple aviso «Toro» puede realizar la misma función que la fórmula más elaborada. Naturalmente, la diferencia está en que si colocamos sólo «Toro»

\* Añado la puntualización del paréntesis porque en castellano, frente al inglés, resulta forzado el utilizar la voz pasiva y se emplea, en cambio, de modo natural, la construcción con *se*: «Se advierte a los pasajeros de que...» y no «Los pasajeros son advertidos de que...». (N. del T.)

\*\* En el original «*is that the little word "hereby"*». (N. del T.)

no quedaría del todo claro que es una advertencia; podría estar allí sólo por interés o información, como «Wallabi» ante la jaula del zoo o «Monumento antiguo». Sin duda que sabríamos por la naturaleza del caso que era una advertencia, pero no sería explícita.

Pues bien, en vista de este derrumbamiento de los criterios gramaticales, lo que nos gustaría suponer —y es bastante suponer— es que cualquier emisión que sea realizativa podría ser reducida a, o desarrollada en, una de estas dos formas estándares que comienzan con «Yo...» tal y cual o que comienzan con «Usted (o él) por la presente...» tal y cual. Si hubiese alguna justificación para esta esperanza, como en alguna medida la hay, entonces podríamos esperar hacer una lista de todos los verbos que pueden aparecer en estas formas estándares, y luego podríamos clasificar los tipos de actos que pueden realizarse con emisiones realizativas. Podríamos hacerlo con la ayuda de un diccionario, usando una prueba como la ya mencionada —el que haya la asimetría característica entre la primera persona del singular del presente de indicativo de la activa y las demás personas y tiempos— a fin de decidir si un verbo entra en nuestra lista o no. Pues bien, si hacemos esta lista encontramos de hecho que caen en determinadas clases bastante bien diferenciadas. Tenemos la clase en que emitimos veredictos y hacemos estimaciones y valoraciones de diversos géneros. Tenemos la clase en que hacemos compromisos, nos comprometemos de diversas formas al decir algo. Tenemos la clase en que al decir algo ejercitamos diversos derechos y poderes, tales como nombrar, votar, etc. Y tenemos una o dos más clases bastante bien diferenciadas.

Supongamos realizada esta tarea. Entonces llamaríamos a los verbos de nuestra lista verbos realizativos explícitos, y a cualquier emisión que se redujese a una u otra de nuestras formas estándares la llamaríamos una emisión realizativa explícita. «Te ordeno cerrar la puerta» sería una emisión realizativa explícita, mientras que «Cierra la puerta» no lo sería —esto es, una emisión realizativa «primaria» o como quiera que deseemos llamarla—. Al usar el imperativo puede que estemos ordenándote cerrar la puerta, pero no queda justamente claro si estamos ordenándote o rogándote o implorándote o solicitándote o incitándote o tentándote, o uno u otro de entre muchos otros actos sutilmente diferentes que, en un lenguaje primitivo poco sofisticado, es muy probable que no sean discriminados. Pero necesitamos no sobreestimar la no sofisticación de los lenguajes primitivos. Hay una gran cantidad de recursos utilizables para aclarar, incluso al nivel primitivo, qué acto estamos realizando cuando decimos algo —el tono de voz, la cadencia, los gestos— y ante todo podemos confiar en la naturaleza de las circunstancias, en el contexto en que la emisión es proferida. Esto muy a menudo hace totalmente inconfundible si se trata de una orden que se está dando o si, pongamos por

caso, estoy simplemente instigándote o rogándote. Puede que, por ejemplo, digamos algo de este tipo: «Viniendo de él yo estaba obligado a tomarlo como una orden.» No obstante, a pesar de todos estos recursos, hay una desafortunada cantidad de ambigüedad y falta de discriminación en ausencia de nuestros verbos realizativos explícitos. Si digo algo como «Estaré allí», puede que no quede determinado si es una promesa, o una expresión de intención, o tal vez incluso una predicción de mi conducta futura, de lo que va a sucederme; y puede que importe bastante, al menos en sociedades desarrolladas, cuál de estas cosas es precisamente. Y por esto es por lo que se ha desarrollado el verbo realizativo explícito —para aclarar exactamente cuál es, hasta qué punto me comprometo y de qué manera, etc.—.

Éste es justamente un modo en que el lenguaje se desarrolla a tono con la sociedad de la cual es el lenguaje. Los hábitos sociales de la sociedad pueden afectar considerablemente qué verbos realizativos se desarrollan y cuáles, a veces por razones más bien irrelevantes, no lo hacen. Por ejemplo, si digo «Eres un cobarde», pudiera ser que te estoy censurando o pudiera ser que te estoy insultando. Ahora bien, dado que al parecer la sociedad aprueba el censurar o reprobar, hemos, en este caso, desarrollado una fórmula «Te repruebo», o «Te censuro», que nos permite llevar a cabo expeditivamente esta deseable empresa. Pero por otro lado, dado que al parecer no aprobamos el insultar, no hemos desarrollado una fórmula simple del tipo «Te insulto», cosa que pudiéramos haber hecho exactamente igual.

Por medio de estos verbos realizativos explícitos y de algunos otros recursos, pues, explicitamos qué acto preciso es el que estamos realizando cuando proferimos nuestra emisión. Pero aquí me gustaría añadir unas palabras de advertencia. Debemos distinguir entre la función de explicitar qué acto es el que estamos realizando, y la muy diferente cuestión de *enunciar* qué acto es el que estamos realizando. Podemos dibujar un valioso paralelo aquí con otro caso en que el acto, el acto convencional que realizamos, no es un acto del habla sino una realización física. Supongamos que aparezco ante usted un día y doblo profundamente el espinazo. Bien, esto es ambiguo. Puedo estar simplemente observando la flora del lugar, atando el cordón de mi zapato, o algo de este tipo; por otro lado, es concebible que pudiera estar haciéndole una reverencia. Pues bien, para esclarecer esta ambigüedad tenemos algunos recursos tales como levantar el sombrero, decir «Salaam», o algo de este tipo, para dejar completamente claro que el acto que se realiza es el convencional de hacer una reverencia en vez de otro acto. Ahora bien, nadie querría decir que levantar el sombrero fue enunciar que usted estaba realizando un acto de reverencia; ciertamente no lo es, pero deja completamente claro que lo estaba haciendo. Y así de la misma manera decir «Te advierto que...» o «Te ordeno que...» o «Prometo que...» no es

enunciar que está usted haciendo algo, pero deja claro que lo está haciendo —constituye su realización verbal, una realización de un género particular—.

Hasta ahora hemos estado avanzando como si hubiese una diferencia totalmente clara entre nuestras emisiones realizativas y aquello con lo que las hemos contrastado, enunciados o informes o descripciones. Pero ahora comenzamos a encontrar que esta distinción no es tan clara como podría ser. Es ahora cuando empezamos a hundirnos un poco. En primer lugar, naturalmente, podemos sentir dudas en cuanto a cuán ampliamente se extienden nuestros realizativos. Si nos paramos a pensar en algunos tipos raros de expresión que usamos en casos raros, pudiéramos muy bien preguntarnos si satisfacen o no nuestros criterios un tanto vagos para ser emisiones realizativas. Supongamos, por ejemplo, que alguien dice «Bravo». Bien, no es verdadero o falso; está realizando el acto de vitorear. ¿La hace esto una emisión realizativa en nuestro sentido o no? O supongamos que él dice «Demonios»; está realizando el acto de jurar, y no es verdadera o falsa. ¿La hace esto realizativa? Sentimos que en cierto modo sí y sin embargo es bastante diferente. Por otro lado, consideremos casos de «dicho y hecho» \*; éstos también pueden hacer que nos preguntemos si tal vez la emisión debiera clasificarse como realizativa. O a veces, si alguien dice «Lo siento», nos preguntamos si esto es justamente lo mismo que «Le pido disculpas» —en cuyo caso, naturalmente, hemos dicho que es una emisión realizativa— o si tal vez ha de ser tomada como una descripción, verdadera o falsa, del estado de sus sentimientos. Si hubiese dicho «Me siento profundamente apenado por ello», entonces pensaríamos que debe entenderse como una descripción del estado de sus sentimientos. Si hubiese dicho «Le pido disculpas», sentiríamos que ésta es claramente una emisión realizativa, que lleva a cabo el ritual de pedir disculpas. Pero si dice «Lo siento» \*\* hay un aleteo entre las dos. Este fenómeno es muy común. Frecuentemente encontramos casos en que hay una obvia emisión realizativa pura y otras obvias emisiones conectadas con ella que no son realizativas sino descriptivas, pero por otro lado muchas en el medio de manera que no estamos enteramente seguros de lo que son. En algunas ocasiones por cierto son obviamente empleadas de una forma, en algunas ocasiones de otra, pero en algunas ocasiones parecen gozarse positivamente en la ambigüedad.

Por otro lado, consideremos el caso del árbitro cuando dice

\* En el original «cases of "suing the action to the word"». (N. del T.)

\*\* Quizá resulte un tanto forzado el ejemplo en castellano; en inglés no resulta forzado porque «*I am sorry*» tiene todas las trazas de una emisión que describe el estado de los sentimientos del que la profiere. (N. del T.)

«Fuera» o «Cambio», o la emisión del jurado cuando dicen que encuentran al detenido culpable. Naturalmente, decimos, éstos son casos de dar veredictos, de realizar el acto de valorar, etc., pero no obstante en cierto modo tienen alguna conexión con los hechos. Parecen tener algo parecido al deber de ser verdaderos o falsos, y parecen no estar por tanto muy alejados de los enunciados. Si el árbitro dice «Cambio», esto seguramente que tiene al menos algo que ver con el hecho de que se hayan lanzado seis pelotas más bien que siete, etc. De hecho podemos en general acordarnos de que «Enuncio que...» no parece muy diferente de «Te advierto que...» o «Prometo...». Aclara seguramente que el acto que estamos realizando es un acto de enunciar, y por tanto funciona exactamente como «Advier-to» u «Ordeno». Por tanto, ¿no es «Enuncio que...» una emisión realizativa? Pero entonces uno puede sentir que las emisiones que comienzan con «Enuncio que...» tienen que ser verdaderas o falsas, que *son* enunciados.

Consideraciones de esta especie, entonces, bien puede que nos hagan sentirnos bastante incómodos. Si volvemos la vista por un momento a nuestro contraste entre enunciados y emisiones realizativas, nos damos cuenta de que estábamos considerando los enunciados fiándonos demasiado en, como dijimos, el tratamiento tradicional. Teníamos que los enunciados habían de ser verdaderos o falsos; las emisiones realizativas por otra parte habían de ser afortunadas o desafortunadas. Consistían en hacer algo, mientras que con todo lo dicho el formular enunciados no era hacer algo. Ahora bien, seguro que este contraste es insatisfactorio si volvemos a examinarlo. Naturalmente los enunciados son susceptibles de ser evaluados en la cuestión de su correspondencia o falta de correspondencia con los hechos, es decir, de ser verdaderos o falsos. Pero son también susceptibles de infortunio al igual que lo son las emisiones realizativas. De hecho puede demostrarse que algunas dificultades que han surgido en el estudio de los enunciados recientemente son sencillamente dificultades de infortunio. Por ejemplo, se ha señalado que hay algo muy extraño en decir algo como esto: «El gato está sobre la alfombra pero yo no creo que esté». Pues bien, el decir esto es una cosa chocante, pero no es contradictorio. No hay ninguna razón por la que el gato no debiera estar sobre la alfombra sin que yo crea que lo está. Por tanto, ¿cómo hemos de clasificar lo que anda mal en este enunciado peculiar? Si recordamos ahora la doctrina del infortunio veremos que la persona que hace esta observación sobre el gato está en gran medida en la misma posición que alguien que diga algo como esto: «Prometo que estaré allí, pero no tengo la menor intención de estar allí.» Una vez más usted puede, por cierto, perfectamente bien prometer estar allí sin tener la menor intención de estar allí, pero hay algo chocante en decirlo, en confesar efectivamente la insinceridad de la promesa

que hace. De la misma manera hay insinceridad en el caso de la persona que dice «El gato está sobre la alfombra pero yo no creo que esté», y está confesando efectivamente esa insinceridad —lo cual forma un tipo peculiar de sinsentido—.

Un segundo caso que ha salido a discusión es el referente a los hijos de Juan —el caso en que se supone que alguien dice «Todos los hijos de Juan son calvos, pero Juan no ha tenido hijos»—. O quizás alguien dice «Todos los hijos de Juan son calvos», cuando de hecho —él no lo dice— Juan no tiene hijos. Pues bien, los que estudian enunciados se han hecho un lío con esto; ¿deben decir que el enunciado «Todos los hijos de Juan son calvos» es carente de significado en este caso? Bueno, si lo es, no se parece ni una pizca a muchísimos otros tipos más estándares de carencia de significado; y vemos, si nos volvemos a mirar nuestra lista de infortunios, que lo que anda mal aquí es en gran medida lo mismo que lo que anda mal en, por ejemplo, el caso de un contrato de venta de un trozo de tierra cuando el trozo de tierra al que se refiere no existe. Pues bien, lo que decimos en el caso de esta venta de tierra, que por cierto sería efectuada por una emisión realizativa, es que la venta es nula —nula por falta de referencia o ambigüedad de referencia—; y así podemos ver que el enunciado sobre todos los hijos de Juan es igualmente nulo por falta de referencia. Y si el hombre en cuestión dice efectivamente que Juan no tiene hijos de la misma sentada que dice que son todos calvos, está haciendo el mismo tipo de emisión chocante que el hombre que dice «El gato está sobre la alfombra y yo no creo que esté», o el hombre que dice «Prometo pero no tengo la intención de hacerlo».

De esta manera, pues, los males que se ha encontrado que afectan a los enunciados pueden ser puestos precisamente en paralelo con los males que son característicos de las emisiones realizativas. Y después de todo, cuando enunciamos algo o describimos algo o informamos de algo, realizamos un acto que es con igual derecho un acto que el acto de ordenar o de advertir. No parece que haya ninguna buena razón por la que debemos darle al acto de enunciar una posición especialmente única. Ciertamente los filósofos han estado acostumbrados a hablar como si usted o yo o cualquiera pudiera justamente ponerse a enunciar cualquier cosa sobre cualquier cosa y esto estuviese perfectamente en orden, sólo que hay justamente una pequeña cuestión: ¿es verdadero o falso? Pero además de la pequeña cuestión, es verdadero o falso, hay seguramente la cuestión: ¿está en orden? ¿Puede usted ponerse precisamente a hacer enunciados sobre cualquier cosa? Supongamos, por ejemplo, que usted me dice «Esta mañana tengo la sensación de estar muy anticuado». Bien, yo le digo «No la tiene»; y usted dice «¿Qué demonios quiere decir, no la tengo?». Yo digo «Oh, nada —sólo estoy enunciando que no la tiene—, ¿es verdadero o falso?». Y usted dice «Espere un momento

sobre si es verdadero o falso, la cuestión es ¿qué quiere usted decir al hacer enunciados sobre los sentimientos de otro? Yo le dije que tengo la sensación de estar muy anticuado. Usted no está precisamente en posición de decir, de enunciar, que no la tengo.» Esto pone de manifiesto que usted no puede justamente hacer enunciados sobre los sentimientos de otras personas (aunque puede hacer conjeturas si gusta); y hay muchas cosas que, no teniendo conocimiento de ellas, no estando en posición de pronunciarse sobre ellas, usted no puede justamente enunciar. Lo que necesitamos hacer con el caso de enunciar, y por la misma regla de tres describir e informar, es bajarlos un poco de su pedestal, darnos cuenta de que son actos del habla no menos que todos esos otros actos del habla que hemos estado mencionando y discutiendo como realizativos.

Entonces consideremos por un momento nuestro contraste original entre el realizativo y el enunciado desde el otro ángulo. Al manejar realizativos hemos estado expresándonos todo el tiempo como si la única cosa que una emisión realizativa tuviera que hacer fuera ser afortunada, salir, no ser un fallo, no ser un abuso. Sí, pero esto no es el final de la cuestión. Al menos en el caso de muchas emisiones que, basándonos en lo que hemos dicho, tendríamos que clasificar como realizativas —casos en que decimos «Te advierto...», «Te aconsejo...», etc.— habrá otras cuestiones además de simplemente: ¿estaba en orden?, ¿fue correcto?, ¿salió? Después de esto seguro que vendrá la cuestión: ¿fue un consejo bueno o sensato? ¿Fue una advertencia justificada? O en el caso, digamos, de un veredicto o una estimación: ¿fue una buena estimación, o un justo veredicto? Y éstas son cuestiones que sólo pueden decidirse considerando cómo se relaciona el contenido del veredicto o de la estimación realizativas en una dimensión general de correspondencia con el hecho. Puede que aún se diga, por cierto, que esto no las hace *muy* parecidas a los enunciados porque todavía no son verdaderas o falsas, y que ésta es una pequeña particularidad en blanco y negro que distingue a los enunciados como una clase aparte. Pero de hecho —aunque nos llevaría mucho proseguir con esto—, cuanto más se piensa en la verdad y la falsedad más se encuentra que muy pocos enunciados de los que emitimos son justamente verdaderos o justamente falsos. Usualmente se plantea la cuestión de si son justos o injustos, de si son adecuados o inadecuados, de si son exagerados o no exagerados. ¿Son demasiado toscos, o son perfectamente precisos, exactos, etc.? «Verdadero» y «falso» son justamente etiquetas generales para toda una dimensión de diferentes valoraciones que tienen una cosa u otra que ver con la relación entre lo que decimos y los hechos. Si, entonces, relajamos nuestras ideas de verdad y falsedad veremos que los enunciados, cuando son evaluados en relación a los hechos, no son tan distintos después de todo de consejos, advertencias, veredictos, etc.

Vemos entonces que enunciar algo es realizar un acto justamente igual que lo es dar una orden o hacer una advertencia; y vemos, por otro lado, que, cuando damos una orden o hacemos una advertencia o damos un consejo, hay la cuestión de cómo esto está relacionado con los hechos, la cual no es quizá muy distinta del tipo de cuestión que surge cuando discutimos cómo está un enunciado relacionado con el hecho. Bien, esto parece significar que en su forma original nuestra distinción entre el realizativo y el enunciado se debilita considerablemente, y en realidad se derrumba. Haré sólo una sugerencia respecto a cómo tratar este asunto. Necesitamos retroceder mucho para considerar todas las maneras y sentidos en que decir algo es hacer esto o aquello —porque por cierto es siempre hacer una buena cantidad de cosas diferentes—. Y una cosa que sale a la luz cuando hacemos esto es que, además de la cuestión que ha sido muy estudiada en el pasado concerniente a lo que una determinada emisión *significa*, hay una cuestión ulterior distinta de ésta concerniente a cuál era la *fuerza*, por así llamarla, de la emisión. Puede que nos resulte totalmente claro lo que «Cierra la puerta» significa, pero no tengamos claro todavía en absoluto el punto ulterior concerniente a si en cuanto emitida en un momento determinado era una orden, un ruego o cosas por el estilo. Lo que necesitamos además de la vieja doctrina sobre los significados es una nueva doctrina sobre todas las posibles fuerzas de emisiones, y nuestra lista propuesta de verbos realizativos explícitos constituiría una gran ayuda para su descubrimiento; y luego, partiendo de allí, una investigación de los diversos términos de valoración que usamos al discutir actos del habla de este, de ese o de aquel tipo preciso —órdenes, advertencias y demás—.

Las nociones que hemos, pues, considerado son el realizativo, el infortunio, el realizativo explícito, y finalmente, un tanto apresuradamente, la noción de las fuerzas de emisiones. No sé si decir que todo esto parece un poco estéril, un poco complicado. Bueno, supongo que en algunos aspectos es estéril, y supongo que debe ser beneficioso. Al menos, sin embargo, creo que si prestamos atención a estas cuestiones podemos aclarar algunas confusiones que se dan en filosofía; y después de todo la filosofía se usa como un chivo expiatorio, exhibe confusiones que son realmente las confusiones de todo el mundo. Podríamos incluso aclarar algunas confusiones que se dan en gramática, que tal vez sea un poco más respetable.

¿Y es complicado? Bueno, es un poco complicado; pero la vida y la verdad y las cosas tienden a ser complicadas. No son las cosas, son los filósofos los que son simples. Supongo que habrán oído decir que la simplificación excesiva es la enfermedad profesional de los filósofos, y en cierto modo se podría estar de acuerdo con ello. Pero por una secreta sospecha de que es su profesión.



## ¿QUÉ ES UN ACTO DE HABLA? \*

JOHN R. SEARLE

### I. INTRODUCCIÓN

En una situación de habla típica que incluye un hablante, un oyente y una emisión del hablante, existen muchos géneros de actos asociados con la emisión del hablante. El hablante habrá movido característicamente su mandíbula y su lengua y habrá producido ruidos. Además, habrá realizado característicamente algunos actos pertenecientes a la clase que incluye informar o irritar o aburrir a sus oyentes; habrá realizado también característicamente algunos actos pertenecientes a la clase que incluye referirse a Kennedy o a Jruschov o al polo Norte; y habrá realizado asimismo actos pertenecientes a la clase que incluye hacer enunciados, plantear preguntas, dar órdenes, emitir informes, saludar y aconsejar. Los miembros de esta última clase son los que Austin<sup>1</sup> llamó actos ilocucionarios, y es de esta clase de la que me ocuparé en este artículo; por consiguiente el artículo podría haberse titulado «¿Qué es un Acto Ilocucionario?». No intento definir la expresión «acto ilocucionario», aunque si mi análisis de un acto ilocucionario particular tiene éxito puede proporcionar las bases para una definición. Algunos de los verbos castellanos y frases verbales asociadas con actos ilocucionarios son: enunciar, aseverar, describir, aconsejar, observar, comentar, mandar, ordenar, suplicar, criticar, pedir disculpas, censurar, aprobar, dar la bienvenida, prometer, dar consentimiento y pedir perdón. Austin afirmaba que existían en inglés más de un millar de expresiones semejantes.

A modo de introducción tal vez pueda decir por qué pienso que tiene interés y es importante en filosofía del lenguaje estudiar los actos de habla, o, como se les denomina algunas veces, actos de lenguaje o actos lingüísticos. Creo que es esencial a cualquier espécimen de comunicación lingüística que incluya un acto lingüístico. La unidad de comunicación lingüística no es, como generalmente se ha supuesto, ni el símbolo ni la palabra ni la oración, ni tan siquiera la instancia del símbolo, palabra u oración, sino más bien lo que constituye

\* Versión castellana de Luis M. Valdés Villanueva.

<sup>1</sup> J. L. AUSTIN, *How to do Things with Words*, Oxford, 1962. Versión castellana de G. R. Carrió y E. A. Rabossi, *Palabras y Acciones*, Paidós, Buenos Aires, 1971.

la unidad básica de la comunicación lingüística es la *producción* de la instancia en la realización del acto de habla. Para establecer más precisamente este punto: la producción de la oración-instancia bajo ciertas condiciones es el acto ilocucionario, y el acto ilocucionario es la unidad mínima de la comunicación lingüística.

No sé cómo *demostrar* que la comunicación lingüística incluye esencialmente actos, pero puedo pensar en argumentos con los cuales se podría intentar convencer a alguien que fuese escéptico. Un argumento consistiría en llamar la atención del escéptico sobre el hecho de que cuando él considera que un ruido o una marca sobre un papel es un caso de comunicación lingüística, como un mensaje, una de las cosas que se incluyen en su considerar así ese ruido o marca es que debe contemplarlo como habiendo sido producido por un ser con ciertas intenciones. No puede contemplarlo como un fenómeno natural, igual que una piedra, una cascada, o un árbol. Para contemplarlo como un caso de comunicación lingüística debe suponer que su producción es lo que yo estoy denominando un acto de habla. Por ejemplo, una presuposición lógica de los intentos corrientes de descifrar los jeroglíficos mayas consiste en que al menos avanzamos la hipótesis de que las marcas que vemos sobre las piedras fueron producidas por seres más o menos parecidos a nosotros mismos y producidas con ciertos géneros de intenciones. Si estuviéramos seguros de que las marcas eran una consecuencia de, digamos, erosión producida por el agua, entonces la cuestión de descifrarlas o incluso de denominarlas jeroglíficos no podría plantearse. Interpretarlas bajo la categoría de comunicación lingüística incluye necesariamente interpretar su producción como actos de habla.

Realizar un acto ilocucionario es tomar parte en una forma de conducta gobernada por reglas. Argüiré que cosas tales como plantear preguntas o hacer enunciados están gobernadas por reglas de maneras completamente semejantes a aquéllas en las que marcar un gol en fútbol o mover un caballo en el ajedrez son formas de actos gobernados por reglas. Por lo tanto, intento explicar la noción de acto ilocucionario enunciando un conjunto de condiciones necesarias y suficientes para la realización de un género particular de acto ilocucionario, y extraer de él un conjunto de reglas semánticas para el uso de la expresión (o dispositivo sintáctico) que distingue la emisión como un acto ilocucionario de ese género. Si tengo éxito al enunciar las condiciones y las reglas correspondientes para tan siquiera un género de acto ilocucionario, esto nos proporcionará un modelo para analizar otros géneros de actos, y consecuentemente para explicar la noción en general. Pero en orden a establecer la plataforma para enunciar efectivamente las condiciones y extraer las reglas para realizar un acto ilocucionario, tengo que discutir otras tres nociones preliminares: *reglas, proposiciones y significado*. Limitaré mi discusión de

estas nociones a aquellos aspectos que son esenciales para mis propósitos principales en este artículo, pero, incluso así, lo que deseo decir respecto de cada una de esas nociones, si hubiera de ser completo, requeriría un artículo para cada una de ellas; sin embargo, algunas veces puede que valga la pena sacrificar la minuciosidad en favor del alcance y, por lo tanto, seré muy breve.

## II. REGLAS

En los años recientes se han producido considerables discusiones en la filosofía del lenguaje respecto a la noción de reglas para el uso de expresiones. Algunos filósofos han dicho incluso que conocer el significado de una palabra es simplemente un asunto de conocer las reglas para su uso o empleo. Una característica inquietante de tales discusiones la constituye el hecho de que ningún filósofo, al menos que yo sepa, ha dado jamás algo parecido a una formulación adecuada de las reglas para el uso de tan siquiera una expresión. Si el significado es un asunto de reglas de uso, seguramente deberíamos ser capaces de enunciar las reglas para el uso de expresiones de una manera que explicase el significado de esas expresiones. Otros ciertos filósofos, desanimados quizás por el fracaso de sus colegas en proporcionar regla alguna han negado el punto de vista de moda de que el significado es un asunto de reglas y han aseverado que no existen en absoluto reglas semánticas del género propuesto. Me inclino a pensar que este escepticismo es prematuro y resulta del fracaso en distinguir diferentes géneros de reglas, en un sentido que ahora intentaré explicar.

Distingo entre dos clases de reglas: algunas regulan formas de conducta existentes anteriormente; por ejemplo, las reglas de etiqueta regulan relaciones interpersonales, pero esas relaciones existen independientemente de las reglas de etiqueta. Algunas reglas, por otra parte, no regulan meramente, sino que crean o definen nuevas formas de conducta. Las reglas del fútbol, por ejemplo, no regulan meramente el juego del fútbol, sino que, por así decirlo, crean la posibilidad de, o definen, esa actividad. La actividad de jugar al fútbol se constituye actuando de acuerdo con esas reglas; el fútbol no tiene existencia aparte de esas reglas. Llamaré al último género de reglas, reglas constitutivas, y al primer género reglas regulativas. Las reglas regulativas regulan una actividad preexistente, una actividad cuya existencia es lógicamente independiente de la existencia de las reglas. Las reglas constitutivas constituyen (y también regulan) una actividad cuya existencia es lógicamente dependiente de las reglas<sup>2</sup>.

<sup>2</sup> Esta distinción aparece en J. Rawls, «Two Concepts of Rules», *The Philoso-*

Característicamente las reglas regulativas toman la forma de imperativos, p. ej., «Cuando cortes los alimentos mantén agarrado el cuchillo con la mano derecha» o «Los oficiales han de llevar corbata en la cena». Algunas reglas constitutivas toman una forma completamente diferente, p. ej., un jaque mate se hace si el rey es atacado de tal manera que ningún movimiento lo dejará inatacado; un *touchdown* se marca cuando un jugador cruza la línea de meta del contrario en posesión de la pelota mientras el juego está en marcha. Si nuestros paradigmas de reglas son reglas regulativas imperativas, tales reglas constitutivas no imperativas han de sorprendernos probablemente como extremadamente curiosas e incluso difícilmente como reglas en absoluto. Obsérvese que tienen casi carácter tautológico, puesto que lo que la «regla» parece ofrecer es una definición parcial de «jaque mate» o «*touchdown*». Pero, naturalmente, este carácter cuasi tautológico es una consecuencia necesaria del hecho de que sean reglas constitutivas: las reglas concernientes a *touchdowns* deben definir la noción de «*touchdown*» del mismo modo que las reglas que conciernen al fútbol definen «fútbol». Que, por ejemplo, un *touchdown* pueda ser marcado de tales y tales maneras y cuente como seis puntos, puede aparecer algunas veces como una regla, otras veces como una verdad analítica; y que pueda ser interpretada como una tautología es una pista para el hecho de que la regla en cuestión es constitutiva. Las reglas regulativas tienen generalmente la forma «Haz *X*» o «Si *Y* haz *X*». Algunos miembros del conjunto de reglas constitutivas tienen esta forma, pero otros tienen también la forma «*X* cuenta como *Y*»<sup>3</sup>.

El fracaso en percibir esto tiene alguna importancia en filosofía. Así, p. ej., algunos filósofos preguntan «¿Cómo puede una promesa crear una obligación?». Una pregunta similar sería «¿Cómo puede un *touchdown* crear seis puntos?». Y tal como están planteadas ambas preguntas solamente pueden responderse enunciando una regla de la forma «*X* cuenta como *Y*».

Me inclino a pensar que tanto el fracaso de algunos filósofos en enunciar reglas para el uso de expresiones como el escepticismo de otros filósofos concerniente a la existencia de reglas tales surge, al menos en parte, del fracaso en reconocer la distinción entre reglas regulativas y constitutivas. El modelo o paradigma de una regla que muchos filósofos tienen es el de una regla regulativa, y si se buscan en semántica reglas puramente regulativas seguramente no se encontrará nada interesante desde el punto de vista del análisis lógico. Existen

*philical Review*, 1955, y J. R. Searle, «How to Derive "Ought" from "Is"», *The Philosophical Review*, 1964.

<sup>3</sup> La formulación de «*X* cuenta como *Y*» me fue originalmente sugerida por Max Black.

sin duda reglas sociales de la forma «No se deben decir obscenidades en las reuniones formales», pero esto difícilmente parece una regla de la clase que es crucial en la explicación de la semántica de un lenguaje. La hipótesis subyacente al presente artículo consiste en que la semántica de un lenguaje puede ser contemplada como una serie de sistemas de reglas constitutivas, y que los actos ilocucionarios son actos realizados de acuerdo con esos conjuntos de reglas constitutivas. Uno de los propósitos de este artículo es formular un conjunto de reglas constitutivas para un cierto género de acto de habla. Y si lo que he dicho respecto de las reglas constitutivas es correcto, no deberíamos sorprendernos si no todas esas reglas toman la forma de reglas imperativas. De hecho, veremos que las reglas pertenecen a diversas y distintas categorías, ninguna de las cuales es completamente igual a las reglas de etiqueta. El esfuerzo para enunciar las reglas de un acto ilocucionario puede también contemplarse como un género de prueba de la hipótesis de que existen reglas constitutivas subyacentes a los actos de habla. Si no somos capaces de dar algunas formulaciones satisfactorias de reglas, nuestro fracaso podría interpretarse como evidencia parcialmente disconforme en contra de la hipótesis.

### III. PROPOSICIONES

Diferentes actos ilocucionarios tienen a menudo características en común unos con otros. Considérense las emisiones de las oraciones siguientes:

- (1) ¿Saldrá Juan de la habitación?
- (2) Juan saldrá de la habitación.
- (3) ¡Juan, sal de la habitación!
- (4) Ojalá Juan saliese de la habitación.
- (5) Si Juan saliese de la habitación, yo también saldría.

Las emisiones de cada una de estas oraciones en una ocasión dada constituirán característicamente realizaciones de diferentes actos ilocucionarios. La primera sería, característicamente, una pregunta, la segunda una aserción sobre el futuro, esto es, una predicción, la tercera una petición o una orden, la cuarta la expresión de un deseo, y la quinta una expresión hipotética de intención. Con todo en la realización de cada uno de ellos el hablante realizará algunos actos subsidarios que son comunes a todos los cinco actos ilocucionarios. En la emisión de cada uno de ellos el hablante *se refiere* a una persona particular, Juan, y *predica* el acto de salir de la habitación de esa persona. En ningún caso es esto todo lo que él hace, pero en todo caso es una parte de lo que hace. Diré, por lo tanto, que en cada uno de esos

casos, aunque los actos ilocucionarios sean diferentes, al menos alguno de los actos no-ilocucionarios de referencia y predicación son el mismo.

La referencia a alguna persona, Juan, y la predicación de la misma cosa de él en cada uno de esos actos ilocucionarios me inclina a decir que hay un *contenido* común en cada uno de ellos. Algo expresable por la cláusula «que Juan saldrá de la habitación» parece ser una característica común de todos ellos. Podríamos, sin demasiada distorsión, escribir cada una de esas oraciones de una manera que aislase esa característica común: «Asevero que Juan saldrá de la habitación», «Pregunto si Juan saldrá de la habitación», etc.

A falta de una palabra mejor propongo llamar a este contenido común una proposición, y describiré esta característica de esos actos ilocucionarios diciendo que en la emisión de cada uno de (1)-(5) el hablante expresa la proposición de que Juan saldrá de la habitación. Obsérvese que no digo que la oración expresa la proposición; no sé cómo podrían las oraciones realizar actos de este género. Pero diré que en la emisión de la oración el hablante expresa una proposición. Obsérvese también que estoy distinguiendo entre una proposición y una aserción o enunciado de esa proposición. La proposición de que Juan saldrá de la habitación se expresa en la emisión de todas las oraciones (1)-(5), pero solamente en (2) se asevera esa proposición. Una aserción es un acto ilocucionario, pero una proposición no es en absoluto un acto, aunque el acto de expresar una proposición sea una parte de la realización de ciertos actos ilocucionarios.

Podría resumirse esto diciendo que distingo entre el acto ilocucionario y el contenido proposicional de un acto ilocucionario. Naturalmente, no todos los actos ilocucionarios tienen un contenido proposicional, por ejemplo una emisión de «¡Hurra!» o «¡Ay!» no lo tienen. En una u otra versión esta distinción es antigua y ha sido señalada por autores tan diversos como Frege, Sheffer, Lewis, Reichenbach y Hare, por mencionar solamente unos pocos.

Desde un punto de vista semántico podemos distinguir en la oración entre el indicador proposicional y el indicador de fuerza ilocucionaria. Esto es, para una extensa clase de oraciones usadas para realizar actos ilocucionarios, podemos decir para los propósitos de nuestro análisis que la oración tiene dos partes (no necesariamente separadas), el elemento indicador de la proposición y el dispositivo indicador de la función<sup>4</sup>. El dispositivo indicador de la función muestra cómo debe ser tomada la proposición, o, dicho de otra ma-

<sup>4</sup> En la oración «Prometo que vendré» el dispositivo indicador de función y el elemento proposicional están separados. En la oración «Prometo venir», que significa lo mismo que la primera y se deriva de ella mediante ciertas transformaciones, los dos elementos no están separados.

nera, qué fuerza ilocucionaria ha de tener, esto es, qué acto ilocucionario está realizando el hablante al emitir la oración. Los dispositivos indicadores de función incluyen en castellano el orden de las palabras, el énfasis, la entonación, la puntuación, el modo del verbo y finalmente un conjunto de los llamados verbos realizativos: puedo indicar el género de acto ilocucionario que estoy realizando comenzando la oración con «Pido disculpas», «Aconsejo», «Enuncio», etc. A menudo en las situaciones efectivas de habla el contexto clarificará cuál es la fuerza ilocucionaria de la emisión, sin que sea necesario apelar al dispositivo indicador de función apropiado.

Si esta distinción semántica tiene alguna importancia real, parece verosímil que deba haber alguna analogía sintáctica, y ciertos desarrollos recientes de la gramática transformacional tienden a respaldar este punto de vista. En el marcador de frase subyacente de una oración existe una distinción entre aquellos elementos que corresponden al dispositivo indicador de la función y aquéllos que corresponden al contenido proposicional.

La distinción entre el dispositivo indicador de función y el dispositivo indicador de la proposición nos resultará muy útil a la hora de proporcionar un análisis de un acto ilocucionario. Puesto que la misma proposición puede ser común a todas las clases de actos ilocucionarios, podemos separar nuestro análisis de la proposición de nuestro análisis de géneros de actos ilocucionarios. Creo que existen reglas para expresar proposiciones, reglas para cosas tales como referencia y predicación, pero estas reglas pueden ser discutidas independientemente de las reglas indicadoras de función. En este artículo no intentaré discutir las reglas proposicionales sino que me concentraré en las reglas para el uso de ciertas clases de dispositivos indicadores de función.

#### IV. SIGNIFICADO

Los actos de habla se realizan característicamente en la emisión de sonidos o en la producción de marcas. ¿Cuál es la diferencia entre *solamente* emitir sonidos o hacer marcas y realizar un acto de habla? Una diferencia consiste en que de los sonidos o marcas que una persona hace en la realización de un acto de habla se dice característicamente que *tienen significado*, y una segunda diferencia relacionada consiste en que característicamente se dice que una persona *quiere decir algo* mediante esos sonidos o marcas. Característicamente cuando se habla se quiere decir algo mediante lo que se dice, y de lo que se dice, de la sarta de morfemas que se emite, se dice característicamente que tienen un significado. Incidentalmente, hay aquí otro punto en el cual nuestra analogía entre realizar actos de habla y jugar se derrumba. De las piezas de un juego como el ajedrez no se dice

característicamente que tengan significado, y además, cuando se hace una jugada no se dice característicamente que se quiere decir algo mediante esa jugada.

Pero ¿qué es para alguien querer decir algo mediante lo que dice, y qué es para algo tener un significado? Para responder a la primera de estas preguntas propongo tomar prestadas y revisar algunas de las ideas de Paul Grice. En un artículo titulado «Meaning»<sup>5</sup>, Grice da el siguiente análisis de un sentido de la noción de «significado». Decir que *A* quiere decir algo mediante *x* es decir que «*A* intentó que la emisión de *x* produjese algún efecto en un auditorio por medio del reconocimiento de esta intención». Esto me parece un útil punto de partida para un análisis del significado, en primer lugar porque muestra la estrecha relación entre la noción de significado y la noción de intención, y en segundo lugar porque captura algo que es, pienso, esencial a hablar un lenguaje: al hablar un lenguaje intento comunicar cosas a mi oyente consiguiendo que él reconozca mi intención de comunicar precisamente esas cosas. Por ejemplo, característicamente, cuando hago una aserción, intento comunicar a, y convencer a mi oyente, de la verdad de cierta proposición; y los medios que empleo para hacer esto son emitir ciertos sonidos, cuya emisión intento que produzca en él el efecto deseado por medio de su reconocimiento de mi intención de producir precisamente ese efecto. Por una parte podría intentar lograr que creyeses que soy francés hablando francés durante todo el tiempo, vistiéndome a la manera francesa, mostrando un frenético entusiasmo por De Gaulle y cultivando amistades francesas. Pero por otra parte podría lograr que creyeses que soy francés diciéndote simplemente que soy francés. Ahora bien, ¿cuál es la diferencia entre esas dos maneras de mi intento de lograr que creas que soy francés? Una diferencia crucial es que en el segundo caso intento lograr que creas que soy francés llevándote a reconocer que mi intención pretende lograr que creas precisamente eso. Esta es una de las cosas que incluye el decirte que soy francés. Pero naturalmente, si intento lograr que creas que soy francés actuando de la manera que he descrito, entonces tu reconocimiento de mi intención de producir en ti la creencia de que soy francés no está constituida por los medios que estoy empleando. En efecto, en este caso, pienso, te volverías más bien receloso si reconocieses mi intención.

A pesar de que este análisis del significado es valioso, me parece que es defectuoso en ciertos aspectos. Primeramente no logra distinguir entre los diferentes géneros de efectos —perlocucionarios *versus* ilocucionarios— que uno puede intentar producir en sus oyentes, y además no logra mostrar la manera en que esos diferentes géneros de

<sup>5</sup> *The Philosophical Review*, 1957. [Versión castellana: «Significado», *Cuadernos de Crítica*, México, 1977. (N. del T.)]



efectos se relacionan con la noción de significado. Un segundo defecto es que no logra dar cuenta de hasta qué punto el significado es un asunto de reglas o convenciones. Esto es, esta explicación del significado no muestra la conexión entre el querer decir algo por parte de alguien mediante lo que dice y lo que eso que alguien dice significa efectivamente en el lenguaje. Para ilustrar este punto quiero presentar ahora un contraejemplo a este análisis del significado. El objeto del contraejemplo será ilustrar la conexión entre lo que un hablante quiere decir y lo que las palabras que él emite significan.

Supongamos que yo soy un soldado americano de la Segunda Guerra Mundial y que soy capturado por las tropas italianas. Y supongamos también que deseo inducir a esas tropas a creer que soy un oficial alemán de modo que consiga que me suelten. Lo que me gustaría hacer sería decirles en alemán o italiano que soy un oficial alemán. Pero supongamos que no sé alemán o italiano suficientes para hacer esto. Entonces yo, por así decirlo, intento representar la pantomima de decirles que yo soy un oficial alemán recitándoles aquellos pequeños trozos de alemán que conozco, confiando en que ellos no saben suficiente alemán como para darse cuenta de mi plan. Supongamos que conozco solamente una línea de alemán que recuerdo de un poema que tenía que memorizar en un curso de alemán de la escuela secundaria. Por lo tanto yo, un prisionero americano, me dirijo a mis capturadores italianos con la siguiente oración: «*Kennst du das Land wo die Zitronen blühen?*» Ahora bien, describamos la situación en términos griceanos. Yo intento producir un cierto efecto en ellos, a saber, el efecto de que crean que yo soy un soldado alemán; e intento producir este efecto por medio de su reconocimiento de mi intención. Intento que ellos piensen que lo que estoy intentando decirles es que yo soy un oficial alemán. Pero ¿se sigue de esta explicación que cuando digo «*Kennts du das Land...*», etc., lo que quiero decir (*mean*) es «Soy un oficial alemán»? No solamente no se sigue sino que en este caso me parece completamente falso que cuando emito la oración alemana lo que quiero decir es «Yo soy un oficial alemán», o incluso «*Ich bin ein deutscher Offizier*», puesto que lo que las palabras significan (*mean*) es «¿Conoces el país donde florecen los limoneros?». Naturalmente, deseo engañar a mis capturadores de modo que piensen que lo que quiero decir es «Yo soy un oficial alemán», pero parte de lo que se incluye en el engaño es lograr que piensen que esto es lo que las palabras que emito significan en alemán. En un punto de las *Philosophical Investigations* Wittgenstein dice «Di “hace frío aquí” queriendo decir “hace calor aquí”»<sup>6</sup>. La razón por la que no somos capaces de hacer esto es que lo que noso-

<sup>6</sup> *Philosophical Investigations*, Oxford, 1953, párrafo 510 [versión castellana, UNAM/Crítica, Barcelona, 1988].

tros podemos querer decir es una función de lo que estamos diciendo. El significado es más que un asunto de intención, es un asunto de convención.

La explicación de Grice puede rectificarse para tratar con contraejemplos de este género. Tenemos un caso donde estoy intentando producir un cierto efecto por medio del reconocimiento de mi intención de producir ese efecto, pero el dispositivo que uso para producir ese efecto se usa convencionalmente, en virtud de las reglas que gobiernan el uso de ese dispositivo, como un medio de producir efectos ilocucionarios completamente diferentes. Debemos por lo tanto reformular la explicación del significado de Grice de una manera tal que haga claro que el querer decir algo por parte de una persona cuando esta dice algo está relacionado más que contingentemente con lo que la oración significa en el lenguaje que esa persona está hablando. En nuestro análisis de los actos ilocucionarios debemos capturar tanto los aspectos convencionales como los intencionales y especialmente las relaciones entre ellos. En la realización de un acto ilocucionario el hablante intenta producir un cierto efecto, logrando que el oyente reconozca su intención de producir ese efecto, y además, si está usando las palabras literalmente, intenta que este reconocimiento se logre en virtud del hecho de que las reglas para el uso de las expresiones que emite asocian las expresiones con la producción de ese efecto. Es esta *combinación* de elementos la que necesitamos expresar en nuestro análisis del acto ilocucionario.

## V. CÓMO PROMETER

Intentaré ahora ofrecer un análisis del acto ilocucionario de prometer. Para llevar a cabo esto preguntaré qué condiciones son necesarias y suficientes para que se haya realizado el acto de prometer en la emisión de una oración dada. Intentaré dar respuesta a esta pregunta enunciando esas condiciones como un conjunto de proposiciones tales que la conjunción de los miembros del conjunto entraña la proposición de que un hablante hizo una promesa, y la proposición de que el hablante hizo una promesa entraña esta conjunción. De esta manera cada condición será una condición necesaria para la realización del acto de prometer, y el conjunto de condiciones tomado colectivamente será una condición suficiente para que el acto haya sido realizado.

Si obtenemos un conjunto tal de condiciones podemos extraer de ellas un conjunto de reglas para el uso del dispositivo indicador de función. Aquí el método es análogo a descubrir las reglas del ajedrez preguntándonos cuáles son las condiciones necesarias y suficientes bajo las cuales puede decirse que un jugador ha movido correcta-

mente un caballo, se ha enrocado, dado jaque mate, etc. Estamos en la posición de alguien que ha aprendido a jugar al ajedrez sin haber tenido jamás formuladas las reglas y que desea tal formulación. Nosotros aprendemos cómo jugar el juego de los actos ilocucionarios, pero en general esto se hace sin una formulación explícita de las reglas, y el primer paso para obtener tal formulación consiste en establecer las condiciones para la realización de un acto ilocucionario particular. Por lo tanto, nuestra investigación servirá a un doble propósito filosófico. Al enunciar un conjunto de condiciones para la realización de un acto ilocucionario particular, habremos ofrecido una explicación parcial de esa noción y también habremos preparado el terreno para el segundo paso, la formulación de las reglas.

Encuentro que el enunciado de las condiciones es muy difícil de hacer, y no estoy enteramente satisfecho con la lista que voy a presentar. Una razón de la dificultad es que la noción de promesa, al igual que muchas nociones del lenguaje ordinario, carece absolutamente de reglas estrictas. Existen toda clase de promesas extrañas, divergentes y dudosas; y pueden presentarse contraejemplos, más o menos bizarros, en contra de mi análisis. Me inclino a pensar que no seremos capaces de obtener un conjunto abrumador de condiciones necesarias y suficientes que reflejen exactamente el uso ordinario de la palabra «promesa». Por lo tanto, limitaré mi discusión al meollo del concepto de prometer e ignoraré los casos límite, dudosos y parcialmente defectuosos. Limito también mi discusión a las promesas abiertas y explícitas e ignoro las promesas hechas mediante giros de frase elípticos, insinuaciones, metáforas, etc.

Otra dificultad surge de mi deseo de enunciar las condiciones sin ciertas formas de circularidad. Deseo proporcionar una lista de condiciones para la realización de cierto acto ilocucionario, que no hagan mención ellas mismas de la realización de ningún acto ilocucionario. Necesito satisfacer esta condición para ofrecer una explicación de la noción de acto ilocucionario en general; de otra manera estaría mostrando simplemente la relación entre diferentes actos ilocucionarios. Sin embargo, aunque no se hará referencia a *actos* ilocucionarios, ciertos *conceptos* ilocucionarios aparecerán en el *analysans* al igual que en el *analysandum*; y creo que esta forma de circularidad es inevitable debido a la naturaleza de las reglas constitutivas.

En la presentación de las condiciones consideraré en primer lugar el caso de una promesa sincera y después mostraré cómo modificar las condiciones para hacer sitio a las promesas insinceras. Dado que nuestra investigación es semántica más que sintáctica, supondré simplemente la existencia de oraciones gramaticalmente bien formadas.

Dado que un hablante *H* emite una oración *O* en presencia de un oyente *A*, entonces, en la emisión de *O*, *H* sincera (y no defectivamente) promete que *p* a *A* si y sólo si:

(1) *Se dan las condiciones normales de input y output.*

Uso los términos *input* y *output* para cubrir un extenso e indefinido rango de condiciones bajo las cuales es posible cualquier género de comunicación lingüística. *Output* cubre las condiciones requeridas para hablar inteligiblemente e *input* cubre las condiciones para la comprensión. Juntos incluyen cosas tales como que el hablante y el oyente conozcan ambos cómo hablar el lenguaje; que ambos sean conscientes de lo que están haciendo; que el hablante no esté actuando bajo coacción o amenazas; que no tengan impedimentos físicos para la comunicación tales como sordera, afasia o laringitis; que no estén actuando en una obra de teatro o hablando en broma, etc.

(2) *H expresa que p en la emisión de O.*

Esta condición aísla el contenido proposicional del resto del acto de habla y nos capacita para concentrarnos en las peculiaridades de prometer en el resto del análisis.

(3) *Al expresar que p, H predica un acto futuro X de H.*

En el caso de prometer el dispositivo indicador de función es una expresión cuyo alcance incluye ciertas características de la proposición. En una promesa debe predicarse un acto del hablante, y éste no puede ser un acto pasado. No puedo prometer haber hecho algo, y no puedo prometer que algún otro hará algo. (Aunque puedo prometer cuidar de que él lo hará.) La noción de acto, tal como lo estoy interpretando para los presentes propósitos, incluye abstenerse de realizar ciertos actos, realizar otras series de actos, y puede también incluir estados y condiciones: puedo prometer no hacer algo, puedo prometer hacer algo repetidamente, y puedo prometer estar o permanecer en cierto estado o condición. Llamaré a las condiciones (2) y (3) *condiciones de contenido proposicional*.

(4) *A preferiría que H hiciese X a que no hiciese X, y H cree que A preferiría que él hiciese X a que no hiciese X.*

Una distinción crucial entre promesas de un lado y amenazas de otro es que una promesa es una garantía de que se hará algo a favor de ti, no en contra de ti, pero una amenaza es una garantía de que se te hará algo en contra de ti, no a favor de ti. Una promesa es defectuosa si la cosa prometida es algo que no desea que se haga la persona a quien se promete; y es adicionalmente defectuosa si el que promete no cree que la persona a quien se promete desea que se haga, puesto que una promesa no defectuosa debe hacerse con la intención de que

sea una promesa y no como una amenaza o consejo. Creo que las dos mitades de esta doble condición son necesarias para evitar contraejemplos bastante obvios.

Se puede, sin embargo, pensar en aparentes contraejemplos a esta condición tal como está enunciada. Supóngase que digo a un estudiante perezoso: «Si no entregas tu trabajo a tiempo, prometo que te daré una mala nota en el curso.» ¿Constituye esta emisión una promesa? Me siento inclinado a pensar que no; la describiríamos más naturalmente como un consejo o posiblemente incluso como una amenaza. Pero ¿por qué es posible usar la locución «Prometo» en tal caso? Creo que en este caso la usamos debido a que «Prometo» y «Por la presente prometo» son dos de los dispositivos indicadores de función más fuertes para *compromiso* que proporciona el idioma castellano. Por esta razón usamos a menudo estas expresiones en la realización de actos de habla que no son, estrictamente hablando, promesas, pero en los cuales deseamos subrayar nuestro compromiso. Para ilustrar esto consideremos otro aparente contraejemplo, de líneas diferentes, al análisis. Algunas veces se oye a la gente decir «Prometo» cuando se hace una aserción enfática. Supongamos, por ejemplo, que te acuso de haber robado el dinero. Digo: «Tú robaste ese dinero, ¿no?» Tú replicas: «No, no lo hice, te prometo que no lo hice.» ¿Has hecho en este caso una promesa? Encuentro muy poco natural describir tu emisión como una promesa. Esta emisión se describiría de manera más apropiada como una negativa enfática, y podemos explicar la aparición del dispositivo indicador de función «Prometo» como derivativo de promesas genuinas y sirviendo aquí como una expresión que añade énfasis a tu negativa.

En general, el punto enunciado en la condición (4) es que si una pretendida promesa ha de ser no defectuosa, la cosa prometida debe ser algo que el oyente desea que se haga, o considera que es de su interés, o preferiría que se hiciese a que no se hiciese, etc.; y el hablante debe ser consciente de, o creer, o saber, etc., que éste es el caso. Pienso que una formulación más exacta y elegante de esa condición requeriría la introducción de terminología técnica.

(5) *No es obvio ni para H ni para A, que H hará X en el curso normal de los acontecimientos.*

Esta condición es una instancia de una condición general de muchos géneros diferentes de actos ilocucionarios al efecto de que el acto debe tener un objeto. Por ejemplo, si pido a alguien que haga algo que es obvio que él está haciendo ya o va a hacer, entonces mi petición carece de objeto, y a ese respecto es defectuosa. En una situación de habla efectiva, los oyentes, conociendo las reglas para realizar actos ilocucionarios, supondrán que esta condición se satisfa-

ce. Supongamos, por ejemplo, que en el curso de una disertación pública digo a un miembro de mi auditorio: «Mira aquí, Pérez, presta atención a lo que estoy diciendo.» Para que esta emisión tenga sentido, el auditorio tendrá que suponer que Pérez no había estado prestando atención o en cualquier caso que no es obvio que él había estado prestando atención, que la cuestión de su prestar atención ha surgido de alguna manera; pues una condición para hacer una petición es que no sea obvio que el oyente esté haciendo o vaya a hacer la cosa pedida.

Lo mismo sucede con las promesas. Está fuera de lugar que yo prometa hacer algo que es obvio que voy a hacer de todos modos. Si parece que estoy haciendo una promesa tal, el único modo en que mi auditorio puede dar sentido a mi emisión es suponer que creo que no es obvio que voy a hacer la cosa prometida. Un hombre felizmente casado que promete a su mujer que no la abandonará la próxima semana está probablemente provocando más ansiedad que alivio.

Entre paréntesis, creo que esta condición es una instancia de la clase de fenómeno enunciado en la ley de Zipf. Pienso que en nuestro lenguaje, como en la mayor parte de las formas de la conducta humana, está operando un principio de mínimo esfuerzo, en este caso un principio de máximos resultados ilocucionarios con mínimo esfuerzo fonético; y creo que la condición (5) es una instancia de esto.

Llamo a condiciones tales como (4) y (5) *condiciones preparatorias*. Ellas son las *sine quibus non* de una promesa feliz, pero no enuncian todavía la condición esencial.

(6) *H tiene la intención de hacer X.*

La distinción más importante entre promesas sinceras e insinceras es que en el caso de las promesas sinceras el hablante tiene la intención de llevar a cabo el acto prometido, y en el caso de las promesas insinceras no tiene intención de llevar a cabo el acto. Además en las promesas sinceras el hablante cree que le es posible llevar a cabo el acto (o abstenerse de hacerlo), pero creo que la proposición de que él tiene intención de hacerlo entraña que él piensa que le es posible hacerlo (o abstenerse de hacerlo), de modo que no enuncio esto como una condición adicional. A esta condición la llamo *condición de sinceridad*.

(7) *H tiene la intención de que la emisión de O le coloque a él bajo la obligación de hacer A.*

La característica esencial de una promesa consiste en asumir la obligación de realizar un cierto acto. Creo que esta condición distingue a las promesas (y a otros miembros de la misma familia, como los votos) de otros géneros de actos de habla. Obsérvese que en el enun-

grado de la condición solamente especificamos la intención del hablante; condiciones adicionales clarificarán cómo ha de realizarse esa intención. Sin embargo resulta claro que tener esta intención es una condición necesaria para hacer una promesa; pues, si un hablante puede demostrar que no tenía esta intención en una emisión dada, puede probar que la emisión no era una promesa. Sabemos, por ejemplo, que Mr. Pickwick no prometió casarse puesto que sabemos que no tenía la intención apropiada.

Llamo a esto la *condición esencial*.

- (8) *H tiene la intención de que la emisión de O produzca en A la creencia de que las condiciones (6) y (7) se dan por medio del reconocimiento de la intención de producir esa creencia, y él tiene la intención de que este reconocimiento se logre por medio del reconocimiento de que la oración se usa convencionalmente para producir tales creencias.*

Esto captura nuestro análisis griceano enmendado de lo que es para el hablante querer decir que hace una promesa. El hablante tiene la intención de producir un cierto efecto ilocucionario por el recurso de lograr que el oyente reconozca su intención de producir ese efecto, y él tiene también la intención de que este reconocimiento se consiga en virtud del hecho de que el carácter léxico y sintáctico del ítem que emite se asocia convencionalmente con la producción de ese efecto.

Estrictamente hablando, esta condición podría formularse como parte de la condición (1), pero tiene un interés filosófico suficiente como para que merezca la pena formularla separadamente. La encuentro difícil por la razón siguiente. Si mi objeción a Grice es realmente válida, entonces seguramente, podría decirse, todas esas intenciones iteradas son superfluas; todo lo que sería necesario es que el hablante emitiera seriamente una oración. La producción de todos esos efectos es simplemente una consecuencia del conocimiento que el oyente tiene de lo que la oración significa, lo cual a su vez es una consecuencia de su conocimiento del lenguaje, que se supone al principio por parte del hablante. Creo que la respuesta correcta a esta objeción es que la condición (8) explica lo que es para el hablante emitir «seriamente» una oración, esto es, emitirla y querer decirla (*mean it*), pero no confío completamente ni en la fuerza de la objeción ni en la de la respuesta.

- (9) *Las reglas semánticas del dialecto hablado por H y A son tales que O se emite correcta y sinceramente si y sólo si se dan las condiciones (1)-(8).*

Esta condición pretende clarificar que la oración emitida es tal

que se usa para hacer una promesa en virtud de las reglas semánticas del lenguaje. Tomada juntamente con la condición (8), elimina contraejemplos semejantes al caso del soldado alemán capturado, que hemos considerado anteriormente. Enseguida veremos cuál es exactamente la formulación de las reglas.

Hasta aquí hemos considerado solamente el caso de una promesa sincera. Pero las promesas insinceras son sin embargo promesas, y ahora tenemos necesidad de mostrar cómo modificar las condiciones para tomarlas en consideración. Al hacer una promesa insincera el hablante no tiene todas las intenciones y creencias que tiene cuando hace una promesa sincera. Sin embargo, él da a entender que las tiene. En efecto, debido a que da a entender que tiene intenciones y creencias que él no tiene describimos su acto como insincero. Así, para tomar en consideración las promesas insinceras necesitamos solamente revisar nuestras condiciones, y enunciar que el hablante asume la responsabilidad de tener las creencias e intenciones, más bien que enunciar que él las tiene efectivamente. Un indicio de que el hablante asume tal responsabilidad lo constituye el hecho de que no podría, sin caer en el absurdo, decir, p. ej., «Prometo hacer *X*, pero no tengo la intención de hacer *X*». Decir «Prometo hacer *X*» es asumir la responsabilidad de tener la intención de hacer *X*, y esta condición vale ya sea la emisión sincera o insincera. Para tomar en consideración la posibilidad de una promesa insincera tenemos solamente que revisar la condición (6) de modo que enuncie, no que el hablante tiene la intención de hacer *X*, sino que él asume la responsabilidad de tener la intención de hacer *X*, y para evitar la acusación de circularidad expresaré esto de la manera siguiente:

(6\*) *H tiene la intención de que la emisión de O le hará a él responsable de tener la intención de hacer X.*

Así rectificado (y con «sinceramente» suprimido de nuestro *analysandum* y de la condición (9)), nuestro análisis es neutral respecto de la cuestión de si la promesa era sincera o insincera.

## VI. REGLAS PARA EL USO DEL DISPOSITIVO INDICADOR DE FUNCIÓN

Nuestra próxima tarea consiste en extraer de nuestro conjunto de condiciones un conjunto de reglas para el uso del dispositivo indicador de función. Obviamente no todas nuestras condiciones son igualmente relevantes para esta tarea. La condición (1) y las condiciones de las formas (8) y (9) se aplican a todos los géneros de actos ilocucionarios normales, y no son peculiares de prometer. Las reglas del dis



positivo indicador de función para prometer ha de hallarse que corresponden a las condiciones (2)-(7).

Las reglas semánticas para el uso de cualquier dispositivo indicador de función *P* para prometer son:

*Regla 1.* *P* ha de emitirse solamente en el contexto de una oración (o trozo de discurso mayor), cuya emisión predica algún acto futuro *X* del hablante *H*. Llamo a esto *la regla de contenido proposicional*. Se deriva de las condiciones de contenido proposicional (2) y (3).

*Regla 2.* *P* ha de emitirse sólo si el oyente *A* prefería que *H* hiciese *X* a que no hiciese *X*, y *H* cree que *A* preferiría que *H* hiciese *A* a que no hiciese *A*.

*Regla 3.* *P* ha de emitirse solamente si no es obvio tanto para *H* como para *A* que *H* no hará *X* en el curso normal de los acontecimientos.

Llamo a las reglas 2 y 3 *reglas preparatorias*. Se derivan de las condiciones preparatorias (4) y (5).

*Regla 4.* *P* ha de emitirse solamente si *H* tiene la intención de hacer *X*.

Llamo a esto *la regla de sinceridad*. Se deriva de la condición de sinceridad (6).

*Regla 5.* La emisión de *P* cuenta como la asunción de una obligación de hacer *X*.

Llamo a esto *la regla esencial*.

Estas reglas están ordenadas: las reglas 2-5 se aplican solamente si la regla 1 es satisfecha, y la regla 5 se aplica solamente si las reglas 2 y 3 son satisfechas también.

Obsérvese que mientras que las reglas 1-4 toman la forma de cuasi imperativos, esto es, son de la forma: emite *P* sólo si *x*, la regla 5 tiene la forma: la emisión de *P* cuenta como *y*. Así, la regla 5 pertenece al género peculiar a los sistemas de reglas constitutivas que he discutido en la sección II.

Obsérvese también que la más bien dificultosa analogía con los juegos se mantiene destacablemente bien. Si nos preguntamos bajo qué condiciones puede decirse que un jugador ha movido correctamente un caballo, encontraríamos condiciones preparatorias, tales como que debe ser su turno para mover, al igual que la condición esencial que enuncia las posiciones de hecho a las que el caballo puede moverse. Creo que incluso existe una regla de sinceridad para los juegos competitivos, la regla de que cada parte intenta ganar. Sugiero que el equipo que en un juego «se vende» está comportándose de una manera estrechamente análoga al hablante que miente o hace promesas falsas. Desde luego, usualmente no existen reglas de contenido proposicional para los juegos, puesto que los juegos, en general, no representan estados de cosas.

Si mi análisis es de algún interés general más allá del caso de pro-

meter, entonces parecería que estas distinciones habrán de trasladarse a otros tipos de actos de habla, y creo que una pequeña reflexión mostrará que es así. Considérese, p. ej., el dar una orden. Las condiciones preparatorias incluyen que el hablante esté en una posición de autoridad sobre el oyente, la condición de sinceridad consiste en que el hablante desea que se lleve a cabo el acto ordenado, y la condición esencial tiene que ver con el hecho de que la emisión es un intento de lograr que el oyente lo haga. Para las aserciones, las condiciones preparatorias incluyen el hecho de que el hablante deba tener algunas bases para suponer que la proposición aseverada es verdadera, la condición de sinceridad consiste en que él debe creer que es verdadera, y la condición esencial tiene que ver con el hecho de que la emisión es un intento de informar al oyente y convencerlo de su verdad. Los saludos son un género muy simple de actos de habla, pero incluso aquí se aplican algunas de las distinciones. En la emisión de «Hola» no existe contenido proposicional ni condición de sinceridad. La condición preparatoria consiste en que el hablante debe haber acabado de encontrarse con el oyente, y la regla esencial es que la emisión indica un cortés reconocimiento del oyente.

Una propuesta para posterior investigación consistiría en llevar a cabo un análisis similar de otros tipos de actos de habla. Esto no solamente nos daría un análisis de conceptos interesantes en sí mismos, sino que la comparación entre diferentes análisis profundizaría nuestra comprensión de la totalidad del tema, e incidentalmente proporcionaría una base para una taxonomía más seria que cualquiera de las usuales y fáciles categorías tales como evaluativo frente a descriptivo, o cognitivo frente a emotivo.

## UNA TAXONOMÍA DE LOS ACTOS ILOCUCIONARIOS \*

JOHN R. SEARLE

### I INTRODUCCIÓN

El principal propósito de este artículo es desarrollar una clasificación razonada de los actos ilocucionarios en ciertas categorías o tipos básicos. Esto es responder a la pregunta: ¿Cuántos géneros de actos ilocucionarios hay? Puesto que cualquier intento de desarrollar una taxonomía debe tener en cuenta la clasificación de los actos ilocucionarios de Austin en sus cinco categorías básicas de veredictivos, expositivos, ejercitativos, comportativos y conmisivos, un segundo propósito de este artículo es valorar la clasificación de Austin para mostrar en qué aspectos es adecuada y en qué aspectos no lo es. Además, puesto que las diferencias semánticas tienen seguramente consecuencias sintácticas, un tercer propósito de este artículo es mostrar cómo esos diferentes tipos ilocucionarios básicos están realizados en la sintaxis de un lenguaje natural como el inglés.

En lo que sigue presupondré familiaridad con el modelo general de análisis de los actos ilocucionarios ofrecido en obras tales como las de Austin, *How to do Things with Words*, Searle, *Speech Acts*, y Searle «Austin on Locutionary and Illocutionary Acts»<sup>1</sup>. En particular presupondré la distinción entre fuerza ilocucionaria de una emisión y su contenido proposicional simbolizada como  $F(p)$ . El propósito de este artículo es clarificar los diferentes tipos de  $F$ .

---

\* Versión castellana de Luis Ml. Valdés Villanueva.

<sup>1</sup> J. L. Austin, *How to Do Things with Words*, Clarendon Press, Oxford, 1962 [existe traducción castellana bajo el título de *Palabras y Acciones*, por G. Carrió y R. Rabossi, Paidós, Buenos Aires, 1971 (N. del T.)]. J. R. Searle, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, Londres 1969 [traducción castellana: *Actos de Habla: Un Ensayo de Filosofía del Lenguaje*, Cátedra, Madrid, 1979 (N. del T.)] y J. R. Searle, «Austin on Locutionary and Illocutionary Acts», *Philosophical Review*, 1968.

## II. DIFERENTES TIPOS DE DIFERENCIAS ENTRE DIFERENTES TIPOS DE ACTOS ILOCUCIONARIOS

Cualquier esfuerzo taxonómico de esta clase presupone criterios para distinguir una (clase de) actos ilocucionarios de otra. ¿Cuáles son los criterios mediante los cuales podemos decir que de tres emisiones efectivas una es un informe, otra es una predicción y otra es una promesa? Para desarrollar géneros de orden superior debemos primero conocer cómo las especies *promesa*, *predicción*, *informe* etc., difieren una de otra. Cuando se intenta responder a esa pregunta se descubre que existen varios principios de distinción completamente diferentes; esto es, existen diferentes géneros de diferencias que nos capacitan para decir que la fuerza de esta emisión es diferente de la fuerza de esa emisión. Por esta razón la metáfora de la fuerza en la expresión «fuerza ilocucionaria» es desorientadora ya que sugiere que las diferentes fuerzas ilocucionarias ocupan posiciones diferentes en un único continuo de fuerza. Lo que efectivamente sucede es que existen varios continuos distintos entrecruzados. Una fuente de confusión relacionada es que estamos inclinados a confundir verbos ilocucionarios con tipos de actos ilocucionarios. Estamos inclinados a pensar, por ejemplo, que siempre que tenemos dos verbos ilocucionarios no sinónimos éstos deben señalar necesariamente dos géneros diferentes de actos ilocucionarios. En lo que sigue intentaré mantener una clara distinción entre verbos ilocucionarios y actos ilocucionarios. Las ilocuciones son una parte del lenguaje como opuesto a los lenguajes particulares. Los verbos ilocucionarios son siempre parte de un lenguaje particular: francés, alemán, castellano, o cualquier otro. Las diferencias en los verbos ilocucionarios son una buena guía pero en absoluto una guía segura para establecer diferencias en los actos ilocucionarios.

Me parece que hay (al menos) doce dimensiones significativas de variación en las que los actos ilocucionarios difieren uno de otro y que enumeraré todas ellas muy de pasada:

1. *Diferencias en el objeto (o propósito) del (tipo de) acto.* El objeto o propósito de una orden puede ser especificado diciendo que es un intento de conseguir que el oyente haga algo. El objeto o propósito de una descripción es que sea una representación (verdadera o falsa, exacta o inexacta) de cómo es algo. El objeto o propósito de una promesa es que sea la asunción de una obligación por parte del hablante de hacer algo. Esas diferencias corresponden a las condiciones esenciales de mi análisis de los actos ilocucionarios en *Speech Acts*<sup>2</sup>. Actualmente, creo, las condiciones esenciales forman las me-

<sup>2</sup> Searle, *Speech Acts*, cap. 3.

jores bases para una taxonomía, como intentaré mostrar. Es importante darse cuenta de que la terminología de «objeto» o «propósito» no implica, ni está basada, sobre el punto de vista de que todo acto ilocucionario tiene por definición un intento perlocucionario asociado. Para muchos —quizás la mayor parte— de los más importantes actos ilocucionarios no existe intento perlocucionario esencial asociado por definición con el verbo correspondiente; por ejemplo, enunciados y promesas no son, por definición, intentos de producir efectos perlocucionarios en los oyentes.

Llamaré al objeto o propósito de un tipo de ilocución su *objeto ilocucionario* (*illocutionary point*). El objeto ilocucionario es parte de, pero no lo mismo que, la fuerza ilocucionaria. Así, por ejemplo, el objeto ilocucionario de una petición es el mismo que el de una orden: ambos son intentos de lograr que los oyentes hagan algo. Pero las fuerzas ilocucionarias son claramente diferentes. En general puede decirse que la noción de fuerza ilocucionaria es la resultante de diversos elementos de los cuales el objeto ilocucionario es sólo uno aunque, creo, el más importante.

2. *Diferencias en la dirección de ajuste entre las palabras y el mundo.* Algunas ilocuciones tienen como parte de su objeto ilocucionario el lograr que las palabras (más estrictamente, su contenido proposicional) encajen con el mundo, otras el lograr que el mundo encaje con las palabras. Las aserciones pertenecen a la primera categoría, las promesas y ruegos a la segunda. La mejor ilustración que conozco de esta distinción es la proporcionada por Miss Anscombe<sup>3</sup>. Supón que un hombre va al supermercado con una lista de compras que le ha dado su mujer en la que están escritas las palabras «habas, mantequilla, *bacon* y pan». Supón que mientras que él anda por allí con su carrito seleccionando esos elementos es seguido por un detective que escribe todo lo que él coge. Cuando salen de la tienda comprador y detective tendrán listas idénticas. Pero la función de ambas listas será completamente diferente. En el caso de la lista del comprador el propósito de la lista es, por así decirlo, lograr que el mundo encaje con las palabras: se supone que la persona al llevar a cabo sus acciones se ajusta a la lista. En el caso del detective el propósito de la lista es hacer que las palabras encajen con el mundo: se supone que la persona, al hacer la lista, se ajusta a las acciones del comprador. Además esto puede ser demostrado observando el papel de un «error» en los dos casos. Si el detective vuelve a casa y de repente se da cuenta que el hombre compró chuletas de cerdo en vez de *bacon*, puede, simplemente, borrar la palabra «*bacon*» y escribir «chuletas de cerdo».

<sup>3</sup> G. M. E. Anscombe, *Intention*, Blackwell, Oxford, 1957.

Pero si el comprador vuelve a casa y su mujer le advierte que ha comprado chuletas de cerdo cuando debía haber comprado *bacon*, no puede corregir el error borrando la palabra «bacon» de la lista y escribiendo «chuletas de cerdo».

En esos ejemplos la lista proporciona el contenido proposicional de la ilocución y la fuerza ilocucionaria determina cómo se supone que el contenido se relaciona con el mundo. Propongo llamar a esta diferencia una diferencia en la *dirección de ajuste*. La lista del detectivo tiene la dirección de ajuste *palabra-a-mundo* (como la tienen los enunciados, descripciones, aserciones y explicaciones); la lista del comprador tiene la dirección de ajuste *mundo-a-palabra* (como la tienen los ruegos, órdenes, votos y promesas). Represento la dirección de ajuste palabra-a-mundo con una flecha hacia abajo, ↓, y la dirección de ajuste mundo-a-palabra con una flecha hacia arriba, ↑. La dirección de ajuste es siempre una consecuencia del objeto ilocucionario. Resultaría más elegante si pudiésemos construir totalmente nuestra taxonomía en torno a esta distinción en la dirección de ajuste pero, aunque figurará extensamente en nuestra taxonomía, no soy capaz de hacer de ella la base total de las distinciones.

3. *Diferencias en los estados psicológicos expresados*. Una persona que enuncia, explica, asevera o afirma que *p*, *expresa la creencia que p*; una persona que promete, hace votos, profiere amenazas o se compromete a hacer *A*, *expresa la intención de hacer A*; una persona que ordena, manda o pide a *O* que haga *A* *expresa un deseo (aspiración o anhelo) de que O haga A*; una persona que pide disculpas por hacer *A* *expresa su sentimiento por haber hecho A*; etc. En general en la realización de cualquier acto ilocucionario con un contenido proposicional el hablante expresa alguna actitud, estado, etc., hacia ese contenido proposicional. Nótese que esto vale incluso si es insincero, incluso si no se tiene la creencia, deseo, intención, sentimiento o placer que se expresa, sin embargo se expresa una creencia, deseo, intención, sentimiento o placer en la realización del acto de habla. Este hecho está marcado lingüísticamente por el hecho de que es lingüísticamente inaceptable (aunque no autocontradictorio) el poner en conjunción el verbo realizativo explícito con la negación del estado psicológico expresado. Así no se puede decir «Enuncio que *p* pero no creo que *p*», «Prometo que *p*, pero no tengo intención de hacer *p*». Adviértase que esto sólo vale en el uso realizativo de la primera persona. Puede decirse «Enunció que *p*, pero no creía realmente que *p*», «Prometí que *p*, pero no tenía realmente la intención de hacerlo», etc. El estado psicológico expresado en la realización del acto ilocucionario es la *condición de sinceridad* del acto, tal como se analizó en *Speech Acts*, capítulo 3.

Si se intenta hacer una clasificación de los actos ilocucionarios ba-

anda enteramente en los diferentes estados psicológicos expresados (diferencias en la condición de sinceridad) puede emprenderse un larguísimo camino. Pues *creencia* no recoge solamente enunciados, aserciones, observaciones y explicaciones, sino también postulaciones, declaraciones, deducciones y argumentos. *Intención* recogerá promesas, votos, amenazas y compromisos. *Deseo o aspiración* recogerá peticiones, órdenes, mandatos, preguntas, súplicas, alegatos, ruegos e imploraciones. *Placer* no recoge tantos —congratulaciones, felicitaciones, bienvenidas y pocos más—.

En lo que sigue simbolizaré el estado psicológico expresado con las letras mayúsculas iniciales del verbo correspondiente; *C* para «creer», *D* para «desear», *I* para «intentar», etc.

Estas tres dimensiones —objeto ilocucionario, dirección de ajuste y condición de sinceridad— me parecen las más importantes y construiré sobre ellas la mayor parte de mi taxonomía, pero existen otras más que necesitan ser tomadas en cuenta.

4. *Diferencias en la fuerza o intensidad con la que se presenta el objeto ilocucionario.* «Sugiero que vayamos al cine» e «Insisto en que vayamos al cine» tienen ambos el mismo objeto ilocucionario pero es presentado con diferente intensidad del mismo modo que «Juro solemnemente que Bill robó el dinero» y «Sospecho que Bill robó el dinero». A lo largo de la misma dimensión de objeto o propósito ilocucionario puede haber diversos grados de intensidad o compromiso.

5. *Diferencias en el «status» o posición del hablante y el oyente en la medida en que afectan a la fuerza ilocucionaria de la emisión.* Si el general pide al soldado que limpie la habitación se trata, con toda probabilidad, de una orden. Si el soldado pide al general que limpie la habitación se tratará seguramente de una sugerencia, una propuesta o un ruego pero no de una orden o de un mandato. Esta característica corresponde a una de las condiciones preparatorias de mi análisis en *Speech Acts*, cap. 3.

6. *Diferencias en la manera en que la emisión se relaciona con los intereses del hablante y del oyente.* Considérense, por ejemplo, las diferencias entre jactancias y lamentos, entre felicitaciones y condolencias. En estas dos parejas la diferencia se percibe entre lo que está o no está en los intereses del hablante y del oyente respectivamente. Esta característica es otro tipo de condición preparatoria de acuerdo con mi análisis de *Speech Acts*.

7. *Diferencias en las relaciones con el resto del discurso.* Algunas emisiones realizativas sirven para poner en relación la emisión con el resto del discurso (y también con el contexto circundante). Considé-

rense, por ejemplo, «replico», «deduzco», «concluyo» y «objeto». Estas expresiones sirven para poner en relación emisiones con otras emisiones y con el contexto circundante. Los rasgos que señalan parecen implicar en gran medida emisiones pertenecientes a la clase de los enunciados. Además de enunciar simplemente una proposición uno puede enunciarla objetando lo que cualquier otro ha dicho, replicando a un punto anterior, deduciéndola de ciertas premisas evidentes, etc. «Sin embargo», «por otra parte» y «por lo tanto» realizan también esas funciones de relación de discurso.

8. *Diferencias en el contenido proposicional que están determinadas por los dispositivos indicadores de fuerza ilocucionaria.* Las diferencias entre, por ejemplo, un informe y una predicción incluyen el hecho de que una predicción debe ser sobre el futuro mientras que un informe puede ser sobre el pasado o el presente. Esas diferencias corresponden a las diferencias en las condiciones de contenido proposicional tal como se explicó en *Speech Acts*.

9. *Diferencias entre aquellos actos que deben siempre ser actos de habla y aquellos que pueden ser, pero no necesitan ser, realizados como actos de habla.* Pueden clasificarse cosas diciendo, por ejemplo, «Clasifico esto como un A y clasifico esto como un B». Pero no es necesario decir nada para clasificarlas; se puede, simplemente, colocar todos los A en el cajón A y todos los B en el cajón B. Lo mismo ocurre con estimar, diagnosticar y concluir. Puedo hacer estimaciones, dar diagnósticos y extraer conclusiones diciendo «estimo», «diagnostico» y «concluyo», pero para diagnosticar, estimar o concluir no es necesario decir nada en absoluto. Puedo, simplemente, estar delante de un edificio y estimar su altura, diagnosticar en silencio que eres un esquizofrénico marginal, o concluir que el hombre que está sentado a mi lado está completamente borracho. En tales casos no es necesario ningún acto de habla, ni siquiera interno.

10. *Diferencias entre aquellos actos que requieren instituciones extralingüísticas para su realización y aquellos que no.* Existe un gran número de actos ilocucionarios que requieren una institución extralingüística y generalmente una posición especial del hablante y el oyente dentro de una institución para que el acto sea realizado. Así para bendecir, excomulgar, bautizar, declarar culpable, declarar fuera al jugador de béisbol, dar jaque mate en ajedrez o declarar la guerra no es suficiente que cualquier hablante experimentado diga a cualquier oyente experimentado «bendigo» o «excomulgo», etc. Debe tenerse una posición dentro de una institución extralingüística. Austin habla algunas veces como si pensara que todos los actos ilocucionarios tuviesen esta característica. Pero claramente esto no es así.



Para hacer un enunciado de que está lloviendo o prometer venir a verte, necesito solamente obedecer las reglas del lenguaje. No se requiere ninguna institución extralingüística. Ese rasgo de ciertos actos de habla, que requieren instituciones extralingüísticas, es necesario distinguirlo del rasgo 5, el hecho de que ciertos actos ilocucionarios requieran que el hablante y posiblemente el oyente también tengan un determinado *status*. Las instituciones extralingüísticas confieren a menudo *status* en un sentido relevante para la fuerza ilocucionaria, pero no todas las diferencias de *status* se derivan de instituciones. Así un ladrón armado, en virtud de su posesión de una pistola, puede *ordenar*, como opuesto por ejemplo a pedir, rogar, o implorar a sus víctimas que levanten las manos. Pero su *status* aquí no deriva de una posición dentro de una institución, sino de su posesión de un arma.

11. *Diferencias entre aquellos actos donde el verbo ilocucionario correspondiente tiene un uso realizativo y aquéllos donde no lo tiene.* Muchos verbos ilocucionarios tienen usos realizativos —por ejemplo, «enunciar», «prometer», «ordenar», «concluir»—. Pero no pueden realizarse actos, por ejemplo, de jactarse o de amenazar diciendo «Por la presente me jacto» o «Por la presente amenazo»<sup>4</sup>. No todos los verbos ilocucionarios son verbos realizativos.

12. *Diferencias en el estilo de realización del acto ilocucionario.* Algunos verbos ilocucionarios sirven para señalar lo que podemos llamar el *estilo* especial en el cual es realizado un acto ilocucionario. Así la diferencia, por ejemplo, entre enunciar y hacer una confidencia no implica ninguna diferencia en el objeto ilocucionario o en el contenido proposicional, sino sólo en el *estilo* de la realización del acto ilocucionario.

### III. DEFICIENCIAS EN LA TAXONOMÍA DE AUSTIN

Austin avanzó sus cinco categorías muy provisionalmente, más como una base de discusión que como un conjunto de resultados es-

<sup>4</sup> Existen otros verbos en castellano que suenan raros en la primera persona del presente. Considérense «ocultarse» y «acechar». Resulta extraño decir como respuesta a la pregunta «¿Qué estás haciendo?», «Estoy ocultándome en los arbustos» o «Estoy acechando hoy». La razón puede ser que ambos verbos implican una valoración negativa y es extraño dar una valoración negativa de lo que uno está haciendo mientras lo hace voluntariamente. Quizás funcione una explicación similar para «jactarse» y «amenazar» puesto que ambos también parecen contener un elemento de valoración negativa. Nótese que son aceptables si están incluidos en alguna forma de disculpa, por ejemplo, «Espero que no te molestarás si me jacto de mi nueva motocicleta».

tablecidos. «No estoy proponiendo, dice él, nada definitivo»<sup>5</sup>. Creo que forman una excelente base de discusión pero también creo que la taxonomía necesita ser seriamente revisada puesto que contiene diversas deficiencias. He aquí las cinco categorías de Austin:

*Veredictivos*: «Consisten en pronunciar un veredicto, oficial o inoficialmente, sobre la base de pruebas o razones respecto a cuestiones de valor o de hecho hasta el punto en que pueda distinguirse entre ambas.» Ejemplos de verbos de esta clase son: *absolver, juzgar, calcular, describir, analizar, estimar, fechar, clasificar, valorar, caracterizar y describir*.

*Ejercitativos*: Algunos de estos consisten en «tomar una decisión en favor o en contra de cierto curso de acción o abogar por ella», «una decisión de que algo debe ser así como algo distinto de juzgar que es así». Algunos ejemplos son, *ordenar, mandar, dirigir (un ruego), abogar por, suplicar, recomendar, rogar y advertir*. Pedit es también un ejemplo obvio pero Austin no lo incluye en la lista. Del mismo modo que los anteriores Austin incluye también en la lista *designar (para un cargo), destituir, dar nombre, vetar, clausurar, declarar abierta (la sesión)*, así como también, *anunciar, aconsejar, proclamar y dar*.

*Conmisivos*: «El objeto total de un comisivo» nos dice Austin «es comprometer al hablante en un cierto curso de acción». Algunos de los ejemplos más obvios son: *prometer, hacer votos, empeñar (la palabra), pactar, contratar, garantizar, abrazar (una causa) y jurar*.

*Expositivos*: «Se usan en actos de exposición que implican la exposición de puntos de vista, la dirección de argumentos y la clarificación de usanzas y referencias.» Austin da muchos ejemplos; entre ellos están: *afirmar, negar, subrayar, ilustrar, preguntar, informar, aceptar, objetar a, conceder, describir, clasificar, identificar y llamar*.

*Comportativos*: Esta clase, con la que Austin estaba muy insatisfecho (*a shocker* la llamó él), «incluye la noción de reacción frente a la conducta y fortuna de las demás personas y la de actitudes y expresiones de actitudes frente a la conducta pasada o inminente de algún otro».

Entre los ejemplos que Austin relaciona están: *pedir disculpas, dar las gracias, deplorar, compadecer, congratular, felicitar, dar la bienvenida, aplaudir, criticar, bendecir, maldecir y brindar por*. Pero también curiosamente *retar, desafiar, protestar y provocar*.

La primera cosa a tener en cuenta en estas líneas es que no son clasificaciones de actos ilocucionarios sino de verbos ilocucionarios castellanos. Austin parece suponer que una clasificación de diferentes verbos es *eo ipso* una clasificación de géneros de actos ilocuciona-

<sup>5</sup> Austin, *How to Do Things With Words*, p. 151.

nos y que cualquier par de verbos sinónimos debe señalar diferentes actos ilocucionarios. Pero no existe ninguna razón para suponer que esto sea así. Como veremos algunos verbos señalan la manera en que se realiza un acto ilocucionario, por ejemplo, «anunciar». Pueden anunciarse órdenes, promesas, e informes pero anunciar no coincide exactamente con ordenar, prometer e informar. Anunciar, para anticipar las cosas, no es el nombre de un tipo de acto ilocucionario, sino el de un modo en el que se realiza algún acto ilocucionario. Un anuncio no es nunca solamente un anuncio porque «anunciar» no es el nombre de un objeto ilocucionario. Un anuncio debe ser también un enunciado, una orden, etc.

Incluso concediendo que las listas son de verbos ilocucionarios y no necesariamente de diferentes actos ilocucionarios me parece que se pueden dirigir las siguientes críticas adicionales contra ellas.

1. En primer lugar una crítica menor aunque no despreciable. Ni siquiera todos los verbos incluidos en la lista son verbos ilocucionarios. Por ejemplo «simpatizar», «considerar como», «proponerse (algo)», «intentar» y «lo haré» (*shall*). Tomemos «intentar»: no es claramente un realizativo. Decir «Intento» no es intentar ni en la tercera persona da nombre a un acto ilocucionario: «Él intentó...» no informa de ningún acto de habla. Naturalmente existe el acto ilocucionario de *expresar una intención* pero la frase verbal ilocucionaria es «expresar una intención» no «intentar». Intentar no es nunca expresar un acto de habla; expresar una intención, usualmente aunque no siempre, lo es.

2. La deficiencia más importante de la taxonomía es simplemente ésta. No existe principio o conjunto de principios claros y consistentes sobre cuya base se construya la taxonomía. Solamente en el caso de los conmisivos Austin utilizó claramente y sin ambigüedades el objeto ilocucionario como base de la definición de una categoría. Los expositivos, hasta el punto en que la caracterización es clara, parecen estar definidos en términos de relaciones de discurso (mi rasgo 7). Los ejercitativos parecen estar, al menos parcialmente, definidos en términos del ejercicio de la autoridad. Tanto las consideraciones de *status* (mi rasgo 5) como las consideraciones institucionales (mi rasgo 10) están ocultas en ellos. Los comportativos no me parece que estén, en absoluto, bien definidos (como Austin, estoy seguro, estaría de acuerdo) pero parecen incluir nociones de lo que es bueno o malo para el hablante y el oyente (mi rasgo 6) así como expresiones de actitudes (mi rasgo 3).

3. Puesto que no existe un principio claro de clasificación y puesto que existe una persistente confusión entre actos ilocuciona-

rios y verbos ilocucionarios, se da un alto grado de superposición de una categoría con otra y una gran heterogeneidad dentro de alguna de las categorías. El problema no es que haya casos dudosos—cualquier taxonomía que tenga que ver con el mundo real está probablemente abocada a encontrarse con casos dudosos— ni es metaforicamente que un puñado de casos inusuales tenga las características definitorias de más de una categoría. Más bien un gran número de verbos se encuentran atrapados en el medio de dos categorías en competición dado que los principios de clasificación no son sistemáticos. Consideremos, por ejemplo, el verbo «describir», un verbo muy importante en cualquier teoría de los actos de habla. Austin lo incluye tanto en la lista de las expositivos como en la de los veredictivos. Dadas sus definiciones es fácil ver por qué: describir puede ser tanto emitir un fallo como un acto de exposición. Pero entonces cualquier «acto de exposición que incluya la exposición de puntos de vista» podría también ser, en su sentido más especial, «pronunciar un veredicto oficial o inoficial, sobre la base de pruebas o razones». Y, en efecto, un vistazo a su lista de expositivos (pp. 161-162) es suficiente para mostrar que muchos de sus verbos se ajustan a su definición de veredictivos lo mismo que «describir». Consideremos «afirmar», «negar», «enunciar», «clasificar», «identificar», «concluir» y «deducir». Todos éstos están incluidos en la lista de los expositivos pero con la misma facilidad podrían haber sido incluidos en la lista de los veredictivos. Los pocos casos que no son claramente veredictivos son casos donde el significado del verbo trata puramente de relaciones de discurso, por ejemplo, «comenzar por», «pasar a», o donde no es un asunto de evidencia o razones, por ejemplo, «postular», «omitir», «llamar» y «definir». Pero esto no es realmente suficiente para justificar una categoría separada, especialmente dado que muchos de estos verbos—«comenzar por», «pasar a», «omitir»— no son en absoluto nombres de actos ilocucionarios.

4. No sólo existe una gran superposición de categorías próximas sino que dentro de alguna de las categorías hay clases de verbos completamente distintas. Así Austin incluye en la lista de los comportativos «retar», «desafiar» y «provocar» al lado de «dar las gracias», «pedir disculpas», «deplorar» y «dar la bienvenida». Pero «retar», «desafiar» y «provocar» tienen que ver con acciones subsecuentes de los oyentes. Estos comparten con «ordenar», «mandar» y «prohibir» los mismos fundamentos semánticos y sintácticos, como argüiré más adelante. Pero si examinamos la familia que incluye «ordenar», «mandar» e «instar» encontramos que están incluidos en la lista de los ejercitativos al lado de «vetar», «contratar» y «degradar». Pero éstos, como argüiré más adelante, pertenecen a dos categorías completamente distintas.

5. En relación con estas objeciones está la nueva dificultad de que no todos los verbos incluidos en las listas dentro de las clases satisfacen realmente las definiciones dadas, incluso si tomamos las definiciones de la manera más bien laxa y sugerente que Austin claramente intenta. Entonces dar nombre, nombrar para un cargo y excolmulgar no son «el tomar una decisión a favor de o en contra de un cierto curso de acción» y mucho menos «abogar» por ella. Más bien son, como Austin mismo podría haber dicho, *realizaciones* de esas acciones y no *defensas* de algo. Esto es, en el sentido en el que podríamos estar de acuerdo en que ordenar, mandar e instar a alguien que haga algo son todos casos de *abogar* por que lo haga, no podemos estar de acuerdo en que sea también abogar por algo el dar nombre o nombrar para un cargo. Cuando yo te nombro presidente no abogo por que seas o llegues a ser presidente; te *hago* presidente.

En suma, existen (por lo menos) las siguientes seis dificultades relacionadas con la taxonomía de Austin. En orden ascendente de importancia, existe una persistente confusión entre verbos y actos; no todos los verbos son verbos ilocucionarios; existe una gran superposición de las categorías; hay una gran heterogeneidad dentro de las categorías; muchos de los verbos relacionados en las categorías no satisfacen la definición dada para la categoría; y, lo más importante, no existe un principio de clasificación consistente.

No creo que haya justificado plenamente todas las seis acusaciones y no intentaré hacerlo en los límites de este artículo que tiene otras miras. Creo, sin embargo, que mis dudas sobre la taxonomía de Austin tendrán mayor claridad y fuerza después que presente una alternativa. Lo que me propongo hacer es tomar el objeto ilocucionario y sus corolarios, dirección de ajuste, condiciones de sinceridad expresadas, como base para construir una clasificación. En tal clasificación, otros rasgos —el papel de la autoridad, las relaciones de discurso, etc.—, se incluirán en sus lugares apropiados.

#### IV. TAXONOMÍA ALTERNATIVA

En esta sección presentaré una lista de lo que considero categorías básicas de los actos ilocucionarios. Al hacerlo discutiré brevemente cómo se relaciona mi clasificación con la de Austin.

*Representativos.* El objeto o propósito de los miembros de la clase de los representativos es comprometer al hablante (en diferentes grados) con que algo es el caso, con la verdad de la proposición expresada. Todos los miembros de la clase de los representativos son evaluables según la dimensión de valoración que incluye *verdadero* y *falso*. Usando el signo de la aserción de Frege para marcar el objeto ilocucionario común a todos los miembros de esta clase y los símbo-

los introducidos más arriba, podemos simbolizar esta clase como sigue:

$$\vdash \downarrow C(p)$$

La dirección de ajuste es palabras-a-el-mundo y el estado psicológico expresado es creencia (de que  $p$ ). Es importante subrayar que palabras tales como «creencia», y «compromiso» se usan aquí con el propósito de señalar dimensiones; son, por así decirlo, determinables más que determinadas. Así existe una diferencia entre *sugerir que  $p$  o presentarlo como una hipótesis de que  $p$*  de un lado e *insistir en que  $p$  o jurar solemnemente que  $p$*  del otro. El grado de creencia y compromiso puede aproximarse o incluso alcanzar cero, pero está claro o llegará a estarlo que *plantear hipotéticamente que  $p$*  y *enunciar llanamente que  $p$*  están en la misma línea de cuestiones en un sentido en el que ninguno de los dos se parece a pedir. Una vez que reconocemos la existencia de los representativos como una clase completamente separada, basada en la noción de objeto ilocucionario, entonces la existencia de un gran número de verbos realizativos que denotan ilocuciones que parecen ser evaluables en la dimensión verdadero-falso y con todo no son solamente «enunciados» será fácilmente explicable en términos del hecho que señalan rasgos de fuerza ilocucionaria que se suman al objeto ilocucionario. Así, por ejemplo, consideremos «jactarse» y «quejarse». Ambos denotan representativos con la característica adicional de que tienen algo que ver con el interés del hablante (rasgo 6 anterior). «Concluir» y «deducir» son también representativos con el rasgo adicional de que señalan ciertas relaciones entre el acto ilocucionario representativo y el resto del discurso o el contexto de emisión (rasgo 7 anterior). Esta clase contendrá la mayor parte de los expositivos de Austin así como también muchos de sus veredictivos por la razón, ahora espero que obvia, de que todos ellos tienen el mismo objeto ilocucionario y difieren solamente en otros rasgos de fuerza ilocucionaria. La prueba más simple de un representativo es ésta: puedes literalmente (*inter alia*) caracterizarlo como verdadero o falso. Sin embargo me apresuro a añadir que esto no proporcionará ni las condiciones necesarias ni las suficientes como veremos cuando lleguemos a mi quinta clase.

Estas observaciones sobre los representativos estarán más claras, espero, cuando discuta mi segunda clase que con alguna reticencia llamaré:

*Directivos.* Su objeto ilocucionario consiste en el hecho de que son intentos (de varios grados, y por eso más precisamente, son determinados del determinable que incluye intentar) del hablante de lograr que el oyente haga algo. Pueden ser «intentos» muy modestos

como cuando te invito a hacer algo o te sugiero que lo hagas, o pueden ser intentos de mayor intensidad, como cuando insisto en que lo hagas. Usando de modo general el signo de la exclamación como dispositivo indicador del objeto ilocucionario para los miembros de esta clase tendremos el simbolismo siguiente:

! ↑ *D* (*O* hace *A*)

La dirección de ajuste es mundo-a-palabras y la condición de sinceridad es desear (o querer o aspirar). El contenido proposicional es siempre que el oyente *O* lleve a cabo alguna futura acción *A*. Los verbos que denotan miembros de esta clase son: *ordenar*, *mandar*, *pedir*, *preguntar*, *interrogar*<sup>6</sup>, *suplicar*, *abogar por*, *rogar*, y también *invitar*, *permitir* y *aconsejar*. Pienso también que está claro que *retar*, *desafiar* y *provocar*, que Austin incluye en la lista de los comportativos, pertenecen a esta clase. Muchos de los ejercitativos de Austin también pertenecen a esta clase.

*Conmisivos*. La definición de Austin de los comisivos me parece intachable y simplemente la haré compatible tal como está con la crítica de que varios de los verbos que incluye en la lista de los comisivos tales como «lo haré» (*shall*), «intentar», «favorecer» y otros no pertenecen en absoluto a esta clase. Los comisivos son, entonces, aquellos actos ilocucionarios cuyo objeto es comprometer al hablante (de nuevo en grados varios) con algún futuro curso de acción. Usando *C* para los miembros de esta clase tenemos el siguiente simbolismo:

*C* ↑ *I* (*H* hace *A*)

La dirección de ajuste es mundo-a-palabras y la condición de sinceridad es intención. El contenido proposicional es siempre que el hablante *H* realice alguna acción futura *A*. Puesto que la dirección de ajuste es la misma para los comisivos y los directivos tendríamos una taxonomía más elegante si pudiésemos mostrar que son realmente miembros de la misma categoría. Soy incapaz de hacer esto ya que mientras que el objeto de una promesa es comprometer al hablante a hacer algo (y no necesariamente inducirse a sí mismo a hacerlo) el objeto de una petición es intentar lograr que el oyente haga algo (y no necesariamente comprometerlo u obligarlo a que lo haga). Para asimilar las dos categorías sería necesario haber mostrado que las promesas son realmente especies de peticiones a uno mismo (esto me

<sup>6</sup> Las interrogaciones son directivos ya que son intentos de lograr que el oyente realice un acto de habla.

ha sido sugerido por Julian Boyd) o alternatively se debería haber mostrado que las peticiones colocar al oyente bajo una obligación (esto me fue sugerido por William Alston y John Kearns). No he podido llevar a cabo ninguno de estos dos análisis y me quedo con la inelegante solución de dos categorías separadas con la misma dirección de ajuste.

Una cuarta categoría que llamaré:

*Expresivos*: El objeto ilocucionario de esta clase es expresar el estado psicológico especificado en la condición de sinceridad sobre el estado de cosas especificado en el contenido proposicional. Los paradigmas de los verbos expresivos son «dar las gracias», «congratularse», «pedir disculpas», «dar el pésame», «deplorar» y «dar la bienvenida». Nótese que en los expresivos no existe dirección de ajuste. Al realizar un expresivo el hablante no está ni intentando hacer que el mundo encaje con las palabras ni que las palabras encajen con el mundo; más bien se presupone la verdad de la proposición expresada. Así, por ejemplo, cuando pido disculpas por haberte pisado no es mi propósito ni afirmar que tu pie ha sido pisado, ni inducir a que se te pise. Este hecho se refleja claramente en las sintaxis (del inglés) por el hecho de que los verbos expresivos paradigmáticos, en su ocurrencia realizativa, no toman las cláusulas *que* sino que requieren una transformación nominal de gerundio (o alguna otra substantivación) \*\*. No se puede decir:

\* *I apologize that I stepped on your toe*

más bien el inglés correcto es

*I apologize for stepping on your toe.*

Del mismo modo no puede decirse:

\* *I congratulate you that you won the race*

ni

\* *I thank you that you paid me the money.*

Debe decirse:

*I congratulate you on winning the race (congratulations on winning the race).*

\*\* En castellano se requiere un infinitivo. Así, no es correcto decir \*«Pido disculpas de que te he pisado», sino «Pido disculpas por haberte pisado». Debido a esta falta de correspondencia entre el inglés y el castellano prefiero dejar el original tal como está. (N. del T.)



*I thank you for paying the money (thanks for paying me the money).*

Estos hechos sintácticos, sugiero, son consecuencias del hecho de que en general no existe dirección de ajuste en los expresivos. La verdad de la proposición expresada en un expresivo se presupone. La simbolización de esta clase debe, por consiguiente, proceder como sigue:

$$E \emptyset (P) (H / O + \text{propiedad})$$

donde **E** indica el objeto ilocucionario común a todos los expresivos,  $\emptyset$  es el símbolo de la clase vacía que indica que no existe dirección de ajuste, (**P**) es una variable cuyo rango son los diferentes estados psicológicos posibles expresados en la realización de los actos ilocucionarios de esta clase, y el contenido proposicional adscribe alguna propiedad (no necesariamente una acción) a **H** o a **O**. Puedo felicitarte no sólo por tu victoria en la carrera sino también por tu buen aspecto o por la victoria de tu hijo en la carrera. La propiedad especificada en el contenido proposicional de un expresivo debe, sin embargo, relacionarse con **H** o con **O**. No puedo sin algunas suposiciones muy especiales felicitarte por la primera ley del movimiento de Newton.

Si pudiéramos incluir todos los actos ilocucionarios en estas cuatro clases resultaría económico y proporcionaría algún apoyo adicional al modelo general de análisis adoptado en *Speech Acts*, pero me parece que no es todavía completo. Queda aún una importante clase de casos donde el estado de cosas representado en la proposición expresada es realizado o traído a la existencia por el dispositivo indicador de la fuerza ilocucionaria, casos donde se trae a la existencia un estado de cosas declarando que existe, casos donde, por así decirlo, «decir es hacer». Ejemplos de estos casos son «Dimi to», «Quedas despedido», «Te excomulgo», «Bautizo este barco como *Acorazado Missouri*» «Te nombro presidente» y «Se declara la guerra». Estos casos se presentaron como paradigmas en las primitivas discusiones sobre los realizativos, pero me parece que están aún descritos inadecuadamente en la literatura sobre el tema y su relación con los otros géneros de actos ilocucionarios es frecuentemente mal comprendida. Llamemos a esta clase:

*Declaraciones:* La característica definitoria de esta clase es que la realización con éxito de uno de sus miembros da lugar a la correspondencia entre contenido proposicional y la realidad. La realización con éxito garantiza que el contenido proposicional corresponde al mundo: si realizo con éxito el acto de nombrarte presidente, entonces eres presidente; si realizo con éxito el acto de proclamarte candi-

dato, entonces eres candidato; si realizo con éxito el acto de declarar el estado de guerra, entonces la guerra ha comenzado; si realizo con éxito el acto de casarte, entonces estás casado.

La estructura sintáctica superficial de muchas oraciones utilizadas para realizar declaraciones nos oculta este punto dado que en ellas no hay distinción sintáctica superficial entre el contenido proposicional y fuerza ilocucionaria. Así «Quedas despedido» y «Dimite» no parece permitir una distinción entre fuerza ilocucionaria y contenido proposicional pero, de hecho, pienso que en su uso para realizar declaraciones su estructura semántica es:

Declaro: tu empleo (por la presente) se termina.

Declaro: mi cargo (por la presente) se termina.

Las declaraciones dan lugar a una alteración en el *status* o condición de lo referido al objeto u objetos en virtud solamente del hecho de que la declaración ha sido realizada con éxito. Esta característica de las declaraciones las distingue de las otras categorías. En la historia de la discusión de estos tópicos desde la primera introducción de esta distinción entre realizativos y constatativos por parte de Austin, este rasgo de las declaraciones no ha sido debidamente comprendida. La distinción original entre realizativos y constatativos se suponía que era una distinción entre emisiones que son dichos (constatativos: enunciados, aserciones, etc.) y emisiones que son hechos (realizativos: promesas, apuestas, avisos, etc.). Lo que llamo declaraciones estaba incluido en la clase de los realizativos. El tema principal de la obra de madurez de Austin, *How to Do Things with Words*, es que esta distinción se derrumba. Lo mismo que decir que ciertas cosas constituyen casarse (un «realizativo») y que decir que ciertas cosas constituyen hacer una promesa (otro «realizativo»), así decir ciertas cosas constituye hacer un enunciado (supuestamente un constatativo). Como Austin vio, pero como muchos filósofos no aciertan aún a ver, el paralelo es exacto. Hacer un enunciado es tanto realizar un acto ilocucionario como hacer una promesa, una apuesta, un aviso o lo que se quiera. Cualquier emisión consistirá en la realización de uno o más actos ilocucionarios.

El dispositivo indicador de fuerza ilocucionaria en la oración opera sobre el contenido proposicional para indicar, entre otras cosas, la dirección de ajuste entre el contenido proposicional y la realidad. En el caso de los representativos la dirección de ajuste es palabras-a-mundo, en el caso de los directivos y conmisivos es mundo-a-palabras, en el caso de los expresivos la fuerza ilocucionaria no conlleva ninguna dirección de ajuste puesto que la existencia del ajuste se presupone. La emisión no puede despegar a menos que haya previamente un ajuste. Pero ahora con las declaraciones descu-

brimos una relación muy peculiar. La realización de una declaración da lugar a un ajuste en virtud de su realización con completo éxito. ¿Cómo es posible tal cosa?

Obsérvese que todos los ejemplos que hemos considerado hasta aquí incluyen una institución extralingüística, un sistema de reglas constitutivas además de las reglas constitutivas del lenguaje, para que la declaración pueda ser realizada con éxito. El dominio de aquellas reglas que constituyen la competencia lingüística del hablante no es, en general, suficiente para la realización de una declaración. Debe existir además una institución extralingüística y hablante y oyente deben ocupar lugares especiales dentro de esa institución. Solamente dadas instituciones tales como la iglesia, la ley, la propiedad privada, el estado y una posición especial del hablante y el oyente dentro de esas instituciones se puede excomulgar, nombrar para un cargo, transmitir y legar las propias posesiones, o declarar la guerra.

Las únicas excepciones al principio de que toda declaración requiere una institución extralingüística son aquellas declaraciones que conciernen al lenguaje mismo<sup>7</sup>, como por ejemplo cuando se dice «defino, abrevio, nombro, llamo o estipulo». Austin habla a menudo como si todos los realizativos (y en la teoría general, todos los actos ilocucionarios) requiriesen alguna institución extralingüística, pero esto no es, evidentemente, el caso. Las declaraciones son una categoría muy especial de actos de habla. Simbolizaremos su estructura como sigue:

$$D \updownarrow \emptyset (p)$$

donde **D** indica el objeto ilocucionario declaracional; la dirección de ajuste es tanto palabras-a-mundo como mundo-a-palabras a causa del carácter peculiar de las declaraciones; no hay condición de sinceridad, de ahí que tengamos el símbolo de la clase vacía en el hueco de la condición de sinceridad y empleamos la variable proposicional acostumbrada *p*.

La razón de que aquí tenga que haber una flecha de relación-de-ajuste es que las declaraciones intentan conseguir que el lenguaje encaje con el mundo. Pero no son intentos de hacerlo ni describiendo un estado de cosas existente (como hacen los representativos) ni intentando lograr que nadie dé lugar a un futuro estado de cosas (como hacen los directivos y los conmisivos).

Algunos miembros de la clase de las declaraciones se superponen a miembros de la clase de los representativos. Esto sucede a causa de que en ciertas situaciones institucionales no solamente averiguamos

<sup>7</sup> Otra clase más especial de excepciones concierne a lo sobrenatural. Cuando Dios dice «Hágase la Luz», esto es una declaración.

los hechos sino que necesitamos una autoridad para asentar una decisión en cuanto a cuáles son los hechos después que se ha llevado a cabo el procedimiento de encontrarlos. La argumentación debe llegar eventualmente a un fin y dar por resultado una decisión y es por esto por lo que tenemos jueces y árbitros. Tanto el juez como el árbitro realizan afirmaciones fácticas: «penalti» o «culpable». Tales afirmaciones son evaluables claramente en la dimensión de ajuste palabra-a-mundo. ¿Tocó, realmente, el jugador la pelota dentro del área? ¿cometió él, realmente, el crimen?, son evaluables en la dimensión palabra-a-mundo. Pero, al mismo tiempo, ambas tienen la fuerza de declaraciones. Si el árbitro te pita un penalti (y sostiene lo dicho después de la apelación), entonces para los propósitos del fútbol tú has cometido un penalti, sin tener en cuenta los hechos del caso y si el juez te declara culpable (después de la apelación) entonces para los propósitos legales eres culpable. No hay nada misterioso en estos casos. Las instituciones requieren característicamente actos ilocucionarios que tienen la fuerza de declaraciones y que son proferidos por autoridades de varios géneros. Algunas instituciones requieren afirmaciones representativas que son proferidas con la fuerza de declaraciones para que la disputa sobre la verdad de la afirmación pueda llegar a un fin alguna vez y los siguientes pasos institucionales que descansan en el establecimiento de la solución fáctica puedan proceder: el prisionero es liberado o enviado a la cárcel, y el penalti se tira. A los miembros de esta clase los podemos denominar «declaraciones representativas». A diferencia de las otras representaciones comparten con los representativos una condición de sinceridad. El juez, jurado y árbitro pueden, lógicamente hablando, mentir, pero la persona que declara la guerra o te nombra no puede mentir en la realización de su acto ilocucionario. El simbolismo para la clase de las declaraciones representativas es, entonces, éste:

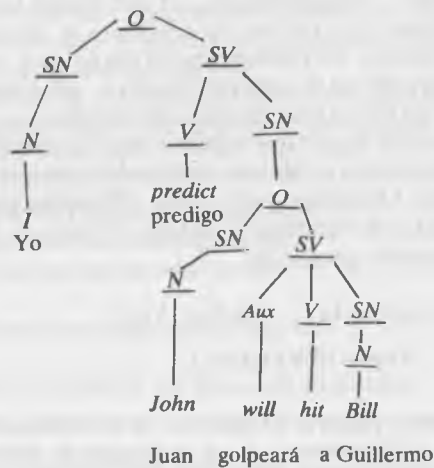
$$D_r \downarrow \uparrow C(p)$$

donde  $D_r$  indica el objeto ilocucionario de emitir un representativo con la fuerza de una declaración; la primera flecha indica la dirección de ajuste representativa, la segunda la dirección de ajuste declarativa, la condición de sinceridad es creencia y la  $p$  representa el contenido proposicional.

## V. ALGUNOS ASPECTOS SINTÁCTICOS DE LA CLASIFICACIÓN

Hasta aquí he clasificado los actos ilocucionarios y he utilizado hechos sobre verbos a modo de prueba e ilustración. En esta sección

quiero discutir explícitamente algunos puntos sobre la sintaxis del inglés \*. Si las distinciones señaladas en la sección IV tienen algún significado real, tendrán, igualmente, diversas consecuencias sintácticas y ahora me propongo examinar la estructura profunda de las oraciones realizativas explícitas en cada una de las cinco categorías; esto es, quiero examinar la estructura sintáctica de oraciones que contienen la ocurrencia realizativa de verbos ilocucionarios apropiados a cada una de las cinco categorías. Puesto que todas las oraciones que consideraremos contendrán un verbo realizativo en la cláusula principal y una cláusula subordinada, abreviaré las estructuras de árbol usuales de la siguiente manera: por ejemplo, la oración: «Yo predigo que Juan golpeará a Guillermo» (*I predict John will hit Bill*) tiene la estructura profunda mostrada en el diagrama introducido a continuación. Abreviaré esto simplemente como: «Yo predigo + Juan golpeará a Guillermo» (*I predict + John will hit Bill*). Los paréntesis se utilizarán para señalar elementos opcionales o elementos que son obligatorios solamente para clases restringidas de los verbos en cuestión. Donde existe una elección entre uno o dos elementos pondré una barra entre ellos, por ejemplo «yo» y «tú».



**Representativos.** La estructura profunda de oraciones representativas paradigmáticas tales como: «Enuncio que está lloviendo» y «Predigo que él vendrá», es simplemente:

Yo verbo (que) + O

\* Que, eventualmente, se aplicarán al castellano. (N del T.)

Esta clase, como tal clase, no presenta restricciones adicionales aunque algunos verbos particulares pueden presentar restricciones adicionales en el nudo inferior *O*. Por ejemplo, «predecir» requiere que en el nudo inferior *O* deba haber un *Aux* de futuro \* o que de cualquier modo no pueda estar en pasado. Verbos representativos tales como «describir», «llamar», «clasificar» e «identificar» toman una estructura sintáctica diferente similar a muchos verbos de declaración que discutiré más adelante.

*Directivos*. Oraciones tales como «Te ordeno salir» o «Te mando prestar atención» tienen la estructura profunda siguiente:

Yo verbo tú + tú Fut.

Verbo (*SN*) (*Adv.*)

«Te ordeno salir» es entonces la realización de estructura superficial de «Yo te ordeno + tú saldrás» con eliminación del *SN* del repetido «tú». Obsérvese que un argumento sintáctico adicional para mi inclusión de «vetar», «desafiar» y «provocar» en mi lista de verbos directivos y objetar a Austin su inclusión al lado de «pedir disculpas», «dar las gracias», «congratularse», es que tienen la misma forma sintáctica que tienen los verbos paradigmáticas directivos «ordenar», «mandar» y «pedir». Del mismo modo «invitar» y «advertir» (en uno de sus sentidos) tienen la sintaxis directiva. «Permitir» tiene también la sintaxis de los directivos aunque dar permiso no es, estrictamente hablando, intentar lograr que alguien haga algo; más bien consiste en eliminar restricciones existentes precedentemente sobre su acción.

*Conmisivos*. Oraciones tales como «Prometo pagarte el dinero», «Juro fidelidad a la bandera» y «Hago el voto de tomar venganza» tienen la estructura profunda:

Yo verbo (tú) + Yo Fut. Vol

Verbo (*SN*) (*Adv.*).

Así «Prometo pagarte el dinero» es la realización de estructura superficial de «Yo prometo + yo te pagaré el dinero» con eliminación del *SN* del repetido «Yo». La diferencia sintáctica entre «Te prometo venir el miércoles» y «Te ordeno venir el miércoles», la percibimos como consistiendo en que en la primera «Yo» es el sujeto de estructura profunda de «venir» y «tú» es el sujeto de estructura profunda en la segunda tal como requieren los verbos «prometer» y «or-

\* Obviamente en inglés. En castellano requeriría simplemente una forma verbal en futuro. (*N. del T.*)

denar”, respectivamente. Nótese que no todos los conmisivos paradigmáticos tienen «tú» como objeto indirecto del verbo realizativo. En la oración «Juro fidelidad a la bandera» la estructura profunda no es «Yo juro a ti bandera + yo seré fiel». Es

Yo juro + yo seré fiel a la bandera

Mientras que existen argumentos puramente sintácticos a favor de que verbos directivos paradigmáticos tales como «ordenar» y «mandar», así como el modo imperativo requieran «tú» como sujeto de estructura profunda del nudo inferior *O*, no conozco ninguna argumentación puramente sintáctica que muestre que los conmisivos requieran «yo» como sujeto de estructura profunda en su nudo inferior *O*. Semánticamente, en efecto, debemos interpretar oraciones tales como «Prometo que Enrique estará aquí el miércoles» como significando.

Prometo que *yo haré* que Enrique esté aquí el próximo miércoles,

hasta el punto en que interpretemos la emisión como una promesa genuina, pero no conozco argumentos puramente sintácticos que muestren que la estructura profunda de la primera oración contiene los elementos en cursiva de la segunda.

*Expresivos*. Como he mencionado anteriormente los expresivos requieren una transformación de gerundio del verbo en el nudo inferior *O*. Decimos \*\*:

*\* I apologize for stepping on your toe,  
I congratulate you on winning the race,  
I thank you for giving me the money*

La estructura profunda de tales oraciones es:

Yo verbo tú + Yo/tú *SV* ⇒ Nom. de gerundio

\*\* En castellano, contrariamente al inglés donde no se permiten en el caso de los *expresivos* oraciones de infinitivo en el nudo inferior *O*, la estructura profunda de una oración con un verbo expresivo sería de modo general:

Yo verbo tú + Yo/Tú *SV* ⇒ Inft. (Nom.) (Adv.).

Así, «Pido disculpas por haberme portado mal» es una oración correcta del castellano mientras que no lo es \*«Pido disculpas de que me he portado mal». Por otra parte, también en castellano es correcto utilizar una transformación de gerundio con los expresivos aunque este empleo no sea demasiado corriente: así, el caso en el que alguien enfáticamente dice: «Habiéndome portado mal, pido disculpas.» (*N. del T.*)

Y, repitiendo, la explicación de la obligatoriedad del gerundio es que no existe dirección de ajuste. Las formas que corrientemente dan lugar a cuestiones concernientes a la dirección de ajuste, cláusulas *que* e infinitivos, no son permisibles. De ahí la imposibilidad de

- \* *I congratulate you that you won the race,*
- \* *I apologize to step on your toe.*

Sin embargo, no todas las transformaciones de nominalización permisibles son gerundios; lo esencial es solamente que no se deben producir cláusulas *que* u oraciones de infinitivo \*\*; así podemos tener o

*I apologize for behaving badly*

o

*I apologize for my bad behaviour.*

Pero no

- \* *I apologize that I behaved badly,*
- \* *I apologize to behave badly.*

Antes de considerar las declaraciones quiero resumir la discusión de aquellos verbos representativos que tienen una sintaxis diferente de los paradigmas anteriores. He dicho que los representativos paradigmáticos tienen la forma sintáctica

Yo verbo (que) + O

Pero si consideramos verbos representativos tales como «diagnosticar», «llamar» y «describir», así como «enclasar», «clasificar» e «identificar», encontramos que no se ajustan completamente a este modelo. Consideremos «llamar», «describir» y «diagnosticar» en oraciones tales como:

*I call him a liar,*  
*I diagnose his case as appendicitis,*  
*I describe John as a Fascist.*

\*\* En castellano también es posible utilizar como complemento de verbos expresivos otras transformaciones distintas del infinitivo y del gerundio. Así, es castellano correcto, p. ej., decir: «Pido disculpas por mi mala conducta.» (N. del T.)



y en general la forma de esto es:

Yo verbo  $SN_1 + SN_1$  ser pred.

No se puede decir:

\* *I call that he is a liar,*

\* *I diagnose that his case is appendicitis*

(perversamente algunos de mis alumnos encuentran aceptable esta forma),

\* *I describe that John is a Fascist* \*\*

Aquí parece, por lo tanto, existir un severísimo conjunto de restricciones sobre una clase importante de verbos representativos que no es compartida por los otros paradigmas. ¿Nos justificaría esto la conclusión de que estos verbos fueron erróneamente clasificados como representativos al lado de «enunciar», «aseverar», «afirmar» y «predecir» y que necesitamos una clase separada para ellos? He oído argüir que la existencia de esos verbos justifica la afirmación de Austin de que se requiere una clase separada de veredictivos distinta de los expositivos, pero esto sería, seguramente, sacar una conclusión muy curiosa ya que Austin incluye la mayor parte de los verbos que mencionamos arriba en la lista de los expositivos. Incluye «describir», «enclasar», «identificar» y «llamar» en los expositivos y «diagnosticar» y «describir» en los veredictivos. Una sintaxis común a muchos veredictivos y expositivos garantizaría difícilmente la necesidad de los veredictivos como una clase separada. Pero dejando a un lado la taxonomía de Austin surge todavía la pregunta, ¿necesitamos una categoría semántica separada para dar cuenta de esos hechos sintácticos? Creo que no. Creo que hay una explicación mucho más simple de la distribución de estos verbos. A menudo, en el discurso representativo, enfocamos nuestra atención sobre algún tópico de discu-

\*\* En castellano los representativos adoptan, en general, las dos formas indistintamente:

Yo verbo (que) + *O* y

Yo verbo  $SN_1 + SN_1$  ser pred.

Es indiferente decir «Diagnosticó su caso como apendicitis» o «Diagnosticó que su caso de apendicitis». No obstante, pueden encontrarse excepciones a esta ambivalencia. Por ejemplo en el caso del verbo «describir» no puede utilizarse la primera forma. Así no puede decirse «Describo que Juan es un fascista», sino «Describo a Juan como un fascista». Tampoco, como se habrá observado, en el caso del verbo «llamar». (*N. del T.*)

sión. La cuestión no es justamente cuál es el contenido proposicional que estamos afirmando sino qué decimos sobre el *objeto(s)* al que nos referimos en el contenido proposicional: no precisamente lo que enunciamos, afirmamos, caracterizamos o aseveramos, sino cómo describimos, llamamos, diagnosticamos o identificamos algún tópico de discusión al que previamente nos hemos referido. Cuando, por ejemplo, es un asunto de diagnosticar o de describir es siempre un asunto de diagnosticar a una persona o su caso, de describir un paisaje, una reunión o una persona, etc. Estos verbos ilocucionarios representativos nos proporcionan un dispositivo para aislar los tópicos de lo que se dice sobre ellos. Pero esta diferencia sintáctica tan genuina no señala una diferencia semántica lo suficientemente grande para justificar la formación de una categoría separada. Obsérvese aquí en apoyo de mi argumentación que las oraciones efectivas en las que se hace la descripción, el diagnóstico, etc., son rara vez del tipo realizativo explícito, sino que son más bien usualmente, en el tipo *ex-tándar*, formas indicativas que son características de la clase representativa.

Emisiones de:

Él es un mentiroso,  
Él tiene apendicitis,  
Él es un fascista

son todas característicamente enunciados al hacer los cuales llamamos, diagnosticamos y describimos al igual que acusamos, identificamos y caracterizamos. Concluyo entonces que hay típicamente dos formas sintácticas para los verbos ilocucionarios representativos: una de ellas se concentra en el contenido proposicional, la otra en el (los) *objeto(s)* a que nos referimos en el contenido proposicional, pero ambas son semánticamente representativas.

*Declaraciones.* Menciono la forma sintáctica

Yo verbo  $SN_1 + SN_1$  ser pred.

tanto para adelantar una argumentación para erigir una categoría semántica separada para ellos como a causa de que muchos verbos de declaración tienen esta forma. En efecto, parece haber diversas y diferentes formas sintácticas para realizativos explícitos de declaración. Creo que las tres clases siguientes son las más importantes:

1. Te declaro culpable (como acusado).  
En este momento os declaro marido y mujer.  
Te nombro presidente.

2. Se declara la guerra (por la presente).  
Declaro clausurada la sesión.
3. Quedas despedido.  
Dimito.  
Te excomulgo.

La estructura sintáctica profunda de las tres es, respectivamente, como sigue:

1. Yo verbo  $SN_1$  +  $SN_1$  ser pred.

Entonces en nuestros ejemplos tenemos:

Yo te declaro + tú eres culpable (como acusado).  
Yo os declaro + vosotros sois marido y mujer.  
Yo te nombro + tú eres presidente.

2. Yo declaro +  $O$ .

Entonces en nuestros ejemplos tenemos:

Yo/nosotros (por la presente) declaro/declaramos + un estado de guerra existe.  
Yo declaro + la reunión es clausurada.

Esta forma es la forma más pura de la declaración: el hablante constituido en autoridad origina un estado de cosas especificado en el contenido proposicional diciendo, en efecto: declaro que el estado de cosas existe. Semánticamente todas las declaraciones tienen este carácter aunque en la clase 1 el centrarse en el tópico produce una alteración en la sintaxis que es exactamente la misma sintaxis que vimos en verbos representativos tales como «describir», «caracterizar», «llamar» y «diagnosticar» y en la clase 3 la sintaxis oculta aún más la estructura semántica.

3. La sintaxis de éstos es la más desorientadora; es simplemente:

Yo verbo ( $SN$ )

como en nuestros ejemplos:

Yo te despido.  
Yo dimito.  
Yo te excomulgo.

Su estructura semántica me parece, sin embargo, que es la misma que la de la clase 2. «Quedas despedido» si es emitida como realización del acto de despedir a alguien y no como un informe significa:

Yo declaro + Tu empleo se termina (por la presente).

Similarmente «Por la presente dimito» significa:

Yo (por la presente) declaro + Mi cargo se termina (por la presente).

«Te excomulgo» significa:

Yo declaro + Tu pertenencia a la iglesia se termina (por la presente).

La explicación de la desorientadoramente simple estructura de estas oraciones parece residir en el hecho de que tenemos algunos verbos que en su ocurrencia realizativa engloban tanto la fuerza de clarativa como el contenido proposicional.

## VI. CONCLUSIONES

Estamos ahora en posición de plantear algunas conclusiones generales.

1. Muchos de los verbos que llamamos verbos ilocucionarios no son marcadores de objeto ilocucionario sino de alguna otra característica del acto ilocucionario. Considérense «insistir» y «sugerir». Puedo insistir en que vayamos al cine o puedo sugerir que vayamos al cine, pero puedo también insistir en que la respuesta se haya en la página 16 o puedo sugerir que se encuentra en la página 16. El primer par es directivo, el segundo par es representativo. ¿Muestra esto que insistir y sugerir son actos ilocucionarios totalmente diferentes de los representativos y los directivos o quizás que son tanto representativos como directivos? Creo que la respuesta a ambas cuestiones es *no*. Tanto «insistir» como «sugerir» se usan para señalar el grado de intensidad con el que es presentado el objeto ilocucionario. No señalan un objeto ilocucionario separado en absoluto. Del mismo modo «anunciar», «presentar» y «hacer confidencias» no señalan objetos ilocucionarios separados sino más bien el estilo o manera de realización de un acto ilocucionario. Por paradójico que pueda parecer tales verbos son verbos ilocucionarios pero no nombres de géneros

de actos ilocucionarios. Es por esta razón, entre otras, por lo que debemos distinguir cuidadosamente una taxonomía de actos ilocucionarios de una de verbos ilocucionarios.

2. En la sección IV intenté clasificar actos ilocucionarios y en la sección V intenté explorar algunas de las características sintácticas de los verbos que denotan miembros de cada una de las categorías. Pero, no intenté clasificar verbos ilocucionarios. Si se hubiese hecho así creo que surgirían las cuestiones siguientes:

a) Primeramente, como he señalado, algunos verbos no designan ningún objeto ilocucionario sino algún otro rasgo; por ejemplo, «insistir», «sugerir», «anunciar», «hacer confidencias», «replicar», «responder», «interrumpir», «hacer observaciones», «exclamar» e «interponer».

b) Muchos verbos señalan un objeto ilocucionario más algún otro rasgo; por ejemplo, «jactarse», «lamentar», «amenazar», «criticar», «acusar» y «aconsejar» añaden todos el rasgo de bondad o maldad a su objeto ilocucionario primario.

c) Unos pocos verbos señalan más de un objeto ilocucionario; por ejemplo una *protesta* incluye tanto una expresión de desaprobación como una petición de cambio.

*Promulgar una ley* tiene tanto un *status* declaracional (el contenido proposicional se convierte en ley) como un *status* directivo (la ley es directiva por lo que respecta a su intención). Los verbos de declaración representativa se incluyen en la clase de los verbos con dos objetos ilocucionarios.

d) Un puñado de verbos puede tomar diferentes objetos ilocucionarios en diferentes emisiones. Considérense «advertir», «avisar» y «aconsejar». Nótese que ambos toman la sintaxis directiva o la sintaxis representativa. Así

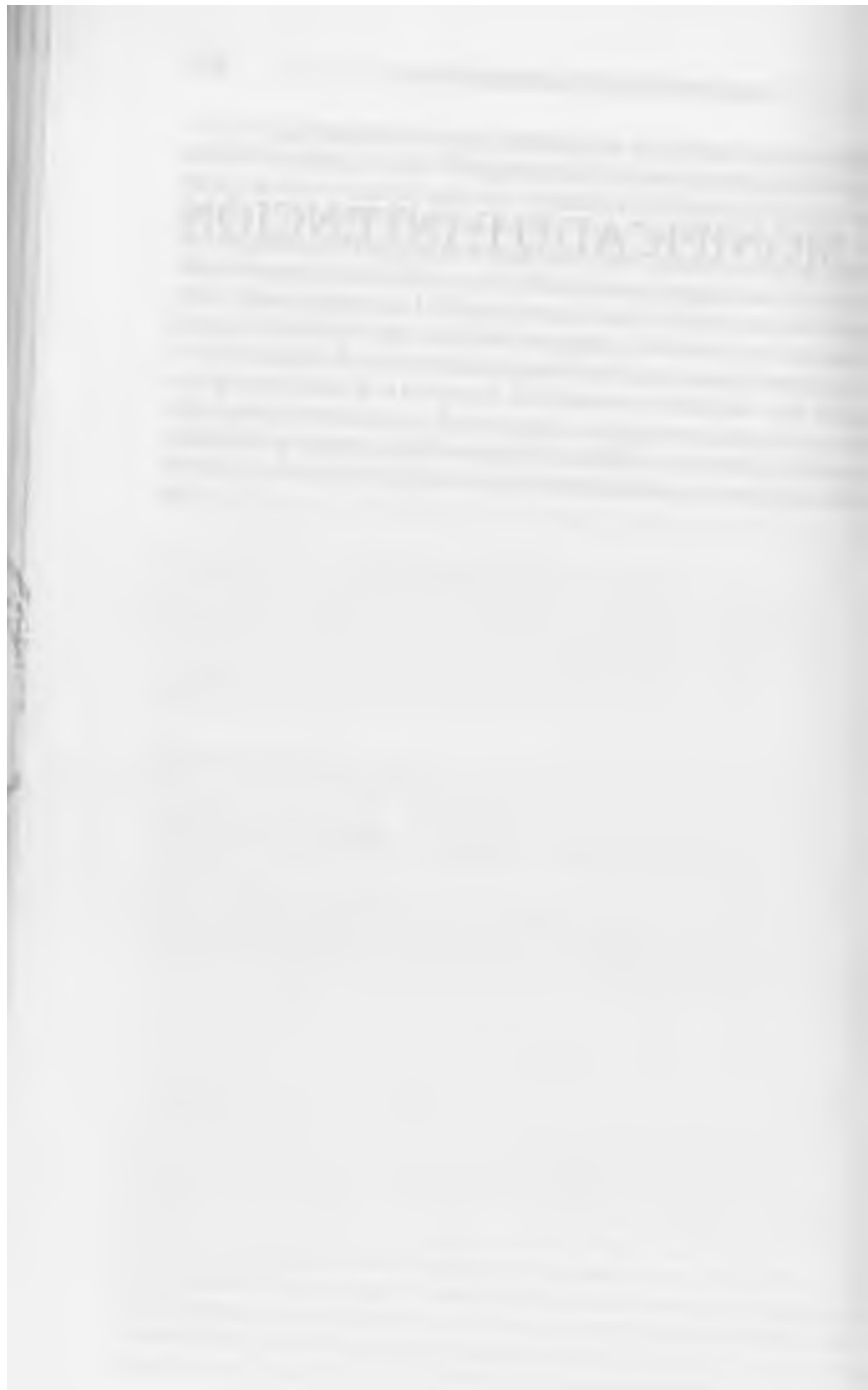
¡Te advierto que te apartes de mi	
esposa!	(directivo)
Te advierto que ese toro va a atacar	(representativo)
Te aconsejo que marches	(directivo)
Por la presente se avisa a los pasajeros	
que el tren llegará con retraso	(representativo).

Del mismo modo me parece que advertir, avisar y aconsejar pueden ser o decirte *que* algo es el caso (con relevancia de lo que es o no es de tu interés) o decirte *que* hagas algo respecto de ello (puesto que es o no es de tu interés). Pueden también, aunque no necesariamente, ser ambas cosas a la vez.

3. La conclusión más importante a extraer de esta discusión es ésta. No existe, como Wittgenstein (en una interpretación posible) y muchos otros han afirmado, un número indeterminado o infinito de juegos o usos del lenguaje. Mas bien la ilusión de los usos ilimitados del lenguaje se origina por una falta de claridad enorme sobre lo que constituyen los criterios para delimitar un juego o uso del lenguaje de otro. Si adoptamos el objeto ilocucionario como noción básica sobre la cual clasificar los usos del lenguaje, entonces existe un número muy limitado de cosas básicas que hacemos con el lenguaje: decimos a la gente cómo son las cosas, intentamos conseguir que hagan cosas, nos comprometemos a hacer cosas, expresamos nuestras creencias o actitudes y damos lugar a cambios mediante nuestras emisiones. Y, a menudo, hacemos más de una de estas cosas a la vez en la misma emisión.

## VI. SIGNIFICADO E INTENCIÓN

[The following text is extremely blurry and illegible. It appears to be a series of paragraphs of text, but the content cannot be transcribed accurately.]





En la filosofía del lenguaje contemporánea hay (al menos) dos enfoques rivales de lo que podemos denominar «el problema del significado». Algunos artículos de Donald Davidson, recogidos bajo el encabezamiento «Significado y verdad» (particularmente «Verdad y significado» e «Interpretación radical»), representan uno de ellos. El otro tiene su paradigma en lo que ha dado en llamarse «teoría causal (del significado)», desarrollado en la obra de H. P. Grice. Pero, mientras que la estrategia de Davidson consiste en investigar la *forma* correcta de una teoría del significado (lo que le lleva, en última instancia, a una explicación del significado en términos de la empresa epistemológica incluida en la doctrina de la interpretación radical), y su interés se centra en la *interpretación* por parte del oyente de una oración, el interés de Grice se dirige, primariamente, hacia la investigación del *significado del hablante*, como opuesto al significado de una emisión.

Ahora bien, ¿cómo se substancia esa investigación del significado del hablante? Grice ha diseñado su teoría del significado en los términos siguientes. Al proferir una emisión un hablante intenta comunicar algo y, a la vez, intenta que su intención comunicativa se reconozca por su oyente: intenta, por ejemplo, inducir en él una creencia o lograr que lleve a cabo determinada acción *mediante el reconocimiento de su intención* (de la del hablante). Pero este reconocimiento no se produce por ensalmo: el hablante intenta que se logre *vía* ciertos rasgos o propiedades de su emisión (y esto, naturalmente, establece restricciones sobre las preferencias comunicativas que un hablante hace: *supuesto que la conducta de éste sea racional*, simplemente no puede querer decir cualquier cosa mediante cualquier cosa e intentar, además, que se reconozca su intención comunicativa). Una de las premisas básicas del enfoque de Grice es, entonces, su énfasis en el carácter eminentemente *racional* que preside la conducta verbal humana.

En «Las intenciones y el significado del hablante» se desarrolla el proyecto expuesto en 1957 («Meaning», *Philosophical Review*, vol. 66), modificando ligeramente la primitiva noción de significado con el objeto de hacer frente a algunos contraejemplos (especialmente los aducidos por Strawson, Schiffer y Searle): los pretendidos efectos de significado no son ahora intentos de que el oyente haga o crea *efectivamente* algo, sino más bien de que el oyente intente hacer algo o de que éste crea que el hablante cree algo. La racionalidad de la conducta verbal como premisa general de la explicación del significado está presente también en «Lógica y conversación». Es una noción común que en numerosas ocasiones los hablantes dicen mucho

más (o algo distinto) de lo que literalmente significan sus emisiones. Partiendo de un conjunto de máximas que guiarían idealmente una conversación y supuesta la racionalidad del intercambio lingüístico, Grice bosqueja —*inter alia*— en este artículo, en términos lo que él denomina «implicaturas conversacionales», cómo es posible que las violaciones patentes de esas máximas produzcan algún beneficio para la comunicación.

## LAS INTENCIONES Y EL SIGNIFICADO DEL HABLANTE \*

H. P. GRICE

### I. INTRODUCCIÓN

Dentro de la gama de usos de la palabra «significar» que se hallan especialmente vinculados a la comunicación (usos, por lo tanto, de la palabra «significar» que responden de una u otra forma a lo que he llamado sentidos *no-naturales*), hay distinciones por hacer. Consideremos la siguiente oración (*O*):

«Si mantengo los circuitos abiertos, podría recibir la noticia»<sup>1</sup>.

1 a) Sería aproximadamente verdadero decir que *O* significa (tiene como uno de sus significados) «Si mantengo abiertas las líneas, podría recibir la noticia». Quizás también sería aproximadamente verdadero decir que (*O*) significa (tiene como otro de sus significados) «Si mantengo abiertas las pistas para competiciones de vehículos a motor, podría recibir la noticia». A semejantes especificaciones

\* Versión castellana de J. José Acero.

Tengo hacia mi colega Stephen Schiffer una deuda mayor de lo que indica el texto de este escrito por los comentarios, críticas y sugerencias que he estado recibiendo de él durante un considerable período de tiempo.

Este escrito fue presentado en el Philosophy Colloquium del Oberlin College en abril de 1968. Una versión corregida de material contenido en él formará parte, espero, de un libro que próximamente publicará Harvard University Press. [Ese libro fue editado en 1989 bajo el título *Studies in the Way of Words*. (*N. del T.*)]

<sup>1</sup> El ejemplo original del autor es «If I shall then be helping the grass to grow. I shall have no time for reading», que reúne un conjunto tal de características que su versión literal castellana ya no sirve a los propósitos que llevan a Grice a ponerlo. En efecto, además del sentido literal propio, contiene la palabra «hierba», que en argot significa también «marihuana», no se basa en ninguna frase hecha o giro reconocido y puede también ser utilizada por alguien para querer decir (en inglés): «Si estoy muerto [sirvo de abono a la hierba], no sabré lo que pasa en el mundo.» Esos cuatro aspectos del caso están presentes en el ejemplo de la versión castellana. Ese ejemplo es una adaptación del siguiente fragmento de diálogo perteneciente a la novela de Ross Macdonald *The Wycherly Woman* (versión castellana de Alicia Steimberg, en Barcelona: Editorial Alfa, 1975): «—¿Para qué sirve rezar? —Mantiene los circuitos abiertos. En caso de que haya alguien al otro lado de la línea.» (*N. del T.*)

de significado las denominaré especificaciones del *significado atemporal* de una preferencia-tipo «completa» (que puede ser una preferencia-tipo de una oración o de algo «como una oración», tal como una señal hecha con una mano).

1 b) Resultaría verdadero decir que la palabra «circuito» significa (aproximadamente) «línea que forman los conductores por donde circula una corriente eléctrica» y también sería verdadero decir que «circuito» significa «pistas para competiciones de vehículos a motor». A tales especificaciones de significado las llamaré especificaciones del *significado atemporal* de una preferencia-tipo «incompleta» (que puede ser una palabra o una frase no-oracionales o una preferencia-tipo no-lingüística análoga a una palabra o frase).

2 a) Puesto que una preferencia-tipo completa  $x$  puede tener más de un significado atemporal, necesitamos poder vincular a una preferencia particular de  $x$  exactamente uno de los significados atemporales de  $x$  y excluir los demás. Necesitamos poder decir, a propósito de una preferencia particular de  $O$ , que  $O$  significaba *aquí* (en esta ocasión) «Si mantengo abiertas las líneas, podría recibir la noticia» y que «mantengo abiertos los circuitos» significaba *aquí* «mantengo abiertas las líneas que forman los conductores por donde circula una corriente eléctrica». A tales especificaciones de significado las denominaré especificaciones del *significado atemporal aplicado* de una preferencia-tipo completa (en una particular ocasión de preferencia).

2 b) De forma similar, necesitamos poder especificar lo que llamaré el *significado atemporal aplicado* de una preferencia-tipo incompleta; necesitamos poder decir, a propósito de la intervención de la palabra «circuito» en una particular preferencia de  $O$  que *aquí*, en esta ocasión, la palabra «circuito» significaba (aproximadamente) «línea que forma el conductor por donde circula una corriente eléctrica» y no «pista para competiciones de vehículos a motor».

3) Podría ser verdadero decir de un cierto usuario ( $H$ ) de  $O$  que cuando  $H$  profirió  $O$ , él quería decir <sup>2</sup> mediante  $O$  (las palabras de  $O$ ):

(i) «Si Dios existe y no dejo de rezarle, me ayudará»

y puede que, además,

(ii) «Una ventaja de no dejar de rezarle a Dios es que, si existe, me ayudará».

<sup>2</sup> El aspecto más conflictivo de la traducción afecta también a la cuestión conceptual más importante del texto. El verbo inglés *to mean* tiene un uso personal del que «significar» carece en castellano. Los principios semánticos del verbo «significar» excluyen precisamente esa posibilidad por la que diríamos de un hablante  $H$  que *significó* que tal-y-cual (en una determinada ocasión de habla). Sucede, además, que dentro

Si fuera verdadero decir de *H* que, al proferir *O*, quería decir (i) por medio de *O*, también sería verdadero decir de *H* que *mediante la expresión* «mantengo los circuitos abiertos» *H quería decir* «Dios existe y no dejo de rezarle».

En el supuesto (que ahora hago) de que la expresión «mantener los circuitos abiertos» no es una frase hecha, a diferencia de «mantener las espadas en alto», ninguna de estas especificaciones de lo que *H* quiso decir con *O* (o con las palabras «mantengo los circuitos abiertos») acabadas de dar, sería admisible como especificación de un significado atemporal o del significado atemporal aplicado de *O* (o de las palabras que componen el consecuente de *O*). Las palabras «mantengo los circuitos abiertos» ni significan ni significan *aquí* «Dios existe y no dejo de rezarle».

La clase de especificación de significado que acabo de citar la denominaré especificación del *significado ocasional de una proferencia-tipo*.

4) La variedad de especificaciones de significado tenidas en cuenta hasta el momento hacen todas ellas uso de comillas al especificar lo que se quiso decir. El cuarto y último tipo por considerar conlleva, en su lugar, el uso del discurso indirecto. Si fuese verdadero

---

del ideario filosófico de Grice es precisamente este uso el que resulta fundamental para analizar aquellos otros usos (o sentidos) en los que decimos que tal o cual palabra, frase u oración significan tal-o-cual cosa. Éstos se definen a partir de aquél. (Cf. el escrito de Grice al que se refiere la nota 3, así como los trabajos que componen la primera parte de *Philosophical Grounds of Rationality*, editado por R. E. Grandy y R. Warner, Clarendon Press, Oxford, 1986.) Esto crea un problema de traducción para el que no existe una solución completamente satisfactoria. La opción aquí seguida consiste en verter esta acepción o uso de *to mean* por «querer decir». La observación de que el ingrediente «querer» introduce un elemento de arbitraria subjetividad del que *to mean* carece —*to mean* resulta ser lo que muchos lingüistas denominan un verbo de *realización* (*achievement verb*)— señala, el inconveniente de esta opción. Sin embargo, esta cruz del problema la compensa sobradamente, en mi opinión, la insistencia de Grice en que no cualquier intención de comunicar una idea da lugar a un caso de querer decir (en el sentido de que Grice tan escrupulosamente se ocupa). Simplemente, es imposible querer decir algo, si nuestra intención no puede ser reconocida por nuestros interlocutores. Como ha escrito David S. Schwarz, «cuando queremos decir algo, damos a nuestro interlocutor una *razón* para [efectuar] una respuesta y no nos limitamos a causarla» (*Naming and Referring*, Walter de Gruyter, Nueva York, 1979, pg. XXXIV.)

La opción, aparentemente obvia, de traducir estos usos de *to mean* por «decir», como suele verse en los diccionarios al uso, tiene el grave inconveniente de que obligaría a distinguir dos sentidos de «decir» para usos en modo indirecto de este verbo. En el trabajo de la nota 3 define Grice la noción de decir de un modo tal que resulta claramente diferente de la de significado ocasional del hablante. Esta salida, además de dar lugar a una distinción inmotivada y completamente artificial, introduciría un serio malentendido en el programa de reconstrucción de los conceptos semánticos abanderado por Grice. (*N. del T.*)

decir de *H* que quiso decir con *O* (i) (y [ii]), también sería verdadero decir de él que cuando profirió *O* (o que al proferir *O*) *quiso decir que*, si Dios existiera y no dejara de rezarle, éste le ayudaría, y que cuando profirió *O* *quiso decir que* (o *parte de lo que quiso decir que*) una ventaja de no dejar de rezarle a Dios es que, de existir, éste le ayudaría. Incluso si, pese a todo, cuando profirió *O* *quiso decir* «Dios existe y no dejo de rezarle» con las palabras «mantengo los circuitos abiertos», no sería verdadero decir que mediante estas palabras *quiso decir que* Dios existe y que no dejaría de rezarle. Para haber querido decir eso, *H* debería haberse comprometido él mismo a afirmar que Dios existe y que no dejará de rezarle; y esto, cuando profirió *O*, no es algo que haya hecho.

Al tipo 4) de especificaciones de significado lo denominaré especificaciones de *significado ocasional del hablante* [«utterer»].

Podemos entonces distinguir cuatro formas principales de especificaciones de significado:

- 1) «*x* (preferencia-tipo) significa "...» [Especificación de significado *atemporal* de una preferencia-tipo que es bien 1 a) completa bien 1 b) incompleta].
- 2) «*x* (preferencia-tipo) significaba aquí "...» [Especificación de *significado atemporal aplicado* de una preferencia-tipo que es bien 2 a) completa, bien 2 b) incompleta].
- 3) «Con *x* (preferencia-tipo) *H* quiso decir "...» [Especificación de *significado atemporal de una preferencia-tipo*].
- 4) «Al proferir *x* *H* quiso decir que...» [Especificación de *significado ocasional del hablante*].

Hay naturalmente, un elemento de legislación en la distinción de las cuatro formas lingüísticas citadas; no se hallan éstas tan reguladas como estoy pretendiendo por conveniencia.

En un trabajo que se publicará en breve en *Foundations of Language*<sup>3</sup>, titulado «Utterer's Meaning, Sentence-Meaning and Word-Meaning» [«El significado del hablante, el significado de la oración y el significado de la palabra»] considero con algún detalle las relaciones entre el significado atemporal, el significado atemporal aplicado y lo que ahora denomino el significado ocasional del hablante. Comenzando con el supuesto de que la noción de un significado ocasional del hablante puede ser elucidada, de cierto modo, en términos de la de las intenciones de un hablante, arguyo en favor de la tesis de que el significado atemporal y el significado atemporal aplicado pue-

<sup>3</sup> Vol. 4 (1968), pp 1-18. Este escrito ha sido reimpresso en J. R. Searle (ed.), *The Philosophy of Language*, Oxford University Press, 1971, pp. 54-70.

den explicarse en términos de la noción de significado ocasional del hablante (y de otras nociones) y así, finalmente, en términos de la noción de intención. En ese escrito no distingo el significado ocasional de la preferencia-tipo del significado ocasional del hablante; pero, una vez realizada la distinción, no debería resultar demasiado difícil explicar el significado ocasional de la preferencia-tipo en términos del significado ocasional del hablante. La siguiente definición provisional, aunque inadecuada, parece proporcionar un prometedor punto de partida en esta dirección.

Denote « $\sigma(x)$ » una preferencia-tipo completa ( $\sigma$ ) que contenga una preferencia-tipo  $x$ .  $x$  puede ser completa o incompleta, y de hecho puede ser idéntica a  $\sigma$ . Denote « $\varphi$ » una preferencia-tipo. Denote « $\sigma(\varphi/x)$ » el resultado de sustituir  $x$  por  $\varphi$  en  $\sigma$ . Someto entonces a consideración la siguiente laxamente articulada definición:

«Con  $x$   $H$  quiso decir  $\varphi$  si, y sólo si,  $\exists \sigma$  [ $H$  profirió  $\sigma(x)$ , y al proferir  $\sigma(x)$   $H$  quiso decir que... [los puntos suspensivos han de colmarse escribiendo  $\sigma(\varphi/x)$ ]].»

Mi tarea es, sin embargo, la de considerar detenidamente el supuesto adoptado en el escrito al que he estado refiriéndome, a saber: que la noción de significado ocasional del hablante es explicable, en cierta forma, en términos de la noción de intención del hablante, y el resto de este escrito tratará de este tema.

## II. DEFINICIÓN INICIAL DE SIGNIFICADO OCASIONAL DEL HABLANTE

Adoptaré como punto de partida la explicación de significado «no-natural» que ofrecí en mi artículo «Meaning» [«Significado»] (*Philosophical Review*, 1957), considerándola un intento de definir la noción de significado ocasional del hablante. Para empezar, adoptaré como mi *definiendum* no la forma expresiva de máximo interés, es decir, (A), «Profiriendo  $x$   $H$  quiso decir que  $p$ », sino mejor la forma expresiva discutida de un modo más prominente en mi artículo de 1957, a saber: (B) «Profiriendo  $x$   $H$  quiso decir *algo*». Mi explicación de 1957 comprendía, naturalmente, la idea de que un adecuado *definiens* de (B) conllevaría una referencia a un efecto buscado [«intended»] por, o en respuesta a, la preferencia de  $x$ , y que una especificación de este efecto o respuesta buscada proporcionaría el material para responder a la pregunta de qué quiso decir  $H$  al proferir  $x$ . En un momento posterior de este escrito me ocuparé del *definiendum* (A) y trataré de aclarar el presunto vínculo entre la naturaleza del efecto buscado y la especificación de lo que  $H$  quería decir al proferir  $x$ .

Comienzo, entonces, considerando la siguiente definición propuesta:

«*H* quiso decir algo profiriendo *x*» es verdadero si, y sólo si, para alguna audiencia *A*, *H* profirió *x* con la intención de que

- 1) *A* llevase a cabo una cierta respuesta *r*,
- 2) *A* pensara (reconociera) que *H* intentaba 1),
- 3) *A* cumpliera 1) sobre la base de su cumplimiento de 2)

Dos comentarios explicativos pueden ser útiles. (i) Uso los términos «proferir» y «proferencia» de un modo artificialmente extendido para referirme a cualquier acto o ejecución que sea o pueda ser candidato a [poseer] significado no-natural. (ii) Suponer que *A* lleve a cabo *r* «sobre la base de» que piense que *H* intenta que él haga *r* es suponer que el que *A* piense que *H* intenta que él haga *r* es al menos parte de sus razones para llevar a cabo *r*, y no meramente la *causa* de que haga *r*. La tercera subcláusula del *definiens* se formula de esta manera a fin de eliminar lo que de otro modo sería un contraejemplo. Si en vez de la subcláusula 3) tuviésemos.

- 3 a) *A* llevase a cabo 1) como resultado de su cumplimiento de 2).

deberíamos admitir, contra lo que dicta la intuición, que *H* quiso decir algo haciendo *x* si (como podría ser el caso) *H* hiciera algo con la intención de que

- 1) *A* se entretuviera,
- 2) *A* pensara que *H* intentaba que él se entretuviera,
- 3) *A* se entretuviera como resultado (al menos en parte) de pensar que *H* intentaba que él se entretuviera.

Pero, aunque el pensamiento de *A* de que *H* intentaba que él se entretuviera podría ser parte de la causa de que estuviese entretenido, no podría ser una parte de su razón para estar entretenido (de hecho, uno no tiene razones para estar entretenido). Así pues, la adopción de 3) mejor que de 3 a) excluye este caso.

Consideraré [ahora] objeciones a este análisis del significado ocasional del hablante bajo dos principales encabezamientos: primero, esas que tratan de mostrar que el *definiens* es demasiado débil, que concede en exceso; y, segundo, esas que tratan de mostrar que el *definiens* es demasiado fuerte, que excluye casos de significado ocasional del hablante. Para hacer frente a estas objeciones proporcionaré redefiniciones de la noción de significado ocasional en diversos esta-



dios; cada tal redefinición ha de considerarse invalidada por su sucesora.

### III. EJEMPLOS ADUCIDOS CONTRA DE LA SUFICIENCIA DEL ANALYSANS SUGERIDO

i) (Urmson)<sup>4</sup> Hay una gama de ejemplos que tienen que ver con la provisión, por parte de *H* (el hablante), de un incentivo, o supuesto incentivo, para que *A* (el receptor, la audiencia) ejecute alguna acción. Supongamos que los capturadores de un prisionero de guerra piensan que éste posee información que desean que revele; el prisionero sabe que ellos desean que él dé esta información. Los capturadores le someten a la tortura de las empulgueras. Los *analysans* apropiados de «Ellos quisieron decir algo sometiéndole a la tortura de las empulgueras (para que les dijera lo que querían saber)» se cumplen todos:

- 1) Le sometieron a la tortura de las empulgueras con la intención de ocasionar una cierta respuesta de la víctima;
- 2) Ellos tenían la intención de que él reconociese (supiera, pensara) que ellos le sometían a la tortura de las empulgueras con la intención de producir esta respuesta;
- 3) Ellos tenían la intención de que el reconocimiento (pensamiento) por parte del prisionero de que tenían la intención mencionada en 2) debería ser parte al menos de las razones de éste para llevar a cabo la respuesta mencionada.

En general, si especificar en 1) la naturaleza de una respuesta buscada es especificar lo que quiso decir, no sólo *debería* ser correcto afirmar que los torturadores quisieron decir algo sometiéndole a la tortura de las empulgueras, sino también afirmar que ellos querían decir que él debía (tenía que) contarles lo que querían saber. Pero, en realidad, no tendríamos que afirmar ninguna de ambas cosas, salvo que los torturadores querían *que* él les informase. Cabe construir un aparente contraejemplo parecido a partir de un caso de soborno (ejemplo original de Urmson).

Una restricción es lo que parece necesitarse [en el ejemplo anterior]; y una que podría servir para eliminar esta gama de contraejemplos puede descubrirse comparando los dos siguientes casos:

- a) Voy a una expendeduría de tabaco, pido un paquete de mis

<sup>4</sup> En conversación con J. O. Urmson.

cigarrillos favoritos y, cuando el inusualmente suspicaz vendedor me hace ver que quiere conocer el color de mi dinero ante de entregarme el producto, pongo sobre el mostrador el precio de los cigarrillos. Aquí nada ha querido decirse.

b) Voy a mi habitual expendeduría de tabaco (en donde también suelo comprar otras cosas) a por un paquete de mi marca de siempre, cuyo precio es característico (digamos que 43 centavos). No digo nada, sino que entrego los 43 centavos. El vendedor reconoce lo que necesito y me entrega el paquete. En este caso, creo, poniendo sobre el mostrador los 43 centavos yo quise decir algo —a saber, que quería un paquete de la marca X—. Al mismo tiempo, he proporcionado un móvil.

El rasgo característico del segundo ejemplo parece ser que en el vendedor reconoció, y se tuvo la intención de que reconociera, lo que con mi «preferencia» (el poner yo el dinero sobre el mostrador) se pretendía que hiciese, mientras que en el primer ejemplo esto no era el caso. Ni tampoco sucede en el ejemplo de la tortura. Así pues, podría proponerse corregir de acuerdo con ello el análisis del significado (Redefinición I):

«*H* quiso decir algo profiriendo *x*» es verdadero si, y sólo si,

- 1) Profiriendo *x* *H* tenía la intención de inducir en *A* una cierta respuesta;
- 2) *H* tenía la intención de que *A* reconociera, *en parte al menos a partir de la preferencia de x*, que *H* tenía la intención de que se produjera esa respuesta;
- 3) *H* tenía la intención de que la satisfacción de la intención mencionada en 2) fuese parte al menos de la razón de *A* para obrar de modo que se satisficiera la intención de 1).

Mientras que con esto se podría hacer frente a esta gama de contraejemplos, hay otros para los que resulta insuficiente.

ii) (Stampe, Strawson, Schiffer) <sup>5</sup>.

a) Un hombre está jugando al *bridge* contra su jefe. Quiere ganarse el favor de éste y por tal razón desea que su jefe gane y, aún más, quiere que su jefe sepa que él quiere que gane (a su jefe le gusta

<sup>5</sup> En conversación con Denis W. Stampe; P. F. Strawson, «Intention and Convention in Speech Acts», *Philosophical Review*, LXXIII (1964), pp. 439-460; en conversación con S. Schiffer. [Hay versión castellana del escrito de Strawson en sus *Ensayos lógico-lingüísticos*, Tecnos, Madrid, 1983. (*N del T.*)]

esta clase de humildad). El hombre no desea, sin embargo, que la cosa sea demasiado obvia: como sucedería, por ejemplo, si se lo dijera de palabra o si llevara a cabo algún acto que pueda entenderse como señal de tal cosa, por miedo a que su jefe pueda sentirse ofendido por la crudeza de su proceder. Por consiguiente, pone en marcha el siguiente plan: cuando tiene una buena mano, sonrío de un cierto modo; la sonrisa se parece *mucho* a, pero no es *todavía* como, una sonrisa espontánea de placer. Su intención es que el jefe detecte la diferencia y arguya así: «Esa sonrisa no es una genuina sonrisa delatora, sino la simulación de una tal. Esa clase de remedo podría ocultar un farol (una mano floja). Sin embargo, esto es el *bridge*, no el póquer, y él no quiere ganarme a mí, su jefe, con semejante incorrección. Así pues, probablemente tiene una buena mano y, deseando que yo venza, esperaba que yo supondría que tiene una buena mano interpretando su sonrisa como si fuese una espontánea. Siendo esto así, no elevaré el envite de mi pareja.»

En un caso así, no pienso que nos gustaría decir que el empleado, con su sonrisa (o sonriendo), había *querido decir* que tenía una buena mano, ni de hecho que habría querido decir nada en absoluto. Sin embargo, se cumplen las condiciones enumeradas hasta ahora. Cuando sonrió,

- 1) El empleado tenía la intención de que su jefe pensara que el empleado tenía una buena mano;
- 2) El empleado tenía la intención de que su jefe pensara, en parte al menos a causa de su sonrisa, que el empleado tenía la intención de que su jefe pensara que se trataba de una buena mano;
- 3) El empleado tenía la intención de que una parte al menos de la razón que tuviera su jefe para pensar que se trataba de una buena mano debería ser que el empleado quería que él pensara precisamente eso.

b) Para hacer frente a un ejemplo parecido al acabado de citar, Strawson propuso que el *analysans* podría restringirse añadiendo una restricción ulterior, a saber, la de que el hablante *H* debería proferir *x* no sólo, como ya se ha exigido, con la intención de que *A* piense que el hablante intenta lograr una cierta respuesta de *A*, sino también con la intención de que *A* piense (reconozca) que *H* tiene la intención acabada de mencionar. En el presente ejemplo, se tiene la intención de que el jefe piense que el empleado desea que aquél piense que [éste] tiene una buena mano, pero *no* se tiene la intención de que [el jefe] piense que [el empleado] *tiene la intención* de que piense que éste desea que [el jefe] piense que el empleado tiene una buena mano; el empleado tiene la intención de que [el jefe] piense

que es sólo como resultado de ser demasiado inteligente para [su] empleado que ha llegado a saber que éste desea que él piense que se está ante una buena mano; [el jefe] ha de pensar que se supuso que interpretaría la sonrisa como una revelación involuntaria.

c) (Schiffer) Cabe elaborar un ejemplo más o menos paralelo en el que la respuesta buscada sea de carácter práctico, que parece mostrar la necesidad de añadir una quinta condición. El hablante *H* se encuentra en una habitación con un hombre *A* cuya avaricia es notoria, aunque tenga también su orgullo. *H* desea deshacerse de *A*. Entonces, bien a la vista de *A*, *H* arroja un billete de cinco libras por la ventana. Su intención es que *A* piense lo siguiente: «*H* quiere que me vaya de la habitación, pensando que correré tras el billete de cinco libras. Quiere también que yo sepa que él desea que yo me vaya (en vista de lo ofensivo de su proceder). Pero yo no me rebajaré yendo tras el billete; me iré, pero no porque desea que me vaya. No me quedo donde no soy bien querido.» En este ejemplo se cumplen las contrapartidas de las cuatro condiciones del *analysans* sugeridas hasta el momento; sin embargo, una vez más, no pienso que debiéramos decir que *H* había querido decir algo arrojando el billete por la ventana; que había querido decir, por ejemplo, que *A* había de (debía) marcharse. Las cuatro condiciones que se cumplen pueden formularse como sigue:

*H* profirió *x* (arrojó el billete) con la intención

- 1) de que *A* se fuese de la habitación,
- 2) de que *A* pensara (en parte al menos sobre la base de *x* que *H* tenía la intención 1),
- 3) de que *A* pensara que *H* tenía la intención 2),
- 4) de que en la satisfacción de la intención 1) una parte al menos de la razón de *A* para obrar debería ser que pensara que *H* tenía la intención 1) (es decir, que [2] se cumpliera).

Así pues, a menos que se juzgue que con esta preferencia se ha querido decir algo, hace falta todavía una restricción más. Parece ser una característica de este ejemplo que aunque *H* tenía la intención de que el que *A* se fuese de la habitación se basara en el pensamiento de *A* de que *H* quería que se fuese de la habitación, *H* no tenía la intención de que *A* reconociese que *H* tenía la intención de que la marcha de *A* dependiera de ello. [*H*] tenía la intención de que *A* pensara que el propósito de *H* era el de conseguir que él se fuera en busca del billete de cinco libras. Por lo tanto, cabe sugerir que la restricción requerida es la de que *H* tenga la intención

- 5) de que *A* piense (reconozca) que *H* tenía la intención de que 4).

Podemos reformular ahora la forma general de estas condiciones sugeridas (Redefinición II, Versión A):

«*H* quiso decir algo mediante *x*» es verdadero si, y sólo si, *H* profirió *x* con la intención

- 1) de que *A* llevase a cabo la respuesta *r*;
- 2) de que *A* pensara, en parte al menos sobre la base de *x*, que *H* tenía la intención 1);
- 3) de que *A* pensara que *H* tenía la intención 2);
- 4) de que la realización de *r* por parte de *A* se basara (en parte al menos) en el pensamiento de *A* de que *H* tenía la intención 1) (es decir, en el cumplimiento de [2] por parte de *A*);
- 5) de que *A* pensara que *H* tenía la intención 4).

Un hecho notable acerca de este *analysans* es que exhibe en diversos lugares el siguiente rasgo: la *n*-ésima «subintención» de *H* se especifica como intención de que *A* deba pensar que *H* tiene la «intención» *n*-1-ésima. La presencia de este rasgo ha llevado a la sugerencia de que el análisis del significado (que siga esta línea) es infinito o indefinidamente regresivo; que siempre podrían encontrarse ulteriores contraejemplos, por muy complejo que resulte el *analysans* sugerido, que forzarán el añadido de otras cláusulas que exhiban este rasgo. Pero tal regreso podría ser/será virtuoso, no vicioso. Podría ser/será tan poco perjudicial como el que conduce de «*Z* sabe que *p*» a «*Z* sabe que *Z* sabe que *p*» a...

No estoy seguro de cuán inocente sería semejante regreso en el *analysans*. No exhibiría ciertamente el tipo de circularidad, que supone a primera vista una fuerte objeción, que surgiría al dar, por ejemplo, un *definiens* de «*H* quiso decir que *p*» que en algún lugar reintrodujera la expresión «*H* quiso decir que *p*» o que introdujera la expresión «*H* quiso decir que *q*». Por otro lado, no sería tan manifiestamente inocuo como suponer que cuando quiera que fuese correcto decir «es verdad que *p*» fuese también verdadero decir «es verdad que es verdad que *p*», etc.; o tan inocuo como lo sería suponer que si *Z* satisface la condición para [atribuirle] saber que sabe que *p*, también satisface la condición para [atribuirle] saber que sabe que *p*. En tales casos, no se precisan condiciones extraordinarias de la verdad de una reiteración como la de, por ejemplo, «sabe que *p*» que vaya más allá de las que se precisan para la verdad de la oración respecto de la que se hace la reiteración. Pero el carácter regresivo del *analysans* de «*H* quiso decir algo mediante *x*» está diseñado para hacer frente en cada estadio a posibles contraejemplos; de modo que cada cláusula adicional impone una restricción, exige que se cumpla una

condición ulterior. Podría preguntarse si, por ejemplo, en el supuesto de que sea siempre posible saber que  $p$  sin saber que uno sabe que  $p$ , sería legítimo definir « $Z$  super-sabe que  $p$ » por medio del conjunto abierto de condiciones:

- 1)  $Z$  sabe que  $p$ ,
- 2)  $Z$  sabe que 1),
- 3)  $Z$  sabe que 2), y así sucesivamente.

Resta, sin embargo, la posibilidad de que no haya que tomar ninguna decisión a propósito de esta cuestión, pues podría suceder que no haya lugar a la amenaza de regreso.

No parece fácil construir ejemplos que fuercen la adición de cláusulas que supongan ulteriores reiteraciones de « $H$  tenía la intención de que  $A$  pensara que...». El siguiente es un intento de Schiffer.  $H$  canta *Tipperary* con una voz estridente con la intención de que  $A$  se vaya de la habitación; se supone que  $A$  reconoce (y sabe que se tiene la intención de que reconozca) que  $H$  desea deshacerse de  $A$ . Además,  $H$  tiene la intención de que  $A$  se vaya porque reconozca la intención de  $H$  de que [ $A$ ] se vaya. El plan de  $H$  es que  $A$  piense (*equivocadamente*) que  $H$  tiene la intención de que  $A$  piense que  $H$  tiene la intención de que  $A$  se marche mediante el reconocimiento de la intención de  $H$  de que  $A$  se marche. Con otras palabras, se supone que  $A$  razona así: « $H$  tiene la intención de que yo piense que él tiene la intención de deshacerse de mí con su estridente forma de cantar. [ $H$ ] tiene realmente la intención de que yo me vaya porque él desea que yo me vaya, no porque yo no pueda soportar su canto.» El hecho de que  $A$  esté realmente ajustándose a los planes de  $H$  mientras piensa que está yendo más allá de ellos se aduce para sugerir que no se le permite afirmar a uno que en este caso  $H$  quiso decir con su canto que  $A$  se fuera.

Sin embargo, cuando trata uno de completar los detalles de esta descripción, el ejemplo se hace frustrante. ¿Cómo se supone que llega  $A$  a la idea de que  $H$  desea que él piense que  $H$  intenta deshacerse de él con su canto? Cabría suponer que  $H$  canta con un particular tono nasal que sabe que no desagrada a  $A$ , aunque sí a la mayor parte de la gente.  $A$  sabe que  $H$  sabe que este tono no le resulta desagradable a  $A$ , pero piensa (*equivocadamente*) que  $H$  no sabe que  $A$  sabe esto. Cabría suponer que  $A$  razona así: «Él no puede querer echarme fuera con su canto, puesto que sabe que este tono nasal no me resulta desagradable. Sin embargo, no sabe que yo sé que él sabe esto, de forma que puede que desee que yo piense que él intenta echarme con su canto.» En este momento, uno esperaría que  $A$  se hallase completamente desbordado para explicar el comportamiento de  $H$ ; no veo ninguna razón en absoluto por la que  $A$  no deba suponer entonces que  $H$  desea realmente deshacerse de él de alguna otra forma.

Al margen de si pudiera lograrse que este ejemplo funcionara, su complejidad es enorme, y cualquier intento de introducir ulteriores restricciones supondría complejidades todavía mayores. En general, es verdad que uno no puede tener intenciones de obtener resultados cuyo logro uno no ve posible; y el éxito de intenciones del tipo involucrado en la comunicación requiere que aquéllos a quienes van dirigidos los mensajes o los cuasi-mensajes sean capaces, en las circunstancias en cuestión, de tener ciertos pensamientos y de extraer determinadas conclusiones. En algún estadio temprano de la pretendida regresión, los cálculos que *H* requiere de *A* serán difícilmente superables; y sospecho que el límite se alcanzó (si es que no se superó) con los ejemplos que dieron lugar al añadido de una cuarta y de una quinta condición. Por lo tanto, *H* no podría tener las intenciones que se le exigían a fin de forzar la adición de ulteriores restricciones. No es sólo que los cálculos que se le pidan a *A* sean demasiado difíciles, sino que resultaría imposible que *H* encontrase señas para indicarle a *A* que debía llevar a cabo cálculos, incluso aunque estuviesen a su alcance.

Pero incluso si esta condición fuese correcta, parece que quedamos en una posición incómoda. Pues, aunque podemos saber que no necesitamos de ninguna serie infinita de subcláusulas «que miran hacia atrás», no podemos decir exactamente cuántas subcláusulas se precisan. De hecho, parece como si expandir la definición de «Profiriendo *x* se quiso decir algo» pudiera tener que variar de caso a caso, dependiendo de cosas como la naturaleza de la respuesta buscada, las circunstancias en las que se hace el intento de obtener una respuesta y la inteligencia del hablante y del interlocutor. Es dudoso que pueda aceptarse semejante variación.

Se evitaría esta dificultad si pudiésemos eliminar posibles contraejemplos, no pidiendo que *H* tenga determinadas intenciones adicionales (de «las que miran hacia atrás»), sino exigiendo mejor que *H* no tenga un cierto tipo de intención o complejo de intenciones. Todos los contraejemplos potenciales de la clase que ahora nos ocupa suponen el diseño de una situación en la cual *H* tiene la intención de que *A* cuente *tanto* con algún «elemento de inferencia» (alguna premisa o algún paso inferencial *E*) *como* que piense también que *H* tiene la intención de que *A* no cuenta con *E* en el proceso de reflexión por el que se supone que *A* arriba a su respuesta. ¿Por qué no eliminar, entonces, semejantes contraejemplos potenciales mediante una simple cláusula que le prohíba a *H* tener semejante clase de intención compleja?

Llegamos así a la Redefinición II, Versión B:

«*H* quiso decir algo mediante *x*» es verdadero si, y sólo si (para algún *A* y para alguna *r*):

- a) *H* profirió *x* con la intención de que:
- 1) *A* llevase a cabo *r*,
  - 2) *A* pensara que *H* tenía la intención 1),
  - 3) la satisfacción por *A* de 1) se basa en la satisfacción por *A* de 2).
- b) No existe ningún elemento de inferencia *E* tal que *H* profiriese *x* con la intención tanto (1') de que la determinación de *r* por parte de *A* contara con *E*, como (2') que *A* pensara que *H* tenía la intención de que (1') fuese falsa.

iii) (Searle)<sup>6</sup> Un soldado americano es capturado por tropas italianas durante la Segunda Guerra Mundial. Quiere que las tropas [italianas] crean que él es un oficial alemán, para que le liberen. Lo que le gustaría hacer es decirles en alemán o en italiano que él es un oficial alemán, pero no conoce ni alemán ni italiano en grado suficiente como para hacer tal cosa. Así pues, «como si dijésemos, trata de montar la farsa de contarles que él es un oficial alemán» recitando la única línea de alemán que conoce, una línea aprendida en la escuela: «*Kennst du das Land, wo die Zitronen blühen.*» Intenta producir una cierta respuesta en sus capturadores, a saber, la de que crean que él es un oficial alemán, e intenta producir esta respuesta por medio del reconocimiento de su intención de producir esta respuesta. No obstante, mantiene Searle, es falso que cuando él dice «*Kennst du das Land*» lo que quiera decir es «Yo soy un oficial alemán» (o incluso la versión alemana de «Yo soy un oficial alemán», porque lo que tales palabras significan es «Conoces la tierra donde florece el limonero»). Usa [Searle] este ejemplo para apoyar la idea de que falta algo en mi análisis del significado; éste (pienso yo que él piensa) mejoraría si (esto conjeturo) se lo completara así: «*H* quiso decir algo mediante *x*» significa «*H* tuvo la intención de producir en *A* un cierto efecto por medio del reconocimiento de la intención de *H* de producir ese efecto, y (si la preferencia de *x* es la preferencia de una oración) *H* tuvo la intención de que el reconocimiento por parte de *A* de la intención de *H* (de producir el efecto) se consiga mediante el reconocimiento de que la oración proferida se usa convencionalmente para producir tal efecto».

Ahora, e incluso si hubiera de enfrentarme a un contraejemplo genuino, debería ser muy renuente a seguir el camino que sospecho

<sup>6</sup> John R. Searle. «What is a Speech Act?» en *Philosophy in America*, ed. por Max Black, Ithaca, Nueva York, 1965, pp. 221-239. [Una versión castellana de este escrito fue publicada como número 15 de la colección *Cuadernos Teorema* que editó el Departamento de Lógica y Filosofía de la Ciencia de la Universidad de Valencia. Se reproduce en pp. 431-48 de este volumen. (N. del T.)]



que se me estaba ofreciendo. (Resulta difícil decir si esto es lo que se me proponía, pues Searle se ocupa propiamente de la caracterización de un acto de habla particular (el de prometer), no de una discusión general de la naturaleza del significado; y fundamentalmente trataba de adaptar mi explicación del significado a su vigente propósito, no de reformularla para que se adapte mejor a su finalidad declarada.) Naturalmente, no quisiera yo negar que cuando el vehículo del significado sea una oración (o la preferencia de una oración) las intenciones del hablante tengan que reconocerse, en un caso normal, en virtud de un conocimiento del uso convencional de la oración (en realidad, mi explicación de «implicatura no-convencional» depende de esta idea). Pero, como indiqué anteriormente, me gustaría, si es que puedo, tratar el querer decir algo mediante la preferencia de una oración sólo como un caso especial del querer decir algo mediante una preferencia (en mi sentido ampliado de preferencia), y tratar una correlación convencional entre una oración y una respuesta específica sólo como una de las maneras en que una preferencia puede ponerse en correlación con una respuesta.

¿Es genuino, sin embargo, el presente contraejemplo? Me parece que la situación imaginaria está infradescripta y que hay quizás tres diferentes casos que considerar.

1) La situación podría ser de tal forma que la única posibilidad real de que los soldados italianos, al oír al soldado americano decir su línea de alemán, supusieran que era un oficial alemán sería la de que argüyeran como sigue: «Acaba de hablar en alemán (quizás en un tono autoritario); no sabemos nada de alemán y no tenemos ni idea de lo que ha querido decirnos; pero si habla alemán lo más probable es que sea un oficial alemán —¿qué otros alemanes podría haber en esta parte del mundo?—.» Si la situación fuese de tal modo que los italianos razonaran de esta guisa y si el americano supiera que así era como había de ser, resulta difícil evitar atribuirle, cuando habló, la intención de que ellos *deberían* argüir de ese modo. Como observé recientemente, en general uno no puede tener la intención de lograr algún resultado, si uno sabe que no hay ninguna posibilidad de conseguirlo. Pero, si la intención del soldado americano es la acabada de describir, entonces ciertamente, de acuerdo con mi explicación, él no estaría queriendo decir que era un oficial alemán; porque tuviera la intención de que los italianos creyeran que era un oficial alemán, su intención no sería la de que ellos creyeran esto por medio del reconocimiento de su intención. Y me parece a mí que, si bien no es ésta la manera en que quería Searle que se interpretase su ejemplo, sería con mucho la situación que con mayor probabilidad se habría dado.

2) Creo que Searle deseaba que supusiéramos que el americano esperaba que los italianos llegarían a la creencia de que él era un ofi-

cial alemán a través de la creencia de que las palabras que profirió eran las palabras alemanas correspondientes a «Soy un oficial alemán» (aunque no es fácil ver cómo pergeñar el contexto de la proferencia de forma que éste le dotase de alguna base para [alimentar] esta esperanza). Ahora se hace dudoso [aceptar] si, después de todo, es correcto afirmar que el americano no quiso decir «Soy un oficial alemán». Consideremos el siguiente ejemplo. El propietario de una tienda de baratijas para turistas en Port-Said se halla de pie a la entrada de su negocio, ve a un inglés y en un tono dulce y con una sonrisa seductora le dice en árabe: «¡Tú, inglés, cerdo!» Debería sentirme bastante inclinado a afirmar que el propietario había querido decirle al visitante que entrara, o algo por el estilo. Obviamente, no me siento tentado en absoluto a afirmar que *mediante las palabras que profirió* había querido decirle al visitante que entrara; y señalar que la línea en alemán no significa «Soy un oficial alemán», sino «Conoces la tierra», no es relevante. Si pudiese afirmarse que el americano había querido decir que él era un oficial alemán, habría querido decir esto mediante esa línea, o profiriéndola de un modo particular; exactamente de igual forma que el comerciante de Port-Said habría querido decirle al visitante que entrara *diciendo* lo que dijo, o hablándole del modo en que lo hizo.

3) Se ha sugerido, no obstante, que existe una diferencia entre que *H* meramente tenga la intención de que *A* piense que una oración en particular tiene un cierto significado que en realidad no tiene y que [*H*] tenga también la intención de que *A* piense que él mismo, según se supone, hace uso de la (equivocada) idea de que [la oración] tiene este significado, para arribar a una creencia acerca de las intenciones de *H*. Quizás se piensa que el comerciante de Port-Said no tiene la intención de que el visitante piense de sí mismo de tal forma; el visitante no ha de suponer que el comerciante piense que el primero puede hablar árabe. Pero, si la intención es que *A* piense que *H* espera que *A* comprenda la oración proferida y se tiene la intención de atribuirle [a ésta] un significado que *H* sabe que no tiene, entonces no debería presentarse [el caso como uno en que] el hablante quiere decir algo mediante su proferencia. No veo la fuerza de esta posición ni en realidad veo fácil aplicar la distinción que hace. Considérese sólo un ejemplo. He estado escuchando una lección de francés que se le ha impartido a la hija pequeña de un amigo. Me di cuenta de que ella piensa que una cierta oración francesa significa «Sírrete un trozo de pastel», aunque en realidad significa algo bastante diferente. Así que a la vista de un trozo de pastel le digo esta oración francesa y, según era mi intención, ella se lo sirve. Mi intención era que ella pensara que (y que pensara que mi intención era que ella pensase que) la oración proferida por mí significaba «Sírrete un trozo de pastel»; y diría que el hecho de que la oración significara, y que yo supiera que

significaba, algo bastante diferente no es obstáculo para que *yo* hubiera querido decir algo con mi preferencia (a saber, que se sirviera un trozo de pastel). Dicho de una forma más general, la idea parece ser como sigue. Resulta característico que un hablante tenga la intención de que su interlocutor reconozca (y piense que se intenta que reconozca) algún rasgo «crucial» *F* y que piense (y que se intente que él mismo piense) que *F* guarda en cierta forma una relación con alguna respuesta que el hablante tiene la intención de producir en su interlocutor. En lo que concierne a la atribución de significado del hablante, no importa si *H* piensa que *F* está o no *realmente* vinculado de tal forma a la respuesta; aunque, naturalmente, en un caso normal *H* pensará que *F* guarda esa relación.

Supongamos, sin embargo, que completamos los detalles del «caso del soldado americano», imaginando que acompaña el «*Kennst du das Land*» de gesticulaciones, golpes de pecho y demás; podría entonces [el soldado americano] esperar tener éxito en transmitir a sus interlocutores su intención de que comprendan la oración alemana, que sepan por esa particular oración alemana que el soldado americano tiene la intención de que ellos piensen que él es un oficial alemán (mientras que en realidad, naturalmente, el americano no espera que ellos sepan de *ese* modo, sino únicamente asumiendo, por la situación y por su tipo de comportamiento, que él debe estar tratando de decirles que es un oficial alemán). Quizás en ese caso no debiéramos sentir la inclinación a afirmar que el americano quiso decir que era un oficial alemán y que sí debiéramos estar dispuestos a afirmar tan sólo que él había querido decirles que pensarán que era un oficial alemán.

¿Cómo puede distinguirse este ejemplo del de «la niña pequeña»? Me gustaría sugerir un conjunto revisado de condiciones para «*H* quiso decir algo mediante *x*» (Redefinición III, Versión A):

Dominios de las variables: *A*: interlocutores.

*f, g*: rasgos de preferencias.

*r*: respuestas.

*c*: modos de relación (por ejemplo, icónico, asociativo, convencional).

( $\exists A$ ) ( $\exists f$ ) ( $\exists r$ ) ( $\exists c$ ) profirió *x* con la intención de que

- 1) *A* pensara que *x* poseía *f*.
- 2) *A* pensara que *H* tenía la intención 1).
- 3) *A* pensara que *f* guardaba la relación *c* con el tipo al que *r* pertenece.
- 4) *A* pensara que *H* tenía la intención 3).
- 5) *A* pensara, sobre la base de la satisfacción de 1) y de 3), que *H* tenía la intención de que *A* llevase a cabo *r*.
- 6) *A* llevase a cabo *r* sobre la base de la satisfacción de 5).
- 7) *A* pensara que *H* tenía la intención 6).

En «el caso de la niña pequeña» hay un solo rasgo  $f$  (el de ser una proferencia de una particular oración francesa) respecto de la cual  $A$  tiene las cuatro primeras intenciones. (El único inconveniente es que esta razón no guarda *en realidad* una relación convencional con las respuestas buscadas, y esto no descalifica la proferencia impidiéndole ser una por medio de la cual  $H$  quiera decir algo.)

En «el caso del soldado americano» no hay ningún simple rasgo  $f$ . Se tiene la intención 1) de que los capturadores reconozcan sin más el rasgo  $f$  (que  $x$  es un poco de alemán y que se profiere acompañado de ciertas gesticulaciones, etc.), pero 2) que piensen que se tiene la intención de que reconozcan que  $x$  tiene  $g$  (que es una *particular* oración del alemán).

El conjunto revisado de condiciones también se hace cargo del anterior ejemplo del *bridge*. Ahí se tiene la intención de que el jefe reconozca que  $x$  tiene  $f$  (ser una sonrisa falsa), pero no que se piense que se tiene tal intención. Así pues, la intención 2) de nuestra lista revisada está ausente. Y, por lo tanto, no necesitamos la condición añadida previamente para eliminar este ejemplo. Creo, no obstante, que la condición 7) (la antigua condición [5]) se precisa aún para eliminar el ejemplo del «billete de cinco libras», a menos que se la pueda reemplazar por una cláusula general de «anti-engaño». Puede que tal sustitución sea posible; puede que quepa omitir las subcláusulas «que miran hacia atrás» 2), 4) y 7) y que se las sustituya por la cláusula prohibitiva que figura en la Redefinición II, Versión B. Hemos de considerar entonces los méritos de la Redefinición III, Versión B, cuyo *definiens* dirá lo siguiente:

( $\exists A$ ) ( $\exists f$ ) ( $\exists r$ ) ( $\exists c$ )  $H$  profirió  $x$  con la intención de que:

- 1)  $A$  pensara que  $x$  posee  $f$ ;
- 2)  $A$  pensara que  $f$  guarda la relación  $c$  con el tipo al que  $r$  pertenece;
- 3)  $A$  pensara, sobre la base de la satisfacción de 1), que  $H$  tenía la intención de que  $A$  llevase a cabo  $r$ ;
- 4)  $A$  llevase a cabo  $r$  sobre la base de la satisfacción de 3);

( $b$ ) no hay ningún elemento de inferencia  $E$  tal que  $H$  tenga la intención de que:

- 1')  $A$  cuente con  $E$  en la determinación de  $r$ ,
- 2')  $A$  piense que  $H$  tenga la intención de que 1') sea falsa.

## IV. EJEMPLOS DIRIGIDOS A MOSTRAR LA EXCESIVA FUERZA DEL ANALYSANS TRIMEMBRE

Volvamos (para mayor simplicidad) al *analysans* original de «*H* quiso decir algo profiriendo *x*», y resumamos

«*H* profiere *x* con la intención de que *A*:

- 1) lleve a cabo *r*,
- 2) piense que *H* tiene la intención de que *A* lleve a cabo *r*,
- 3) piense que *H* tiene la intención de que la satisfacción de 1) se base en la de 2)»

en «*H* profiere *x* *M*-intentando que *A* lleve a cabo *r*».

En mi artículo original supuse que la identificación de *lo que H* quería decir mediante *x* dependería de la identificación de la respuesta o efecto *M*-buscado. Supuse, en particular, que las diferencias genéricas en el tipo de respuesta se corresponderían con diferencias genéricas en lo que se quiere decir. Por considerar dos ejemplos centrales, supuse que *a*) «*H* quiso decir mediante *x* que tal-y-tal es el caso» se elucidaría mediante «*H* profirió *x* *M*-intentando producir en *A* la creencia de que tal-y-tal»; *b*) que «*H* quiso decir mediante *x* que *A* debería hacer tal-y-tal cosa» se elucidaría por medio de «*H* profirió *x* *M*-intentando dar lugar a que *A* hiciera tal-y-tal cosa». Las preferencias indicativas o cuasi-indicativas se hallan ligadas a las generación de creencias; las preferencias imperativas o cuasi-imperativas lo están a la generación de acciones.

Deseo someter a su consideración la [siguiente] enmienda a esta idea: que en la explicación de las preferencias imperativas o cuasi-imperativas se haga constar, como la respuesta directamente *M*-buscada «la intención de *A* de hacer tal-y-tal cosa» (en lugar de «que *A* hiciera tal-y-tal cosa»). Esto tiene las ventajas 1) de que se logra la simetría, ya que en ambos casos (indicativo e imperativo) la respuesta *M*-buscada será una actitud proposicional; y 2) de que acomoda el hecho de que el acuerdo («sí», «perfectamente») en el caso de «La máquina se ha detenido» significa creencia, y en el caso de «¡Detén la máquina!» significa *intención*. Naturalmente que la acción es el objetivo *último* del hablante. Los casos en que la acción es respuesta inmediata pueden tratarse —a saber, la intención con la que el agente actúa—. Los imperativos reclaman siempre acción *intencional*.

La mejor forma de entender los contraejemplos que se han aducido es considerándolos intentos de crear dificultades no al análisis dado de «*H* quiso decir *algo* profiriendo *x*» que fue sugerido [más arriba], sino a este análisis completado con el género de detalle que se acaba de mencionar, para ofrecer un bosquejo de explicación de

«Profiriendo  $x$   $H$  quiere (quiso) decir que...». En particular, se sugiere que explicar «Profiriendo  $x$   $H$  quiso decir que es el caso que tal-y-tal» por medio de « $H$  profirió  $x$   $M$ -intentando dar lugar a que  $A$  hiciera tal-y-tal cosa» es seleccionar como *explicans* una condición demasiado fuerte. Tenemos que poder decir que en una cierta ocasión  $H$  quiso decir que tal-y-tal sin comprometernos nosotros mismos con la proposición de que  $H$   $M$ -intentó producir en  $A$  una creencia de que tal-y-tal.

Los siguientes ejemplos parecen presentar dificultades:

*Examinando:* P: «¿Cuándo tuvo lugar la batalla de Waterloo?»  
R: «1815».

Aquí el examinando quiso decir que la batalla de Waterloo tuvo lugar en 1815, pero difícilmente  $M$ -intentó inducir semejante creencia en su examinador. Para el examinando es natural pensar que las creencias del examinador (sean las que fueren) son independientes de las respuestas propias. El efecto  $M$ -buscado es (quizás) que el examinador sepa o piense que el examinando piensa que la batalla de Waterloo tuvo lugar en 1815; *o bien* (quizás) que el examinador sepa si el examinando conoce la respuesta correcta a la pregunta. (Quizás sea el primero el efecto directo, y que el efecto indirecto, buscado, sea el segundo.)

*Confesión* (en algunos casos):

Madre: «¡No lo niegues! Rompiste el cristal de la ventana, ¿verdad?»  
Hijo: «Sí».

Aquí el hijo sabe que su madre ya piensa que él rompió el cristal de la ventana; lo que ella desea es que él le diga que lo hizo. Quizá el efecto  $M$ -buscado sea, entonces, que la madre piense que su hijo desea decirle que lo hizo (¿qué significa «decir» aquí? ¿cómo debería analizarse?); o que la madre piense que su hijo no quiere seguir sosteniendo que no rompió el cristal de la ventana (que ni dice ni hace nada con la intención de causar la creencia de que *no* rompió el cristal de la ventana). (La confesión es quizás un caso ritual y sofisticado.)

*Hacer recordar:* P: «Veamos, ¿cómo se llamaba la muchacha?»  
R: «Rosa» (o muestra una rosa).

De quien aquí pregunta ya se presume que sabe (al menos en un sentido disposicional) que la muchacha se llama Rosa; sólo que el nombre se le ha ido de la cabeza. El efecto buscado parece ser el de que  $A$  recuerde que se llama Rosa.

*Reseña de los hechos:* Tanto el hablante como el oyente se supone que creen que  $p$  ( $q$ , etc.). El efecto buscado, una vez más, parece ser el que  $A$  (y puede que también  $H$ ) tenga presentes (juntos) «los hechos» en su mente.

*Conclusión de un argumento:*  $p$ ,  $q$ , luego  $r$  (de premisas ya sentadas).

Mientras que  $H$  tiene la intención de que  $A$  piense que  $r$ , no espera (y tampoco tiene la intención de) que  $A$  arribe a la creencia de que  $r$  a partir de la intención de  $H$  de que ha de alcanzarla. Se supone que son las premisas, no la confianza en  $H$ , lo que importa.

*El hombre contrasugestionable:*  $A$  considera a  $H$  casi invariablemente equivocado en ciertos temas, o bien una persona con la que no soportaría estar de acuerdo.  $H$  sabe esto.  $H$  dice: «Mi madre te tiene en gran estima» con la intención de que  $A$  (por la fuerza de lo que  $H$  dice) piense que la madre de  $H$  tiene una pobre opinión de él [ $A$ ]. En este caso existe la inclinación a afirmar que, pese a la intención de  $H$  de que  $A$  piense que la madre de  $H$  piensa mal de él, lo que  $H$  quiso decir fue que la madre de  $H$  piensa bien de  $A$ .

Estos ejemplos conllevan dos dificultades relacionadas.

1) Hay alguna dificultad en suponer que la forma indicativa se halla *convencionalmente ligada* a la indicación de que el hablante está  $M$ -intentando inducir una cierta creencia en su interlocutor, si es que hay casos en los que el modo indicativo va acompañado de diferentes intenciones del hablante, casos en los que éste no  $M$ -intenta (ni puede interpretarse que  $M$ -intente) inducir una creencia (como, por ejemplo, en el de hacer recordar [algo a alguien]). Sin embargo, por otro lado, parece difícil suponer que la función del modo indicativo *no* tenga *nada que ver* con el inducimiento de creencias. El indicio de que el hablante tiene la intención de que su interlocutor actúe (o se forme una intención de actuar) ha de considerarse, plausible si no forzosamente, por convención la función del modo imperativo; con seguridad la función del modo indicativo debería ser análoga. ¿Qué alternativa queda a la relación sugerida en el caso de la intención de inducir una creencia?

El problema podría resolverse en este caso distinguiendo las preguntas relativas a lo que significa una oración en modo indicativo de las preguntas relativas a lo que un *hablante* quiere decir. Cabría sugerir que una especificación completa del significado de una oración (en el caso de oraciones en modo indicativo) supone hacer referencia al hecho de que la forma indicativa significa convencionalmente una intención del hablante de inducir una creencia; pero puede muy bien suceder que el significado del hablante no coincida con el significado

de la oración que profiere. Puede estar claro que, si bien usa un recurso que indica convencionalmente una intención suya de inducir una creencia, carece *en este caso* de esa, pero no de otra, intención. Quizás quepa reforzar esto señalando que *cualquier* recurso cuya función primaria (estándar) sea la de indicar la intención del hablante de inducir la creencia de que *p podría* emplearse fácil y comprensiblemente en circunstancias apropiadas para propósitos próximos, por ejemplo (como en «el caso del examinando») para indicar que el *hablante* cree que *p*. El problema sería entonces poner de manifiesto que los contraejemplos aducidos son adaptaciones naturales de un mecanismo o forma que está primariamente vinculada a la indicación de intención de inducir una creencia.

Pienso que, si fuese posible, nos gustaría no tener que considerar los contraejemplos como usos extendidos de la forma indicativa y hallarle a esa forma una función aplicable de modo más general. En cualquier caso, la segunda dificultad es más seria.

2) Incluso pudiendo preservar la idea de que la forma indicativa se halla ligada por convención a la indicación de la intención del hablante de inducir una creencia, tendríamos que permitir que el significado del hablante difiera en distintos usos (*occurrences*) de la misma oración en modo indicativo. (De hecho, esto lo precisa la solución que sugerimos a la dificultad [1].) Habrá que conceder esto si es que diferencias en la respuesta buscada conllevan diferencias en el significado del hablante. Pero no resulta muy plausible afirmar que si *H* profiere «La batalla de Waterloo tuvo lugar en 1815»:

- 1) como maestro (y con la intención de inducir una creencia),
- 2) como examinando,
- 3) como maestro *en una clase de repaso*,

*H* querría decir algo diferente en cada uno de estos casos al profiere esta oración. Incluso si el examinando *M* intenta inducir la creencia de que él piensa que la batalla de Waterloo tuvo lugar en 1815, no resulta atractivo afirmar que cuando profirió «La batalla de Waterloo tuvo lugar en 1815» quería decir que *él pensaba* que la batalla de Waterloo tuvo lugar en 1815 (a diferencia del maestro que enseña ese período histórico por primera vez).

Podríamos tratar de hacer frente a algunos de los ejemplos (como los de hacer recordar o reseñar hechos) suponiendo que el efecto *M*-buscado estándar no sea sólo una creencia, sino una «creencia activada» (que *A* crea que *p* y que piense que *p*). Hay tres formas de fracasar aquí: puede que uno

- 1) ni crea que *p* ni tenga en mente que *p*;



- 2) crea que  $p$ , pero no tenga en mente que  $p$ ;
- 3) no crea que  $p$ , pero tenga en mente que  $p$ .

Es así que quien se acuerda de algo tiene la intención [de obtener] la misma respuesta que quien informa, pero se intenta remediar deficiencias diferentes.

Esto (incluso en los ejemplos para los que parece prometedor) desemboca en una nueva dificultad. Si  $H$  dice (acordándose): «La batalla de Waterloo tuvo lugar en 1815», se cumplen dos de mis condiciones:

- 1)  $H$  tiene la intención de inducir en  $A$  la creencia *activada* de que la batalla de Waterloo tuvo lugar en 1815,
- 2)  $H$  tiene la intención de que  $A$  reconozca que 1).

Pero, si la fecha de la batalla de Waterloo «estaba en la punta de la lengua de  $A$ » (como bien pudiera ser),  $H$  no puede esperar (y, por lo tanto, no puede tener la intención) de que la creencia activada de  $A$  se produzca a través del reconocimiento por  $A$  de que  $H$  tiene la intención de ocasionarla. Si  $A$  cree ya (aunque haya momentáneamente olvidado) que la batalla de Waterloo tuvo lugar en 1815, entonces la mención de esta fecha dará lugar a la creencia activada, al margen de la intención de  $H$  de ocasionarla.

Sugiere esto abandonar la exigencia (para el significado del hablante) de que  $H$  tenga la intención de que la respuesta de  $A$  se base en el reconocimiento de  $A$  de la intención de  $H$  de que  $A$  realice su respuesta; ello sugiere retener meramente las condiciones 1) y 2) de más arriba. Pero esto no servirá: hay ejemplos que requieren esta condición.

- a) Pienso que no puede decirse de Herodes, al mostrar a Salomé la cabeza de San Juan Bautista, que ha querido decir que San Juan Bautista estaba muerto.
- b) Exhibir una pierna vendada (en respuesta a una invitación a jugar al *squash*).

En *b*) el exhibidor podría querer decir: 1) que no puede jugar al *squash*, o bien (aunque dudosamente), 2) que tiene una pierna en mal estado (el vendaje podría ser falso), pero no, 3) que su pierna está vendada.

La tercera condición parece ser necesaria a fin de protegernos en estos casos de resultados contraintuitivos.

*Remedios posibles*

i) Cabría retener la idea de que el efecto o la respuesta buscados (en casos de querer decir que es el caso que  $p$  —tipo indicativo—) es una creencia activada, conservando a la vista la distinción entre arribar a este estado 1) desde una deficiencia en la seguridad (*assurance deficiency*) y [hacerlo] 2) desde una deficiencia en la atención (*attention-deficiency*); y estipular que la tercera condición (que  $H$  tenga la intención de provocar la respuesta sobre la base del reconocimiento de la intención de provocar esa respuesta) sea operativa *sólo* cuando  $H$  tenga la intención de ocasionar una respuesta activada eliminando la *deficiencia en la seguridad*, *no* cuando tenga la intención de hacer tal cosa eliminando la deficiencia de atención. Esta idea podría quizás extenderse también a los tipos de casos imperativos, dado que podemos hallar ejemplos de recordarle algo a alguien que haga algo (restituyéndole una intención *activada*), en los que la intención de  $H$  de que  $A$  acceda al estado es igualmente ociosa, casos en los que no se espera que el que  $A$  arribe a la intención activada dependa de su reconocimiento de que  $H$  tiene la intención de que llegue a ella. Por lo tanto, la definición podría decir aproximadamente esto:

(\* $\mu$  es un marcador de modo, un elemento auxiliar vinculado a la actitud proposicional  $\mu$  de entre una gama dada de actitudes proposicionales.)

«Profiriendo  $x$   $H$  quiere decir que \* $\mu$ » = « $H$  profiere  $x$  con la intención

- 1) de que  $A$   $\mu$  activamente que  $p$ ,
- 2) de que  $A$  reconozca que  $H$  tiene la intención 1) y (a menos que  $H$  tenga la intención de que la preferencia de  $x$  se limite a remediar deficiencias de atención)
- 3) que la satisfacción de 1) se base en la de 2)».

Este remedio no puede, sin embargo, con 1) el ejemplo del «examen», 2) los casos de «confesión» o 3) el caso del hombre contrasugestionable.

ii) Puesto que cuando  $H$  tiene la intención, profiriendo  $x$ , de promover en  $A$  la creencia de que  $p$ , es una exigencia estándar la de que  $A$  piense (y se intenta que piense) que  $H$  piensa que  $p$  (de otra forma  $A$  no pensará que  $p$ ), ¿por qué no hacer que el efecto buscado *directamente* no sea el de que  $A$  piense que  $p$ , sino el de que  $A$  piense que  $H$  piensa que  $p$ ? En muchos, aunque no en todos los casos,  $H$  tendrá la intención de que  $A$  pase de pensar que  $H$  piensa que  $p$  a

pensar él mismo que *p* (casos de «informar»). Pero semejante efecto ha de considerarse indirecto (aunque frecuentemente de interés primordial).

Podemos retener ahora la tercera condición, puesto que incluso en los casos en los que se hace recordar [algo a alguien] cabe esperar que *A* piense que la intención de *H* de que *A* piense que *H* piensa que *p* resulte pertinente con vistas a la pregunta de si *A* ha de pensar que *H* piensa que *p*. No sólo hemos salido al paso del ejemplo en que se hace recordar [algo a alguien], sino también al paso de los casos del examinando y del hombre contrasugestionable (que se intenta que piense que *H* piensa que *p*, aunque no que piense que *p* él mismo). Y aunque todavía no nos hemos hecho cargo del ejemplo de la reseña de hechos (puesto que puede considerarse que *A* sabe ya que *H* piensa que *p*), si entendemos «*H* cree que *p*» como «*H* tiene la creencia activada de que *p*», este ejemplo puede también acomodarse. Si bien se supone que *A* sabe que *H* cree que *p*, *A* no sabe, hasta que *H* habla, que *H* tiene en mente que *p*.

Pero, mientras que una solución que siga estas directrices puede resultar aceptable en los casos de tipo indicativo, no puede generalizarse a los no-indicativos. Comparemos:

- a) «No cruzarás la barrera.»
- b) «¡No cruces la barrera!»

Cuando se profiere a) *H* tiene, de forma característica, la intención de que *A* piense que *H* tiene la intención de que *A* no cruce la barrera; pero parece que una especificación del significado de *H*, en una preferencia normal de b), no quedaría completamente explicada a menos que se diga que *H* tiene la intención de que *A* no se limita a pensar que *H* tiene la intención de que *A* no cruce la barrera, sino también a formar la intención de que [*A*] no la cruce.

Tracemos entonces una distinción entre lo que podría denominarse preferencias «puramente exhibitivas» (preferencias mediante las cuales el hablante *H* tiene la intención de impartir la creencia de que él tiene una determinada actitud proposicional) y preferencias que no son sólo exhibitivas, sino también «protrépticas» (es decir, preferencias por medio de las cuales *H* trata de inducir una correspondiente actitud en el oyente, impartiendo la creencia de que él [*H*] tiene una determinada actitud proposicional).

Llegamos, entonces, a la Redefinición IV, Versión A:

«Profiriendo  $x$   $H$  quiso decir que  $*\mu p$ » es verdadero si, y sólo si,  $(\exists A)$   
 $(\exists f) (\exists c) H$  profirió  $x$  con la intención de que:

- |    |   |   |
|----|---|---|
| 1) | } | [como en III (A), con « $r$ » reemplazada por<br>« $\mu$ -er»] <sup>7</sup> |
| 2) |   |   |
| 3) |   |   |
| 4) |   |   |
| 5) |   |   |
| 6) |   |   |
| 7) |   |   |

y (en algunos casos)

- 8)  $A$   $\mu$ -se él mismo que  $p$ , sobre la base de la satisfacción de 6).

El que una instancia de sustitución de la subcláusula 8) haya de aparecer en la expansión de un enunciado de la forma representada en el *definiendum* dependerá de la naturaleza de « $*\mu$ » que ese enunciado incorpore.

También podemos obtener la Redefinición IV, Versión B, añadiendo al *definiens* de III(B), como subcláusula a)5), la subcláusula 8) de más arriba, junto con una modificación de la cláusula b) de III(B), a fin de tener en cuenta que la respuesta buscada  $r$  se especifica ahora en términos de la idea de  $\mu$ -er que  $p$ .

El que una versión de la Redefinición IV sea o no sea correcta dependerá crucialmente de la actitud que se adopte ante una versión en modo imperativo del ejemplo del hombre contrasugestionable. El Sr.  $A$ , deseando librarse de la presencia de la Sra.  $A$ , pero considerando que ella es, en lo que a él le concierne, contrasugestionable le dice: «Ahora, querida, hazme compañía un rato.» ¿Sería correcto afirmar que el Sr.  $A$ , que es obvio que no quiso pedirle a la Sra.  $A$  que le hiciese compañía, quiso decir con sus palabras que le hiciera (o que debía hacerle) compañía un rato? Si la respuesta es «sí», la Redefinición IV es inadecuada, puesto que, según ella, para haber querido decirle a la Sra.  $A$  que le hiciera compañía, el Sr.  $A$  habría tenido que tener la intención de que ella se formase la intención de hacerle compañía, intención que él ciertamente no tuvo. Sin embargo, no sería difícil corregir esto; transformamos la nueva subcláusula « $A$   $\mu$  él

<sup>7</sup> Siendo  $\mu$  una actitud proposicional (como creer, saber, temer, esperar, etc), « $\mu$ -er» representa la forma verbal infinitiva del verbo que expresa la correspondiente actitud. Por lo tanto, « $\mu$ -se» representaría el pretérito imperfecto de subjuntivo del correspondiente verbo. Si  $\mu$  fuese la actitud proposicional de creer, « $\mu$ -er» sería la forma verbal «creer» y « $\mu$ -se» sería entonces «creyese». (*N. del T.*)

mismo que  $p$  sobre la base de la satisfacción de 6)» en «*A pensara que H intentaba que A  $\mu$ -se que  $p$* ». Si, no obstante, la respuesta es «no», entonces la Redefinición IV queda intacta.

#### V. EL SIGNIFICADO OCASIONAL DEL HABLANTE EN AUSENCIA DE INTERLOCUTOR

Hay varios ejemplos de preferencias mediante las cuales, podría considerarse correctamente, el hablante ha querido decir algo (ha querido decir que tal-y-tal), sin que haya ninguna persona real ni ningún conjunto de personas a quienes se dirija el hablante y en quienes intente inducir una respuesta. La gama de estos ejemplos incluye, o podría pensarse que incluye, casos como los de colocar avisos como «No entre» o «Este puente se halla en mal estado», apuntes en un diario, notas cuya finalidad es la de esclarecer los propios pensamientos cuando se encuentra uno reflexionando acerca de un problema, soliloquios, ensayos de parte de una conversación prevista en silencio. Algunos de estos ejemplos al menos no han sido tenidos en cuenta en las definiciones propuestas hasta el momento.

Los ejemplos que deberá cubrir mi explicación se dividen en tres grupos:

a) Preferencias de las que piensa el hablante que pueden tener (ahora o más adelante) un interlocutor.  $H$  puede pensar que alguna persona en particular, por ejemplo él mismo en una fecha futura, en el caso de un apunte en un diario, puede (pero puede también no) encontrarse con la preferencia de  $H$ ; o cabe que  $H$  piense que sea posible que haya alguien que sea (o que vaya a ser) interlocutor de su preferencia.

b) Preferencias de las que el hablante sabe que no están dirigidas a ningún interlocutor efectivamente dado, pero que pretende dirigirlas a alguna o a algún tipo de personas en particular, o que piensa que van dirigidas a algún interlocutor o tipo de interlocutor imaginado (como en el caso de un discurso o de la parte que le corresponde en una conversación prevista).

c) Preferencias (incluyendo preferencias «internas») que el hablante ni piensa que pueda dirigir a ningún interlocutor efectivamente dado ni se imagina a sí mismo dirigiéndolas a nadie pero que, no obstante, tiene la intención de que induzcan un cierto tipo de respuesta en un interlocutor convenientemente indefinido, en el caso de que tal interlocutor se hallase presente. En un ejemplo de pensamiento silencioso, la idea de un interlocutor presente habrá de interpretarse liberalmente como la de un interlocutor en una contrapartida pública del discurso interno del hablante. A este respecto resulta

quizá digno de mención que algunos casos de pensamiento verbal quedan fuera del alcance de mi explicación. Cuando los pensamientos verbales se limitan a pasar por mi cabeza, y no a que yo los «forje», es inapropiado afirmar que yo haya querido decir algo mediante ellos; en semejantes casos soy, quizás, mejor un interlocutor que un hablante.

Propondré una redefinición final que espero explique los ejemplos que necesitan ser explicados y que sancione favorablemente los casos especiales de la gama de ejemplos en que hay, y que el hablante sabe que hay, un interlocutor efectivamente dado. Esta redefinición será relativamente informal; podría presentar una versión formal que ganaría en precisión a costa de una mayor dificultad de comprensión.

Tome « $\varphi$ » (y « $\varphi'$ ») como valores propiedades de personas (interlocutores posibles); entre las expresiones por las que puede reemplazarse « $\varphi$ » (y « $\varphi'$ ») se incluirán algunas tan distintas como «es un viajero», «es un viajero que lee este aviso», «es un inglés nativo», «es idéntico a Jones». Según se verá, para que  $H$  quiera decir algo habrá de poderse identificar el valor de « $\varphi$ » (que puede muy bien hallarse indeterminado) que  $H$  tenga en mente; pero *nosotros* no hemos de determinar el dominio del que  $H$  efectúa su elección.

#### *Redefinición V*

«Profiriendo  $x$   $H$  quiso decir que  $*\mu p$ » es verdadero si, y sólo si,  $(\exists \varphi) (\exists f) (\exists c)$ :

- I.  $H$  profirió  $x$  con la intención de que  $x$  sea de tal forma que quien tenga  $\varphi$  piense que:
  - 1)  $x$  tiene  $f$ ,
  - 2)  $f$  guarda la relación  $c$  con  $\mu$ -er que  $p$
  - 3)  $\exists \varphi$ :  $H$  tenga la intención de que  $x$  sea de tal forma que quien tenga  $\varphi'$  piense que  $H \mu$  que  $p$ , por la vía de pensar que 1) y que 2),
  - 4)  $H \mu$  que  $p$ , en vista de que 3);
 y
- II. ([cláusula] vigente tan sólo para ciertos substituyendos de « $*\mu$ »)  $H$  profirió  $x$  con la intención de que, *en el caso de que efectivamente hubiera* alguien que tuviera  $\varphi$ , este mismo  $\mu$  que  $p$ , por la vía de pensar que 4);
- y
- III. No es el caso que, para algún elemento de inferencia  $E$ ,  $H$  tenga la intención de que  $x$  sea de tal forma que quien tenga  $\varphi$  1') cuente con  $E$  para llegar a  $\mu$ -er que  $p$  y 2') piense

que  $\exists\varphi'$ :  $H$  tenga la intención de que  $x$  sea de tal forma que quien tenga  $\varphi'$  llegue a  $\mu+$  que  $p$  sin contar con  $E$ .

- Notas:* 1) « $\mu+$ » ha de leerse como « $\mu$ » si la Cláusula II está vigente, y como «pensar que  $H \mu$ » si la Cláusula II no está vigente.  
2) Necesitamos usar tanto « $\varphi$ » como « $\varphi'$ », ya que no deseamos exigir que  $H$  tenga la intención de que su posible interlocutor piense en el posible interlocutor de  $H$  bajo la misma descripción que  $H$  emplee.

*Comentarios explicativos:*

1) Resulta esencial que la intención que se expone en la Cláusula II deba especificarse como la intención de  $H$  «de que en el caso de que efectivamente hubiera alguien que tuviera  $\varphi$ , éste ...», mejor que adoptar el mismo procedimiento en las Cláusulas I y II, atribuye a  $H$  la intención «de que  $x$  deba ser de tal forma que, si hubiera alguien que fuese  $\varphi$ , él ...». Si optamos por la última especificación, dejaremos sin respuesta una objeción puesta por Schiffer, como cabe mostrar con ayuda de un ejemplo del mismo tipo que uno suyo. Supongamos que, enfurecido después de pasar una tarde con mi suegra y tras encontrarme a solas, me desahogo gritando: «¡No vuelva por aquí!» Sin duda resultará esencial a mi momentáneo bienestar que yo diga esta palabras de tal forma que, de hallarse presente mi suegra, ella se haga el propósito de no venir por aquí más. Sería, sin embargo, inaceptable si se representara la situación como si de mi intención se siguiera que yo *quise decir* que ella no volviera más por aquí; pues ese falso que, en esas circunstancias, yo quisiera decir tal cosa con mi observación. Según se la ha reformulado, la redefinición evita esta dificultad.

2) Supongamos que, de acuerdo con el *definiens* de la última redefinición  $\exists\varphi$ :  $H$  tiene la intención de que quien tenga  $\varphi$  piense que... y supongamos que el valor de « $\varphi$ » que  $H$  tiene presente es la propiedad de ser idéntico a una cierta persona  $A$ . Se seguirá entonces que  $H$  tiene la intención de que  $A$  piense que ...; y dada la ulterior condición, que se cumple en un caso normal, de que  $H$  tiene la intención de que  $A$  piense que él [ $A$ ] es el interlocutor buscado, se nos asegura la verdad de un enunciado del que resulta inferible el *definiens* de IV(B) por la regla de generalización existencial (asumiendo la legitimidad de esta aplicación de GE a un enunciado cuya expresión verbal contiene verbos «intensionales» como «intentar» y «pensar»). Creo que puede mostrarse que, para el caso en que haya un interlocutor que sepa que él es el interlocutor al que el hablante trataba de

dirigirse, si el *definiens* de IV(B) es verdadero, el *definiens* de V será verdadero. Si eso es así, dado que la redefinición V es correcta en cualquier caso en que haya un interlocutor efectivamente dado, la satisfacción del *definiens* de IV(B) constituirá una condición necesaria y suficiente del haber querido decir que \* $\mu$ p.

## VI. CONCLUSIÓN

Juzgo que hay algunas razones para esperar que, prestando una debida atención a la relación entre significado no-natural y significado natural, no sólo podría alcanzarse una explicación simplificada del significado ocasional del hablante, sino mostrar también que cualquier institución humana cuya función sea la de proporcionar sustitutos artificiales de signos naturales debe incorporar, como su concepto clave, uno que posea aproximadamente los rasgos que yo adscribo al concepto de significado ocasional del hablante. Pero semejante empresa trasciende el alcance de este escrito.



## LÓGICA Y CONVERSACIÓN \*

H. P. GRICE

Es un lugar común de la lógica filosófica la afirmación de que hay, o de que parece haber, diferencias en los significados de al menos algunos de los que denominaré expedientes formales, como « $\neg$ », « $\wedge$ », « $\vee$ », « $\rightarrow$ », « $(x)$ », « $\exists x$ », « $\iota x$ » (cuando se les confiere la interpretación canónica estándar), por un lado, y lo que se considera que son sus contrapartidas o análogos del lenguaje natural —expresiones como «no», «o», «si», «todo», «algún», (o «al menos un»), «él»—, por otro. Ocasionalmente, algunos lógicos pueden haber sido partidarios de la idea de que, de hecho, no hay diferencias de significado entre unos y otras; pero semejantes afirmaciones, en el caso de haber sido enunciadas, han sido llevadas a cabo algo imprudentemente, y sus presuntos autores no han sido objeto de un muy buen trato.

Aquellos que admiten la existencia de las susodichas diferencias se adhieren, fundamentalmente, a uno u otro de dos grupos rivales, a los que denominaré, a efectos de este ensayo, el grupo formalista y el grupo informalista. El perfil característico de la posición formalista podría trazarse del siguiente modo: En la medida en que los lógicos se ocupan de formular pautas muy generales de inferencia válida, los expedientes formales poseen una ventaja decisiva frente a sus contrapartidas naturales. Porque en términos de los primeros cabe construir sistemas de fórmulas muy generales de entre las cuales un número considerable de ellas pueden considerarse pautas de inferencia, o muy cercanas a pautas así, cuya expresión involucra algunos o todos de dichos expedientes. Semejantes sistemas pueden consistir en un cierto conjunto de fórmulas simples que habrán de aceptarse si los expedientes tiene el significado que se les ha asignado, así como también de un número indefinido de fórmulas ulteriores muchas de las cuales no serán tan manifiestamente aceptables, pero que habrá que aceptar si eso fue lo que se hizo con cada uno de los miembros del conjunto inicial. Disponemos, por lo tanto, de una manera para habérnoslas con pautas de inferencia cuya aceptabilidad puede no ser obvia; y si pudiésemos aplicar un procedimiento decisorio, como de

---

\* Versión castellana de Juan José Acero.

hecho a veces resulta posible, dispondríamos entonces de una estrategia todavía mejor.

Más aún, desde un punto de vista filosófico se considera que el que las contrapartidas tengan como parte de su significado elementos que no comparten con sus correspondientes expedientes formales es una imperfección de nuestras lenguas; pues los elementos en cuestión no serán sino excrecencias indeseables. La razón de ser de esta opinión estriba en que la presencia de tales elementos da como resultado que los conceptos dentro de los cuales aparecen no podrán definirse con precisión y claridad, y que al menos a algunos de los enunciados en que intervengan no será posible asignarles, en algunas circunstancias, un valor de verdad definido. No sólo cabrá poner reparos al carácter indefinido de estos conceptos, sino también que este abre las puertas a la metafísica; no podemos tener la certeza de que ninguna de estas expresiones del lenguaje natural no se encuentre metafísicamente «podrida». Por estas razones, a las expresiones, tal y como se las usa en el discurso cotidiano, no se las puede aceptar en definitiva, y cabe incluso que, a la larga, no sean del todo inteligibles. La estrategia adecuada consiste en concebir y erigir un lenguaje ideal, que incorpore los expedientes formales, cuyas oraciones sean claras, con un valor de verdad determinado inequívocamente y manifiestamente libres de secuelas metafísicas; con ello, los fundamentos de la ciencia se hallarán ahora filosóficamente seguros, pues los enunciados del científico podrán expresarse (aunque de hecho no se proceda necesariamente así) dentro de este lenguaje ideal. (No pretendo sugerir que todos los formalistas aceptarían la totalidad de este resumen, pero sí que creo que todos ellos darían el visto bueno a alguna parte suya.)

A esto, un informalista podría replicar con el siguiente espíritu. La exigencia filosófica de un lenguaje ideal descansa en ciertos supuestos que deberían rechazarse; éstos son los siguientes: que el patrón con el que habría que medir la adecuación de un lenguaje es su capacidad para servir a las necesidades de la ciencia; que no es posible garantizar la inteligibilidad de una expresión a menos que se haya proporcionado una explicación o un análisis de significado; y que toda explicación o todo análisis debe adoptar la forma de una definición precisa, que sea la expresión o la aserción de una equivalencia lógica. El lenguaje sirve a muchos propósitos importantes, además de los propios de la investigación científica. Podemos conocer perfectamente bien lo que significa una expresión (y, por lo tanto, *a fortiori* que es inteligible) sin saber cómo analizarla, de manera que el ofrecer un análisis de ella pueda (y de hecho así sucede) consistir en la especificación, tan general como quepa, de las condiciones bajo las cuales es o no posible usar la expresión que estamos analizando. Más aún, mientras que es verdad sin duda que los expedientes formales se

prestan a un tratamiento sistemático por parte del lógico, resta todavía el caso de muchas inferencias y argumentos, expresados en el lenguaje natural, y no en términos de los susodichos expedientes, cuya validez hay no obstante que admitir. Por consiguiente, debe quedar un lugar para una lógica no formalizada, y por lo tanto más o menos asistemática, de las contrapartidas naturales de aquellos expedientes. La lógica simplificada de los expedientes formales puede servir de ayuda y de guía a esta otra, pero no puede ocupar su lugar. En realidad, no es sólo que las dos lógicas difieren la una de la otra, sino que alguna veces pueden entrar en conflicto: las reglas propias de un expediente formal pueden no valer para su contrapartida natural.

Ahora bien, en lo que respecta al problema general de cuál es el lugar de la filosofía en la reforma del lenguaje natural, nada tengo que decir en este ensayo. Me limitaré a mediar en la disputa en lo que hace a las presuntas diferencias de significado mencionadas al principio. Además, no tengo la intención de mediar en ella ocupando el puesto de ninguno de los dos contendientes. Deseo, en realidad, defender que el supuesto, común a las dos partes en disputa, de que las diferencias de significado existen es (hablando en términos generales) un error compartido, y que este error deriva de haber prestado poca atención a la naturaleza y a la importancia de las condiciones que gobiernan la conversación. Por lo tanto, procederé sin más preámbulos a investigar las condiciones generales que, de una u otra forma, se aplican a la conversación, al margen de cuál pueda ser su tema.

## IMPLICATURA

Supongamos que A y B están hablando acerca de un amigo común, C, que está ahora trabajando en un banco. A le pregunta a B cómo le va a C en su empleo y B responde: «¡Muy bien! Se siente a gusto con sus compañeros y, además, no le han metido todavía en la cárcel.» Justo en este momento, A podría preguntarse qué era lo que B le estaba dando a entender (*implying*), lo que le estaba sugiriendo o lo que le quería decir (*meant*) al afirmar que a C no le habían metido todavía en la cárcel. La respuesta podría ser, por ejemplo, que C es el tipo de persona proclive a sucumbir a las tentaciones derivadas del empleo que tiene; o que los compañeros de C son, en realidad, gente muy desagradable y traicionera; etc. Naturalmente, podría no ser necesario en absoluto que A le preguntase tal cosa, por dejar muy claro el contexto de antemano cuál iba a ser la respuesta que recibiría. Creo que es obvio que, fuese lo que fuese aquello que B le dio a entender, sugirió o quiso decir, en este ejemplo ello diferiría de lo que B dijo, que fue simplemente que a C no le habían metido todavía

en la cárcel. Quiero introducir ahora como término técnico el verbo «implicar» (*implicate*) y los sustantivos «implicatura» —cf. «lo que implica»— e *implicatum* —cf. «lo implicado»—. La razón de ser de esta maniobra es que con ella se soslaya tener que recoger cada vez uno u otro de los miembros de la familia a la que «implicar» se encuentra vinculado; éste hará la labor de todos ellos. De momento, habré de asumir que, en una medida no desdeñable, se comprende el significado intuitivo de «decir» en semejantes contextos, así como que se es capaz de identificar los verbos concretos de la familia a la que pertenece «implicar». Puedo, sin embargo, hacer una o dos observaciones que ayuden a aclarar el más problemático de todos esos supuestos: a saber, el que tiene que ver con el significado de la palabra «decir».

1) En el sentido en que estoy empleando la palabra «decir», lo que una persona cualquiera ha dicho se halla íntimamente relacionado con el significado convencional de las palabras (de la oración) que ha proferido. Supongamos que alguien ha proferido la oración «desde aquella experiencia le tenía auténtica aversión al potro». Dado un mínimo de conocimientos de la lengua española, e incluso ignorando las circunstancias en que se proferieron dichas palabras, podría decirse que conoceríamos algo de lo que el hablante dijo, en el supuesto de que estuviese hablando español liso y llano y de que, además, estuviera hablando seriamente. Uno sabría que la persona había dicho, del hombre o mujer en cuestión, o bien que (i) su aversión hacia un cierto caballo de menos de cuatro años de edad, aproximadamente, tenía su origen en una cierta experiencia o bien que (ii) su aversión hacia un cierto tipo de instrumento de tortura derivaba de una muy desagradable experiencia previa. (No hace falta decir que esta explicación tiene tan solo un valor aproximado.) Pero para identificar de lleno lo que el hablante dijo haría falta saber (a) a qué persona se refería, (b) cuál es la experiencia previa, de tan nefasta memoria, a la que se alude, (c) el momento de tiempo en que se proferió la citada oración y (d) el significado de la frase «el potro» en ese particular momento de proferencia (es decir, decidirnos entre (i) y (ii)) \*. Esta escueta indicación sobre mi modo de usar la palabra

\* He optado en este lugar por un ejemplo distinto del que el autor ofrece, que gira en torno a la oración «*He is in the grip of a vice*», y he adaptado consiguientemente las observaciones que hace al respecto. Éstas, que aclaran los dos sentidos distintos de la citada oración inglesa, son las siguientes: «Uno sabría que había dicho, sobre la persona (varón) o el animal *X*, que en el momento de la proferencia (cualquiera que fuese) o bien (i) *X* era incapaz de desprenderse de una cierta clase de vicio, o bien (ii) alguna parte del cuerpo de *X* había quedado atrapada en una cierta herramienta o instrumento.» Grice saca partido aquí de la ambigüedad del término «*vice*», que puede significar tanto «vicio» como «torno». (*N. del T.*)

«decir» deja abierta la cuestión de si una persona que dijera (hoy) «Harold Wilson es un gran hombre» y otra que optara (también hoy) por «el Primer Ministro británico es un gran hombre» dicen lo mismo, aun cuando ambos supieran que los dos términos singulares relevantes tienen la misma referencia. Sin embargo, al margen de la decisión que tomemos al respecto, el aparato teórico que voy a describir ahora puede dar cuenta de cualquier implicatura que dependa de uno de estos términos singulares en la oración proferida. Semejantes implicaturas tendrían tan sólo que ver con dos máximas diferentes.

2) En algunos casos, el significado convencional de las palabras usadas determinará qué es lo que se implicó, además de ayudarnos a identificar lo que se dijo: Si digo (con un gesto de autosuficiencia) «Es un latino; luego es muy temperamental», yo mismo me comprometo ciertamente, en virtud del significado de mis palabras, con la idea de que él (la persona en cuestión) sea muy temperamental es una consecuencia (se sigue) de que sea latino. Pero mientras que he dicho que es un latino y que es temperamental, no me gustaría defender la tesis de que he *dicho* (en el sentido deseado) que del hecho de que alguien sea un latino se sigue que es muy temperamental, si bien ciertamente lo he indicado o implicado. No pretendo sostener que mi preferencia de la mencionada oración sea, *estrictamente hablando*, falsa, pese a que lo primero no fuese una consecuencia de lo segundo. Así pues, *algunas* implicaturas son convencionales, a diferencia de lo que acontece con aquella otra con la que inicié la presente discusión del fenómeno de la implicatura.

Voy a referirme ahora a una cierta subclase de las implicaturas no-convencionales a las que daré el nombre de implicaturas *conversacionales*, por hallarse esencialmente vinculadas a ciertos rasgos generales del discurso; así pues, mi próximo paso consistirá en decir cuáles son esos rasgos.

Lo que viene inmediatamente a continuación puede valer como primera aproximación a un principio general. Nuestras conversaciones no son habitualmente sucesiones de observaciones inconexas, y no sería racional que así fuese. Hasta cierto punto, son esfuerzos cooperativos de forma característica. Cada participante se percibe de que hay en ellas, en alguna medida, un propósito común o conjunto de propósitos comunes, o al menos una dirección mutuamente aceptada. Este propósito, o dirección puede fijarse desde el principio (por ejemplo, proponiendo un tema de discusión) o puede evolucionar durante su transcurso; cabe que esté perfectamente definido o que se halle tan poco delimitado que deje a las partes un margen de movimientos considerable (como sucede en las conversaciones casuales).

Pero en cada estadio se excluirán *algunas* contribuciones conversacionales por inadecuadas. Cabría formular entonces un principio general aproximado que, puede esperarse, las partes implicadas observarán (*ceteris paribus*): a saber, «Haga usted su contribución a la conversación tal y como lo exige, en el estadio en que tenga lugar, el propósito o la dirección del intercambio que usted sostenga.» A este principio podríamos bautizarlo el Principio Cooperativo (PC).

En el supuesto de que resulten aceptables principios generales como el citado, podría distinguirse quizás cuatro categorías a una u otra de las cuales pertenecerán máximas o submáximas más específicas. De entre todas ellas, las siguientes darán pie a resultados que están de acuerdo con el Principio Cooperativo. Haciéndonos eco de Kant, denominaré a estas categorías Categorías de Cantidad, Cualidad, Relación y Modo. La categoría de *Cantidad* tiene que ver con la cantidad de información a proporcionar, y a ella pertenecen las máximas

- 1) «Haga usted que su contribución sea tan informativa como sea necesario» (teniendo en cuenta los objetivos de la conversación), y puede que también
- 2) «No haga usted que su contribución resulte más informativa de lo necesario».

(Esta segunda máxima es discutible; podría decirse que el que una contribución a una conversación sea más informativa de lo necesario no constituye una transgresión de PC, sino tan solo una pérdida de tiempo. Sin embargo, podría responderse a esto afirmando que semejante suprainformatividad puede resultar confundente al poner sobre el tapete cuestiones marginales; y puede que también cause el efecto indirecto de equivocar a los hablantes, al inducirles a pensar que existía una *razón* particular por la que se suministraba un exceso de información. Al margen de todo esto, cabe apuntar una explicación distinta de por qué podemos sentir dudas a la hora de admitir esta segunda máxima: la de que lo que se logre con ella pueda conseguirse también por medio de una máxima ulterior que tiene que ver con la pertinencia de las contribuciones conversacionales.)

A la categoría de *Cualidad* pertenece una supermáxima: «Trate usted de que su contribución sea verdadera», y dos máximas más específicas:

- 1) «No diga usted lo que crea que es falso»,
- 2) «No diga usted aquello de lo cual carezca de pruebas adecuadas».

Dentro de la categoría de *Relación* situó una sola máxima: «Vaya usted al grano» \*. Si bien la máxima es algo lacónica, hay que decir que su formulación oculta una serie de problemas que me preocupan considerablemente; problemas relativos a las clases y los focos de relevancia que pueda haber, a cómo cambian unas u otros a lo largo del desarrollo de una conversación y a cómo dar cuenta del hecho de que los temas que se debaten pueden legítimamente variar, etc. Considero que la tarea de habérselas con estas cuestiones es sumamente difícil, y espero tratarlas en una conferencia posterior [que no se incluye aquí].

Finalmente, a la categoría de *Modo*, la cual concibo de manera que no tiene que ver (como sucede con las categorías precedentes) con lo que se dice, sino con cómo se dice lo que se dice, pertenece la supermáxima: «Sea usted perspicuo», así como diversas máximas:

- 1) «Evite usted ser oscuro al expresarse»,
- 2) «Evite usted ser ambiguo al expresarse»,
- 3) «Sea usted escueto (y evite ser innecesariamente prolijo)»,
- 4) «Proceda usted con orden».

Incluso podrían ser necesarias otras máximas distintas de éstas.

Es obvio que la observancia de algunas de estas reglas es tarea menos urgente que la de algunas otras; una persona que se ha expresado con excesiva prolijidad se haría a sí mismo blanco de comentarios más suaves que los que merecería una persona que ha dicho algo que cree ser falso. De hecho, puede uno apercibirse de que la importancia de la primera máxima de cualidad, al menos, es tal que no se la debería incluir en un esquema del género que estoy elaborando; otras máximas entran en consideración sólo sobre la base de que se satisface esta máxima de Cualidad. Aun cuando esto puede ser correcto, en tanto en cuanto nos limitemos a la generación de las implicaturas, la máxima de Cualidad no parece tener un papel totalmente diferente del de otras máximas; y al menos de momento será conveniente concebirla como un miembro más de entre los que componen la lista de máximas.

Hay, naturalmente, todo tipo de máximas (estéticas o morales), tal como «Sea usted educado», que los sujetos de una conversación observan normalmente, y también éstas puedan generar implicaturas no-convencionales. No obstante, las máximas conversacionales, así como las implicaturas relacionadas con éstas, guardan un vínculo especial (así lo espero) con los objetivos particulares a cuyo servicio se

---

\* Ésta es una forma coloquial de traducir la expresión «*Be relevant*», formulación que da al autor de la máxima de Relación. (N. del T.)

adapta la conversación (y los intercambios habidos en ella) y con vistas a los cuales primariamente se emplea. He anunciado mis máximos como si el objetivo central fuera el de intercambiarse información de forma máximamente efectiva; esta percepción es demasiado restringida, y el esquema ha de ampliarse hasta que tengan cabida en él objetivos generales tales como el de gobernar o influir en la conducta de los demás.

Puesto que uno de mis fines es el de hacer ver que la conversación es una variedad o caso especial de la conducta guiada por propósitos, racional de hecho, puede que valga la pena notar que las expectativas y presupuestos específicos asociados a algunas de las máximas tienen sus contrapartidas en la esfera de las transacciones distintas de los intercambios de palabras. Para cada categoría de máximas, daré ahora una lista de ejemplos de semejantes contrapartidas.

I. *Cantidad*. Si me está usted ayudando a reparar un automóvil yo espero que su contribución no sea ni mayor ni menor de lo necesario; si, por ejemplo, necesito cuatro tornillos, espero que usted me dará cuatro, y no dos ni seis.

II. *Cualidad*. Yo espero que su contribución sea genuina y no espúrea. Si necesito azúcar como ingrediente del pastel que usted me está ayudando a cocinar, no espero que me alcance la sal; si necesito una cuchara, no espero que lo que usted me entregue sea una de esas cucharas que tienen truco.

III. *Relación*. Y espero que la contribución de mi compañero de faenas resulte apropiada a los fines inmediatos de cada uno de los estadios de la transacción; si estoy mezclando los ingredientes de un pastel, no espero que se me entregue un buen libro, ni siquiera un portaollas (si bien esto último podría ser una contribución apropiada en un momento posterior).

IV. *Modo*. Yo espero que mi compañero deje bien explícita la contribución que está llevando a cabo, y que la realice con una celeridad razonable.

Estas analogías resultan relevantes con respecto a lo que yo considero que es una cuestión fundamental que tiene que ver con PC y sus máximas asociadas: a saber, cuál es la base del supuesto que hacemos, y del que dependen un gran número de implicaturas (como espero que se haga manifiesto, en virtud del cual los hablantes se conducen en general tal y como prescriben estos principios (*ceteris paribus*, y en ausencia de indicaciones que nos hagan suponer lo contrario). Una respuesta sosa, aunque no hay duda que adecuada a un cierto nivel, es la de que es un hecho empírico bien constatado que la gente *sí* que se comporta de estas formas; ha aprendido a conducirse así en su niñez y no ha perdido el hábito de hacerlo; y supondría en realidad un gran esfuerzo apartarse radicalmente de semejantes há-



bitos. Es mucho más fácil, por ejemplo, decir la verdad que inventar mentiras.

No obstante, soy lo suficientemente racionalista como para desear encontrar una base que subyazca a estos hechos, innegables como son; desearía poder concebir el tipo usual de práctica conversacional, no meramente como algo que todos o muchos de nosotros seguimos *de hecho*, sino como algo que es *razonable* que nosotros sigamos, como algo de lo que *no tendríamos que* apartarnos. Durante algún tiempo, me atrajo la idea de que la observancia de PC y de las restantes máximas, durante una conversación, podía considerarse materia cuasi-contractual, con muchos ejemplos paralelos fuera del ámbito del discurso. Si usted pasa por delante mío cuando estoy luchando con mi automóvil, que me ha dejado tirado, sin duda que me invadirá la esperanza de que me ofrezca su ayuda; pero una vez que usted se me una al trabajo, mis expectativas se harán más fuertes y adoptarán una forma específica (al menos mientras no me dé yo cuenta de que usted es tan sólo un entrometido incompetente). Análogamente, las conversaciones me parece que exhiben, de forma típica, ciertos rasgos que conjuntamente caracterizan a las transacciones cooperativas: 1) que las partes tienen algún objetivo común inmediato, como antes era el de arreglar un cierto automóvil (sus objetivos últimos pueden ser, naturalmente, independientes los unos de los otros, e incluso entrar en conflicto; cada uno de ellos desea arreglar el automóvil para poder irse a continuación en él, abandonando al otro a su propia suerte. En los intercambios propios de una conversación, hay un objetivo común, aunque se trate incluso de un objetivo de segundo orden, como en las conversaciones de sordos, en donde cada parte debe identificarse momentáneamente con los intereses conversacionales transitorios de la otra parte); 2) que las contribuciones de los participantes han de encajar unas con otras, siendo mutuamente dependientes; 3) que de alguna manera (que puede que sea explícita, pero que es a menudo tácita) se reconoce por ambas partes que, permaneciendo constantes todas las condiciones restantes, la transacción habrían de proseguir con un estilo adecuado, a menos que unos y otros se pongan de acuerdo en darle fin. No se trata ni de irse sin más ni de comenzar de improviso a hacer otra cosa.

Pero, mientras que es posible aplicar a algunos casos semejantes base cuasi-contractual, existen muchas formas de intercambio conversacional, como pelearse o escribir cartas, a las cuales no se ajusta con comodidad. Sea como fuere, uno siente que el hablante que no va al grano o que es oscuro, se defrauda primariamente a sí mismo, mejor que a sus interlocutores. Por lo tanto, me gustaría poner de manifiesto que la observancia de PC y de las otras máximas es algo razonable (racional) cuando se la analiza teniendo en cuenta las siguientes pautas: que es de esperar que cualquiera que se preocupe

por los fines que son centrales en la conversación/comunicación (tales como dar y recibir información, influir en los demás y ser influidos por ellos) tenga interés, dadas las circunstancias oportunas, en tomar parte en conversaciones que le resulten de provecho, sólo en el caso de que éstas se conduzcan de acuerdo con PC y las restantes máximas. No estoy seguro de si es o no lícito arribar a semejante conclusión; en cualquier caso, sí que lo estoy de que no cabe tomar ninguna decisión hasta que se haga mucho más clara la naturaleza de lo que es relevante en una conversación, así como de las circunstancias en que eso ha de exigirse.

Ha llegado el momento de mostrar la conexión existente entre PC y las otras máximas, por un lado, y las implicaturas, por otro.

Una persona que participe en una conversación puede dejar de cumplir una máxima de varias formas, entre las que se incluyen las siguientes:

1) Puede *violar* una máxima tranquilamente y sin ostentación ninguna; si así sucede, podrá en algunos casos inducir a equívocos.

2) Puede *dejar en suspenso* tanto la máxima de marra como PC; puede decir, indicar o hacer manifiesto que no desea cooperar de la manera que la máxima exige. Cabe, por ejemplo, que diga: «No puedo decir más cosas: mis labios están sellados.»

3) Puede desembocar en una *situación conflictiva*: puede ser incapaz, por ejemplo, de cumplir la máxima de Cantidad («Haga usted que su contribución sea tan informativa como sea necesario») sin violar con ello la segunda máxima de Cualidad («No diga usted aquello respecto de lo cual carece de pruebas adecuadas»).

4) Puede *pasarse olímpicamente por alto* una máxima; es decir, dejar *manifiestamente* de cumplirla. En el supuesto de que el hablante pueda cumplir con lo que la máxima exige y de que haga esto sin violar ninguna otra (es decir, sin llegar a un punto conflictivo), de que tampoco está dejándola en suspenso y de que no trata de inducir a nadie a equívocos, a la vista de lo manifiesto de su contribución conversacional, al oyente se le plantea un problema menor: ¿Cómo puede reconciliarse el haber dicho lo que dijo con el supuesto de que está observando plenamente PC? Esta situación es la que da característicamente lugar a una implicatura conversacional; y cuando se genera una implicatura tal de esta manera, diré que *se está explotando* una máxima.

Estoy ahora en situación de caracterizar la noción de implicatura conversacional. Una persona que al (o por el hecho de) decir (o de hacer como si dijera) que *p* ha implicado que *q* ha implicado conversacionalmente que *q*, *supuesto que*: 1) se supone que está observando las máximas conversacionales, o al menos el principio cooperativo;

2) es preciso el supuesto de que dicha persona es consciente de, o piensa que  $q$ , a fin de que el que haya dicho, o haya hecho como si dijera, que  $p$  sea consistente con este supuesto; y 3) que el hablante piense (y espere que el oyente piense que el hablante piensa) que está dentro de la competencia del oyente determinar, o captar intuitivamente, que el supuesto mencionado en 2) es necesario. Apliquemos todo esto a mi ejemplo inicial; es decir, a la observación de B de acuerdo con la cual a C no le habían metido todavía en la cárcel. En un contexto adecuado, A podría razonar de esta guisa: «1) B ha violado aparentemente la máxima “Vaya usted al grano”, de manera que puede pensarse que se ha saltado a la torera una de las máximas definitorias de la perspicuidad; sin embargo, no tengo ninguna razón para suponer que está dejando en suspenso PC; 2) dadas las circunstancias, puedo considerar lo irrelevante de su observación como cosa meramente aparente si, y sólo si, supongo que piensa que C es potencialmente deshonesto; 3) B sabe que soy capaz de inferir el paso 2). Por lo tanto, B implica que C es una persona potencialmente deshonesto.»

La presencia de una implicatura conversacional ha de poderse inferir; porque incluso en el caso de que se la pueda captar intuitivamente, y a no ser que la intuición pueda reemplazarse por un argumento, la implicatura (aun estando ahí) no podrá considerarse *conversacional*; será una implicatura *convencional*. Para inferir que está ante una implicatura conversacional, el oyente habrá de apoyarse en los siguientes datos: 1) el significado convencional de las palabras proferidas, junto con la identidad de las referencias implicadas; 2) PC y sus máximas; 3) el contexto lingüístico o extralingüístico de la proferencia; 4) otra información de fondo; 5) el hecho (o presunto hecho) de que ambos interlocutores conocen, o suponen que conocen, cada uno de los detalles relatados en los apartados precedentes, y que éstos están a su alcance. Una pauta general para diagnosticar la presencia de una implicatura conversacional puede ser la siguiente: «Él ha dicho que  $p$ ; no hay ninguna razón para suponer que no está observando las máximas, o al menos PC; podría estar cumpliéndolas si pensase que  $q$ ; sabe (y sabe que yo sé que él sabe) que yo me apercibo de la necesidad del supuesto de que piensa que  $q$ ; no ha hecho nada para impedirme pensar que  $q$ ; por lo tanto, pretende que yo piense, o al menos desea hacerme posible que piense, que  $q$ ; y consiguientemente ha implicado que  $q$ .»

Daré ahora unos cuantos ejemplos divididos en tres grupos.

*Grupo A:* Ejemplos en los que no se viola ninguna máxima, o al menos no es obvio que alguna máxima quede violada.

1) A está de pie ante un automóvil obviamente paralizado, y B se le aproxima. Entonces se produce el siguiente intercambio:

A: «Me he quedado sin gasolina.»

B: «Hay una gasolinera al volver la esquina.»

(*Glosa:* B estaría infringiendo la máxima «Vaya usted al grano», a no ser que piense o crea posible, que la gasolinera está abierta y que hay en ella gasolina a la venta; así pues, implicó que la gasolinera está, o puede estar, abierta, etc.)

En este ejemplo, a diferencia del relativo al comentario «No le han metido todavía en la cárcel», el nexo tácito entre las observaciones de B y de A es tan obvio que, incluso interpretando la supermáxima de Modo «Sea usted perspicuo» de manera que no sólo se aplique a la expresión de lo que se dijo, sino también a la relación de lo dicho con otros factores adyacentes, no parece que estemos ante una infracción de esa supermáxima en este ejemplo. El siguiente es, quizás, a este respecto algo menos claro.

2) A: «No parece que Enrique tenga ahora ninguna amiga.»

B: «Pues ha ido últimamente muchas veces a Veracruz.»

B implica que Enrique tiene, o puede tener, una amiga en Veracruz. Nada hay que glosar aquí, en vista de cómo comentamos nuestro ejemplo anterior.

En ambos casos, el hablante implica eso que hay que asumir que él cree, a fin de preservar el supuesto de que está observando la máxima de relación.

*Grupo B:* Un ejemplo en el cual se viola una máxima, pero en el que esta infracción ha de explicarse suponiendo un conflicto con otra distinta. A está planeando con B un itinerario para sus vacaciones en Francia. Ambos saben que A desea visitar a su amiga C, si esto no supone que su viaje se prolongue demasiado:

A: «¿Dónde vive C?»

B: «En algún lugar del sur de Francia.»

(*Glosa:* No hay ninguna razón para suponer que B está dejando ninguna máxima en suspenso; su respuesta es, como muy bien sabe, menos informativa de lo necesario para satisfacer los intereses de A; esta infracción de la primera máxima de Cantidad sólo puede explicarse suponiendo que B es consciente de que proporcionará una mayor información supondría decir algo que atentaría contra la máxima de Cualidad «No diga usted aquello para lo cual carece de pruebas adecuadas»; por lo tanto, B implica que no sabe en qué ciudad vive C.)

*Grupo C:* Ejemplos que llevan consigo que se está explotando alguna máxima, es decir, un procedimiento por medio del cual se pasa manifiestamente por alto una máxima con vistas a dar en una implicatura conversacional utilizando algo que tiene la naturaleza de un tropo o de una figura del discurso. En estos ejemplos, si bien se viola una máxima desde la perspectiva de lo que se dice, el oyente tiene derecho a suponer que esa máxima, o al menos el Principio Cooperativo, se observa desde el punto de vista de lo que se implica.

1 a) (Una infracción manifiesta de la primera máxima de Cantidad): A está escribiendo un informe sobre un alumno suyo que aspira a un empleo en un departamento de Filosofía, y su escrito dice lo siguiente: «Distinguido señor: La preparación del Sr. X en Literatura Inglesa es excelente, y se ha atendido regularmente a las indicaciones de sus tutores. Suyo afectísimo, etc.» (*Glosa:* A no puede estar dejando en suspenso ninguna máxima, puesto que si no deseara cooperar, ¿por qué escribir informe alguno? Tampoco se trata de que, por ignorancia, no pueda decir nada más, pues X ha sido alumno suyo; aún más, él sabe que se desea que proporcione más información. Por consiguiente, está pretendiendo dar datos que se resiste a proporcionar por escrito. Este supuesto es consistente tan sólo sobre la base de que piensa que el Sr. X deja mucho que desear desde el punto de vista de su formación filosófica. Por lo tanto, esto es lo que está implicando.)

Ejemplos extremos en los que se pasa uno manifiestamente por alto la primera máxima de Cantidad los proporcionan aquellas preferencias de tautologías patentes como «Las mujeres son las mujeres», «La guerra es la guerra», etc. Quiero sostener que en lo que respecta a lo dicho, según el sentido que doy a este término, semejantes palabras no ofrecen información ninguna y, de aquí que, vistas así, infringen la primera máxima de Cantidad en cualquier contexto conversacional. Obviamente, son informativas cuando se las juzga desde la perspectiva de lo que se implica, y la información que el oyente extraiga de ellas depende de su capacidad para explicar la selección que el hablante hace de la *particular* tautología proferida.

1 b) (Una infracción de la segunda máxima de Cantidad: «No proporcione usted más información de lo necesario», en el supuesto de que admitamos semejante máxima): A desea saber si *p*; y B no sólo le da la información de que *p*, sino también información que pone de manifiesto la certeza de que *p*, pruebas que muestran la verdad de que *p*, y así sucesivamente.

La locuacidad de B podría no ocultar segunda intención ninguna; y si A lo juzga así, podría despertar en él la duda de si B está tan seguro como dice que está. («Me parece que la dama protesta demasia-

do.») \* Pero, si A piensa que debajo de ella se esconde algo más, podría sospechar que la preferencia de B no es sino una manera oblicua de transmitir que resta todavía mucho que hablar acerca de si *p* o de si *no-p*. No obstante, puede argüirse que semejante implicatura puede ser explicada apelando a la máxima de Relación, sin tener que invocar una presunta segunda máxima de Cantidad.

2 a) Ejemplos en los que uno pasa conspicuamente por alto la primera máxima de Cualidad.

*Ironía.* X, con quien A ha mantenido hasta la fecha relaciones muy estrechas, ha traicionado un secreto de A al pasar cierta información a un rival de éste en los negocios. A y sus interlocutores saben ambas cosas. A dice «X es un magnífico amigo». (*Glosa:* Es del todo obvio, tanto para A como para sus interlocutores, que lo que A ha dicho o aparentado que decía es algo que no cree; y estos segundos saben que A sabe que esto es patente para ellos. Así pues, y a menos que su preferencia carezca de toda razón de ser, A debe estar tratando de que se fijen en una proposición manifiestamente relacionada con la primera; de entre las presuntas candidatas la primera a tener en cuenta es la contradictoria de la que parece haber expresado.)

*Metáfora.* Por ejemplo, «Eres la niña de mis ojos» \*\*. Ejemplos como éste son indesligables de alguna falsedad rotunda, de manera que lo contradictorio de lo que el hablante ha amagado que decía ha de ser, estrictamente hablando, una perogrullada; por consiguiente, no puede ser *eso* lo que el hablante trataba de transmitir. Lo más probable es que el hablante esté atribuyendo a su interlocutor una o más características que expliquen por qué lo identificaba con las pupilas de sus órganos visuales (recurriendo para ello, en mayor o menor medida, a la imaginatividad de éste).

Cabe combinar metáfora e ironía forzando al oyente a dar dos pasos sucesivos en su proceso de interpretación. Yo digo: «Eres la niña de mis ojos», tratando de que dé a mi metáfora el sentido de «A nadie le tengo mayor aprecio que a ti», e ironizo a renglón seguido: «Ojos, por cierto, con los que ya no veo nada.»

*Meiosis.* Dicho de una persona que ha destrozado todo el mobiliario de una vivienda: «Estaba ligeramente bebido.»

*Hipérbole.* «Tiene una novia en cada puerto» \*\*\*.

\* Grice recurre aquí a Shakespeare: *Hamlet*, acto III, escena II. Recojo la versión de Luis Astrana Marín: William Shakespeare, *Obras completas*, Madrid, Aguilar, 1969<sup>15</sup>, p. 1365. El texto original en inglés es «*Methinks the lady doth protest too much*». (N. del T.)

\*\* El ejemplo original de Grice es «*You are the cream in my coffee*». Las observaciones subsiguientes se han adaptado al ejemplo de la versión. (N. del T.)

\*\*\* El ejemplo original de Grice es «*Every nice girl loves a sailor*». (N. del T.)

2 b) Los ejemplos en los que se viola flagrantemente la segunda máxima de Cualidad («No diga usted aquello respecto de lo cual carece de pruebas adecuadas») son quizás difíciles de encontrar, si bien el siguiente parece ser un caso así. Digo de la esposa de X: «Esta tarde le está probablemente engañando.» En un contexto adecuado, o si acompaño mi preferencia de un gesto o de un tono de voz apropiado, puede quedar claro que carezco de buenas razones para suponer que pasa justamente eso. Mi interlocutor, a fin de dejar incólume el supuesto de que sigo el juego conversacional, asume que mis tiros van encaminados a una proposición relacionada con la anterior para cuya aceptación sí que dispongo de una base razonable. La proposición en cuestión podría ser muy bien la de que la esposa de X es una mujer dada a engañar a su marido, o posiblemente la de que es el tipo de persona que no podría desterrar de sí tal conducta.

3) Los ejemplos en los que se realiza una implicatura mediante una violación real, y no aparente, de la máxima de Relación son posiblemente raros, pero el siguiente parece ser un buen candidato. En una reunión de buen tono, A dice «La Sra. X es una vieja bruja». Por un momento el silencio puede oírse, y entonces B dice «Ha hecho un magnífico tiempo este verano. ¿Verdad?». B ha rehusado ostensivamente la posibilidad de que sus palabras tuvieran algo que ver con el precedente comentario de A. Por lo tanto, B implica que la observación hecha por A no es cosa que haya que discutirse, y quizás, más específicamente todavía, que A ha cometido un desliz social.

4) Ejemplos en los que se pasa conspicuamente por alto varias de las máximas que caen bajo la supermáxima «Sea usted perspicuo».

a) *Ambigüedad*. Debemos recordar que aquí nos concierne tan sólo la ambigüedad deliberada, la que el hablante pretende o espera que su interlocutor reconozca. El problema que tiene que resolver éste es el de por qué habría de apartarse un hablante de su camino al seguir el juego conversacional, optando por una preferencia ambigua. Debemos distinguir al respecto dos tipos de casos:

(i) Ejemplos en los que no hay ninguna diferencia, o ninguna diferencia chocante, entre dos interpretaciones que haga que una de ellas resulte más natural o adecuada que la otra; ninguna interpretación es más rebuscada o menos conspicua, más recóndita o traída por los pelos que la otra. Podríamos detenernos en los versos de Blake: «Nunca pretendas revelar tu amor, / el Amor que revelado nunca puede ser» \*. Para soslayar las complicaciones derivadas del modo

\* Éstos son los dos primeros versos de un poema de W. Blake que pertenece a sus

imperativo, ciñámonos a la oración, emparentada con la anterior: «Pretendí revelar mi amor, el Amor que revelado nunca puede ser». Aquí puede haber una doble ambigüedad. «Mi amor» puede referirse tanto a un estado emocional como al objeto de mi emoción, y «Amor que revelado nunca puede ser» puede significar lo que «Amor que es imposible expresar» o lo que «Amor que dejaría de existir si fuese revelado». En parte a causa de lo complejo que es el poeta mismo y en parte a causa de pruebas de carácter interno (a saber: que la ambigüedad sigue ahí viva), no parece haber más solución que la de suponer que las ambigüedades son deliberadas y que el poeta transmite tanto una interpretación como la otra; aunque no esté *diciendo* explícitamente ninguna de estas cosas, las está sugiriendo o las está dando a entender (cf. «[Pero] puesto que ella» —la Naturaleza— «te ha consagrado al placer de las mujeres, que tu amor sea mío y que su tesoro» —el de ellas, las mujeres— «pertenezca el goce de tu amor» \*).

(ii) Ejemplos en los que una interpretación es notablemente más natural o adecuada que la otra. Consideremos el complicado ejemplo del general inglés que capturó la ciudad de Sind y que envió, para comunicar dicho evento, el siguiente mensaje: «Peccavi». La ambigüedad de marras («I have Sind» — «I have sinned» \*\*) es fonética, no morfológica, y la expresión empleada no es en realidad ambigua; pero puesto que pertenece a una lengua que no es ni la del hablante ni la del oyente tampoco, su traducción se hace necesaria, y la ambigüedad hace acto de presencia en el proceso de verterla al inglés común.

Al margen de que la interpretación más natural («*I have sinned*»)

---

*Poems from MSS c. 1793.* Las líneas originales rezan así: «*Never seek to tell thy love / Love that never told can be.*» Véase *William Blake*, Penguin, 1958, p. 60. (N. del T.)

\* He aquí otro ejemplo tomado de Shakespeare. Los versos originales son éstos: «*But since she pricked thee out for women's pleasure, / Mine be thy love, and thy love's use their treasure.*» Pertenecen al soneto XX, al que cierran en realidad. Lo que importa al caso es no perder de vista la muy posible ambigüedad del término «*thy love*», que puede significar una cosa distinta en cada una de sus apariciones. En el primer caso, el poeta parece pedirle al conde de Southampton, noble de la corte isabelina, que le reserve a él tan sólo ciertos sentimientos —de corte platónico, nos dicen los críticos—; en el segundo, parece instarle a que dedique sus favores sexuales a las mujeres: «*thy love*» referiría en ese caso a sus atributos sexuales, los del susodicho noble. Esta interpretación explicaría muy bien por qué recurre Grice a estos versos en el análisis de la ambigüedad en el contexto de su teoría de las implicaturas. Mi agradecimiento al profesor Ian MacCandless, del Departamento de Lengua Inglesa de la Universidad de Granada, por la ayuda prestada durante el tortuoso recorrido por el shakesperiano soneto XX. (N. del T.)

\*\* Se explota aquí el hecho de que la fonética de «*I have Sind*» —«he capturado Sind»— y la de «*I have sinned*» —«he pecado»— es muy similar. (N. del T.)



sea la transmitida, la restante habrá de serlo también. Podría haber razones estilísticas que favorecieran la interpretación más natural, pero no tendría ningún sentido desde el punto de vista de la comunicación que nos tomásemos la molestia de dar con una expresión que nos permitiera comunicar que *p* de manera retorcida, obligando con ello a nuestros interlocutores a esforzarse por descubrir semejante interpretación, si ésta fuese ociosa. El que la interpretación más natural se imponga por sí sola es cosa que depende de si semejante supuesto entra en conflicto con otras exigencias conversacionales; por ejemplo, la de si resulta pertinente, la de si se trata de algo que podríamos suponer que el hablante acepta, y así sucesivamente. Si estas exigencias no se satisfacen, entonces no es la interpretación más natural la que se está comunicando. Pero si se cumplen, sí que lo es. Si cabe suponer que el autor de «peccavi» había cometido algún tipo de transgresión, había desobedecido órdenes al capturar Sind, pongamos como ejemplo, y si fuese oportuno referirse a esta transgresión al tener en cuenta los presuntos intereses de sus interlocutores, entonces habría transmitido ambas interpretaciones; en caso contrario, habríamos optado únicamente por una de ellas.

b) *Oscuridad*. ¿Cómo saco partido, con vistas a mis objetivos en la conversación, de una violación manifiesta y deliberada del requisito de que debería soslayar toda oscuridad en mi forma de expresarme? Obviamente, si el Principio Cooperativo ha de seguir incólume, mi intención será la de que mi interlocutor comprenda lo que yo estoy diciendo, a pesar de la oscuridad con que lastro mi preferencia. Supongamos que A y B están conversando en presencia de una tercera persona, de un niño, por ejemplo; A podría expresarse entonces de manera deliberadamente oscura, si bien no demasiado, con la esperanza de que B entienda lo que él dice sin que el niño se entere de ello. Aún más: si A espera que B se aperciba de que A está siendo deliberadamente oscuro, parece razonable suponer que, al efectuar su contribución conversacional de esta manera, A está implicando que el contenido de su mensaje no tendría que ser compartido por nadie más.

c) *Cuando no se expresa uno escueta o sucintamente*. Comparemos las observaciones:

- 1) «La Srta. X cantó *Hogar dulce hogar*.»
- 2) «La Srta. X emitió una serie de sonidos que guardaban alguna correspondencia con lo que es la canción *Hogar dulce hogar*.»

Supongamos que un testigo del caso ha optado por proferir 2), en lugar de 1). (*Glosa*: ¿Por qué ha preferido semejante charada en

lugar de la palabra, más concisa y casi sinónima suya, «cantó»? Es de presumir, que para indicar alguna diferencia chocante entre la interpretación que hizo la Srta. X de la citada canción y otras distintas a las que «cantar» se aplicaría habitualmente. La suposición más obvia es la de que la interpretación de la Srta. X adolecía de algún defecto horrible. El testigo sabe que es muy probable que este supuesto le venga a uno rápidamente a la mente; así pues, eso era lo que estaba implicando.)

Hasta el momento he considerado únicamente ejemplos de lo que podrían denominarse implicaturas conversacionales particularizadas; es decir, ejemplos en los que se lleva a cabo una implicatura al decir que *p* es una ocasión concreta y en virtud de rasgos específicos del contexto; ejemplos en los que se descarta de antemano la idea de que diciendo que *p* se lleve a cabo *normalmente* una implicatura de este tipo. Pero hay otros casos distintos: los de las implicaturas generalizadas. A veces puede uno decir que el uso de una cierta forma verbal daría usualmente lugar (al proferirla) a tal y cual implicatura o tipo de implicatura, *en ausencia* de circunstancias especiales. Ejemplos claros de este fenómeno son quizás difíciles de explicar, ya que muy fácilmente se tratan las implicaturas conversacionales generalizadas como si fuesen implicaturas convencionales. Doy un ejemplo que espero que pueda ser ampliamente aceptado.

Quienquiera que use una oración de la forma «X tiene una cita con una mujer esta tarde» implica normalmente que la persona con la que X va a verse no es ni su mujer, ni su madre, ni su hermana y ni siquiera, incluso, una amistad platónica. Análogamente, si uno dijese «X entró ayer en una casa y se encontró con que había una tortuga tras la puerta delantera», causaría una cierta sorpresa en su interlocutor si revelara, algo más tarde, que la casa en cuestión era la del propio X. Sería posible producir fenómenos lingüísticos similares mediante expresiones como «un jardín», «un colegio», «un automóvil», etc. No obstante, a veces no tendría lugar semejante implicatura («he estado toda la mañana en un automóvil») y a veces tendría lugar la inversa («pagué la factura de mis muebles con un cheque bancario») \*. Soy de la opinión de que no confiaríamos en el filósofo que sugiriera que la forma verbal «un X» tiene tres sentidos: uno de acuerdo con el cual dicha expresión significara aproximadamente «algo que cumple con las condiciones definitorias de la palabra X»; otro por el que equivaliera a algo así como «un X (en el primer sentido) que guarda una relación remota con alguna persona que el con-

\* El ejemplo original del autor es «*I broke a finger yesterdary*», que no puede traducirse con un mínimo de literalidad sin afectar con ello a la implicatura generalizada que se ejemplifica. (*N. del T.*)

texto indica quién es»; y un tercero por el que significara «un X (en el primer sentido) que guarda una relación estrecha con alguna persona que el contexto indica quién es». ¿No preferiríamos mucho antes una explicación que se atuviese a las siguientes pautas (que, naturalmente, pueden ser erróneas en algunos de sus detalles)? Cuando alguien implica, al usar la forma verbal «un X», que el X no es parte o no guarda una relación muy próxima con alguna persona susceptible de ser identificada, la implicación ha tenido lugar, pues el hablante no ha sido lo suficientemente explícito que cabía esperar de él; y eso tiene la consecuencia de que hay probablemente que asumir que no se encuentra en posición de serlo más. En lo que hace a las implicaturas, esta situación es familiar y se la puede calificar de malogro, pues la primera máxima de Cantidad no ha sido cumplida por una u otra razón. La única cuestión difícil es la de por qué habría de suponerse, en ciertos casos problemáticos y al margen de toda información relativa a los contextos de preferencia particulares, que existe una conexión próxima o remota entre una persona u objeto concretos y otra persona, la indicada o mencionada en la preferencia, que es probablemente de interés en sí misma. La respuesta ha de buscarse en lo siguiente: las transacciones habidas entre una persona y otras personas o cosas íntimamente relacionadas con la primera pueden diferir de otras transacciones del mismo tipo, en las que tan solo de una manera remota se hallen implicadas personas o cosas, en lo que hace a sus circunstancias concomitantes y a sus resultados; por ejemplo, las circunstancias concomitantes y los efectos de mi descubrimiento de un agujero en el tejado de *mi* casa serán probablemente muy distintos de los de mi descubrimiento de un agujero en el tejado de la casa de otra persona. Al igual que con el dinero, la información se da muy a menudo sin que el que la da sepa exactamente qué va a hacer con ella el receptor. Si alguien a quien se menciona una transacción reflexiona sobre ella, es muy probable que de pronto se vea a sí mismo deseando conocer las respuestas a ulteriores preguntas que el hablante puede no ser capaz por sí solo de hacerse; si una especificación apropiada de estas cuestiones capacita probablemente al oyente a darles respuesta, entonces está ahí presente el supuesto de que el hablante ha contado con la susodicha especificación al llevar a cabo su preferencia; si lo primero no es el caso, entonces no hay que presumir nada.

Finalmente, podemos mostrar ahora que, siendo una implicatura conversacional lo que es, ha de poseer ciertos rasgos.

1) Puesto que para asumir que estamos ante una implicatura conversacional hemos de suponer que se observa el Principio Cooperativo, y puesto que es posible optar por no hacer esto, se sigue de ahí que una implicatura conversacional generalizada puede cancelarse

en un caso particular. Se la puede cancelar explícitamente por medio de una cláusula que establezca o que implique que el hablante ha preferido no guardar el citado principio; o se la puede cancelar contextualmente, si la forma verbal proferida que habitualmente lleva consigo semejante implicatura es usada en un contexto tal que no deja duda ninguna de que el hablante *está* dejando en suspenso la aplicación del principio.

2) En la medida en que para inferir que se está ante una implicatura conversacional haga falta, además de información contextual y de fondo, tan solo el conocimiento de lo que se ha dicho (o del compromiso convencional de la preferencia), y en tanto en cuanto el modo de la expresión no juegue función alguna en la inferencia no será posible dar con otra forma de decir la misma cosa, por carecer de la implicatura en cuestión, a no ser que algún rasgo especial de la nueva versión sea relevante por sí solo para la determinación de una implicatura (en virtud de alguna de las máximas de modo). Si le damos a esta característica el nombre de «indesligabilidad», cabe esperar que toda implicatura conversacional generalizada que se lleve a cabo mediante una locución familiar, y nada especial, tenga un alto grado de indesligabilidad.

3) Hablando en términos aproximados, puesto que para inferir la presencia de una implicatura conversacional se presupone un conocimiento previo de la fuerza convencional de la expresión cuya preferencia lleva consigo la implicatura, el *implicatum* conversacional será una condición a no incluir a la especificación primera de la fuerza conversacional de la expresión. Aunque puede que no sea imposible que lo que se inicie en la vida, por decirlo así, como implicatura conversacional adquiera la naturaleza de implicatura convencional, suponer que en un caso dado esto sucede así requeriría de una justificación especial. Así pues, inicialmente al menos, los *implicata* conversacionales no son parte del significado de las expresiones a cuyo uso se adhieren.

4) Puesto que la verdad de un *implicatum* conversacional no se sigue de la verdad de lo que se dice (lo que se dice puede ser verdadero y lo que se implica puede ser falso), la implicatura no va unida inseparablemente a lo que se dice, sino al decir lo que se dice o al «expresarlo de esa manera».

5) Puesto que inferir una implicatura conversacional es inferir lo que se ha supuesto para salvaguardar la observancia del Principio Cooperativo, y dado que pueden darse varias explicaciones específicas posibles, quedando la lista de éstas abierta, en tales casos el *implicatum* conversacional será la disyunción de tales explicaciones específicas; y si la lista queda abierta, el *implicatum* tendrá el carácter de indeterminación que muchos *implicata* reales parecen de hecho poseer.

## VII. ALGUNAS PROPUESTAS RECIENTES



Recientemente ha surgido una serie de nuevas propuestas en filosofía del lenguaje de las que se recogen aquí tres de las más significativas.

En «Situaciones y actitudes» Barwise y Perry presentan un avance de lo que ha dado en llamarse «semántica de situaciones». Tal género de semántica se presenta como una alternativa a la semántica de mundos posibles, aun compartiendo con ella algunas de sus aspiraciones y técnicas. De acuerdo con la «concepción heredada», el valor semántico de un enunciado es (excepto en casos muy particulares) un valor de verdad bien sea en el mundo actual o en cada uno de los distintos mundos posibles. Este punto de vista, casi indiscutido desde Frege, es desafiado por Barwise y Perry al mantener que lo que designa un enunciado al ser usado en una ocasión particular es un conjunto de situaciones abstractas. Tales situaciones, clasificadas en tipos, se presentan como colecciones de secuencias de *objetos* que tienen *propiedades* o que mantienen *relaciones* entre sí, siendo los objetos, las propiedades y las relaciones ontológicamente básicos. El significado es entonces una relación entre diferentes tipos de situaciones.

En el artículo «Semántica del rol conceptual» G. Harman, con el propósito confesado de presentar de manera menos vaga de lo que es habitual «lo que podría querer decirse mediante el eslogan “El significado es el uso”», combina una serie de problemas que son centrales a la filosofía del lenguaje y a la filosofía de la mente. Dadas las tesis de que: *a*) el significado de las expresiones lingüísticas viene determinado por los conceptos y pensamientos para cuya expresión se usan, y *b*) el contenido de los conceptos y pensamientos está determinado por su «rol funcional» en la psicología de una persona, Harman concluye que el uso de símbolos en el «cálculo» y en el pensamiento es más básico que su uso en la comunicación. En una posición completamente distinta se sitúa el artículo de Deirdre Wilson y Dan Sperber «Sobre la definición de relevancia»; éstos intentan proporcionar una explicación del concepto (máxima) griceano(a) de relevancia que resulte adecuado dentro de la armazón de una teoría pragmática del significado. El foco de interés de este artículo se dirige a mostrar que existe una estrecha conexión entre relevancia e inferencia y que la explicación de aquélla en términos inferenciales puede producir resultados fructíferos para la teoría pragmática.

## SITUACIONES Y ACTITUDES \*

JON BARWISE Y JOHN PERRY

Al leer los trabajos pioneros de lógica de Frege y Russell, difícilmente puede uno dejar de impresionarse por la medida en que la preocupación de entender los verbos de las actitudes cognitivas, es decir, verbos como *preguntarse*, *creer* y *saber* configuró su pensamiento. A pesar de esta preocupación, y de todo el progreso subsiguiente en lógica, no hay todavía ninguna explicación sistemática de la lógica de las actitudes que sea satisfactoria.

En este artículo esbozamos un enfoque que creemos va a llevar a una explicación satisfactoria y sistemática. Como muestra de los verbos en cuestión tomamos *ver*, *saber*, *creer* y *decir*. Llamamos *semántica de situaciones* a nuestra teoría; su espíritu está más cerca de Russell que de Frege en algunos aspectos fundamentales. Comenzaremos con algunas de las características clave de la semántica de situaciones para pasar a continuación a la discusión de cuestiones filosóficas y semánticas que giran alrededor de las actitudes. En preparación está una semántica rigurosa para un fragmento del inglés que incluya estos verbos, además de expresiones indicadoras, demostrativos y descripciones definidas, nombres propios, pronombres y condicionales.

---

\* Versión castellana de Daniel Quesada.

Este artículo será presentado en el simposio de la *American Philosophical Association* sobre «La lógica de la percepción y la creencia», el 30 de diciembre de 1981. Será comentado por Richmond Thomason; su artículo no está disponible en este momento.

Este artículo representa el trabajo en común de los autores: el orden de los nombres es puramente alfabético. El primer autor expresa su agradecimiento a la *National Science Foundation*. El artículo se completó mientras el segundo autor era investigador del *Center for Advanced Study in the Behavioral Sciences* [Centro para Estudios Avanzados en las Ciencias de la Conducta], y está agradecido por el apoyo del Centro, así como del *National Endowment for the Humanities*, del *Andrew Mellon Foundation* y de la Universidad de Stanford. Ambos autores le están agradecidos al Centro por proporcionar generosamente los colaboradores necesarios.

Estamos en deuda con John Etchemendy, Michael Turvey y otras personas del Centro y de Stanford por las conversaciones mantenidas con ellos. Nuestro punto de vista experimentó la influencia profunda de Turvey y otras personas que trabajan en la tradición del realismo ecológico. Para una introducción de este punto de vista, el lector puede consultar el libro de Claire Michaels y Claudia Carello, *Direct Perception*, Prentice-Hall, Englewood Cliffs, N.J., 1980.



## SITUACIONES

Las situaciones son básicas y omnipresentes. Siempre estamos en una u otra situación. La actividad cognitiva humana categoriza estas situaciones en términos de objetos que poseen atributos y están en relaciones unos con otros en ciertas ubicaciones (regiones espacio-temporales conectadas). Los lenguajes humanos reflejan (e intensifican) esta actividad cognitiva al proporcionarnos un modo de comunicar información relativa a las situaciones, tanto aquéllas en que nos encontramos como las que están alejadas de nosotros en el espacio y en el tiempo.

En el intento de desarrollar una teoría del significado lingüístico que se centre en las situaciones, reconocemos la primacía epistemológica de las situaciones, pero seguimos el ejemplo del lenguaje y tomamos, como primitivos de nuestra teoría, objetos, relaciones y ubicaciones, reconstruyendo las situaciones a partir de ellos. Tenemos así como primitivos:

- (i) un conjunto  $A$  de individuos  $a, b, c, \dots$ ;
- (ii) un conjunto  $R$  de relaciones,  $R = R_0 \cup R_1 \cup \dots \cup R_n \cup \dots$ , donde  $R_n$  consta de las relaciones  $n$ -arias; y
- (iii) un conjunto  $L$  de ubicaciones espacio-temporales  $l, l_1, \dots$

Una *situación*  $s$  se caracteriza por su ubicación  $l$  y su *tipo*  $s$ ,  $s = \langle l, s \rangle$ . El tipo representa cuáles son los objetos y cuáles las relaciones en que se encuentran en la ubicación de que se trate. Representamos estos tipos por medio de funciones parciales de relaciones  $r \in R_n$  y secuencias de objetos  $\langle a_1, \dots, a_n \rangle$  a 1 (verdadero) y 0 (falso)<sup>1</sup>. La función parcial  $s_0$  definida por:

$$\begin{aligned} s_0(\text{despierta, Jackie}) &= 1 \\ s_0(\text{despierta, Molly}) &= 0 \end{aligned}$$

se realizará en las situaciones  $s$  en las que la perra del primer autor está despierta y la del segundo dormida, independientemente de lo que esté haciendo la del lector, si es que ella o él tiene una. ( $s_0$  se realiza en  $s = \langle l, s \rangle$  si  $s_0 \subseteq s$ .) Utilizamos la letra  $S$  para el conjunto de los tipos de situaciones ( $\sigma$  situaciones-tipo)  $s, s_0, s_1, \dots$  y  $\mathbf{S} (= L \times S)$  para el conjunto de situaciones  $s, s_0, s_1, \dots$

Un *curso de eventos*  $\sigma$  es una función parcial del conjunto  $L$  de ubicaciones en  $S$ . Así todo curso de eventos es también un conjunto

<sup>1</sup> Consideramos a los valores veritativos como algo que se desliza en el universo en el proceso de abstracción que va de situaciones a objetos que están o no están en varias relaciones.

de situaciones, una como máximo para cada ubicación  $l$  dada. Si  $l$  es el dominio ( $\sigma$ ), escribimos  $\sigma_l$  para la situación-tipo  $\sigma(l)$ . Utilizamos  $\Sigma$  para el conjunto de todos los cursos de eventos. Un *curso de eventos total* es un curso de eventos que está definido para todas las ubicaciones. Señalamos uno entre ellos como el *curso de eventos real*  $\sigma^*$ . Una situación  $s = \langle l, s \rangle$  es *real* si  $s \subseteq \sigma_l$ ; es decir, si el tipo de  $s$  es parte de o todo lo que sucede realmente en  $l$ .

Una proposición (realista) es un conjunto  $P \subseteq \Sigma$  que satisface:  
(Monotonía)  $\sigma \in P$  y  $\sigma \subseteq \sigma'$  implica  $\sigma' \in P$ .

El adjetivo «realista» se utiliza aquí para hacer hincapié en que se trata de constructos de objetos, propiedades y ubicaciones reales, no de cosas que estén en la cabeza de alguien. [Una proposición no ubicada (realista) es un conjunto  $P \subseteq S$  que satisface una restricción de monotonía similar:  $s \in P$  y  $s \subseteq s'$  implican  $s' \in P$ .]

En inglés se representan tres relaciones básicas entre ubicaciones espacio-temporales:

$l_1 \circ l_2$	$l_1$ se solapa temporalmente con $l_2$
$l_1 < l_2$	$l_1$ precede temporalmente y totalmente a $l_2$
$l_1 @ l_2$	$l_1$ se solapa espacialmente con $l_2$

Consideramos a éstas como relaciones extensionales en  $L$ , relaciones a partir de las cuales pueden construirse «instantes» de tiempo y «puntos» del espacio al modo de Whitehead y Russell.

El punto de partida de la semántica de situaciones es que los enunciados indicativos sin marcador de tiempo describen o designan tipos de situaciones o, dicho de otro modo, situaciones-tipo, y que los enunciados indicativos con marcador temporal designan proposiciones, conjuntos de cursos de eventos. Empleamos aquí el término *enunciado* deliberadamente, pues una oración como *Yo estoy sentado* puede utilizarse para hacer tantos enunciados como hablantes y momentos de tiempo haya para proferirla. La oración tiene un «significado» fijo, pero los diferentes enunciados describirán eventos distintos. Es decir, los diferentes enunciados tendrán «interpretaciones» distintas. Esta distinción entre significado e interpretación es el tema del epígrafe que sigue a continuación.

## SIGNIFICADO E INTERPRETACIÓN

Pueden desarrollarse varios aspectos importantes de la semántica de situaciones comentando las siguientes oraciones simples:

- 1) *Yo estoy sentado.*
- 2) *Sandy está sentado.*
- 3) *Ella estaba sentado.*

Vamos a empezar por la palabra *yo*. Algo que es razonable decir sobre esta expresión es que, en cada ocasión en que un hablante del inglés (o del español) la utiliza, está en lugar de, o designa a, esa persona. Creemos que eso es todo lo que hay que saber sobre el significado de *yo* en inglés (o en español) y que ello sirve como regla paradigmática para el significado.

Considérese la relación:

La expresión  $\alpha$  (del lenguaje  $L$ ) al ser usada por  $x$ , está en lugar de  $y$ ,

que nosotros escribimos así:  $[\alpha](x, y)$ . Nuestro candidato a teoría del significado para el lenguaje  $L$  es una teoría que nos diga todas las condiciones en que se da  $[\alpha](x, y)$ . Lo que se nos dice sobre *yo* nos proporciona una de las condiciones:

$$[\text{yo}](a, y) \text{ syss } a = y.$$

Esta concepción relacional del significado exige que se preste sistemáticamente atención a los valores adecuados de cada una de las coordenadas. Nuestro punto de partida en semántica de situaciones es que, cuando la primera coordenada  $\alpha$  sea una oración de indicativo con marcador temporal, los cursos de eventos constituyen la adecuada tercera coordenada  $y$ .

Esta decisión tiene algunas repercusiones para la segunda coordenada  $x$ . Muestra que sería demasiado simple escoger para esta coordenada a los hablantes. La oración (1) puede ser usada por una misma persona en lugares espacio-temporalmente distintos para describir eventos distintos. Análogamente, la designación de *tú*, *ahora*, *ella*, *esto*, *estaba* varía de preferencia a preferencia, dependiendo de a quién le está hablando el hablante, cuándo, acerca de quién, de qué le habla y cuándo ocurría lo dicho. De modo que representamos los hechos específicos de la preferencia haciendo referencia a *situaciones del discurso* y a *conexiones*.

Una *situación del discurso*  $d$  representa la situación en que se encuentran el hablante y la persona o animal al que éste se dirige. Consiste de una situación  $s_d = \langle l_d, s_d \rangle$  con un individuo señalado  $a_d$  tal que  $s_d$  (habla,  $a_d$ ) = 1. Modificamos de este modo la regla para *yo* que dimos anteriormente:

$$[\text{yo}](d, y) \text{ syss } y = a_d.$$

Análogamente, *ahora* restringe el tiempo al que se hace referencia de manera que se solape con el momento de tiempo de la preferencia; por ello definimos

$$[\text{ahora}](d, y) \text{ syss } y \in L \text{ e } y \circ l_d.$$

Similarmente,

$$[aquí] (d, y) \text{ syss } y \in L \text{ e } y @ l_d.$$

Sin embargo, frecuentemente existen hechos específicos de la preferencia que poseen constituyentes que no están presentes en la situación de discurso de que se trate. Considérense, por ejemplo las oraciones 2) y 3) anteriores. Es razonable suponer que en una preferencia de 2) [o de 3)] susceptible de interpretación, *Sandy* está en lugar de *Sandy* (o que *ella* está en lugar de una hembra). Pero ¿qué *Sandy*? ¿*Sandy Koufax*, *Sandy Dennis* o la perra de la pequeña huérfana *Anita*? Lo que aquí queda por explicar es que cuando se usa 2) con significado, se trata de un uso acerca de algún individuo determinado, *Sandy*, [y que cuando se usa 3) con significado, se trata de un uso acerca de alguna hembra determinada]. Puesto que no es necesario que estos individuos estén presentes en la situación del discurso de que se trate, no tenemos más remedio que reconocer otro componente de nuestra segunda coordenada, componente que representa las conexiones *c* entre ciertas palabras y las cosas del mundo, que están implícitas en cualquier uso de esas palabras que tenga significado. De este modo, una preferencia de 2) en la que el hablante estuviera hablando sobre *Sandy Koufax* se representaría mediante la expresión 2), un determinado discurso *d*, y una función parcial *c* tal que  $c(\textit{Sandy}) = \textit{Sandy Koufax}$ . Podemos, por consiguiente, representar los significados de *Sandy* y de *ella* mediante

$$\begin{aligned} [Sandy] (d, c, y) \text{ syss } c(Sandy) = y \text{ e } y \text{ se llama } Sandy. \\ [Ella] (d, c, y) \text{ syss } c(ella) = y \text{ e } y \text{ es una hembra.} \end{aligned}$$

(Incluso esto es demasiado simple, pero es suficiente por ahora.)

Con esto hemos concluido con los sintagmas nominales de 1) a 3) y tenemos a mano las herramientas para habérmolas con los sintagmas verbales *estoy sentado*, *está sentado* y *estaba sentado*. Todos ellos constituyen formas del verbo *sentar(se)*. Como la mayoría de los verbos, su interpretación es i. s. d. —insensible a la situación del discurso en que se profiere—. Sin embargo, *sentar(se)* puede utilizarse para designar bien sea una actividad,  $\textit{sentarse}_a \in R_1$ , la actividad de sentarse, o un estado,  $\textit{sentado}_e \in R_1$ , el estado de estar sentado. Cuál de las dos cosas, depende del hablante \*. De este modo, aparecen de nuevo las conexiones:

$$[sentar(se)] (d, c, y) \text{ syss } c(sentar(se)) = y \text{ e } y = \textit{sentarse}_a \text{ o } y = \textit{sentado}_e.$$

\* La ambigüedad del verbo inglés *sit* que los autores señalan no tiene paralelo en castellano. En el original inglés las oraciones 1)-3) son ambiguas. Al traducir las ora-

Vamos ahora a ocuparnos del tiempo de 2)-3). Como con *ahora*, las formas de presente de 1) y 2) indican que el sentarse o estar sentado tiene lugar en un momento de tiempo que se solapa temporalmente con el de la preferencia. Las formas de pasado se utilizan para indicar que el sentarse o el estar sentado tuvo lugar en el pasado. Pero, del mismo modo en que parte del significado de *ella estaba sentada* estriba en que para enunciar algo con esta oración debe utilizársela aludiendo a una hembra determinada, también ocurre que para enunciar algo debe utilizársela aludiendo a una determinada ubicación espacio-temporal. Para interpretar correctamente mi afirmación de que ella estaba sentada, debe usted interpretar correctamente mis usos de *ella* y de *estaba* como usos relativos a una hembra y una ubicación espacio-temporal del pasado. Para representar las conexiones entre marcadores temporales y ubicaciones espacio-temporales, dejamos que nuestras conexiones asignen ubicaciones espacio-temporales a los marcadores temporales. Tenemos así lo siguiente:

Si  $\alpha$  es *estoy/estás/está*, entonces  $|\alpha|(d, c, y)$  syss  $c(\alpha) = y \in L e y \circ l_d$ .

Si  $\alpha$  es *estaba/estabas/estaba*, entonces  $|\alpha|(d, c, y)$  syss  $c(\alpha) = y \in L e y < l_s$ .

Cuando fijamos todo lo que es específico de una determinada preferencia de una expresión  $\alpha$  obtenemos lo que denominamos las *interpretaciones* de la preferencia. De este modo, si fijamos una cierta expresión  $\alpha$ , una situación de discurso  $d$  y una conexión  $c$ , obtenemos los  $y$  tales que se da  $|\alpha|(d, c, y)$ , lo cual escribiremos también así:  $y \in_{d,c} |\alpha|$ . Si hay un único  $y$  tal, a ese  $y$  le llamamos *la* interpretación de la preferencia  $\alpha$ ,  $d$ ,  $c$ ,  $y$  escribimos  $_{d,c} |\alpha| = y$ . Así, por ejemplo,  $_{d,c} |\text{yo}| = a_d$  y  $_{d,c} |\text{estaba}| = c(\text{estaba})$ , una cierta ubicación  $l < l_d$ .

Podemos ahora asignarles significados a todas las oraciones  $\alpha$  de la forma

#### SN AUX SV

donde  $SN \in \{\text{yo, Sandy, ella}\}$ ,  $AUX \in \{\text{estoy, estás, está, estaba, estabas}\}$ , y  $SN = \text{sentar}(se)$ , a saber:

$$|\alpha|(d, c, \sigma) \text{ syss } \sigma_i(P, a) = 1$$

ciones al castellano, esta ambigüedad no puede conservarse. En el texto posterior a los ejemplos he tratado de sugerir la ambigüedad traduciendo por *sentar(se)*. Pero en castellano el análisis gramatical del verbo habría de ser diferente (difícilmente podría mantenerse que se trata de un solo verbo ambiguo). Dado que el ejemplo lo han elegido los autores a efectos expositivos, la diferencia entre los dos idiomas no es pertinente para lo que ellos pretenden, pero ha de tenerse bien presente para entender el texto ligado a este ejemplo. (*N. del T.*)

donde  ${}_{d,c}[AUX] = I$ ,  ${}_{d,c}[SV] = P$  y  ${}_{d,c}[SN] = a$ . Una preferencia de  $a$  describe un curso de eventos  $\sigma$  en el caso de que  $a$  esté sentado (o se esté sentado <sup>(\*)</sup>) en la ubicación propuesta  $l$  del curso de eventos  $\sigma$ . Nótese que  ${}_{d,c}[a]$  es una proposición, es decir, un conjunto monótono de cursos de eventos.

No puede exagerarse la importancia que la distinción significado/interpretación tiene para entender las actitudes. Esa importancia descansa en parte en los dos hechos siguientes, hechos que están relacionados entre sí: 1) la eficiencia y 2) la relatividad de perspectiva.

*Eficiencia:* Una expresión  $\alpha$  con un solo significado  $[\alpha]$  puede utilizarse en circunstancias diferentes con interpretaciones distintas.

Una palabra como *yo*, por ejemplo, puede usarse para designar a cualquiera de nosotros. Aunque esto hace que el lenguaje sea eficiente, permitiendo que una expresión dada pueda usarse una y otra vez para fines diferentes, también tiene un corolario. Una oración que describa una situación dada desde la perspectiva de una persona, no describirá, en general, la misma situación desde una perspectiva distinta. A fin de que puedan alcanzarse las mismas situaciones, un lenguaje humano satisfará el siguiente principio:

*Relatividad:* Expresiones diferentes con significados distintos pueden usarse en circunstancias diversas con una misma interpretación.

Así, por ejemplo, para expresar el hecho de que yo estoy sentado (ahora), usted, el futuro lector, podría decir: «Él estaba sentado»<sup>2</sup>.

Alguien podría decir que el significado es una función de la situación de discurso y las conexiones a la interpretación. De este modo la proposición,  ${}_{d,c}[a]$ , es el único curso de eventos  $\sigma$  tal que se da  $[\alpha](d, c, \sigma)$ . Esto no es totalmente erróneo, pero puede inducir a engaño. La interpretación, el conjunto que obtenemos cuando fijamos la expresión, la situación del discurso y las conexiones, es muy importante. Constituye un tipo de uniformidad en las preferencias, y es una uniformidad muy importante; el reconocimiento de su importancia está integrado directamente en el lenguaje. Pero al recordar que el significado es una relación, se nos llama la atención sobre otras varias

<sup>2</sup> El lector sagaz se dará cuenta de que sus conexiones al describir mi situación son:

$c(\text{él}) = \text{yo}$                        $c(\text{estaba}) = \text{aquí y ahora}$

Estas conexiones están objetivamente determinadas por quién fuera, de los dos autores, el que escribió esta parte del artículo, por dónde y cuándo lo hizo, y son independientes de la capacidad del lector de especificarlas de alguna manera más completa. Por otra parte, la lectura que el lector realiza está conectada, causalmente, con la escritura.

uniformidades importantes, uniformidades que son cruciales para la comprensión de las actitudes.

Junto a la interpretación, podemos construir varias «interpretaciones inversas», manteniendo fija la última coordenada del significado y dejando que las otras varíen. Estas interpretaciones inversas se utilizan en la vida diaria, y se las involucra frecuentemente cuando se habla de «condiciones de verdad» y de «cuándo una oración es verdadera». Supóngase, por ejemplo, que decimos que una niña *b* entiende *esto es leche*, porque lo dice sólo cuando es verdad. Lo que queremos decir es que lo dice sólo cuando su atención está fijada en un vaso de leche. Estamos recurriendo a

$$\{ \langle d, c \rangle \mid \text{esto es leche} \langle d, c, \sigma^* \rangle \text{ y } a_d = b \}.$$

Y cuando pensamos que ésta es una buena prueba de la comprensión de ese tipo de oración, ello se debe a que pensamos que hay alguna uniformidad, que  $a_d$  puede descubrir por medio de la percepción, a través de ese conjunto, o de un subconjunto significativo del mismo. La uniformidad no es la interpretación, si suponemos que la niña interacciona con botellas de leche diferentes en momentos de tiempo distintos<sup>3</sup>.

Algunos filósofos sostienen que los verdaderos vehículos del significado no pueden ser ni relativos ni eficientes; esta concepción conduce al atomismo lógico. Otros piensan que los verdaderos vehículos pueden ser relativos, pero no eficientes. Esto conduce a la idea de que las oraciones eficientes del lenguaje natural deben estar respaldadas por sentidos o representaciones mentales que «tensen la cuerda», es decir, que sean «completas en todos los respectos». Nosotros pensamos que las oraciones del lenguaje natural son verdaderos vehículos del significado, y que son otros factores de la preferencia los que «tensan la cuerda». El aprendizaje del lenguaje necesita de la coordinación del lenguaje con partes del entorno más o menos remotas, no con sentidos o representaciones mentales.

## ACTITUDES INOCENTES

A los enunciados que se hacen con oraciones del tipo siguiente les llamamos informes de actitud:

<sup>3</sup> Para un análisis de algunos de estos temas, véase «Perception, Action and the Structure of Believing». [«La percepción, la acción y la estructura del creer»], de próxima aparición en un volumen de homenaje a Paul Grice recopilado por Richard Grandy y Richard Warner. [Se trata de *Philosophical Grounds of Rationality*, Clarendon Press, Oxford, 1986. (*N. del T.*)]

- 4) *Inés nos vio saltar en la fuente.*
- 5) *Inés vio que yo estaba sentado en la fuente.*
- 6) *Inés sabía que yo estaba nervioso y cansado.*
- 7) *Inés dijo que yo estaba bebido.*
- 8) *El policía creyó lo que dijo Inés.*

Los verbos de actitud se combinan con oraciones para producir sintagmas verbales que se utilizan para clasificar individuos. Nótese, sin embargo, que en la interpretación de un determinado informe en el que se utilice cualquiera de las oraciones 4)-8) interviene la interpretación de la oración subordinada, y *no* su significado. Un enfoque semántico sencillo, que constituye una primera aproximación al nuestro, estriba en considerar que los verbos de actitudes expresan una relación entre un individuo y la interpretación del enunciado subordinado. Por ejemplo, una preferencia de 6) expresaría una relación entre Inés y el hecho de que yo estaba nervioso y cansado, un complejo que me incluye a mí, a dos propiedades y a una ubicación.

Este enfoque de las actitudes ejemplifica lo que Donald Davidson llama «inocencia semántica»:

Si pudiéramos recuperar nuestra inocencia semántica pre-fregeana, pienso que nos parecería sencillamente increíble que las palabras «la tierra se mueve», proferidas tras las palabras «Galileo dijo que», signifiquen algo diferente, o refieran a algo distinto de lo que acostumbran cuando aparecen en otros entornos <sup>4</sup>.

En el enfoque que se acaba de esbozar, los enunciados subordinados y sus constituyentes tienen exactamente el mismo significado e interpretación que cuando no aparecen subordinados.

Las objeciones tradicionales a un enfoque inocente se han basado en la creencia de que la única interpretación (referencia) plausible de una oración es su valor veritativo, el cual elimina el contenido de la oración. Esta creencia de que un valor de verdad constituye la única interpretación plausible de una oración se ha apoyado mediante un argumento formal al que nosotros llamamos «la honda». En otro lugar hemos mostrado que este argumento depende de que se deje de lado desde el comienzo mismo la posibilidad de una semántica que se base en situaciones <sup>5</sup>.

<sup>4</sup> «On Saying That» [«Sobre el decir que»], reimpreso en Donald Davidson y Gilbert Harman, (eds.), *The Logic of Grammar*, Dickenson, Encino, Cal., 1975, p. 152. Publicado originalmente en *Synthese*, XIX (1968-1969), pp. 130-146. [Version en castellano en D. Davidson, *De la verdad y de la interpretación*, Gedisa, Barcelona, 1990.]

<sup>5</sup> «Semantic Innocence and Uncompromising Situations» [«Inocencia semántica y situaciones no comprometidas»], *Midwest Studies in Philosophy*, VI (1981), pp. 387-403.



Una vez eliminada esa objeción, un enfoque inocente parece completamente natural. Al centrar la atención en la interpretación (no en el significado) de la oración subordinada, se nos permite de una manera sencilla dar razón de la relatividad de perspectiva de la oración subordinada.

En el caso de informes de actitud en primera persona y en tiempo presente, una oración expresiva se utiliza en el informe como oración subordinada. Con la oración *Yo creo que (yo) estoy sentado*, informo de la creencia que yo expresaría con *Yo estoy sentado*. Pero, en general, las oraciones que utilizamos para informar de las actitudes de los demás, o las que ellos usan para informar de sus propias actitudes pretéritas, no son las oraciones que ellos utilizarían o que habrían utilizado para expresar esas actitudes. Así, usted informará de mi creencia con *Él creía que (él) estaba sentado*, no con *Él creía que yo estaba sentado*. Lo mismo vale para los demás verbos de actitud (VA). Las actitudes son actitudes dirigidas a proposiciones. Un informe de actitud SN VA $\alpha$  apunta a una proposición  $P$  utilizando una oración subordinada  $\alpha$  cuya interpretación, desde la perspectiva del hablante,  $(d, c)$ , es  $P$ . El agente  $a$  ( $= {}_{a,c}[\text{SN}]$ ) tendría que utilizar alguna oración expresiva  $\alpha'$  cuya interpretación relativa a su propia perspectiva,  $(d', c')$ , sería también  $P$ :

$$P = {}_{a,c}[\alpha] = {}_{a',c'}[\alpha']$$

Lo que podría denominarse «teoría heredada» de las actitudes (dejando a un lado verbos como *ver* y otros verbos de percepción) es, más o menos, la siguiente. Las actitudes son relaciones que se dirigen a oraciones, significados de oraciones, sentidos de oraciones o representaciones mentales. Un informe de actitud SN VA $\alpha$  informa sobre la actitud de un agente hacia la oración  $\alpha$  o hacia una representación mental asociada con  $\alpha$  de algún modo. Este es un informe de actitud «*de dicto*». La oración no se usa inocentemente para referir a lo que usualmente refiere, sino para referirse a sí misma, a su significado o sentido o a una representación mental.

Para explicar el fenómeno que hemos estado comentando, la disparidad entre la oración subordinada del hablante y la oración expresiva del agente, la teoría heredada admite que algunas veces *se informa* de las actitudes de una manera distinta, pero mantiene que la actitud misma es una actitud que se dirige hacia uno de los tipos mencionados de objetos. De este modo, en los llamados informes «*de re*», algunas partes de la oración subordinada no se utilizan para aportar su significado (sentido, etc.), sino para identificar, por ejemplo, a un cierto individuo,  $b$ . Tal informe *de re*, se afirma, significa que el agente tiene una actitud dirigida hacia una oración o significado que tiene a  $b$  como referencia de una de sus partes. Usualmente se

dejan de lado los problemas relativos al tiempo, pero presumiblemente se tratarían del mismo modo.

Hay graves problemas con todas las versiones de la teoría heredada, siendo estos problemas los que nos impulsan a trabajar en una semántica inocente. En primer lugar, lo que a nosotros nos parece un fenómeno sencillo, tiene en la teoría heredada una explicación extremadamente complicada que nunca ha sido elaborada en detalle [considérese 8), por ejemplo].

La idea de que las actitudes son relaciones que se dirigen a oraciones es plausible en el caso del decir, poco plausible en el caso de la creencia y del saber y totalmente extravagante en el caso de la percepción. Incluso en el caso del decir, la teoría no funciona sin problemas, como subraya Davidson.

Cuando pasamos de las oraciones a los significados, los sentidos o las representaciones mentales, nos vemos acosados por varios problemas. Se apela frecuentemente a la noción fregeana de sentido como si fuera una herramienta técnica bien desarrollada, cuando no lo es. Los intentos de elaborar una teoría bien desarrollada de los sentidos se encuentran con graves problemas técnicos, problemas que reflejan las objeciones filosóficas a la noción misma de sentido.

La semántica de índices o «mundos posibles» que se desarrolló para la lógica modal, nos ofrece, una vez adaptada a las actitudes, una alternativa más para la designación de la oración subordinada, a saber, su «intensión», el conjunto de mundos posibles en los que la oración es verdadera. Pero, aun si uno piensa que la idea primitiva de mundo posible tiene algún sentido, se suscita el problema de la equivalencia lógica. Considérense, por ejemplo,

9) *Quico ve entrar a Isabel.*

10) *Quico ve entrar a Isabel y (fumar a Sara o no fumar a Sara).*

No podemos de ningún modo pasar de 9) a 10), no importa lo dotado que esté Quico para la lógica. Si lo hiciéramos, tendríamos que admitir que Quico o bien vio fumar a Sara o bien vio no fumar a Sara, aunque no haya puesto nunca sus ojos en Sara. Esta admisión la forzarían los principios siguientes:

Si Quico ve  $P$  y  $Q$ , entonces Quico ve  $Q$ .

Si Quico ve  $P$  o  $Q$ , entonces Quico ve  $P$  o Quico ve  $Q$ .

(No esperaríamos la «omnipercipiencia» ni entre los lógicamente omniscientes<sup>6</sup>.)

<sup>6</sup> Véase el artículo «Scenes and Other Situations» [«Escenas y otras situaciones»] del primer autor, en este *Journal*, LXXVIII, 7 (julio 1981), pp. 369-397.

La semántica de situaciones y la inocencia semántica resuelven el problema de la equivalencia lógica. No se asigna la misma proposición a oraciones lógicamente equivalentes, incluso respecto a la misma situación de discurso; asuntos diferentes proporcionan conjuntos distintos de situaciones-tipo. Este es el dividendo que se obtiene usando libremente las funciones parciales. En realidad, desde el punto de vista de la semántica de situaciones, la expresión *lógicamente equivalentes* debería utilizarse para oraciones que sean verdaderas en las mismas situaciones-tipo o cursos de eventos, no para las que satisfagan la condición más débil de ser verdaderas en los mismos tipos o cursos de eventos totales. Sin embargo, al poseer la expresión *equivalencia lógica* un uso muy arraigado, nos referiremos a esta relación más estricta como *equivalencia fuerte*.

Por consiguiente, nuestro enfoque inocente tiene una sencillez natural, resuelve algunos problemas y evita otros. Sin embargo, hay algunas dificultades.

#### LA INOCENCIA AMENAZADA

En esta sección señalamos cuatro problemas que constituyen una amenaza para nuestra explicación de las actitudes, problemas que indican que en nuestra teoría falta un cierto elemento.

*La lógica de las actitudes.* Existen varios hechos que implican a las actitudes y que parecen requerir una explicación semántica, hechos especialmente claros en el caso de los informes de percepción epistémicamente neutros [*ve*, por contra de *ve que*, como en la oración 4) de más arriba]<sup>7</sup>. Anteriormente hemos enunciado dos de ellos:

- (i) si  $a$  ve  $\phi$  y  $\psi$ , entonces  $a$  ve  $\phi$  y  $a$  ve  $\psi$ .
- (ii) si  $a$  ve  $\phi$  y  $\psi$ , entonces  $a$  ve  $\phi$  y  $a$  ve  $\psi$ .
- (iii) si  $a$  ve  $\phi$ , entonces  $\phi$ .
- (iv) si  $a$  ve  $\phi(t_1)$  y  $t_1 = t_2$ , entonces  $a$  ve  $\phi(t_2)$ .

Hasta aquí nuestro enfoque proporciona una explicación tan sólo de (iv)<sup>8</sup>.

*La opacidad.* El lector sagaz habrá notado que nuestra explicación viola la sabiduría original sobre las actitudes que debemos a Frege y a Russell —la afirmación de que la sustitución de expresiones

<sup>7</sup> Véase Barwise, *op. cit.*, para más detalles sobre estas cuestiones.

<sup>8</sup> Las limitaciones de espacio impiden considerar las relaciones lógicas entre las actitudes, como, por ejemplo, la afirmación de que saber implica creer. Esperamos que el lector pueda ver más o menos lo que nosotros diríamos, basándose en nuestro análisis de las diferentes actitudes.

correferenciales no preserva el valor veritativo en los informes de actitud—. Después de todo, si con nuestra explicación se predice (iv), entonces habrá de hacerse una afirmación similar para todas las actitudes. Al considerar que las actitudes son relaciones con objetos, propiedades y relaciones reales, nos comprometemos a sostener que son, en algún sentido, transparentes.

*La falta de objetos para las actitudes.* En nuestra explicación se toman proposiciones como objetos de las actitudes. Esto no es filosóficamente satisfactorio cuando se trata del ver y el decir. Con el ver se pierde la conexión con lo que realmente se vio, y con decir, se pierde la conexión con lo que realmente se profirió. Supóngase que un hablante  $a$ , en una situación de discurso  $d$  y con conexiones  $c$  dice

(11)  $b$  dijo que  $\phi$ .

Hemos visto que  $\phi$  mismo no puede, en general, servir para lo que  $b$  profirió realmente. Pero de la verdad de (11) se sigue con toda seguridad que  $b$  profirió algo, que hay alguna oración  $\psi$  proferida por  $b$  tal que, desde la ubicación  $d'$  del discurso de  $b$  y con las conexiones  $c'$  de  $b$ ,  $d, c[\phi] = d', c'[\psi]$  y  $c(\text{dijo}) = l_{d'}$ . Similarmente, si  $a$  dijo con verdad

(12)  $b$  vio  $\phi$ .

entonces lo que  $b$  vio en realidad es una escena en la que se daba  $d, c[\phi]$ , y no una proposición.

*El contenido cognitivo de las actitudes.* En nuestra explicación falta el hecho de que las actitudes tienen algo que ver con las mentes (o los cerebros) y los fenómenos cognitivos. Del mismo modo que el decir requiere un agente que profiera algo con un significado, y que el ver requiere que el agente vea algo con sus ojos (que son parte del cerebro), también el creer y el conocer requieren que el agente esté en un estado cognitivo con significado. Parte de lo que nos dan los informes de actitud constituye información relativa al estado cognitivo del agente. Esto es lo que hace que los informes de actitud sean útiles en la explicación y en la predicción de lo que la gente hará. Las personas que tienen percepciones, creencias y deseos similares se comportan de un modo similar.

El concentrarse en este aspecto del problema hace que los sentidos fregeanos parezcan atractivos. Interpretando una actitud como una relación con un sentido, o con una «representación mental», como se tendría en algunas versiones de la teoría, se ve al objeto de la actitud como si clasificara los estados cognitivos. Según esta teoría la similitud de las actitudes apunta a la similitud de los estados, expli-

cando aparentemente la similitud de las acciones <sup>9</sup>. Sin embargo, según nuestra teoría, diferentes personas podrían creer exactamente la misma cosa de innumerables modos distintos. Hasta aquí la teoría no refleja ninguna similitud cognitiva.

Estos cuatro problemas están relacionados entre sí, y todos ellos requieren que meditemos en qué es lo que estamos haciendo cuando intentamos formular una teoría semántica de los verbos de actitud, o, para el caso, de cualesquiera otras palabras «no lógicas». Y es en ello en lo que se apoya nuestra defensa.

### LA DEFENSA DE LA INOCENCIA

Si los enunciados simples describen situaciones, entonces los informes de actitudes deben describir situaciones que incluyan actitudes, es decir, situaciones de percepción *p* en los casos de *ve* y *ve que*, situaciones epistémicas *k* en el caso de *sabe que*, situaciones doxásticas *c* con *cree que*, y preferencias *u* en el caso de *dice que*. Pero ¿qué es exactamente lo que decimos sobre una situación cuando decimos que *a ve que φ* o que *b dice que ψ*? ¿Qué es lo que falta, algo relativo al agente, en nuestra anterior explicación? ¿Y qué es lo que tienen estas situaciones de actitud que hace que sean clasificables con oraciones subordinadas y, por lo tanto, de acuerdo con una semántica inocente, con proposiciones realistas, proposiciones que por regla general no son verdaderas acerca de la situación de actitud? Para responder estas preguntas debemos hacer una ligera digresión.

*Restricciones estructurales.* Las cosas no pueden ocurrir de cualquier manera. Existen todo tipo de restricciones sobre las situaciones-tipo que puede producirse realmente y sobre el curso que los eventos pueden realmente tomar. Algunas restricciones se originan en propiedades de objetos y relaciones entre objetos bastante obvias. (Besar implica tocar, ser abuelo implica ser padre.) Otras se originan en leyes naturales. Y otras son más bien temporales y algo accidentales (escribir a máquina solía implicar el hacer que las teclas se movieran). Un hablante nativo de una lengua entiende normalmente muchas de estas restricciones y utiliza este conocimiento en su discurso. La adecuación de intercambios verbales como los siguientes puede explicarse sólo en relación con tales restricciones.

«¿Me besaste?» «No te toqué.»  
 «¿Hace calor afuera?» «Bueno, está nevando.»  
 «¿Por qué no escribes a máquina?» «Las teclas se enganchan.»

<sup>9</sup> Pero véase Perry, *op. cit.*, donde se muestra que esta teoría no funciona.

correferenciales no preserva el valor veritativo en los informes de actitud—. Después de todo, si con nuestra explicación se predice (iv), entonces habrá de hacerse una afirmación similar para todas las actitudes. Al considerar que las actitudes son relaciones con objetos, propiedades y relaciones reales, nos comprometemos a sostener que son, en algún sentido, transparentes.

*La falta de objetos para las actitudes.* En nuestra explicación se toman proposiciones como objetos de las actitudes. Esto no es filosóficamente satisfactorio cuando se trata del ver y el decir. Con el ver se pierde la conexión con lo que realmente se vio, y con decir, se pierde la conexión con lo que realmente se profirió. Supóngase que un hablante *a*, en una situación de discurso *d* y con conexiones *c* dice

(11) *b* dijo que  $\phi$ .

Hemos visto que  $\phi$  mismo no puede, en general, servir para lo que *b* profirió realmente. Pero de la verdad de (11) se sigue con toda seguridad que *b* profirió algo, que hay alguna oración  $\psi$  proferida por *b* tal que, desde la ubicación *d'* del discurso de *b* y con las conexiones *c'* de *b*,  $d, c[\phi] = d', c'[\psi]$  y  $c(\text{dijo}) = l_{d'}$ . Similarmente, si *a* dijo con verdad

(12) *b* vio  $\phi$ .

entonces lo que *b* vio en realidad es una escena en la que se daba  $d, c[\phi]$ , y no una proposición.

*El contenido cognitivo de las actitudes.* En nuestra explicación falta el hecho de que las actitudes tienen algo que ver con las mentes (o los cerebros) y los fenómenos cognitivos. Del mismo modo que el decir requiere un agente que profiera algo con un significado, y que el ver requiere que el agente vea algo con sus ojos (que son parte del cerebro), también el creer y el conocer requieren que el agente esté en un estado cognitivo con significado. Parte de lo que nos dan los informes de actitud constituye información relativa al estado cognitivo del agente. Esto es lo que hace que los informes de actitud sean útiles en la explicación y en la predicción de lo que la gente hará. Las personas que tienen percepciones, creencias y deseos similares se comportan de un modo similar.

El concentrarse en este aspecto del problema hace que los sentidos fregeanos parezcan atractivos. Interpretando una actitud como una relación con un sentido, o con una «representación mental», como se tendría en algunas versiones de la teoría, se ve al objeto de la actitud como si clasificara los estados cognitivos. Según esta teoría la similitud de las actitudes apunta a la similitud de los estados, expli-

cando aparentemente la similitud de las acciones <sup>9</sup>. Sin embargo, según nuestra teoría, diferentes personas podrían creer exactamente la misma cosa de innumerables modos distintos. Hasta aquí la teoría no refleja ninguna similitud cognitiva.

Estos cuatro problemas estan relacionados entre sí, y todos ellos requieren que meditemos en qué es lo que estamos haciendo cuando intentamos formular una teoría semántica de los verbos de actitud, o, para el caso, de cualesquiera otras palabras «no lógicas». Y es en ello en lo que se apoya nuestra defensa.

### LA DEFENSA DE LA INOCENCIA

Si los enunciados simples describen situaciones, entonces los informes de actitudes deben describir situaciones que incluyan actitudes, es decir, situaciones de percepción *p* en los casos de *ve* y *ve que*, situaciones epistémicas *k* en el caso de *sabe que*, situaciones doxásticas *c* con *creo que*, y preferencias *u* en el caso de *dice que*. Pero ¿qué es exactamente lo que decimos sobre una situación cuando decimos que *a ve que φ* o que *b dice que ψ*? ¿Qué es lo que falta, algo relativo al agente, en nuestra anterior explicación? ¿Y qué es lo que tienen estas situaciones de actitud que hace que sean clasificables con oraciones subordinadas y, por lo tanto, de acuerdo con una semántica inocente, con proposiciones realistas, proposiciones que por regla general no son verdaderas acerca de la situación de actitud? Para responder estas preguntas debemos hacer una ligera digresión.

*Restricciones estructurales.* Las cosas no pueden ocurrir de cualquier manera. Existen todo tipo de restricciones sobre las situaciones-tipo que puede producirse realmente y sobre el curso que los eventos pueden realmente tomar. Algunas restricciones se originan en propiedades de objetos y relaciones entre objetos bastante obvias. (Besar implica tocar, ser abuelo implica ser padre.) Otras se originan en leyes naturales. Y otras son más bien temporales y algo accidentales (escribir a máquina solía implicar el hacer que las teclas se movieran). Un hablante nativo de una lengua entiende normalmente muchas de estas restricciones y utiliza este conocimiento en su discurso. La adecuación de intercambios verbales como los siguientes puede explicarse sólo en relación con tales restricciones.

«¿Me besaste?» «No te toqué.»

«¿Hace calor afuera?» «Bueno, está nevando.»

«¿Por qué no escribes a máquina?» «Las teclas se enganchan.»

<sup>9</sup> Pero véase Perry, *op. cit.*, donde se muestra que esta teoría no funciona.

Las teorías semánticas tradicionales, reconociendo la importancia de esas restricciones, intentan imponérselas a las expresiones lingüísticas, vía «postulados de significado». Esto nos parece hacer las cosas al revés. Nosotros creemos que estas restricciones sobre los cursos de eventos son (excepto en los casos más especiales) independientes de cuáles sean las expresiones del lenguaje natural (si es que las hay) que designen los objetos, relaciones y ubicaciones que los constituyen.

Cuando los constituyentes pertinentes están individualizados claramente, es posible representar las restricciones con bastante claridad:

Si  $\sigma_l(\text{besar}, a, b) = 1$  entonces  $\sigma_l(\text{tocar}, a, b) = 1$ .

Si  $\sigma_l(\text{soltero}, a) = 1$  entonces  $\sigma_l(\text{casado}, a) = 0$ .

Si  $\sigma_l(\text{dar una patada}, a, b) = 1$  y  $l \circ l'$  entonces  $\sigma_r(\text{dar una patada}, a, b) \neq 0$ .

Si  $\sigma_l(\text{nevar}) = 1$  entonces  $\sigma_l(\text{hacer calor}) \neq 1$ .

En otros casos puede ser bastante difícil. Por ejemplo, sería imposible enunciar todas las restricciones sobre  $\sigma^*$  que impone  $\sigma_l^*$  (caminar,  $a$ ) = 1.

Los sistemas de restricciones pueden utilizarse para varios fines. Un curso de eventos  $\sigma$  es *estructuralmente completo* relativamente a un conjunto  $C$  de tales restricciones, si  $\sigma$  satisface cada una de las restricciones de  $C$ . Un curso de eventos  $\sigma$  es *estructuralmente coherente* con respecto a  $C$  si  $\sigma$  es parte de algún  $\sigma'$  completo<sup>10</sup>. Si no es coherente, es *incoherente*.

Una restricción es correcta si  $\sigma^*$ , el curso real de eventos, satisface la restricción. Un conjunto  $C$  de restricciones es correcto si cada restricción de  $C$  es correcta, es decir, si  $C$  capta correctamente restricciones relativas al modo en que las cosas pueden ocurrir realmente, de modo que  $\sigma^*$  es completo con respecto a  $C$ . Si  $C$  es correcto, entonces cualquier parte de  $\sigma^*$  es coherente con respecto a  $C$ . Ninguna parte del curso de eventos real puede ser estructuralmente incoherente, aunque podría ser estructuralmente incompleta.

Si  $a_d$  es un organismo situado en el mundo, su dotación biológica y lo que haya aprendido de su experiencia anterior le llevarán a actuar de acuerdo con ciertas restricciones correctas —a *sintonizar* con estas restricciones—. Como hemos visto anteriormente, las personas sintonizan con todo tipo de restricciones que, en realidad, no pueden enunciar —por ejemplo, todas las cosas que el caminar implica—. Esto no es sorprendente, puesto que los peces están en sintonía con ciertas leyes naturales del agua y del nadar, y no pueden decir ni una sola palabra.

<sup>10</sup>  $\sigma_1$  es parte de  $\sigma_2$  si  $\text{dom}(\sigma_1) \subseteq \text{dom}(\sigma_2)$  y, para cada  $l \in \text{dom}(\sigma_1)$ ,  $\sigma_1(l) \subseteq \sigma_2(l)$ .



Pero cuando estamos haciendo la semántica de palabras como *besar* o *caminar*, nos vemos forzados a reflexionar sobre las restricciones relativas al besar y al caminar con que están en sintonía los hablantes nativos del inglés y que se reflejan y se explotan en su uso del inglés.

Lo mismo se aplica a los verbos de actitud que hemos estado considerando aquí. Las restricciones estructurales entran en los informes de actitud de dos modos distintos. En primer lugar, existen todo tipo de restricciones estructurales en las situaciones de actitud, exactamente igual que con besar y caminar. En segundo lugar, sin embargo, existen también restricciones estructurales correctas con las que el agente de una situación de actitud sintoniza y que afectan a sus actitudes. Estamos interesados primordialmente en formular restricciones del primer tipo, dejando las segundas para otras partes de la ciencia.

Cuando nos concentramos en las restricciones del primer tipo que se reflejan claramente en el lenguaje, encontramos una diferencia notable entre los factivos (*ve*, *ve que*, *sabe que*) y los no factivos (*creo que*, *dice que*). La diferencia se muestra con la máxima claridad en los extremos, a saber, el *ve* epistémicamente neutro (*Guillermo vio ganar a Juana*) y el *dice que* (*Guillermo dijo que Juana ganó*); por consiguiente, comenzaremos por analizar estos dos y compararlos a continuación.

*Ver*. Existen múltiples uniformidades a través de situaciones de percepción visual. Las de un determinado tipo se integran directamente en la estructura de los informes de percepción, pero para explicar los usos que podemos hacer de los informes de percepción se necesitan otras uniformidades. Con enunciados de *ver* epistémicamente neutros, tratábamos el ver como una relación entre un agente *a* y una proposición no ubicada *P*. Así es como funciona el lenguaje: *a ve  $\phi$*  se centra en *a*, en el ver, y en *lo que es verdad* acerca de aquello que *a ve*,  $P =_{d,c}[\phi]$ . Pero las propiedades semánticas de oraciones como las que dimos en el primer problema, reflejan una información distinta —a saber, la escena que *a* aprehende visualmente—. El ver lleva consigo una escena aprehendida visualmente. Una escena es una situación real  $\langle l, s \rangle$ , pero su tipo no incluye todo lo que ocurre en *l*, sino únicamente la parte visible bajo las condiciones pertinentes. Estas condiciones incluyen la dirección y la distancia del agente con respecto a *l*, las condiciones de iluminación y muchas otras cosas. Podemos, en términos de escenas, formular la siguiente restricción:

$$\sigma_1(\text{ve}, a, P) = 1$$

sys hay una escena  $s = \langle l, s \rangle$  tal que  $\sigma_l(\text{ve}, a, s) = 1$  y  $s \varepsilon P$ .

Todos los principios semánticos relativos a *ve* que se dieron en la sección anterior se desprenden de esta restricción estructural. La restricción extrae otra uniformidad presente en las situaciones visuales, a saber, la escena visualmente aprehendida. Al hacerlo nos proporciona un modo indirecto alternativo de clasificar individuos, según lo que éstos vean. Ésta es la razón de que podamos decir

*María vio un camión pararse frente a ella. Guillermo lo vio también.*

Parece que los informes de situaciones visuales epistémicamente neutros informan primordialmente sobre lo que es verdad acerca de una escena aprehendida visualmente. No es difícil imaginar por qué el lenguaje nos ha de proporcionar un mecanismo para tales informes. Basta pensar en los exploradores, cuya misión es escrudinar el horizonte buscando signos de pioneros hostiles. Aquí utilizamos el informe de percepción como dato sobre cómo es el mundo que ha sido inspeccionado, dado que lo que nos interesa es el mundo inspeccionado, no el agente.

Pero también usamos informes de percepción para *explicar* las actividades de los agentes, como, por ejemplo, cuando decimos que María apretó los frenos porque vio a un camión pararse frente a ella.

Para explicar y predecir actividades de los agentes, necesitamos encontrar principios de clasificación que sean proyectables sobre actividades, es decir, similitudes entre los agentes que lleven a acciones similares. Dada la complejidad de las causas de la acción, estas conexiones no serán simples. Pero la idea es que cualquier teoría adecuada o casi adecuada —como seguramente lo es la teoría que apoya nuestra explicación de por qué María apretó los frenos— debe funcionar con toda una serie de estados del agente que se relacionen sistemáticamente con otros estados y, en último término, con actividades.

Podemos ver ahora el impacto de la relatividad y la eficiencia. *Las uniformidades seleccionadas no proporcionan por sí solas todos los estados o principios de clasificación que utilizamos.* Consideremos nuestra explicación de la acción de María apretando los frenos. Considérese la clase de situaciones de percepción determinadas por *vio un camión parándose frente a ella*. Claramente, existen muchas diferencias relevantes. Si uno ve un camión parándose a un kilómetro no apretará los frenos, y tampoco lo hará alguien que vea a un camión deteniéndose en el arcén.

En los dos ejemplos últimos hemos estrechado la clasificación de dos maneras distintas. En el segundo caso ampliábamos la oración subordinada; en el primero considerábamos la distancia entre el agente y ciertas partes de la escena. El panorama general que surge

de todo esto es el siguiente. La uniformidad seleccionada —la verdad de una proposición dada en la escena aprehendida visualmente— no es un principio de clasificación que apoye por sí solo la explicación. Pero es parte de un sistema tal. La uniformidad seleccionada, junto con otros factores, nos proporciona un sistema de estados (abstractos) que es útil en la predicción y explicación de las actividades del agente. Incluso cuando explicamos informando de una actitud, nos apoyamos en una comprensión de los demás factores. Así, en la explicación anterior de la acción de María apretando los frenos, el oyente limita los demás factores de modo que la explicación funcione —es decir, supone que el camión estaba delante de María y no muy lejos de ella—.

*Dice.* En el ver, la situación real aprehendida visualmente juega un papel crucial en el esquema clasificatorio. Con un no factivo, como *dice*, no es necesario que haya una situación real para apoyar la clasificación. Cuando decimos, *Guillermo dijo que Juana ganó, Juana ganó* no sirve para clasificar una situación real con la que Guillermo esté relacionado, por ejemplo una «aprehensión asertiva». ¿Cómo funciona entonces la clasificación?

La respuesta es fácil de ver o de oír. Utilizamos *dice* para clasificar preferencias. Las uniformidades que se dan a través de las preferencias son precisamente las uniformidades para las que hemos elaborado anteriormente una explicación. Las preferencias constan de situaciones de discurso, conexiones y expresiones. La uniformidad seleccionada es la interpretación de la preferencia.

Es cierto que utilizamos la palabra *dice* de modos distintos, uno que se concentra en lo que se dice, en el sentido de la interpretación, y el otro en términos de las palabras proferidas. Para este último utilizamos *dice* con la oración subordinada entrecomillada. Estos dos usos de *dice* se centran en dos modos distintos de clasificar las situaciones de preferencia. Uno se centra en la uniformidad de interpretación, el otro en la uniformidad de la oración con significado.

María dijo que yo corría peligro.  
 María dijo «Tú corres peligro».  
 María dijo «Él corre peligro».  
 María dijo «¡Cuidado!».

Nótese que ninguna de estas uniformidades puede ser determinada de modo único por la otra. La proposición que enuncian es absoluta:

$\{a/ \sigma_1 (\text{correr peligro}, a_d) = 1\}$ .

Las expresiones *Tú estás en peligro* y *Él está en peligro* no lo son.

La eficiencia y la relatividad del lenguaje hacen que sea imposible pasar de una de ellas a la otra de un modo unívoco.

Nótese que las dos maneras diferentes de clasificar agentes que proporcionan los dos sentidos distintos de *decir*, suministran clases de agentes muy diferentes, clases que son pertinentes para tipos distintos de generalizaciones. Supóngase que Hulk dice *Yo soy un asesino*. Pertenecen entonces a dos clases distintas, aunque éstas se solapan, la de los que dicen *Yo soy un asesino* y las de los que dicen que Hulk es un asesino.

Aunque estas uniformidades no se determinen de un modo unívoco la una a la otra, con información adicional un tipo de clasificación puede llevarnos, más o menos sin problemas, al otro. Decir *Hulk dijo que él era un asesino* no le clasifica automáticamente entre los que han proferido *Yo soy un asesino*, pero lo sugiere muy fuertemente, puesto que es la manera normal de que Hulk diga que lo es.

Las explicaciones de acciones que hacen referencia a «lo que se dijo» llevan consigo este tipo de implicaciones. Considérese, por ejemplo,

Guillermo se apartó de un salto porque le oyó a María decir que él corría peligro.

Como explicación, esto sólo tiene sentido si María dijo que Guillermo corría peligro de manera que le comunicara una sensación de peligro —es decir, si ella utilizó alguna expresión que sea uniforme a través de situaciones en las que la persona a la que uno se dirige está en peligro—. *¡Cuidado!* y *Tú corres peligro* son expresiones de ese tipo. *Él corre peligro* no serviría, en general. Lo que explica la conducta de Guillermo es la existencia de una manera de decir que él está en peligro que se relaciona sistemáticamente con situaciones en las que la persona a la que uno se dirige está en peligro. No importa solamente la relatividad del lenguaje, es decir, el hecho de que hay muchas maneras de decir lo mismo, sino la eficiencia del lenguaje. Una expresión como *¡Cuidado!* puede utilizarse en muchas situaciones para prevenir de un peligro.

Así pues, para completar nuestra teoría del decir, necesitamos únicamente explotar la semántica de situaciones (y el otro uso de *dice*) al formular una restricción estructural:

Si  $\sigma_I$  (dice que,  $b$ ,  $P$ ) = 1  
entonces hay una preferencia  $u = (\psi, d', c')$  tal que  $\sigma_I$  (dice,  $d'$ ,  $c'$ ,  $\psi$ ) = 1, donde  
 $b = a_d \cdot l = l_{d'}$ , y  $a_c \mid \psi = P$ .

Comparemos ahora ver y decir. La motivación que hay detrás de la uniformidad seleccionada es bastante distinta. En el caso de ver, la

proposición realista clasifica directamente la situación real, y de ese modo clasifica indirectamente al agente que aprehende visualmente la situación. (Nótese que tenemos una teoría de la percepción *directa*, y una clasificación indirecta de los perceptores.) Pero en el *decir*, la proposición no puede funcionar de ese modo, pues puede que no haya una situación real a la que se ajuste la proposición.

Para que funcione este segundo esquema, debe haber algo que «llene el hueco» que deja la ausencia de una situación real clasificable. Este algo es la relación proferida, o, más plausiblemente, la preferencia de una oración con significado con un determinado conjunto de intenciones. En lugar de una relación con una situación real que la proposición caracteriza, tenemos una relación con una entidad que tiene significado, la cual tiene como interpretación, en la preferencia, la proposición.

Sin embargo, hay una diferencia importante entre la relación con la situación real y el papel de la oración proferida.

En el ver, la proposición se conecta con el agente «a través de» la escena aprehendida. En el decir, la proposición se conecta con el agente mediante la oración producida. Pero la proposición es verdadera en la situación de la cual es una parte la escena, con total independencia de la ubicación del agente, de las conexiones, la historia, etc. Podríamos decir que la situación nos proporciona toda una serie de proposiciones; los otros factores únicamente influyen el modo en que el agente puede aprehender la escena acerca de la cual es verdadera la proposición.

Pero en el decir la proposición no se conecta en general (si es que lo hace alguna vez) con la oración con significado aislada, sino sólo con toda la preferencia. La proposición no sirve para clasificar *uno* de los factores, y con ello el todo, sino que caracteriza el todo mediante la relación entre los factores. (Un caso especial sería la preferencia de una oración cuyo significado determina unívocamente su interpretación, si es que hay alguna.)

Cuando pasamos a la cuestión del interés por la uniformidad a través de las preferencias que hayan de seleccionarse encontramos un parecido con el ver. *Dice que* está diseñado para mostrarnos cómo es el mundo si lo que dice el agente es verdadero. Pero, como sucede con el ver, otras son las uniformidades cruciales cuando utilizamos *dice* para explicar y predecir actividades del agente, o de los que oyen o leen la preferencia.

*Ve que, sabe que.* Imaginemos que un par de gemelas idénticas, Julia y Juana, participaron en una maratón y que Guillermo vio que una de las dos ganaba. De hecho fue Julia, pero Guillermo no puede distinguir la una de la otra. Si se le preguntara cuál de ellas, no podría decirlo. Casos como éste hacen patente la diferencia entre el *ve* no

epistémico y el *ve que* y el *sabe que*. Si Guillermo vio ganar a una de las gemelas, y ganó Julia, entonces Guillermo vio ganar a Julia. Pero si bien vio que una de ellas ganaba y sabe que una de ellas ha ganado, no vio que Julia ganara, ni sabe que Julia ha ganado.

El *ver que* involucra escenas (o, más en general, cursos de eventos), pero la relación es menos directa que en el caso de *ve*. Las restricciones estructurales son (aproximadamente):

$$\sigma_I(\text{ve que}, a, P) = 1$$

si hay un evento  $\sigma_0$  (posiblemente una escena) tal que:

$$(i) \sigma(\text{ve}, a, \sigma_0) = 1$$

(ii) hay un sistema  $C$  de restricciones estructurales correctas con el que sintoniza visualmente  $a$  tal que todo  $\sigma$  que contenga  $\sigma_0$  y que sea completo con respecto a  $C$  está en  $P$ .

En el ejemplo anterior existe una cierta propiedad visual  $p$  con la que Guillermo sintoniza:

$$\text{Si } \sigma_I(p, x) = 1, \text{ entonces } \sigma_I(p_{\text{Julia}}, x) = 1 \text{ ó } \sigma_I(p_{\text{Juana}}, x) = 1.$$

donde  $p_{\text{Julia}}$  es la propiedad de llamarse Julia. Cualquier escena  $\sigma$  estructuralmente completa que abarque el evento que vio Guillermo tendrá  $\sigma_I(p, x) = 1$  y  $\sigma_I(p_{\text{Julia}}, x) = 1$ , pero esta  $\sigma_0$  no era completa.

La diferencia entre los informes que utilizan el verbo *ver* y los que utilizan *ver que* estriba, en parte, en estas restricciones, pero también en diferentes estrategias de interpretación para los sintagmas nominales y los sintagmas verbales de las oraciones subordinadas. Hablaremos de esto brevemente en la sección siguiente, bajo el rötulo «evaluar».

Una restricción estructural similar puede darse para *sabe*. La idea básica es que saber es estar en sintonía. Simplificando para los fines de esta exposición, vamos a hacer como si la visión fuera la única forma de percepción. Podríamos decir entonces

$$\sigma_I(\text{sabe que}, a, P) = 1$$

si hay un curso de eventos  $\sigma_0$  tal que

$$(i) \sigma_I(\text{ve}, a, \sigma_0) = 1$$

(ii) si  $C'$  es el conjunto de todas las restricciones estructurales correctas con las que sintoniza  $a$ , entonces todo  $\sigma$  que contenga a  $\sigma_0$  y que sea completo con respecto a  $C'$  está en  $P$ .

Notará el lector que la única diferencia que hay aquí entre *ve que* y *sabe que* estriba en el hecho de que se admite un conjunto más amplio  $C'$  de restricciones correctas para *sabe que*. Esto tiene presumi-

blemente algo que ver con la tendencia a decir que vemos que algo es el caso cuando queremos decir que lo sabemos.

Como ocurre con el ver, el análisis de *ve que* y *sabe que* explica los muchos tipos de uniformidades, distintas de las «seleccionadas», a través de situaciones epistémicas. En particular, la estructura del conocimiento es tal que podemos saber o conocer la misma cosa sobre la base de conjuntos distintos de experiencias anteriores, diferencia que podría ser relevante para la cuestión de cómo le afecta a uno el conocimiento.

*Creer.* Usualmente se piensa que saber que  $P$  implica creer que  $P$ . Y en el lenguaje suponemos usualmente que, si el hablante dice que  $P$ , entonces cree que  $P$ . En este respecto, creer parece algo más o menos intermedio entre saber y decir. Pero, desde una perspectiva realista, la creencia es con mucho la más enigmática de nuestras cuatro actitudes. En efecto, ¿dónde está o qué es lo que en la realidad no varía en distintas situaciones doxásticas, algo que apoye el que las clasifiquemos juntas como situaciones en las que un agente cree que  $P$  (donde  $P$  es una proposición realista)?

Con el ver, el ver que y el saber que existe un curso de eventos real con respecto al cual  $P$  es verdadera. Con la creencia no tiene por qué haberlo; de modo que creer parece más similar a decir que. Pero, en el caso de decir que, existe la expresión proferida realmente, algo real que, junto con la situación del discurso y las conexiones del hablante, daba lugar a la proposición. ¿Qué es lo análogo en el caso de la creencia?

Parece que el realista, si cree en la creencia, se ve abocado o bien a una metafísica que incluya situaciones que tengan algún tipo de realidad pero que, en el sentido estricto en que hemos utilizado el término, no sean reales (un paso esotérico que seguramente nadie defendería) o a una metafísica que contemple «estados de creencia» reales, es decir, una especie de invariantes abstractos pero reales a través de situaciones doxásticas reales, invariantes que apoyen su clasificación mediante proposiciones realistas, al modo en que las oraciones apoyan proposiciones en el caso del decir.

Aquí es donde uno podría pensar que entra en escena algo afín a los sentidos fregeanos, «pensamientos completos y eternos» captados por las mentes. Podríamos modificar la explicación fregeana de manera que la referencia de un pensamiento  $T$  fuera una proposición realista  $P = \text{ref}(T)$  y utilizar la restricción estructural:

$$\sigma_1 (\text{cree que, } a, P) = 1$$

syss hay un pensamiento  $T$  tal que  $\sigma_1$  (capta doxásticamente,  $a, T$ ) y

$$\text{ref}(T) = P.$$

Esto nos permitiría reproducir la relatividad de la creencia, el hecho

de que diferentes personas puedan creer la misma cosa de modos distintos, al tener pensamientos distintos. ¡Pero esto constituiría un grave error! Como en el caso de las oraciones, tan importante es que los estados de creencia sean relativos como que sean eficientes. Existen pues otros factores que desempeñan un papel en el paso del estado  $S_i$  a la proposición:  $P = F(S_i, \dots ?)$ . ¿Cuáles son estos otros factores?

El agente, naturalmente, es uno de tales factores. Cuando está en lo que podríamos llamar el estado de creencia «Yo corro peligro», sus creencias versan sobre él mismo y su ubicación en el momento, tanto en el tiempo como en el espacio. Y no hay razón alguna para suponer que otras propiedades del agente, digamos su altura o su cultura, no puedan desempeñar un papel en la interpretación de su estado de creencia. Metamos todo esto dentro de una situación del agente  $d = \langle s_d, a_d \rangle$ , donde  $s_d = \langle l_d, s_d \rangle$ , siendo  $a_d$  el agente,  $l_d$  su ubicación, y  $s_d$  los hechos relativos a  $a_d$  que se necesitan para la interpretación. Esta  $d$  es análoga a la situación del discurso en el caso del decir.

Pero, naturalmente, podemos tener creencias relativas a cosas distintas de nosotros mismos y de nuestra presente ubicación. Tenemos conexiones con los objetos, relaciones y ubicaciones que se originan en la percepción, y estas conexiones ayudan a determinar cuáles son las cosas sobre las que versan nuestras creencias.

Así pues, hablar de creencias presupone un sistema abstracto de clasificación  $S_1, S_2, \dots$  de *estados* y una relación *cr* entre estados, situaciones del agente, conexiones y cursos de eventos:

$$cr(S_i, d, c, \sigma)$$

Imponemos entonces la restricción de que  $\sigma_i(\text{cree que}, a, P) = 1$  si y sólo si hay  $d, c$ , y  $S_i$  tales que  $l = l_d, a = a_d, \sigma_i(S_i, d, c) = 1$ , y  $P = \{\sigma / cr(S_i, d, c, \sigma)\}$

La relación *cr* es análoga a  $| \cdot |$ . Del mismo modo en que  $| \cdot |$  identifica una proposición relativamente a una expresión, situación del discurso y conexiones, *cr* identifica una proposición relativamente a un estado de creencia, situación doxástica y conexiones: la proposición que cree una persona que esté en ese estado, en tal situación doxástica y con tales conexiones. El postular tal relación y tal sistema de estados «con significado» se presupone en el modo en que usamos *cree*. Nos parece que este enfoque de la creencia encaja bien con varios enfoques existentes en filosofía de la mente, enfoques que hacen hincapié en el modo en que las atribuciones de estados mentales se conectan con las actividades de la especie y del individuo.

En el caso de agentes que hablen un lenguaje  $L$ , es muy tentador suponer que las oraciones con significado de  $L$  pueden estar incrusta-



das en la estructura de los estados de creencia, es decir, que hay una función  $S(\phi)$  de las oraciones de  $L$  en el sistema de los estados de creencia. Esto sugiere que los estados de creencia tienen una cierta «sintaxis» análoga a la sintaxis de  $L$  y que esta «sintaxis» es importante en el análisis de  $c_r$ , igual que la sintaxis del lenguaje es importante en el análisis de  $d_c[\phi] = P$ . Esta tentación debería distinguirse de algo

que no encontramos muy tentador: la idea de que creer consiste en tener una relación con algunas oraciones de un cierto lenguaje.

### LA OPACIDAD

La interacción de los verbos de actitud con los términos singulares (nombres propios y descripciones definidas, por ejemplo) fue una de las fuerzas rectoras tras las teorías de Russell y Frege —aunque estas teorías fueran distintas—. Es ciertamente posible que Jorge IV se preguntara si Scott era el autor de *Waverly* sin preguntarse si Scott era Scott, y es posible creer que el lucero del alba es un planeta sin creer que el lucero vespertino sea un planeta. Estos hechos llevaron a Frege a decir que, dentro del alcance de un verbo de actitud, una expresión se refiere a su «sentido» usual, no a lo que comúnmente es su referencia. Russell introdujo la «forma lógica» y argumentó que las descripciones definidas no denotan, sino que lo que hacen es aportar las propiedades definitorias a la proposición cuando a ésta se la entiende adecuadamente. (Muchas teorías contemporáneas apelan tanto al sentido como a la forma lógica.)

Nosotros no recurrimos ni al sentido ni a la forma lógica, sino que manejamos estos problemas básicamente con los recursos de que ya disponemos. No tenemos el espacio suficiente para explicar nuestro análisis de los nombres, aunque el lector podrá adivinar probablemente el modo en que nociones como interpretación inversa y conexión nos permiten sustituir antiguas preguntas, como por ejemplo «¿Tienen los nombres propios un sentido?» y «¿Cuáles son las condiciones de verdad de una oración con nombres propios?», por preguntas más tratables. Explicaremos las ideas básicas que hay tras nuestro análisis de las descripciones, análisis que tiene muchas características russellianas, pero que no nos amenaza con el atomismo ni requiere apelar a la forma lógica.

Para simplificar la discusión, ignoraremos cuestiones de tiempo y lugar, de manera que podamos habérmolas con situaciones-tipo en lugar de con cursos de eventos. Nos restringiremos también a descripciones  $\alpha$  que no sean sensibles a la situación del discurso  $d$  o a la conexión  $c$ , de modo que podremos escribir  $|a|$  para la interpretación  $d_c[\alpha]$ , siempre con el fin de simplificar la exposición.

La interpretación de una descripción definida es una relación entre situaciones-tipo  $s$  e individuos  $a$ :

$$[El \beta](s, a) \text{ syss } \{a\} = \{x / \beta\}(s, x)$$

Esta relación puede verse también como una función *parcial* de tipos de situaciones  $s$  a individuos  $a$ . Utilizando notación estándar para función y argumento, podemos escribir  $a = [el \beta](s)$ . Esta función impone una restricción mutua entre  $s$  y  $a$ . Dada una  $s$  en su dominio, podemos utilizar el  $\beta$  para referirnos a  $a = [el \beta](s)$ . O bien, dado un  $a$ , podemos utilizar el  $\beta$  para afirmar que la situación  $s$  es una situación en la que  $[el \beta](s) = a$ . O también puede ser utilizada simplemente para decir que, cualesquiera que sean  $s$  y  $a$ ,  $[el \beta](s) = a$ .

#### EJEMPLOS

(i) Entro en el estudio de Alfredo, en el que está él sentado con su perro *Clariso*. Dice él, «Cuidado. El perro tiene pulgas». En la situación  $s_0$  en que estamos está muy claro que se está refiriendo a *Clariso* (=  $[el \text{ perro}](s_0)$ ).

Alfredo ha aseverado la proposición realista:

$$\{s_1 / s_1 (\text{tiene pulgas, Clariso}) = 1\}$$

Nótese que, si le creo, entonces lo que yo creo no es que haya un único perro que tenga pulgas, sino que este perro en particular tiene pulgas.

(ii) Ahora hay varios perros en la habitación. Alfredo dice, señalando a *Clariso*, *Éste es el perro que muerde*. Aquí la descripción definida *el perro que muerde* no se utiliza para identificar a *Clariso*, sino que más bien se le atribuye la propiedad de ser el único perro que muerde. La proposición es

$$\{s_1 / \text{Clariso} = [el \text{ perro que muerde}](s_1)\}$$

(iii) Estamos ahora en la situación en que Inés una vez me dijo, refiriéndose a una cierta persona  $a$ , *Ella es imprudente*. Inés es perspicaz para juzgar caracteres. De acuerdo con ello, le desaconsejo a usted invertir en el banco de  $a$  con la advertencia, *Inés cree que la presidenta del Banco Interfederal es imprudente*. Aquí, la interpretación de mi preferencia es, esencialmente,

$$\{s_1 / s_1 (\text{cree que, Inés, } P_{s_1}) = 1\}$$

donde

$$Ps_1 = \{s_2/s_2 \text{ (imprudente, [la presidenta del Banco Interfederal](s_1))} = 1\}$$

La descripción definida restringe  $s_1$  de manera que contenga una única presidenta del Banco Interfederal,  $a_{s_1}$ , y asevera que Inés cree que  $a_{s_1}$  es una imprudente ( $P_{s_1}$ ).

Utilizamos la notación: *a dice (cree/sabe/ve)que(-(el)<sub>j</sub>...)* con  $j = 0, 1, \text{ ó } 2$  para indicar, respectivamente, las lecturas en que [el  $\beta$ ] se evalúa en una situación-tipo accesible  $s_0$ , como en (i); para restringir la situación-tipo  $s_1$  designada por el todo (como en iii); o para restringir las situaciones-tipo  $s_2$  descritas por la oración subordinada. Si  $j = 0$ , tenemos la correspondiente al uso referencial de Donnellan. Si  $j = 2$ , tenemos su uso atributivo. El caso  $j = 1$  está en medio.

Podría parecer que estas distintas lecturas coinciden con diferentes lecturas del alcance. Pero en realidad reflejan un fenómeno distinto, que muchos confunden con el del alcance. La manera más fácil de ver la distinción es con descripciones indefinidas, como *miembro de la familia*. La interpretación de una descripción indefinida *un  $\beta$*  (v. gr., *un perro, un elefante*) es también una relación entre situaciones e individuos:

$$\{un \beta (s, b) \text{ sys } \beta (s, b)\}.$$

Considérese el caso de que Jack haya sido asesinado. Holmes reúne a los miembros de la familia y les dice «Uno de ustedes ha asesinado a Jack». «¿Qué es lo que ha dicho?», pregunta la vieja tía Agnes, que padece sordera. «Ha dicho que un miembro de la familia asesinó a Jack», grita Jill, la viuda de Jack.

Desde luego, no hay nada erróneo en el informe de Jill, pero no se puede dar cuenta de él únicamente con la distinción entre el alcance largo y el corto. Holmes no afirmó acerca de ningún miembro de la familia en particular que él fuera el asesino, de modo que no se trata del alcance largo. Pero tampoco dijo nada en absoluto sobre la pertenencia a la familia, de modo que no se trata del alcance corto. Se trata de lo que nosotros escribiríamos como *El dijo que (un miembro de la familia)<sub>1</sub> asesinó a Jack*. La interpretación es:

$$\{s_1/s_1 \text{ (dice que, Holmes, } Ps_1) = 1\}$$

donde

$$Ps_1 = \{s_1/\text{para algún } a \text{ tal que [un miembro de la familia]}(s_1, a), \\ s_2 \text{ (asesinó, } a, \text{ Jack)} = 1\}.$$

## CUESTIONES DE FUNDAMENTACIÓN

Cualquiera que trate de desarrollar una teoría semántica de las actitudes con el suficiente cuidado habrá de enfrentarse con ciertas cuestiones de fundamentación. No podemos hablar aquí del modo exacto en que éstas se presentan en la semántica de situaciones, ni de los detalles de nuestra solución. La idea básica es la de limitarnos a los objetos conjuntistas hereditariamente finitos construidos a partir de los objetos, relaciones y ubicaciones de que disponemos. En último término, esto nos exige ser más realistas en cuanto a las oraciones, estados y otros factores involucrados en las actitudes. Por ejemplo, en último término definimos

$$\sigma_I(\text{dice que, } a, P) = 1$$

mediante

$$\exists d, e, \psi [\sigma_I(\text{dice, } d, c, \psi) = 1 \ \& \ \sigma_{d,c}[\psi] = P] \text{ (con } \sigma_I(\text{dice que, } a, P) = 0, \text{ en caso contrario).}$$

Esto nos permite evitar tener proposiciones como argumentos de situaciones-tipo. Nuestra teoría nos ha hecho subir por una espiral para llegar hasta aquí. Partimos de una actitud realista hacia las situaciones del mundo, lo cual nos forzó a ser realistas sobre objetos, propiedades, relaciones y ubicaciones. Esto nos obligó a imponernos un realismo filosófico respecto de las actividades y estados cognitivos. Al final, ello permite un cierto abandono de la inocencia pura en favor de un tipo de inocencia mundana, que tenemos la esperanza de que resulte atractiva para muchos de nuestros lectores.

## SEMÁNTICA DEL ROL CONCEPTUAL \*

GILBERT HARMAN \*\*

En este artículo intentaré decir algo sobre cierto número de cuestiones que surgen en conexión con la «semántica del rol conceptual», el enfoque de la semántica por el que siento más simpatía. En esta ocasión sólo podré bosquejar algunos aspectos de este punto de vista sin que pueda dar argumentos detallados.

### 1. ¿QUÉ ES LA SEMÁNTICA DEL ROL CONCEPTUAL?

Una teoría de la semántica del rol conceptual incluye las dos afirmaciones siguientes:

1. El significado de las expresiones lingüísticas está determinado por los contenidos de los conceptos y pensamientos para cuya expresión éstas pueden usarse.
2. Los contenidos de los conceptos y pensamientos están determinados por su rol funcional en la psicología de una persona.

«Pensamientos» incluye aquí creencias, esperanzas, deseos, temores y otras actitudes, además de los pensamientos así llamados propiamente. «Rol funcional» incluye cualesquiera roles especiales que pueda desempeñar un concepto en la percepción y en la inferencia o razonamiento, incluyendo el razonamiento práctico que lleva a la acción.

La semántica del rol conceptual representa lo que podría querer decirse mediante el eslogan «el significado es el uso». Pero una apreciación adecuada de lo que quiere decirse requiere distinguir (al menos) dos usos de símbolos: su uso en el cálculo, como sucede al sumar una columna de números, y su uso en la comunicación, como sucede al decirle a alguien el resultado.

---

\* Versión castellana de Luis Ml. Valdés.

\*\* Estoy en deuda con Scott Soames por sus comentarios sobre este artículo.

### 1.1. DOS USOS DE SÍMBOLOS: COMUNICACIÓN Y CÁLCULO

Los símbolos que están siendo usados en el cálculo no se usan típicamente en ese momento para la comunicación. Cuando se procede a sumar una columna de números no se está normalmente comunicando nada, ni siquiera a uno mismo. Una observación similar vale en el caso inverso. Normalmente se comunican los resultados del cálculo que uno ha hecho a otra persona distinta sólo después de que se ha hecho el cálculo. Hay, desde luego, casos mixtos. Se puede proceder a un cálculo en una pizarra, intentando que el auditorio vea cómo se produce el proceso.

La semántica del rol conceptual puede verse como una versión de la teoría de que el significado es el uso, donde el uso básico de los símbolos se considera que está en el cálculo, no en la comunicación, y donde los conceptos se consideran como símbolos en un «lenguaje del pensamiento». Claramente, el uso relevante de tales símbolos —el uso que determina sus contenidos— es su uso en el pensamiento y en el cálculo más bien que en la comunicación. Si el pensamiento es semejante a hablar con uno mismo, es la suerte de habla incluida en resolver algún cálculo, no la suerte de habla incluida en la comunicación. Pensar no es comunicarse con uno mismo.

Sin embargo, sería más exacto decir que el contenido es el uso que decir que el significado es el uso; estrictamente hablando, los pensamientos y los conceptos tienen contenido, no significado.

### 1.2. EL SIGNIFICADO DE «SIGNIFICADO»

Supongo, siguiendo a Grice [6], que podemos distinguir lo que él llama significado natural (el humo significa fuego) de lo que llama significado no-natural (la palabra alemana «*Feuer*» significa «fuego»), y que podemos también distinguir significado no-natural del hablante o usuario (lo que un hablante o usuario de ciertos símbolos quiere decir) de lo que ciertas palabras, expresiones u otros símbolos significan.

Grice propone analizar el significado de una expresión en términos del significado del hablante; y propone, de manera más controvertida, analizar el significado del hablante en términos de las intenciones de un hablante para comunicar algo. Esta última propuesta parece pasar por alto el uso significativo de símbolos en el cálculo. Se podría inventar una notación especial para solucionar un cierto tipo de problema. Resultaría completamente apropiado decir que mediante un símbolo dado se quería decir tal-y-tal, aunque no se tenga intención alguna de usar esos símbolos en ningún tipo de comunicación.

Parece haber algún tipo de conexión entre significado del hablante o usuario y las intenciones del hablante o usuario. Supóngase que se usa esa notación especial para solucionar un problema específico. Se formulan las suposiciones del problema en esa notación, se hacen algunos cálculos, y se acaba con un resultado significativo en esa notación. Sería correcto decir de la persona que hace eso, cuando escribe una suposición particular en su notación, que quería decir tal y tal mediante lo que escribió: pero sería incorrecto decir de esa persona que, cuando escribió la conclusión que alcanzó en su notación, quería decir tal y tal mediante lo que escribió. Esto parece conectado con el hecho de que, al formular la suposición tal como se hizo en la notación especial, esa persona intentaba expresar tal y tal suposición; mientras que, al alcanzar la conclusión que se alcanzó en esa notación, la intención de la persona en cuestión no era expresar tal y tal conclusión, sino más bien alcanzar en su notación cualquier conclusión que se siguiese de los pasos anteriores por medio de las reglas de su cálculo. Esto sugiere que esa persona quiso decir que tal y tal al usar ciertos símbolos si y sólo si ella usó tales símbolos para expresar el pensamiento de que tal y tal, con la intención de expresar tal pensamiento.

Los pensamientos inexpressados (creencias, temores, deseos, etc.) no incluyen significado. No diríamos ordinariamente que al pensar como se ha hecho se quería decir que tal y tal. Si los pensamientos están en el lenguaje del pensamiento, no se expresan (normalmente) en ese lenguaje. (Digo «normalmente» para permitir la posibilidad de pensamientos que se piensan en un lenguaje usado en la comunicación, por ejemplo, pensamientos en castellano.)

Los conceptos y otros aspectos de la representación mental tienen contenido pero no (normalmente) significado (a menos que sean también expresiones en un lenguaje usado en la comunicación). No diríamos normalmente que el concepto de rojez que alguien tiene significaba algo de la manera en que la palabra «rojo» en castellano hacía lo propio. Ni tampoco diríamos que alguien quiso decir algo mediante ese concepto en una ocasión particular en la que se sirvió de él.

## 2. CONTENIDO Y ROL INFERENCIAL

### 2.1. TIPOS DE ROL

Suponiendo que la semántica del rol conceptual es un almacén básico, resulta plausible que todos los conceptos tengan una función en el razonamiento que es relevante para sus contenidos. Sin duda, algunos conceptos tienen primariamente el contenido que tienen a

causa de un rol especial que desempeñan en la percepción —por ejemplo, los conceptos para color—. Pero incluso el contenido de esos conceptos depende hasta cierto punto de un rol inferencial. Un concepto para color dado es el concepto de una característica de objetos del mundo, normalmente persistente, una característica que puede usarse tanto para seguirles la pista a los objetos, como un signo de otras cosas. Por ejemplo el verdor es un signo de inmadurez de ciertas frutas. Sin embargo, hay varias relaciones internas entre colores. A partir de la premisa de que un objeto es rojo en un cierto lugar y en un cierto tiempo puede inferirse que el objeto no es también verde en ese tiempo y lugar.

En el caso de los conceptos de forma y número, las conexiones inferenciales desempeñan un rol más amplio. Las conexiones perceptivas son todavía relevantes; hasta cierto punto el concepto que se tiene de un triángulo incluye la noción de aquello a lo que se parece un triángulo y el concepto que se tiene de diversos números naturales está conectado con la capacidad de contar objetos que se perciben. Pero el papel que esas nociones desempeñan en la inferencia parece más abultado.

El contenido expresado por la palabra «porque» desempeña un rol importante en la propia comprensión de fenómenos y tiene (mantengo) un rol central en la inferencia, ya que la inferencia es a menudo inferencia a la mejor explicación. Este rol hace del concepto expresado por «porque», según creo, el concepto que es. ¿Es aquí relevante la percepción? Quizás. Puede ser que algunas veces se perciba directamente la causalidad o ciertas relaciones explicativas, y puede ser que esto ayude a determinar el contenido del concepto que se expresa mediante la palabra «porque». O quizás no. Puede ser que la percepción de la causalidad y otras relaciones explicativas esté mediada siempre por una inferencia.

Las palabras lógicas «y», «no», «todo» y «alguno» expresan conceptos cuya función en la inferencia parece claramente muy importante para sus contenidos; ésta es la razón por la que parece tan plausible decir que esas palabras no significan en las lógicas intuicionista y cuántica lo que significan en la denominada lógica clásica, aunque incluso aquí puede haber funciones perceptivas cruciales. Por ejemplo, puede ser central al concepto que alguien tenga de negación el que algunas veces pueda percibirse que ciertas cosas no son así, como cuando se percibe que Pedro no está en el café. Puede ser central al concepto que se tenga de generalidad o cuantificación universal el que algunas veces se perciba que todo lo de un cierto género es de tal y cual manera, por ejemplo, que todo el mundo que está en la habitación lleva sombrero.

Es posible que haya ciertos géneros de términos teóricos tales como «*quark*», que no desempeñen en absoluto ningún rol en la per-



cepción, de modo que el contenido de los conceptos que expresan esté determinado enteramente por el rol inferencial. (¡Incluso aquí puede ser importante para el concepto de un *quark* que el concepto desempeñe un rol en la percepción de ciertas figuras o diagramas!)

## 2.2. INFERENCIA E IMPLICACIÓN

Las palabras lógicas tienen una función en la inferencia y en el razonamiento porque ciertas implicaciones dependen de ellas. La inferencia es, desde luego, un proceso de pensamiento que culmina típicamente con un cambio de punto de vista, un cambio en las creencias si se trata de razonamiento teórico, un cambio en los planes e intenciones en el caso del razonamiento práctico. (Existe también el caso límite en el que no se hace ningún cambio.)

No hay hasta ahora ninguna teoría substancial de la inferencia o del razonamiento. Ciertamente, la lógica está bien desarrollada; pero la lógica no es una teoría de la inferencia o del razonamiento. La lógica es una teoría de la implicación o de la argumentación, donde un argumento se concibe como una secuencia de implicaciones.

La lógica es relevante para el razonamiento porque la implicación lo es. La implicación es relevante para el razonamiento porque la implicación es una relación explicativa y, al razonar, se intenta entre otras cosas incrementar la coherencia explicativa del punto de vista propio (ver [13]).

## 2.3. FORMA LÓGICA

Las explicaciones de la forma lógica de un lenguaje natural pueden arrojar luz sobre el significado en la medida en que indican aspectos del lenguaje de los que pueden depender las implicaciones, ya que esto es indicar algo sobre el rol inferencial desempeñado por los conceptos expresados por esos aspectos del lenguaje.

Presumiblemente, tales explicaciones de la forma lógica deberían ser relevantes para, o quizás incluso parte de, un análisis gramatical de las oraciones relevantes. He aquí un área donde puede haber una útil interacción entre lo que hacen los filósofos y los lingüistas. Sin embargo, como Chomsky [3] observa, las distinciones que son importantes para la lingüística pueden no coincidir con las que son importantes para los filósofos. O, dicho de otra manera, los factores que determinan relaciones de implicación pueden no ser todos de la misma suerte. Algunos pueden ser aspectos de lo que Chomsky llama «gramática de oración», otros pueden no serlo. Y algunos aspectos de la «gramática de oración» que funcionan sintácticamente

como los rasgos lógicos pueden no estar directamente conectados con la implicación. Por ejemplo, Chomsky sugiere que las reglas de gramática que determinan cómo han de entenderse los cuantificadores, que son desde luego cruciales para determinar cuáles son las implicaciones lógicas de la oración, pueden ser las mismas que las reglas que determinan cosas tales como el «foco» de una oración, algo que parece no afectar a las implicaciones lógicas de una oración, sino solamente a sus «implicaturas conversacionales».

#### 2.4. INDETERMINACIÓN

Aparentemente hay análisis de las formas lógicas de oración que compiten entre sí. Donde un análisis ve lógica modal o lógica temporal, otro ve referencia a mundos posibles o tiempos. Donde un análisis ve referencia a eventos, otro análisis invoca una lógica adverbial. Y así sucesivamente. De modo similar, hay teorías gramaticales que compiten entre sí: la gramática de Montague, la armazón actual de Chomsky en términos de reglas de rección y ligamiento, y muchas otras variantes. ¿Qué hemos de hacer con ésto?

Quine [22] argumenta plausiblemente que incluso todas las observaciones posibles sobre un lenguaje pueden no decidir entre varias «hipótesis analíticas» localmente incompatibles (donde por «hipótesis analítica» Quine quiere decir hipótesis sobre la forma lógica o gramatical de la suerte que se acaba de mencionar). Ha habido disputas considerables respecto de cómo debería interpretarse exactamente la afirmación de Quine, respecto de si es verdadera, y acerca de cuáles podrían ser las implicaciones de su verdad. Se ha dicho (falsamente) que todo a lo que la tesis de Quine equivale es a la afirmación de que una teoría está subdeterminada por la evidencia. También se ha dicho (correctamente, creo), cualquiera que sea la observación válida que Quine está haciendo, que ésta no involucra ninguna diferencia significativa entre las «ciencias duras», como la física, y el estudio del lenguaje.

Una cuestión sugerida por la argumentación de Quine es ésta. Supóngase que tenemos una teoría, de la realidad física o del lenguaje, que pensamos que es verdadera. Incluso si pensamos que la teoría es verdadera, se puede proseguir con la consideración de qué aspectos de la teoría se corresponden con la realidad y qué aspectos son, por otra parte, meros artefactos de la notación en que se presenta la teoría. Una descripción geográfica verdadera de la Tierra mencionará longitudes, así como ciudades y montañas, pero las longitudes no tienen realidad geográfica en el modo en que las ciudades y las montañas la tienen. Es verdad que Greenwich (Inglaterra) está a cero grados de longitud, pero esta verdad es un artefacto de nuestro modo de describir la Tierra, puesto que hay otras maneras igualmente verda-

deras de describir la geografía de la Tierra que asignarían a Greenwich otras longitudes. Del mismo modo, hay varias descripciones físicas del mundo que son verdaderas, que asignan a un punto espacio-temporal dado diferentes coordenadas. Puede ser verdadero que bajo una descripción particular un punto particular tenga las coordenadas especiales  $(0, 0, 0, 0)$ , pero esto es un artefacto de la descripción que, por sí mismo, no se corresponde a nada en la realidad. Y lo mismo es verdadero de las gramáticas y de las teorías de la forma lógica. Incluso si una explicación dada de gramática o de forma lógica es verdadera, queda aún la cuestión de qué aspectos de la explicación corresponden a la realidad y qué aspectos son meramente artefactos de esa descripción particular. Es completamente posible que diversas explicaciones incompatibles localmente diferentes puedan ser todas ellas verdaderas, lo mismo que diversas asignaciones incompatibles localmente diferentes de longitudes y latitudes a lugares de la Tierra podrían ser todas ellas verdaderas.

Esto podría decirse de otra manera. La realidad es lo que es invariante entre teorías verdaderas. La realidad geográfica es lo que es invariante en diferentes descripciones geográficas verdaderas del mundo. La realidad física es lo que es invariante en diferentes descripciones físicas verdaderas del universo. Lo que despista a Quine es el hecho de que él tiene un bastante buen sentido de la realidad física y geográfica, pero poco o ninguno de la realidad gramatical o de la realidad descrita por explicaciones de la forma lógica. Es más, Quine está inclinado a pensar que hay aquí solamente dos posibilidades:

1. la realidad física en el nivel de la neurofisiología,
2. la realidad de la conducta, incluyendo disposiciones a comportarse de diferentes maneras.

Un punto de vista alternativo es que hay otros niveles de realidad definidos funcionalmente entre los dos que Quine reconoce. No veo otra manera de investigar este problema excepto el ver dónde lleva en última instancia la investigación actual sobre la gramática y la forma lógica. Desde luego, desde un punto de vista heurístico, es probablemente mejor suponer que diferentes explicaciones de gramática y forma lógica hacen afirmaciones en conflicto sobre la realidad a menos que haya alguna razón para pensar de otro modo (ver [16]).

## 2.5. SIGNIFICADO Y CONDICIONES DE VERDAD

Davidson [4], Lewis, [17] y otros han argumentado que una explicación de las condiciones de verdad de las oraciones de un lenguaje puede servir como una explicación de los significados de esas oracio-

nes. Pero esto parece erróneo. Desde luego, si uno sabe el significado en su lenguaje de la oración *O*, y sabe lo que significa la palabra «Verdadero», entonces también sabrá algo de la forma «*O* es verdadera si y sólo si ...»; por ejemplo: «“La nieve es blanca” es verdadera si y sólo si la nieve es blanca» o «“Estoy enfermo” es verdadera si y sólo si el hablante está enfermo en el tiempo de la emisión». Pero éste es un punto trivial sobre el significado de «verdadero», no algo profundo sobre el significado (veáse [10]).

Esto no es negar que los intentos de desarrollar teorías de la verdad que sean adecuadas para ciertos aspectos del lenguaje natural puedan muy bien arrojar luz sobre el significado. Algunos ejemplos podrían incluir el análisis veritativo-funcional de «y», «no», y «o»; el análisis tipo Frege-Tarski de la cuantificación; el análisis davidsoniano de las oraciones de acción; y la explicación de la modalidad en términos de mundos posibles. Pero en todos los casos los análisis ayudan a especificar implicaciones entre oraciones. Su relación con el significado puede deberse enteramente a esto, aparte de cualquier cosa adicional necesaria para tener una teoría de la verdad —aunque esto, desde luego, permite también que haya alguna cuestión de tipo heurístico en el intento de desarrollar teorías de la verdad (veáanse [9] y [16])—.

## 2.6. SEMÁNTICA PROBABILÍSTICA

Field [5] sugiere que el rol inferencial podría capturarse en términos de una distribución de probabilidad. Esto proporcionaría en el mejor de los casos una teoría de la implicación probabilística o de la coherencia, no una teoría de la inferencia en el sentido relevante, que incluya (normalmente) un cierto tipo de cambio en el punto de vista. Además, la gente ni opera ni podría operar probabilísticamente, puesto que seguir la pista a las probabilidades incluye capacidades de memoria y de cálculo que son funciones exponencialmente explosivas del número de proposiciones lógicamente no relacionadas que están incluidas (veáse Harman [11] y [14]).

En la mayor parte de los casos han de aceptarse proposiciones en un sentido de todo o nada. El conservadurismo es importante. Se debe continuar creyendo lo que se cree en ausencia de alguna razón especial para dudar de los propios puntos de vista, y al razonar debe intentarse minimizar el cambio en las propias opciones iniciales al alcanzar otras metas del razonamiento. Tales otras metas incluyen la coherencia explicativa y, desde luego, el éxito práctico al satisfacer las propias necesidades y deseos (véanse [11] y [12]). Pero estas cuestiones son vagas y no nos llevan muy lejos. Además, necesitamos decir algo en última instancia sobre el razonamiento práctico (véase [15]).

### 3. ROL CONCEPTUAL Y MUNDO EXTERNO

La semántica del rol conceptual no incluye una teoría «solipsista» del contenido de los pensamientos. No hay ninguna sugerencia de que el contenido dependa sólo de relaciones funcionales entre pensamientos y conceptos, tal como el rol que un concepto particular desempeña en la inferencia. (Field [5] no capta esto.) También son relevantes las relaciones funcionales con el mundo externo en conexión con la percepción, por una parte, y con la acción, por otra. Lo que hace que el concepto de rojo sea algo, es en parte el modo en que el concepto se incluye en la percepción de objetos rojos en el mundo externo. Lo que hace que el concepto de peligro sea algo, es en parte el modo en que el concepto se incluye en pensamientos que afectan a la acción de maneras diversas.

#### 3.1. LA DIVISIÓN DEL TRABAJO LINGÜÍSTICO

El contenido de ciertos conceptos parece depender crucialmente de relaciones funcionales entre esos conceptos y ciertas palabras de un lenguaje público. Se puede tener un concepto de roble en virtud de que se tienen pensamientos sobre robles donde la relación funcional crucial es una relación entre el concepto y la palabra «roble» en castellano. Uno podría preguntarse, por ejemplo, si había robles en su jardín incluso si no puede distinguir robles de hayas y no conoce ninguna de las propiedades que distinguen a esos dos tipos de árboles (véase Putnam [21]).

La semántica del rol conceptual asevera que una explicación del contenido de los pensamientos es más básica que una explicación del significado comunicado y de la significación de los actos de habla. Desde este punto de vista, el contenido de las expresiones lingüísticas se deriva de los contenidos de los pensamientos para cuya expresión pueden usarse aquéllas. Pero debe hacerse también un lugar para casos en los que el contenido de los pensamientos de uno depende, en parte, del contenido de ciertas palabras, tales como «roble» y «haya».

Desde luego, en este caso, hay otras personas que pueden reconocer robles y distinguirlos de las hayas y que conocen diversas propiedades que distinguen a los árboles. Esas otras personas pueden tener un concepto de roble que tiene roles funcionales que son suficientes para hacer de él el concepto de un roble independientemente de cualesquiera relaciones que el concepto tenga con la palabra «roble». Es plausible:

1. Que su concepto adquiriera su contenido a partir de este as-

pecto de su rol funcional, esto es, aparte de su relación con la palabra «roble».

2. Que la palabra «roble» tal como la usan tiene el contenido que tiene a causa de su conexión con su concepto de roble.
3. Que la palabra «roble» tal como es usada por una persona más ignorante puede tener el mismo contenido en virtud de conexiones entre el uso de la palabra por parte de esa persona ignorante y el uso del experto.
4. Que el contenido del concepto de roble de la persona más ignorante se derive de su conexión con el uso por parte de él o de ella de la palabra.

Esto aún nos permitiría decir que los significados de las palabras se derivan en última instancia de los contenidos de los conceptos para cuya expresión se usan las palabras, donde los contenidos de esos conceptos no se derivan ellos mismos de los significados de las palabras; sin embargo, los significados de las palabras de una persona particular pueden no derivarse de esta manera de los contenidos de los conceptos de esa persona.

Esto sugiere una cuestión interesante. ¿Existe alguna palabra para la que hay una división real del trabajo lingüístico de modo que ninguna persona aislada tenga un concepto correspondiente cuyo contenido esté determinado funcionalmente aparte de su relación con el uso por parte de esa persona de esa palabra? Es ciertamente imaginable que esto sea así en conexión con alguna suerte de investigación de grupo. Gente diferente podría investigar aspectos diferentes de un fenómeno que cada uno podría identificar como «cualquier cosa que todos nosotros estamos investigando y que tiene tales y cuales efectos cuando se investiga en el modo en que lo he investigado». Incluso en tal caso, el significado de la palabra derivaría del rol que el concepto correspondiente desempeña en el pensamiento, aunque diferentes aspectos de ese rol se cumplirían por diferentes instancias que la gente tuviese del concepto.

### 3.2. EL CONTENIDO COMO RELATIVO A LA ELECCIÓN DE UN CONTEXTO NORMAL

Putnam imagina un mundo, que él llama la «Tierra Gemela», que es justamente la Tierra excepto por lo que respecta a ciertas diferencias menores. En la Tierra Gemela hay duplicados de toda la gente de la Tierra, y la gente de la Tierra Gemela habla el mismo lenguaje que la de la Tierra, usando expresiones de la misma manera, excepto en que, a causa de diferencias menores entre la Tierra y la Tierra Gemela, se refieren mediante sus palabras a cosas diferentes. En parti-

cular, la principal diferencia entre la Tierra Gemela y la Tierra es que, donde hay agua en la Tierra, en la Tierra Gemela hay un líquido con las mismas propiedades macroscópicas que el agua pero con una diferente estructura química, que Putman llama «XYZ».

Ahora bien, comparando la Tierra en 1750 (antes de que fuese investigada la microestructura del agua) con la Tierra Gemela en el tiempo correspondiente, encontramos que la palabra castellana «agua» significa algo distinto en los dos lugares, simplemente porque la palabra se usa en la Tierra para referirse a lo que es de hecho H<sub>2</sub>O y se usa en la Tierra Gemela para referirse a lo que es de hecho XYZ. Similarmente allí donde los terrícolas piensan sobre H<sub>2</sub>O, los nativos de la Tierra Gemela piensan sobre XYZ. Esta diferencia no se refleja en 1750 en ninguna diferencia en las disposiciones a reaccionar a diversas situaciones perceptivas, en ninguna diferencia en inferencias que la gente podría hacer en los lugares respectivos, ni en ninguna diferencia en las acciones que la gente emprende como resultado de pensamientos que incluyen el concepto relevante.

La diferencia es también no solamente una diferencia en contexto de emisión o contexto de pensamiento. Supóngase que un terrícola viajase en una nave espacial y aterrizase en un océano de XYZ en la Tierra Gemela. El terrícola mira a su alrededor y llega a creer que todo lo que hay en torno suyo es agua. Esta creencia es falsa, puesto que el concepto de agua que tiene el terrícola es un concepto de algo que es de hecho H<sub>2</sub>O. El concepto de agua del terrícola continúa siendo un concepto de la misma cosa a la que se hace referencia mediante «agua» en la Tierra incluso cuando el terrícola está colocado ahora en un contexto distinto. El contexto de los pensamientos del terrícola y el contexto de los pensamientos de los habitantes de la Tierra Gemela son ahora los mismos; pero sus pensamientos son sobre XYZ mientras que los de aquél son todavía sobre agua. Así pues, esta diferencia en el contenido de los pensamientos de los terrícolas y de los habitantes de la Tierra Gemela no puede ser simplemente una diferencia en el contexto en el que tienen sus pensamientos.

La diferencia se debe más bien, me atrevo a sugerir, al hecho de que el contenido del concepto de una persona está determinado por su rol funcional en algún contexto normal. El contexto normal para los pensamientos de un terrícola sobre lo que él o ella llaman «agua» está aquí en la Tierra, y el contexto normal para los pensamientos de un habitante de la Tierra Gemela sobre lo que él o ella llaman «agua» está en la Tierra Gemela.

El contexto normal puede cambiar. Si el viajero de la Tierra a la Tierra Gemela permanece en ella, después de un tiempo se considerará que el contexto normal para los conceptos que él o ella usan es el contexto de la Tierra Gemela. Los pensamientos sobre lo que él o ella llaman «agua» se considerarán propiamente pensamientos sobre

XYZ más bien que sobre  $H_2O$ . Hay, desde luego, un cierto grado de arbitrariedad en cualquier decisión sobre cuándo ha ocurrido este cambio. Algunas veces será posible considerar con igual justicia que un pensamiento dado es un pensamiento sobre  $H_2O$  o un pensamiento sobre XYZ.

Una arbitrariedad similar surgiría respecto de una persona creada espontáneamente en el espacio exterior como resultado improbable de eventos fortuitos en el nivel cuántico, suponiendo que la persona fuese salvada de la muerte en el espacio por una nave espacial que pasase por allí por casualidad, y suponiendo que la persona hablase algo que se pareciese mucho al castellano. Es más, supóngase que esta persona fuese un duplicado de cualquiera de nosotros y también (por supuesto) de nuestra contrapartida de la Tierra Gemela. Cuando la persona tiene pensamientos que expresaría usando el término «agua», ¿son esos pensamientos sobre agua ( $H_2O$ ) o pensamientos sobre XYZ? Si interpretamos los pensamientos de esta persona en el trasfondo normal de la Tierra, interpretaremos los pensamientos relevantes como pensamientos sobre agua. Si consideramos que el trasfondo normal es la Tierra Gemela, entonces son pensamientos sobre XYZ. Claramente es del todo arbitrario lo que decimos aquí.

### 3.3. QUALIA

De acuerdo con la semántica del rol conceptual, entonces, el contenido de un pensamiento no es un asunto de la «naturaleza intrínseca» ni de ese pensamiento ni de otros estados mentales y experiencias, sino que es más bien un asunto de cómo están relacionados los estados mentales entre sí, con cosas del mundo externo, y con cosas en un contexto entendido como normal. Hay una objeción familiar a esto (Block y Fodor [2] y Nagel [19]) que afirma que el contenido no está determinado siempre por tales funciones o relaciones. Desde este punto de vista las cualidades intrínsecas o *qualia* de ciertas experiencias son relevantes algunas veces. Se dice que el concepto que uno tiene de rojo incluye su noción de lo que se parece a ver algo rojo, donde lo que se parece a ver algo rojo no es solamente un asunto de las características funcionales o relacionales de la experiencia relevante, sino también de su carácter intrínseco.

Un argumento a favor de esto es que es posible imaginar una persona cuya espectro se haya invertido respecto del nuestro, de modo que la cualidad de la experiencia que tú tienes al ver algo rojo es la cualidad que esa otra persona tiene al ver algo verde, la cualidad de la experiencia que tú tienes al ver algo azul es la cualidad que esta otra tiene al ver algo naranja, y similarmente para otros colores, aunque en todos los aspectos relevantes tus experiencias de color funcio-



nen similarmente, de modo que cada uno de nosotros sea tan bueno como cualquier otro a la hora de aplicar las palabras públicas de color a los objetos coloreados. De acuerdo con esta argumentación, esa persona y tú tendríais conceptos diferentes que tú expresarías usando la palabra «rojo», aunque sería difícil o incluso imposible descubrir esta diferencia, puesto que no es una diferencia funcional.

Hablo aquí de una argumentación, aunque (como D. Lewis [18] ha observado en un contexto similar) la «argumentación» se reduce en realidad simplemente a negar la explicación funcionalista del contenido de los conceptos y pensamientos, sin ofrecer ninguna razón para esa negación. Esto hace que la argumentación sea difícil de responder. Todo lo que puede hacerse es mirar más estrechamente a una explicación funcionalista del contenido de los conceptos de color con el objeto de encontrar el camino en el cual, de acuerdo con el funcionalismo, este contenido no depende del carácter intrínseco de las experiencias de color.

¿Cómo se podría imaginar a alguien cuyo espectro estuviese invertido respecto del nuestro? Un modo sería imaginar que esto te sucede a ti mismo. Supóngase que hubiese lentes de contacto que invertiesen los colores. Uno se pone un par de lentes y los colores de las cosas parecen haberse dado la vuelta. El cielo parece ahora naranja más bien que azul, las manzanas maduras parecen verdes, las manzanas que no están en sazón parecen rojas, y así sucesivamente. Supóngase que tú sigues con esas lentes y adaptas tu conducta. Aprendes a decir «verde» más bien que «rojo» cuando ves algo que tiene el aspecto que suelen tener las cosas rojas; aprendes a tratar lo que solías considerar como aspecto verde de las manzanas como un signo de que están en sazón, y así sucesivamente. Los años pasan y tu adaptación se convierte en habitual. ¿Sería esto un caso inteligible en el que alguien, el imaginado tú futuro, tiene una noción de a qué se parece tener la experiencia de ver algo a lo que se aplica el término «rojo», donde la noción funciona exactamente de la misma manera en que funciona tu noción de a qué se parece tal experiencia, aunque vuestras nociones fuesen diferentes? El funcionalista tendría que negar esto y decir que el tú imaginado asocia con la palabra «rojo» el mismo concepto que el tú efectivo ahora e incluso que ve el mundo tal como tú lo ves ahora.

Considérese un caso análogo. Actualmente existen lentes que invierten el espacio. Con esas lentes puestas, las cosas que están hacia arriba parecen estar hacia abajo, y viceversa. En principio es muy difícil moverse si se llevan esas lentes, puesto que las cosas no están donde parecen estar. Pero, después de un rato, uno comienza a adaptarse. Si se quiere agarrar algo que parece estar arriba, lo alcanzas abajo, y viceversa. Si se quiere mirar directamente hacia algo que parece estar en la parte baja del campo visual, se mira hacia arriba, y así

sucesivamente. Eventualmente una adaptación de este tipo se convierte en más o menos habitual.

Ahora bien, el funcionalismo implica que, si uno se adapta a tales lentes que invierten el espacio, entonces su experiencia será la misma que la de alguien que no lleva las lentes inversoras (que se ha adaptado a no llevarlas en el caso necesario), puesto que ahora el contexto normal en relación con el cual funcionan sus conceptos se habrá convertido en un contexto en el que uno lleva dichas lentes, y, de hecho, la gente que ha llevado tales lentes dice que, a medida que se adaptan a las lentes, el mundo tiende a parecer de nuevo al derecho (véanse [23], [20] y [24]).

Similarmente el funcionalismo implica que, si una persona se adapta perfectamente a las lentes que invierten colores, el mundo comenzará a parecerse como le parecía antes en el sentido de que, dada tal adaptación perfecta, el contexto normal en el que funcionan sus conceptos de color será un contexto en el que llevas las lentes que invierten colores. De acuerdo con el funcionalismo, el modo en que las cosas le parecen a uno es una característica relacional de su experiencia, no parte de su carácter intrínseco.

Para obtener una visión de este aspecto de la semántica del rol conceptual, puede ser útil considerar ciertos casos adicionales. Considérese la Tierra Invertida, un mundo completamente igual al nuestro, con duplicados de nosotros, con la única diferencia de que los colores efectivos de los objetos son los opuestos de los que en realidad son. El cielo es naranja. Las manzanas maduras son verdes, etc. Los habitantes de la Tierra Invertida hablan algo que suena de una manera parecida al castellano, excepto que ellos dicen que el cielo es «azul», llaman «rojas» a las manzanas que están maduras, y así sucesivamente. Pregunta: ¿de qué color les parece a ellos su cielo? Respuesta: les parece naranja. El concepto que ellos expresan con la palabra «azul» desempeña un rol especialmente relevante en la percepción normal de las cosas que son efectivamente naranja.

Supóngase que hay una base física distinta para cada experiencia de color distinta. Supóngase también que la base física para la experiencia de rojo es la misma para toda la gente normal no adaptada a las lentes que invierten colores, y similarmente para los demás colores. De acuerdo con la semántica del rol conceptual, este hecho es irrelevante. La persona que se ha adaptado perfectamente a las lentes que invierten los colores será diferente de cualquier otra por lo que respecta a las bases físicas de su experiencia de rojo, pero esto no afectará a la cualidad de su experiencia.

Considérese una persona en la Tierra Invertida que se adapta perfectamente a las lentes que invierten colores. Mirando hacia el cielo de la Tierra Invertida, esta persona tiene una experiencia de color cuya base física es la misma que la de una persona normal que

mira hacia el cielo de la Tierra. Pero el cielo, para la persona de la Tierra Invertida, tiene el aspecto de ser naranja y, para las personas normales de la Tierra, tiene el aspecto de ser azul. Lo que hace de una experiencia la experiencia de algo que tiene el aspecto de tener el color que tiene no es su carácter intrínseco y/o su base física, sino más bien sus características funcionales dentro de un contexto normal supuesto.

Considérese un cerebro creado en el espacio espontáneamente como resultado improbable de eventos azarosos en el nivel cuántico. Los eventos físicos del cerebro resultan ser los mismos que los que suceden en el de cualquiera de nosotros cuando miramos al cielo en la Tierra, y también los mismos que tienen lugar en el de una persona adaptada a las lentes que invierten colores que está mirando al cielo de la Tierra Invertida. ¿Qué pasa en el cerebro? ¿Está teniendo una experiencia de naranja o de azul? De acuerdo con la semántica del rol conceptual, no hay ningún modo no arbitrario de responder a esta pregunta; depende de lo que se considere como contexto normal para valorar el rol funcional de eventos en ese cerebro. Si se considera que el contexto normal es el contexto normal para la percepción de color en la Tierra, el cerebro está teniendo una experiencia de azul. Si se considera que el contexto normal es el contexto normal para un portador de lentes invertidas en la Tierra Invertida, el cerebro está teniendo una experiencia de naranja.

#### 4. SIGNIFICADO Y ACTOS DE HABLA

Puede ser una exageración el decir que todos los aspectos del significado lingüístico se derivan del uso del lenguaje para expresar conceptos con los contenidos correspondientes. Ciertos aspectos del significado parecen derivarse directamente del uso del lenguaje y de los actos de habla en la comunicación. La semántica del rol conceptual puede, por lo tanto, tener esto en cuenta, intentando (desde luego) tratarlo como un fenómeno menor y relativamente periférico. Por otra parte, puede ser que no se necesite hacer tal concesión. He aquí algunas de las reflexiones tentativas y preliminares de algunos de los problemas planteados.

##### 4.1. SALUDOS

Los casos que aquí pueden parecer más claros se refieren a expresiones que se usan primariamente en saludos y despedidas, palabras tales como «hola» o «adiós». El significado de esas palabras no parece derivarse de ningún uso que puedan tener al expresar conceptos

únicos que desempeñen un papel distintivo en el cálculo y en el pensamiento. Es cierto que se podría, en alguna ocasión, «saludar» una nueva idea o percepción con el pensamiento, «¡Hola! ¿Qué es esto?». Pero esto parece ser un caso en el que el contenido de tu pensamiento deriva del uso de la palabra «hola» para saludar a personas y comenzar conversaciones más bien que a la inversa.

Sin embargo, puede suceder que tales palabras puedan ser analizadas como expresando combinaciones de conceptos que, individualmente, tienen contenidos conectados con roles conceptuales distintivos. Por ejemplo, «hola» significa quizás algo semejante a «Reconozco tu presencia», o algunas veces puede que «¡hablemos!», y análogamente para «adiós» y otras palabras y frases de este tipo. Si esto es así, el asunto se convierte en si el aspecto del significado expresado por el imperativo en «¡hablemos!» y el aspecto realizativo del significado de «Reconozco tu presencia» se derivan irreductiblemente del uso de las palabras en actos de habla o se derivan en última instancia del uso del lenguaje para expresar conceptos cuyo contenido está determinado por su rol en el cálculo y el pensamiento. Es éste un problema complejo que tenemos que considerar dentro de un momento. Por lo que respecta a la cuestión de si éste es el modo correcto de analizar «hola», «adiós» y así sucesivamente, no estoy muy seguro de qué pensar. Los análisis que he sugerido parecen dejar algo fuera, pero esto podría acomodarse por medio de análisis mejores.

#### 4.2. PALABRAS DE CORTESÍA

Antes de volver a los imperativos y a los realizativos podríamos considerar palabras y frases que funcionan como formas de deferencia y cortesía: por ejemplo, «por favor» y «gracias». El uso de tales palabras y frases parece presuponer alguna suerte de interacción social, más bien que simplemente la expresión de concepto con un contenido determinado por el modo en que esos conceptos funcionan en el pensamiento y en el cálculo. Es verdad que uno podría decirse a sí mismo, para sus adentros, «por favor» o «gracias», si estuvieras hablando contigo mismo como podría hacerlo un instructor o un entrenador. Pero esto parece un caso muy especial, que presupone el uso social de esas expresiones.

Nuevamente, podría argumentarse que en este uso «por favor» significa lo mismo que alguna expresión más larga, tal como «si te apetece» o «si te apetece hacerlo», y «gracias» significa algo parecido a «Por la presente te doy las gracias». Si esto es así el problema está determinado por sus funciones distintivas en el cálculo o en el pensamiento, un problema que necesitamos discutir. También, está otra vez la cuestión de si el elemento realizativo de «Por la presente te doy

las gracias» comporta un significado que está, al menos en parte, irreductiblemente conectado con actos de habla, un problema al que nos aplicaremos dentro de un momento. Por lo que respecto a la cuestión de si tales frases de cortesía pueden analizarse siempre de ese tipo de manera, no estoy seguro de qué decir.

#### 4.3. «TÚ»

¿Qué sucede con el significado de la palabra «tú»? Es plausible que esta palabra signifique algo parecido a «la persona a la que me estoy dirigiendo ahora» y es plausible mantener que cada una de las palabras de esta frase tiene un significado que depende del concepto para cuya expresión se usa la palabra, donde el concepto en cuestión tiene el contenido que tiene en virtud del modo en que funciona en el pensamiento y en el cálculo. Así pues, aunque la palabra «tú» tiene la función distintiva en los actos de habla y la comunicación de designar el auditorio hacia el que uno se dirige, esta función puede explicarse plausiblemente en términos de los roles funcionales en el pensamiento de los conceptos para cuya expresión se usa la palabra.

#### 4.4. MODO IMPERATIVO E INTERROGATIVO

El uso del modo imperativo en castellano parece comportar un cierto significado, conectado con el dar instrucciones de algún género. Esto no implica por sí mismo que este significado no se derive del contenido de ciertos conceptos con un uso distintivo en el pensamiento y en el cálculo. Es más, algo funciona en tus pensamientos para distinguir tus creencias de tus planes e intenciones, que son instrucciones de alguna clase. Así, hay un sentido en el que el resultado del razonamiento práctico es una modificación de ciertas instrucciones que uno intenta seguir. Quizás el modo imperativo sirve para expresar el concepto que funciona en el pensamiento para así distinguir los pensamientos prácticos o directivos de los pensamientos teóricos.

Observaciones similares se aplican al modo interrogativo. Es más, las preguntas no son distintas de las peticiones de información, de manera que el modo interrogativo se analiza plausiblemente en términos del modo imperativo. En cualquier evento, las preguntas tienen obviamente una función en el pensamiento. Uno se plantea un problema a sí mismo y trabaja sobre la respuesta, planteando quizás varias subpreguntas y respondiéndolas.

#### 4.5. REALIZATIVOS

Considérense los siguientes realizativos explícitos, como «Prometo estar allí» y «Por la presente pido disculpas por mi mala educación». Parece parte del significado de tales oraciones el que se usen no para describir al hablante como prometiendo o pidiendo disculpas, sino como haciendo efectivamente la promesa o pidiendo efectivamente disculpas. Además, es plausible que prometer o pedir disculpas a uno mismo no sea un caso típico o normal de prometer o pedir disculpas, y es más bien el tipo de caso que ha de entenderse en relación con casos más típicos o normales en los que prometes algo a alguien distinto o pides disculpas a otra persona.

Por otra parte, cada una de las palabras de una oración como «Prometo estar allí» tiene un significado que expresa un concepto cuyo contenido, puede argumentarse, está determinado por su rol funcional en el cálculo y en el pensamiento. Y es posible que el significado de la oración completa, incluyendo cualquier cosa que da a la oración su función realizativa de ser apropiada para prometer efectivamente, surge del significado de las palabras usadas de una manera regular. Dado lo que las palabras de la oración significan y dado el modo en que esas palabras se combinan, puede ser predecible que la oración tenga un uso realizativo (véase [1]).

Supóngase que adoptamos la convención de que las promesas han de hacerse de alguna manera especial: por ejemplo, escribiendo el contenido de la promesa con tiza de color morado en un tablero especial para promesas que no se usa para ningún otro propósito. La convención sería que nada más contará como una promesa. En tal caso, las palabras «Prometo estar allí» no podrían usarse para prometer que estaré allí. ¿Sería ésta una manera de separar el significado realizativo de «Prometo» de ese aspecto de su significado que se deriva de su uso para expresar conceptos cuyo contenido está determinado por su rol funcional en el pensamiento? No, obviamente. En un aspecto, esto podría cambiar el concepto de prometer de una manera significativa. La palabra «promesa» podría no significar lo que significaba cuando una promesa podía hacerse diciendo «Prometo». Podría argumentarse que si «Prometo» significa lo que ordinariamente significa, entonces se sigue de los conceptos expresados por la palabra «Prometo» que esa palabra puede usarse para prometer.

Alternativamente, podría decirse que, incluso si la palabra «promesa» retuviese (bastante de) su significado usual cuando el prometer se restringe a esa convención, el ejemplo se parece a aquél en el que se adopta una convención especial de que una emisión de la oración «El cielo es azul» no ha de ser interpretada como una aserción de que el cielo es azul, sino más bien como una pregunta que plantea si ha llovido la semana pasada. Esto no mostraría que hay algún as-

pecto del significado de una aserción ordinaria de la oración «El cielo es azul», tal como la usamos ahora sin tal extraña convención, que no se derive del modo en que las palabras de la oración se usan para expresar conceptos que tienen el contenido que tienen a causa de su rol funcional en el pensamiento.

#### 4.6. IMPLICATURA CONVERSACIONAL

Grice ([7] y [8]) argumenta que las aplicaciones de una emisión no se corresponden siempre directamente con los significados de las expresiones lingüísticas usadas, incluso en los casos completamente ordinarios. En particular sugiere que lo que parecen ser aspectos del significado pueden deberse a «implicaturas conversacionales», esto es, a conclusiones que se intenta que el auditorio alcance al reflexionar sobre las razones que el hablante tiene para decir lo que se dice, suponiendo que el hablante está intentando cooperar. Por ejemplo, si se usa «o...o» como en «Alberto está o en Nueva York o en Boston», normalmente se implica que no se sabe dónde. Esto no ha de considerarse que muestra algo especial sobre el significado de «o...o», una diferencia en el significado de esta expresión ordinaria en tanto que comparada con lo que los lógicos consideran que significa. En lugar de esto debemos suponer que esta implicación se debe a la suposición natural de que normalmente, si sabes en qué ciudad está Alberto, dirás en qué ciudad está. Para tomar otro ejemplo, la diferencia aparente de significado entre «María cerró la puerta y apagó la luz» y «María apagó la luz y cerró la puerta» no ha de explicarse suponiendo que la palabra castellana «y» significa algunas veces «y a continuación». La sugerencia es que la diferencia puede explicarse suponiendo que un hablante cooperativo relatará normalmente eventos de una manera ordenada, de modo que un oyente está justificado normalmente al suponer que el orden en el que el hablante relata los eventos ha de entenderse como el orden en el que ocurrieron. Similarmente, Grice sugiere que ciertos aspectos de la presuposición podrían explicarse considerando las expectativas normales de hablantes y oyentes.

Sin embargo, alguno de esos fenómenos pueden ocurrir pensando con uno mismo, allí donde no se deben presumiblemente a una implicatura conversacional. El cálculo y el razonamiento incluyen a menudo varias presuposiciones. Normalmente uno querrá que las descripciones usadas en el razonamiento relaten eventos de una manera ordenada, de modo que puede ocurrir el mismo fenómeno que veíamos con la palabra «y». Por otro lado, es dudoso que el uso de una proposición «o...o» en el pensamiento comporte normalmente la implicación de que uno no sabe qué alternativa es el caso; así pues, *este*

fenómeno puede realmente ocurrir sólo en el nivel de la conversión.

#### 4.7. LENGUAJE FIGURATIVO

Las metáforas y los símiles ocurren en el pensamiento. Estoy inclinado a pensar que no sucede lo mismo con la ironía. No estoy seguro sobre qué pasa con la hipérbole.

#### 5. CONCLUSIÓN

Resumiendo: hay dos usos de símbolos, en la comunicación y en los actos de habla, y en el cálculo y en el pensamiento. La semántica del rol conceptual considera que el segundo uso es el básico. La fuente última de significado o contenido es el rol funcional que los símbolos desempeñan en el pensamiento.

El contenido de un concepto depende de su rol en la inferencia y algunas veces en la percepción. Son particularmente importantes las implicaciones de un término. La implicación es relevante para la inferencia y, por lo tanto, para el significado, puesto que la implicación es explicativa y la inferencia aspira a coherencia explicativa. Las explicaciones de las condiciones de verdad pueden arrojar luz sobre el significado en la medida en que den lugar a implicaciones; es dudoso si tales explicaciones tienen algún peso adicional sobre el significado, aunque puedan tener un valor heurístico para estudios de la forma lógica. La semántica probabilística no proporciona una semántica del rol conceptual adecuada, puesto que la gente no hace ni puede hacer mucho uso del razonamiento probabilístico.

Deben tenerse en cuenta diversas conexiones entre conceptos y el mundo externo. Algunos conceptos tienen el contenido que tienen a causa de las palabras con las que están asociadas, aunque (de acuerdo con la semántica del rol conceptual) este contenido siempre se deriva en última instancia del uso que alguien hace de él en conceptos. El contenido de los conceptos es a menudo relativo a una elección de un contexto normal de funcionamiento. Esto es verdad de los conceptos de color, a pesar del punto de vista no argumentado que mantienen algunos filósofos de que esos conceptos dependen del carácter intrínseco de la experiencia.

Finalmente no está claro si algunos aspectos del significado se derivan directamente del uso del lenguaje en actos de habla de un modo no reducible a las expresiones de conceptos cuyo contenido está determinado independientemente. En cualquier evento, muchos fenó-



menos que se consideran a menudo que están conectados particularmente con actos de habla y conversación ocurren también en el cálculo y en el pensamiento.

## REFERENCIAS

- [1] BACH, K.: «Performatives are statements, too», *Philosophical Studies*, vol. 28 (1975), pp. 229-236.
- [2] BLOCK, N., y FODOR, J.A.: «What psychological states are not», *Philosophical Review*, vol. 81 (1972), pp. 159-181.
- [3] CHOMSKY, N.: *Reflections on Language*, Columbia University Press, Nueva York, 1980. [Versión castellana: Ariel, Barcelona, 1979.]
- [4] DAVIDSON, D.: «Truth and meaning», *Synthese*, vol. 17 (1967), pp. 304-323. [Versión castellana: pp. 314-334 de este volumen.]
- [5] FIELD, H.: «Logic, meaning, and conceptual role», *Journal of Philosophy*, vol. 74, 1977, pp. 379-409.
- [6] GRICE, H.P.: «Meaning», *Philosophical Review*, vol. 68 (1959), pp. 377-388. [Versión castellana: *Cuadernos de Crítica*, UNAM, México, 1977.]
- [7] GRICE, H.P.: «The causal theory of perception», *Proceedings of the Aristotelian Society*, Supplementary Volume 35 (1961).
- [8] GRICE, H.P.: «Logic and conversation», en D. Davidson y G. Harman (eds.), *The Logic of Grammar*, Dickenson Press, Encino, California, 1975. [Versión castellana: pp. 511-530 de este volumen.]
- [9] HARMAN, G.: «Logical form», *Foundations of Language*, vol. 9 (1972), pp. 38-65.
- [10] HARMAN, G.: «Meaning and semantics», en M.K. Munitz y P. Unger (eds.), *Semantics and Philosophy*, New York University Press, Nueva York, 1974.
- [11] HARMAN, G.: «Reasoning and explanatory coherence», *American Philosophical Quarterly*, vol. 17 (1980), pp. 151-157.
- [12] HARMAN, G.: «Reasoning and evidence one does not possess», *Midwest Studies in Philosophy*, vol. 5 (1980), pp. 163-182.
- [13] HARMAN, G.: «Logic and reasoning», que aparecerá en las actas de una conferencia sobre lógica, celebrada en Nueva York en 1980.
- [14] HARMAN, G.: «Problems with probabilistic semantics», que aparecerá en A. Orenstein y otros (eds.), *Developments in Semantics*, Haven Press, Nueva York.
- [15] PUTNAM, H.: «Willing and intending», que aparecerá en 1985 en un *Festschrift* a H.P. Grice, preparado por R. Grandy y R. Warner, Oxford University Press.
- [16] HARMAN, G.: «Quine's grammar», que aparecerá en P. Schilpp (ed.), *The Philosophy of W.V. Quine*, Open Court Publishing Company, La Salle, Illinois.
- [17] LEVIS, D.: «General semantics», en D. Davidson y G. Harman (eds.), *Semantics of Natural Language*, D. Reidel Publishing Company, Dordrecht, (Holanda), 1972.
- [18] LEVIS, D.: «Mad pain and Martian pain», en N. Block (ed.), *Readings in Philosophy of Psychology*, Harvard University Press, Cambridge, Massachusetts, 1980.
- [19] NAGEL, T.: «What is it like to be a bat?», *Philosophical Review*, vol. 83 (1974), pp. 435-450.
- [20] PITCHER, G.: *A Theory of Perception*, Princeton University Press, Princeton, Nueva Jersey, 1971.
- [21] PUTNAM, H.: «The meaning of meaning», en *Mind, Language, and Reality: Philosophical Papers*, vol. 2, Cambridge University Press, Cambridge, Massachusetts, 1975. [Versión castellana: pp. 131-194 de este volumen.]

[22] QUINE, W.V.: *Word and Object*, Massachusetts Institute of Technology Press, Cambridge, Massachusetts, 1960. [Versión castellana: Labor, Barcelona, 1968.]

[23] TAYLOR, J.G.: *The Behavioral Basis of Perception*, Yale University Press, New Haven, Connecticut, 1962.

[24] THOMAS, S.: *The Formal Mechanics of Mind*, Cornell University Press, Ithaca, Nueva York, 1978.

## SOBRE LA DEFINICIÓN DE RELEVANCIA \*

DEIRDRE WILSON Y DAN SPERBER <sup>1</sup>

### 1. INTRODUCCIÓN

Bajo la categoría de RELACIÓN, coloco una única máxima, a saber: «Sé relevante». Aunque la máxima es ella misma bastante lacónica, su formulación esconde un buen número de problemas que me preocupan bastante... Encuentro que el tratamiento de tales cuestiones es extraordinariamente difícil, y espero volver a ellas en una obra posterior. (Grice, 1975, p. 46.)

Grice introdujo con estos cautos términos en sus *William James Lectures* una máxima de relevancia que, junto con las máximas de veracidad, informatividad y perspicuidad, iba a formar la base para su estimulador e influyente enfoque de la pragmática. En los años transcurridos desde entonces, a pesar de la considerable cantidad de investigación llevada a cabo dentro de y sobre la armazón pragmática griceana, no ha surgido ninguna elaboración explícita de la relevancia que sea adecuada para su uso en la teoría pragmática <sup>2</sup>. Es más, ha habido un escepticismo considerable sobre si tal explicación es, en principio, posible <sup>3</sup>. Allí donde se han intentado dar definiciones, se

\* Versión castellana de Luis M. Valdés Villanueva.

<sup>1</sup> Nos gustaría agradecer a los participantes en el *1980 Pragmatics Workshop*, en Asilomar, California, y en particular a Herb Clark, Paul Grice, Larry Horn, Jerry Morgan, Ellen Prince, Ivan Sag y Bonnie Webber, sus estimulantes comentarios a una primera versión de este artículo. Estamos agradecidos también a Diane Brockway, Robyn Carston, Jerry Katz, Ruth Kempson y Yuji Nishiyama por su ayuda.

<sup>2</sup> Lo que sigue es una muestra representativa de comentarios:

«...ningún intento de aplicar la teoría de la semántica formal a esta noción (relevancia) ha tenido suficiente éxito para proporcionar un modelo que pueda usarse en pragmática» (Thomason, 1973, p. 12).

«El que la relevancia es relevante para la descripción lingüística es dolorosamente aparente... Igualmente aparente es la casi completa ausencia de cualquier género de tratamiento lingüístico formal de la noción» (Gazdar, 1979, p. 45).

«...las explicaciones actuales de la interacción conversacional dependen de manera crucial de la noción no definida de "relevancia"» (Werth, 1981, p. 30).

<sup>3</sup> Veáanse, por ejemplo, los comentarios siguientes:

«La relevancia no es una categoría lógica precisa... la palabra se usa para transmitir una idea esencialmente vaga» (Berlin, 1938-1939, p. 21).

«Convencionalmente, las consideraciones de relevancia son más propias de ser te

las ha enfocado en términos retóricos, emocionales o estéticos más bien que lógicos<sup>4</sup>. Con todo, puesto que la comprensión de una emisión incluye un elemento inferencial substantivo, no es poco razonable esperar que la extracción de inferencias contribuya de alguna manera a los juicios de relevancia. En este artículo queremos argumentar que existe una íntima conexión entre inferencia y relevancia y que un enfoque de la relevancia en términos inferenciales puede proporcionar resultados fructíferos para la teoría pragmática.

## 2. LA NATURALEZA DE LA RELEVANCIA

Resulta usual tratar a la relevancia como una propiedad de emisiones o como una relación entre una emisión y un texto o discurso. Sin embargo, la información relevante puede derivarse no sólo de las emisiones y otros actos comunicativos, sino también de la memoria, la observación y la inferencia. Trataremos la relevancia en primera instancia como una propiedad de proposiciones (unidades de información, combinaciones de sentido y referencia); las definiciones subsidiarias de relevancia para emisiones o discursos son relativamente fáciles de construir. Similarmente, trataremos en primera instancia una proposición como relevante, no para un texto o discurso, sino para un contexto, donde un contexto es un *stock* de información derivado no sólo del discurso precedente, sino también de la memoria, la observación y la inferencia. Desde el punto de vista formal, un contexto es simplemente un conjunto de proposiciones de tamaño y contenido arbitrarios. Desde el punto de vista psicológico, tanto el tamaño como el contenido están sujetos a constricciones considerables; las ignoraremos por el momento y las discutiremos en una sección posterior.

La interpretación de una emisión incluye, por una parte, la identificación de la proposición que ha expresado el hablante y, por otra,

---

legadas a la retórica que a la dimensión lógica de valoración de argumentos» (Haack, 1978, pp. 16-17).

«...la noción de conexión o dependencia a la que se apela aquí es demasiado vaga para ser un concepto formal de la lógica» (Suppes, 1957, p. 8).

«La dificultad de tratar la relevancia con el mismo grado de sofisticación y exactitud matemáticas características de los tratamientos de la lógica extensional ha llevado a muchos influyentes lógico-filósofos a creer que era *imposible* encontrar un tratamiento satisfactorio del tópico» (Anderson y Belnap, 1975, p. XXI).

<sup>4</sup> Para algunos tratamientos recientes, véanse Van Dijk, 1979; Dascal, 1977, y Wert, 1981. Van Dijk hace equivaler la relevancia con el grado de importancia en el discurso; Dascal la relaciona con la satisfacción de una demanda conversacional, y Werth la hace equivaler con pertinencia para el significado de la emisión previa, juntamente con el contexto.

el procesamiento de esta proposición en un contexto proporcionado por el oyente y que consta, como hemos visto, de información derivada de una variedad de fuentes. Procesar una proposición es simplemente extraer información de ella; procesarla en un contexto es proporcionar una información de trasfondo adicional que contribuye de alguna manera a su procesamiento. El procesamiento de una proposición puede tratarse en gran medida en términos inferenciales. Por ejemplo, el procesamiento de 1) podría incluir el uso de 1) como premisa, junto con la información de trasfondo de 2), para producir la conclusión 3):

- 1) Juan acaba de comprar un *Rolls-Royce*, pero su mujer se niega a conducir coches caros.
- 2) El *Rolls-Royce* es un coche caro.
- 3) La mujer de Juan se niega a conducir en su *Rolls-Royce*.

En este caso, el procesamiento de 1) incluye el dar los pasos necesarios para extraer un cierto conjunto de inferencias, y el contexto será, o al menos incluirá, el conjunto de suposiciones de trasfondo tales como 2), usadas como premisas suplementarias del proceso inferencial.

Cuando una proposición  $P$  se añade a un contexto  $C_1 \dots C_n$ , pueden tener lugar dos tipos distintos de procesos de inferencia que incluyen  $P$  como premisa. Por una parte,  $P$  puede considerarse como la única premisa y obtenerse sus implicaciones lógicas. Éstas son, desde luego, invariantes de contexto a contexto, pero podemos decir que  $P$  es *informativa* en el contexto  $C_1 \dots C_n$  si y sólo si  $P$  tiene al menos una implicación lógica no implicada por  $C_1 \dots C_n$ .

El procesamiento de  $P$  en un contexto  $C_1 \dots C_n$  puede también proporcionar un conjunto adicional de implicaciones, esta vez dependientes del contexto: el conjunto de proposiciones que están lógicamente implicadas no por  $P$  sola, ni por  $C_1 \dots C_n$  solos, sino por la unión de  $P$  y  $C_1 \dots C_n$ . Llamemos a éstas las *implicaciones contextuales* de  $P$  en el contexto  $C_1 \dots C_n$ . Por ejemplo, considérese un contexto que conste de las dos proposiciones 4 a) y 4 b):

- 4) a. Si el presidente dimite, Juan asumirá sus funciones.  
b. Si Juan asume las funciones del presidente, la compañía quebrará.

Si a este contexto se añade la proposición 5), las conclusiones 6) y 7) pueden obtenerse como implicaciones contextuales de 5) en el contexto 4 a)-4 b):

- 5) El presidente ha dimitido.
- 6) Juan asumirá las funciones del presidente.
- 7) La compañía quebrará.

Las proposiciones 6) y 7) son implicaciones contextuales de 5) en este contexto: resultan implicadas lógicamente por la unión de 4) y 5), pero no por 4) sola ni por 5) solas.

Puesto que la informatividad en un contexto es definible en términos de implicación lógica, la relevancia en un contexto podría enfocarse en términos de implicación contextual. Diríamos que una proposición es relevante en un contexto  $C_1...C_n$  si y sólo si tiene al menos una implicación contextual en  $C_1...C_n$ . Intuitivamente, ser relevante en un contexto es un asunto de conectar con el contexto de alguna manera. De acuerdo con esta propuesta, la relevancia en un contexto es un asunto de conectar con el contexto de un modo altamente específico: de modo que se tengan implicaciones contextuales en este contexto. Mediante esta definición, puesto que 5) tiene implicaciones contextuales en el contexto 4), sería relevante también en ese contexto <sup>5</sup>.

Este enfoque inferencial de la relevancia encajaría de manera natural en una teoría inferencial de la comprensión. Al procesar una proposición en un contexto, el oyente derivaría automáticamente sus implicaciones contextuales sobre cuya base se establecería la relevancia de la proposición. De hecho, podría irse más allá y afirmar que el propósito de procesar una proposición es precisamente establecer su relevancia, haciendo de la relevancia la piedra angular de la teoría pragmática. Cualquiera que sea su plausibilidad inicial esta propuesta no puede, sin embargo, ser tomada en serio sin que se proporcionen respuestas a un determinado número de cuestiones. Consideraremos aquí brevemente las tres que nos parecen más urgentes. Vale quizá la pena mencionar que ninguna es exclusiva de nuestro enfoque particular. La primera, que tiene que ver con la naturaleza de las reglas de inferencia usadas, surge en cualquier teoría pragmática inferencial. La segunda, que tiene que ver con el hecho de que la relevancia es un asunto de grado, surge en cualquier teoría en la que la relevancia desempeña algún papel. La tercera, que tiene que ver con la naturaleza e identificación de los contextos, surge en cualquier teoría pragmática en la que el contexto desempeña algún papel. Las respuestas que daremos, al igual que las cuestiones mismas, son a menudo posibles solamente dentro de la armazón que proponemos.

### 3. EL SISTEMA DEDUCTIVO

Supongamos que la lógica usada al derivar implicaciones contextuales es una lógica estándar, digamos un sistema estándar de deduc-

<sup>5</sup> Goodman, 1961, define una noción de «aboutness relativa» en términos algo similares, aunque con propósitos más bien diferentes.

ción natural. Entonces, para cualquier contexto  $Q$  y cualquier proposición  $P$ , la conjunción de  $P$  y  $Q$  será una implicación contextual de  $P$  en el contexto  $Q$ :  $P$  y  $Q$  resulta implicada lógicamente por la unión de  $P$  y  $Q$ , pero ni por  $P$  sola ni por  $Q$  sola. Toda proposición tendrá al menos una implicación contextual en todo contexto y, por lo tanto, de acuerdo con nuestra definición de relevancia, será relevante en todo contexto, algo que es absurdo. Por lo tanto, o la lógica usada al derivar implicaciones contextuales no es la lógica estándar, o nuestro enfoque de la relevancia es incorrecto.

De hecho, de manera completamente independiente de este enfoque, hay buenas razones para pensar que la lógica usada en la comprensión de una emisión no es una lógica estándar. Por una parte, tiene que ser mucho más extensa, debe proporcionar reglas para todo concepto que pueda desempeñar un papel en el procesamiento inferencial de proposiciones, incluyendo muchas que no son de interés particular para los lógicos. Por otra parte, tiene que, en ciertas maneras, ser más restrictiva. Por ejemplo, las implicaciones lógicas estándares de una única proposición  $P$  incluyen muchas proposiciones que, de hecho, jamás se derivarían durante la comprensión de una emisión que expresase  $P$ :

- 8) a.  $P \& P$   
 b.  $P \vee P$   
 c.  $\neg P$   
 d.  $P \vee Q$   
 e.  $\neg P \rightarrow Q$   
 f.  $Q \rightarrow P$ .

En general, es simplemente falso que un oyente, dada una emisión que expresa  $P$ , podría pensar que cualquiera de las implicaciones de 8) era parte del mensaje que intentaba transmitir el hablante, y cualquier explicación inferencial de la comprensión debe incluir algún método para excluirlas.

La necesidad de alguna restricción está compuesta por el hecho de que las reglas que dan lugar a 8) pueden volverse a aplicar a sus propios productos, proporcionando conjuntos infinitos de implicaciones lógicas de acuerdo con las líneas siguientes:

- 9) a.  $(P \& P) \& P \dots$   
 b.  $(P \vee P) \vee P \dots$   
 c.  $\dots \neg \neg P$   
 d.  $(P \vee Q) \vee R \dots$   
 e.  $\dots \neg(\neg P \rightarrow Q) \rightarrow R \dots$   
 f.  $\dots R \rightarrow (Q \rightarrow P)$ .

El conjunto total de implicaciones lógicas de una proposición no podría entonces en principio extraerse de ninguna manera del tipo paso-a-paso durante la cantidad de tiempo finita que ocupa el procesar una proposición.

Las inferencias de 8) y 9) son, en algún sentido, triviales. Las reglas que dan lugar a ellas son bastante fáciles de caracterizar: son reglas que pueden aplicarse a cualquier proposición, independientemente de su forma o contenido, y que pueden después volverse a aplicar indefinidamente a sus propios resultados. Hay dos reacciones posibles al problema de las inferencias triviales y de las reglas que dan lugar a ellas. La mayor parte de las teorías de la comprensión que consideran el problema en su conjunto retienen algo equivalente a la lógica estándar, pero intentan restringir su funcionamiento de alguna manera. Se tiene la sensación de que en ciertas circunstancias sería necesaria una regla trivial, y la maniobra que eliminase enteramente todas las reglas triviales sería demasiado drástica<sup>6</sup>. Por el contrario, tenemos la sensación de que el enfoque más drástico es correcto.

Nuestra hipótesis es que el sistema deductivo usado en el procesamiento de información espontáneo es un sistema puramente interpretativo, en el sentido de que cada una de sus reglas exige, para su aplicación, la presencia de un concepto particular en la proposición o proposiciones que están siendo procesadas. En otras palabras, cada regla es esencialmente una regla de eliminación. Tal sistema podría contener algo parecido a la regla estándar de eliminación de *y*, que se aplica solamente a proposiciones en las que *y* está presente, pero podría no tener equivalente en la regla estándar de introducción de *y*, que no impone condiciones sobre la forma o contenido de las proposiciones a las que se aplica. Tal sistema, aunque podría contener reglas de eliminación para un amplio rango de conceptos no tratados en todas las lógicas estándares, podría no contener regla alguna que permitiese la derivación de cualquiera de las proposiciones de 8) y 9) a partir de la única premisa *P*.

En un sistema de este tipo, el problema de las inferencias triviales no se plantearía. En particular, puesto que no habría ninguna regla de introducción de *y*, no sería posible, dada una proposición *P* y un contexto arbitrario *Q*, derivar la implicación trivial *P* & *Q* como una implicación contextual de *P* en el contexto *Q*. La eliminación del problema de la inferencia trivial proporciona entonces automáticamente una solución al primero de los problemas planteados por la explicación propuesta de la relevancia<sup>7</sup>.

<sup>6</sup> Véase, por ejemplo, Johnson-Laird, 1975.

<sup>7</sup> Éste es solamente uno de los problemas que han de ser tratados por los lógicos de la relevancia. El sistema que contemplamos tiene alguna similitud con el sistema de Parry de «implicaciones analíticas», tal como es bosquejado por Anderson y Bel-



El procesamiento inferencial de una proposición podría ahora concebirse como siendo llevado a cabo por un autómata que, dada una proposición, un contexto finito, y un conjunto de reglas de inferencia no triviales como *input*, derivase el conjunto completo y finito de implicaciones lógicas no triviales de la unión de la proposición con el contexto como *output*. Estas incluirían las implicaciones contextuales de la proposición en el contexto. Uno podría comparar entonces las implicaciones contextuales que una proposición dada tuviese en diferentes contextos, o que tuviesen diferentes proposiciones en el mismo contexto. También se podría comparar la *cantidad* de procesamiento que diferentes proposiciones requerirían en un contexto dado, o que una proposición dada requiriese en diferentes contextos, donde la cantidad de procesamiento es el número de pasos por los que algún autómata tendría que pasar para estar seguro de derivar todas las implicaciones contextuales de una proposición en un contexto. Dejando de lado los casos especiales y los asuntos técnicos, la cantidad de procesamiento está determinada por, de un lado, el número de implicaciones lógicas no triviales de la proposición que está siendo procesada y, de otro, por el número de implicaciones lógicas no triviales del contexto. El autómata efectivo más simple examinaría meramente cada uno de los miembros del producto cartesiano de esos dos conjuntos para comprobar si se aplica alguna regla de inferencia. Dicho aproximadamente: a mayor complejidad semántica de la proposición que está siendo procesada y cuanto más extenso sea el contexto, mayor cantidad de procesamiento se requerirá<sup>8</sup>. Argumentaremos que los grados o comparaciones de relevancia se basan, por una parte, en valoraciones de números de implicaciones contextuales y, por otra, en la cantidad de procesamiento.

nap, 1975. Se sugiere algunas veces que una regla de introducción de  $y$  sería necesaria para derivar  $R$  en las circunstancias siguientes: el contexto contiene  $P$  y  $(P \& Q) \rightarrow R$ , y la nueva información que tiene que procesarse es  $Q$ . Sin embargo, no estamos afirmando que el sistema de eliminación de reglas más adecuado empíricamente contenga solamente aquellas que contemplan las lógicas estándares. Hay una buena razón para pensar que en las circunstancias anteriores, el oyente convertiría automáticamente, en un único paso,  $(P \& Q) \rightarrow R$  en  $P \rightarrow (Q \rightarrow R)$ , que se combinaría con  $P$  para dar  $Q \rightarrow R$  antes de que se presentase  $Q$ . Al presentársele  $Q$  el oyente podría entonces proceder directamente a la conclusión  $R$ . Para una discusión adicional, véase «Reply to Gazdar and Good», en Sperber y Wilson, 1892.

<sup>8</sup> Hay, desde luego, un tercer factor que afecta a la cantidad de procesamiento que cierta proposición requiere en un contexto dado, a saber: el número mismo de implicaciones contextuales: cada implicación contextual añade un paso a la derivación. Sin embargo, podemos ignorar este factor, puesto que no contribuye a una valoración de la cantidad de procesamiento independientemente de la valoración del número de implicaciones contextuales, que es tras lo que andamos.

#### 4. GRADOS DE RELEVANCIA

La máxima de Grice dice simplemente «Sé relevante». Si, como hemos sugerido, ser relevante es meramente un asunto de expresar una proposición que tiene al menos una implicación contextual en un contexto accesible al oyente, tal máxima difícilmente constreñiría al hablante y, por lo tanto, difícilmente guiaría al oyente. Sin embargo, si los grados de relevancia pudieran darse, o pudiera darse alguna base para comparaciones de relevancia, entonces podría proponerse alguna máxima mucho más restrictiva, como, por ejemplo, «Trata de ser tan relevante como sea posible en las circunstancias», y la interpretación tendría lugar bajo constricciones correspondientemente más severas. En esta sección proporcionaremos una base para comparar la relevancia de proposiciones diferentes en un contexto fijo, continuando en la sección siguiente con el tratamiento de la selección y valoración contextual en contextos variables. Los ejemplos que consideremos serán altamente irreales, en parte porque la comprensión de una emisión no tiene lugar en contextos fijos, y en parte porque los contextos efectivos es poco probable que sean tan pequeños como los usados aquí. Haremos también la suposición simplificadora de que los hablantes no sólo aspiran a la máxima relevancia, sino también que tienen éxito en sus aspiraciones, de modo que la tarea del oyente es meramente elegir la interpretación que resulta más relevante. En una explicación más completa esta suposición simplificadora podría eliminarse. Así pues, los ejemplos y la discusión deberían, pues, tomarse solamente como una ilustración de los criterios alrededor de los cuales podría construirse una explicación más completa.

Si la relevancia está ligada a las implicaciones contextuales, parece razonable que cuantas más implicaciones contextuales tenga una proposición en un contexto dado, más relevante es. Diremos que, siendo las demás cosas igual, la más relevante de dos proposiciones en un contexto dado será aquella con más implicaciones contextuales. Sin embargo, debe tomarse en cuenta un factor adicional. Dos proposiciones pueden tener el mismo número de implicaciones contextuales en un cierto contexto, pero una puede ser más compleja semánticamente que la otra, y contener trechos completos de información que no conectan con el contexto y que no hacen contribución alguna a la relevancia. Intuitivamente hablando: la existencia de esta extraña información resta atractivo a la relevancia de la proposición. Puesto que también incrementa la cantidad de procesamiento que se requiere, diremos, por razones que resultarán claras en la próxima sección, que, siendo las demás cosas iguales, la más relevante de dos proposiciones en un contexto dado es aquella que requiere menor procesamiento. En las valoraciones de relevancia hay, entonces, dos

factores que tomar en cuenta: de una parte, el número de implicaciones contextuales y, de otra, la cantidad de procesamiento que se necesita para obtenerlas.

Para ilustrar esto, tomemos el contexto de 10):

- 10) *a.* Los décimos cuestan 500 pesetas cada uno.
- b.* La gente puede comprar más de un décimo.
- c.* La persona que compró el décimo verde ganó el premio.
- d.* El premio es de 100.000.000 de pesetas.
- e.* Cualquiera que gane 100.000.000 de pesetas puede llevar a cabo los sueños de su vida.

Si un oyente tuviese sólo la información de 10), y un hablante supiese que todos los enunciados 11)-13) fuesen verdaderos, ¿cuáles de esas proposiciones elegiría un hablante que aspirase a la máxima relevancia?

- 11) Santiago compró el décimo verde.
- 12) Carlos compró el décimo azul.
- 13) Santiago compró el décimo verde y hoy es jueves.

11) y 12) requerirían aproximadamente la misma cantidad de procesamiento<sup>9</sup>, puesto que tienen estructuras lógicas paralelas y están siendo procesadas en el mismo contexto. Aquella que contiene más implicaciones contextuales es, por lo tanto, de esas dos proposiciones, la que será más relevante. Mientras que 12) tiene sólo la única implicación contextual 14), 11) tiene el conjunto de implicaciones contextuales 15)-18):

- 14) El décimo azul le costó a Carlos 500 pesetas.
- 15) El décimo verde le costó a Santiago 500 pesetas.
- 16) Santiago ganó el premio.
- 17) Santiago ganó 100.000.000 de pesetas.
- 18) Santiago puede llevar a cabo los sueños de su vida.

Por consiguiente 11) es más relevante que 12) en este contexto.

Al comparar 11) y 13), deben tomarse en cuenta los mismos dos factores: número de implicaciones contextuales y cantidad de procesamiento. En este caso, 11) y 13) tienen exactamente las mismas implicaciones contextuales, pero requieren cantidades diferentes de procesamiento. Puesto que 13) entraña 11), 13) requiere todos los pasos de procesamiento que requiere 11), e incluso más. Por los cri-

<sup>9</sup> Con la poco importante puntualización sugerida en la nota 8.

terios propuestos, es la proposición que requiere la menor cantidad de procesamiento —esto es, 11)—, la que es la más relevante en el contexto. De las tres proposiciones 11)-(13), es entonces 11) la que se selecciona como la más relevante por los criterios propuestos, y esto está de acuerdo con los juicios intuitivos de relevancia.

Estos criterios no proporcionan siempre, y no necesitan hacerlo, resultados tan claros y nítidos. Por ejemplo, no dicen cuál sería el resultado de comparar (12) y (13), que difieren ambos en la cantidad de procesamiento requerido y en sus números de implicaciones contextuales. Están diseñados para producir resultados claros y nítidos justamente en aquellos casos en los que los seres humanos pueden hacer juicios claros y nítidos de relevancia relativa. La suerte de comparaciones entre proposiciones distintas que un hablante necesita hacer durante la comprensión de una emisión son generalmente más semejante a las de (11) y (12), y (11) y (13) que a las de (12) y (13). Por ejemplo, un oyente puede tener que elegir entre varios candidatos para eliminar la ambigüedad y varias asignaciones de referencia para una emisión dada. En la mayor parte de los casos de asignación de referencia, y en muchos de eliminación de ambigüedad, las proposiciones que son candidatas a ser comparadas compartirán gran parte de sus estructuras lógicas, como sucede en (11) y (12) y pueden entrar en relaciones de entranamiento, como lo hacen (11) y (13). Sin embargo, en el caso normal, los candidatos alternativos diferirán tanto en los números de implicaciones contextuales, que las comparaciones serán completamente simples y sólo merecerá la pena considerar una única proposición. Cuando no surge claramente ningún candidato único, predecimos que el oyente será incapaz de decidir qué interpretación es la que se intenta <sup>10</sup>.

Hasta ahora hemos supuesto que los contextos son conjuntos arbitrarios de proposiciones, fijados de antemano. Al mirar los ejemplos (11)-(13) anteriores, el lector puede haberse sentido tentado a expandir el contexto (10) en una dirección u otra para incrementar la relevancia de la proposición que está siendo procesada: introduciendo, por ejemplo, suposiciones que se combinarían con la extraña información de (13) para permitir que se deriven implicaciones contextuales adicionales. En la vida real, los contextos no se fijan de antemano, sino que se eligen, al menos en parte, en función de la proposición que está siendo procesada. Vamos ahora a la cuestión de la selección, o construcción, del contexto.

---

<sup>10</sup> En una explicación más completa podría mostrarse que un hablante que no creyese que habría de surgir un candidato claro, o que causase que el oyente dudase entre interpretaciones alternativas, no habría observado el principio de máxima relevancia. Véase Sperber y Wilson, 1986, para discusiones adicionales.

## 5. SELECCIÓN DEL CONTEXTO

El problema de la selección del contexto no es algo único de nuestra armazón. Cualquier teoría adecuada de la comprensión tiene que describir el papel de las suposiciones de trasfondo en la interpretación de emisiones, y los principios mediante los cuales se seleccionan para desempeñar este papel. En nuestra armazón, el papel de las suposiciones de trasfondo es proporcionar premisas que se combinarán con la proposición que está siendo procesada para proporcionar implicaciones contextuales. La meta del procesamiento es, en general, maximizar en lo posible la relevancia de la proposición que está siendo procesada, esto es, obtener el máximo de implicaciones contextuales a título de devolución por cualquier cantidad dada de procesamiento. Queremos mostrar que, puesto que las variaciones en el contexto pueden incrementar o hacer decrecer la relevancia de la proposición que está siendo procesada, la meta de maximizar la relevancia puede guiar simultáneamente la elección del contexto.

El oyente de una emisión tiene a su disposición un conjunto de contextos potenciales a partir de los cuales puede elegir un contexto efectivo. Supongamos que hay un contexto pequeño, inmediatamente accesible, fijado de antemano, que consta de proposiciones que se han procesado muy recientemente, junto con sus implicaciones contextuales. Si el contexto inicial es (19) y la proposición expresada por la emisión es (20), se logra inmediatamente algún grado de relevancia:

- (19) Si la tasa de interés ha subido, la compañía quebrará.
- (20) La tasa de interés ha subido.

Puede, sin embargo, suceder que, a menos que el contexto inicial se extienda de alguna manera, no pueda lograrse ningún grado de relevancia. Esto sería así si, con el mismo contexto inicial, la proposición expresada no fuese (20), sino una de las (21)-(23):

- (21) La tasa de interés ha hecho lo que tú dijiste que haría.
- (22) La tasa de interés ha hecho lo que el sol hace cada mañana.
- (23) La tasa de interés ha hecho esto. [El hablante señala.]

Si la meta del procesamiento es maximizar la relevancia de la información que se está procesando, el oyente estará forzado a añadir al contexto inicial (19) información adicional, que puede recordarse de un discurso anterior (en el caso de (21)), recuperada de memoria enciclopédica (en el caso de (22)), o derivada de percepción sensorial (en el caso de (23)). La meta será encontrar, a partir de la fuente más

inmediatamente accesible, unas premisas que se combinarán con la proposición que está siendo procesada, para proporcionar el máximo de implicaciones contextuales a título de devolución por la cantidad disponible de procesamiento.

Como muestran los ejemplos, la accesibilidad de contextos potenciales puede alterarse por el contenido de la proposición que está siendo procesada, lo que puede dirigir la atención del oyente en un caso al entorno físico, en otro al discurso precedente, en otro a la memoria enciclopédica. Los ejemplos más complejos requieren extensiones simultáneas o secuenciales del contexto en una variedad de direcciones diferentes y no hay en principio un límite en el número de extensiones que pueden necesitarse para establecer la relevancia de una proposición dada. Cada extensión proporciona un nuevo contexto potencial en que podría procesarse la proposición, hasta un contexto de todos los contenidos de la memoria accesible del oyente.

Si hubiese sólo una dirección de extensión, el conjunto de contextos potenciales para procesar una proposición dada estaría estrictamente ordenado desde el contexto menor, el contexto inicial más fácilmente accesible, pasando por las extensiones mayores, menos accesibles, hasta el contexto que incluyese más, que fuese el menos accesible y que contuviese todo lo que el oyente sabe. A causa de la variedad de las direcciones posibles de extensión, la situación efectiva es más bien compleja; sin embargo, hay una ordenación parcial de contextos potenciales, con una serie de contextos que incluyen cada vez más, extendiéndose a partir del contexto inicial en varias direcciones.

Si el conjunto de contextos potenciales está estructurado de esta manera, toda extensión del contexto inicial tendría un costo doble en términos de procesamiento. En primer lugar, cada extensión incrementa el tamaño del contexto, y cuanto mayor es el contexto, mayor es la cantidad de procesamiento. En segundo lugar, las extensiones adicionales tienen un más difícil acceso, y el acceso es presumiblemente un proceso paso-a-paso con costes asociados en el procesamiento. Por consiguiente, toda extensión del contexto incrementa el coste del procesamiento y, tiene que esperarse, permaneciendo las demás cosas igual, hace decrecer la relevancia de la proposición que está siendo procesada.

Diremos que una proposición es relevante en un conjunto de contextos  $K_1 \dots K_n$  si y sólo si es relevante en al menos un contexto que es un miembro de ese conjunto, y que es más relevante en  $K_1$  que en  $K_2$  si, permaneciendo las demás cosas igual, la cantidad de procesamiento requerida para obtener las implicaciones que tiene en  $K_1$  es menor que la cantidad requerida para obtener las implicaciones que tiene en  $K_2$ . El maximizar la relevancia de una proposición dada es entonces un asunto de elegir un contexto que maximice sus implicaciones con-

textuales y mininice la cantidad de procesamiento; en otras palabras: es un asunto de procesamiento máximamente eficiente.

Este enfoque de la valoración de la relevancia en contextos variables puede ilustrarse con el siguiente (altamente simplificado) ejemplo. Supóngase que hay un contexto inicial (24C1), al que pueden añadirse (24C2), (24C3) y (24C4), en este orden:

- (24) C1: Juan ha elegido la fecha de la reunión.
- C2: Si la fecha de la reunión es el 1.º de febrero, el presidente no podrá asistir.
- C3: Si el presidente no puede asistir, la propuesta de Juan será aceptada.
- C4: Si las propuestas de Juan se aceptan, la compañía quebrará.

Un oyente que desee maximizar la relevancia de (25) o (26) en este conjunto de contextos debería querer solamente causar los costos adicionales de extender el contexto como contrapartida por un incremento en compensación en el número de las implicaciones contextuales:

- (25) La fecha de la reunión es el 1.º de febrero.
- (26) La fecha de la reunión es el 5 de febrero.

En el contexto inicial C1, (25) tiene la implicación contextual única (27) y (26) la implicación contextual única (28):

- (27) Juan ha elegido el 1.º de febrero como fecha de la reunión.
- (28) Juan ha elegido el 5 de febrero como fecha de la reunión.

En este contexto ambas proposiciones son relevantes y, es más, a causa de sus similitudes en estructura lógica, contexto e implicaciones contextuales, ambas serán igualmente relevantes.

Si el contexto inicial se extiende para incluir C2, (25) gana la implicación contextual suplementaria (29):

- (29) El presidente no podrá asistir.

La extensión merecería la pena si el coste de acceder a C2 no fuese tan grande que pesase más que la ganancia en implicaciones contextuales. Si el contexto se extiende adicionalmente para incluir C3, y a continuación C4, las implicaciones contextuales adicionales (30) y (31) se obtendrían a su vez:

- (30) Las propuestas de Juan serán aceptadas.
- (31) La compañía quebrará.

De nuevo, si los costos del acceso no fuesen demasiado altos, esas extensiones merecerían ambas la pena. Así pues, alguien que esté procesando (25) en el conjunto de contextos C1-C4 podría tener una buena razón para continuar expandiendo el contexto hasta su tamaño máximo, puesto que toda expansión incrementaría la relevancia de (25).

En el caso de (26), los resultados de alguna extensión más allá del contexto inicial C1 serían completamente diferentes. Cada extensión origina un costo adicional en el procesamiento, pero sin ningún incremento compensatorio en las implicaciones contextuales. Sea el que sea el contexto que se elija, solamente se obtiene la implicación contextual (28). Así pues, alguien que esté procesando (26) en el conjunto de contextos C1-C4 no tendrá razón alguna para extender el contexto más allá del estado inicial C1, puesto que toda expansión haría decrecer la relevancia de (26).

Dentro de esta armazón inferencial, es entonces posible ver en principio cómo podría tener lugar la selección de contexto y, más específicamente, cómo podría ayudar una proposición a determinar su propio contexto, sujeto a restricciones de accesibilidad y relevancia. En una explicación más realista debería prestarse una atención más seria a la elección de instrucciones particulares para la extensión del contexto. Por ejemplo, el oyente de (26) no abandonaría probablemente el contexto C1, pero buscaría la relevancia en otra dirección. Similarmente, debería haberse prestado más atención a la cuestión de cómo los costes de procesamiento se equilibran con números de implicaciones contextuales en las situaciones de la vida real. Por ejemplo, ¿en qué punto debería decidir el oyente de (25) que el coste de nuevas extensiones del contexto se ha convertido en demasiado alto? Éstas son cuestiones empíricas, que no hemos tratado de resolver directamente y para las cuales creemos que, sin embargo, habrá respuestas diferentes para diferentes personas en diferentes circunstancias.

La relación de nuestro trabajo con estas y otras cuestiones empíricas en la teoría de la comprensión es indirecta. No creemos, por ejemplo, que un oyente, al eliminar la ambigüedad de una emisión, compare efectivamente todas las interpretaciones posibles y las ordene por lo que respecta a la relevancia antes de decidir lo que se ha dicho. Lo que hemos intentado describir no son los procedimientos usados para eliminar la ambigüedad, sino la meta para cuyo logro están diseñados esos procedimientos, la propiedad para cuyo diagnóstico están diseñados. Nuestra afirmación es que la interpretación que el hablante intenta que se haga y la que el oyente debe elegir, es aquella que satisface un principio de relevancia máxima. Conociendo la meta, resultaría más fácil describir los procedimientos, y es de este modo cómo nuestro trabajo puede contribuir al trabajo descriptivo sobre la comprensión.



Las suposiciones cruciales de nuestra armazón son, primero, que la comprensión es en gran medida un proceso inferencial, siendo esencialmente el papel del contexto el proporcionar premisas para el cálculo de las implicaciones contextuales y, en segundo lugar, que las implicaciones contextuales de una proposición en un contexto finito son ellas mismas finitas. Con esas suposiciones es posible ver cómo podría tener lugar en principio la selección del contexto; sin ellas la armazón se derrumbaría. Lo que es en algunos aspectos menos importante es nuestra elección de la terminología. Para nosotros «relevancia» es un término técnico, diseñado para desempeñar un papel en una teoría pragmática general más bien que para aproximarse al uso cotidiano. Creemos que un rango bastante substancial de intuiciones ordinarias queda capturado dentro de esta teoría y, por lo tanto, que esta elección de terminología es completamente apropiada. Sin embargo, la terminología podría cuestionarse sin que la armazón general resultase afectada en absoluto.

## 6. CONCLUSIÓN

Grice considera la práctica de observar las máximas de conversación como no meramente «algo que todos o la mayor parte seguimos DE HECHO, sino como algo que resulta RAZONABLE que lo hagamos, algo que NO DEBERÍAMOS abandonar» (Grice 1975, p. 48). Él no está seguro de por qué esto debe de ser así, y sugiere que una respuesta tiene que esperar una clarificación «de la naturaleza de la relevancia y de las circunstancias en las que se requiere» (p.49). La respuesta sugerida por nuestro intento de clarificación es la siguiente.

En términos cognitivos, lo que los seres humanos buscan en la información que procesan es relevancia. El procesar incluye esfuerzo y sólo se emprenderá cuando se espera alguna recompensa en términos de implicaciones contextuales; cuanto mayor sea la recompensa que se espera, tanto mayor será el esfuerzo que uno está preparado para emprender. Las expectativas de relevancia no son constantes a través de todos los individuos y circunstancias, y un hablante o escritor que frustra constantemente las expectativas de un auditorio causará un reajuste a la baja de sus expectativas subsiguientes, y puede perder enteramente la atención de ese auditorio. Si alguien quiere que se lea su obra, es interés suyo el aspirar a un nivel de relevancia tan alto como sea posible, creando expectativas que inducirán al lector a continuar leyendo la obra con la atención que el autor piensa que merece. Dados los hechos brutos de la psicología cognitiva, esto es conducta racional, o al menos razonable, aunque no particularmente altruista o virtuosa.

Se sigue que la información que ha sido comunicada deliberada-

mente, al igual que otra información proveniente de distintas fuentes, llega con una garantía de que se ha intentado, sino logrado, un cierto nivel de relevancia. Es esta garantía la que constituye la diferencia entre emitir meramente oraciones delante de alguien y decir algo a alguien, lo que justifica la eliminación de la ambigüedad del significado y el cálculo de las implicaturas, lo que define un campo separado: el de la pragmática. La pragmática así concebida es sobre el modo en que se le hace justicia a la búsqueda universal de la relevancia a la vez que se la explota en la comunicación verbal, y el estudio de la relevancia es su piedra fundacional <sup>11</sup>.

## REFERENCIAS

- ANDERSON, A. R., y BELNAP, N. D.: *Entailment*, Princeton University Press, 1975.
- BERLIN, I.: «Verification», en *Proceedings of the Aristotelian Society*, 1938-1939.
- DASCAL, M.: «Conversational Relevance», en A. Margalit (ed.), *Meaning and Use*, D. Reidel, 1979.
- VAN DIJK, T.: «Relevance Assignment in Discourse Comprehension», en *Discourse Processes*, 2 (1979), pp. 113-126.
- GAZDAR, G.: *Pragmatics: Implicature, Presupposition and Logical Form*, Academic Press, 1979.
- GOODMAN, N.: «About», en *Mind*, LXX, 277 (1961), pp. 1-24.
- GRICE, H. P.: «Logic and Conversation», en P. Cole and J. Morgan (eds), *Syntax and Semantics 3: Speech Acts*, Academic Press, 1975. [Versión castellana: pp. 511-530 de este volumen.]
- HAACK, S.: *Philosophy of Logics*, Cambridge University Press, 1979. [Versión castellana: Cátedra, Madrid, 1982.]
- JOHNSON-LAIRD P. N.: «Models of Deduction», en R. J. Falmagne (ed.), *Representation and Process*, Lawrence Erlbaum, Hillsdale, NJ, 1975.
- SPERBER, D., y WILSON, D.: «Mutual Knowledge and Relevance in Theories of Comprehension», en N. V. Smith (ed.) *Mutual Knowledge*, Academic Press, 1982, pp. 61-131.
- *Relevance: Communication and Cognition*, Blackwell, 1986.
- SUPPES, P.: *Introduction to Logic*, D. van Nostrand, 1957.
- THOMASON, R.: «Semantics, Pragmatics, Conversation and Presupposition» mimeo inédito, Philosophy Dept. University of Pittsburgh, 1973.
- WERTH, P. N.: «The Concept of Relevance in Conversational Analysis», en P.N. Werth (ed), *Conversation and Discourse*, Croom Helm, 1981, pp. 129-154.
- WILSON, D., y SPERBER D.: «On Grice's Theory of Conversation», en P.N. Werth (ed.), *Conversation and Discourse*, Croom Helm, 1981, pp. 155-178.

<sup>11</sup> Para una discusión de la contribución de Grice a la pragmática, véanse Wilson y Sperber, 1981; para discusiones adicionales de la armazón descrita en este artículo, véase Sperber y Wilson, 1982; para el intento de un bosquejo de una teoría pragmática, véase Sperber y Wilson, 1986.

## ORIGEN DE LOS ARTÍCULOS

— «Sobre sentido y referencia» fue publicado originalmente en la *Zeitschrift für Philosophie und philosophische Kritik*, nueva serie, núm. 100, 1982, pp. 25-50. La presente traducción castellana, reproducida aquí con permiso de Editorial Ariel, apareció en Gottlob Frege, *Estudios sobre Semántica*, Barcelona, 1971.

— «Descripciones» apareció por vez primera en B. Russell, *Introduction to Mathematical Philosophy*, George Allen & Unwin, Londres, 1919. Se reproduce aquí con permiso de la editorial Allen & Unwin.

— «Sobre el referir» se publicó en *Mind*, vol. LIX, 1950. Aquí se reproduce la versión castellana recogida en P.F. Strawson, *Ensayos lógico-lingüísticos*, Tecnos, Madrid, 1983.

— «Nombres propios y descripciones» se recogió en P. Edwards (ed.), *The Encyclopedia of Philosophy*, MacMillan, Nueva York, 1967. Se reproduce aquí con permiso del autor.

— «Identidad y necesidad» apareció originalmente en la recopilación de M. K. Munitz, *Identity and Individuation*, New York University Press, Nueva York, 1971. Los derechos de traducción al castellano fueron cedidos a *Crítica*, que ha tenido la gentileza de autorizar su reproducción aquí.

— «El significado de "significado"» fue publicado por vez primera en K. Gunderson (ed.), *Lenguaje, Mind and Knowledge, Minnesota Studies in the Philosophy of Science*, VII, University of Minnesota Press, 1975. La versión castellana aquí recogida apareció en *Teorema*, vol. XIV/3-4, y se reproduce con permiso del editor.

— «Problemas y cambios en el criterio empirista de significado» apareció originalmente en la *Revue Internationale de Philosophie*, vol. 11, 1950, que ha concedido el correspondiente permiso para su inclusión en este libro. La traducción se ha tomado de la *Antología semántica* de Mario Bunge (ed.), Nueva Visión, Buenos Aires, 1960. Agradecemos a dicha editorial el permiso que nos ha otorgado para su reproducción.

— «Dos dogmas del empirismo» se publicó por vez primera en W.v.O. Quine, *From a Logical Point of View*, Harvard University Press, 1953. La traducción aquí recogida es la publicada por Editorial Ariel de Barcelona en 1962, a la que agradecemos su amable permiso. El autor y Harvard University Press han concedido los correspondientes permisos para la inclusión del artículo citado en este libro.

— «Significado y traducción» apareció en Brower (ed.), *On Translation*, Harvard University Press, 1959. El autor y Harvard University Press han autorizado su traducción e inclusión aquí.

— «La concepción semántica de la verdad y los fundamentos de la semántica» fue publicado en *Philosophy and Phenomenological Research*, 4, 1944, pp. 341-375. Los editores de dicha revista han concedido el pertinente permiso de publicación. La tra-

ducción que aquí aparece se toma de M. Bunge (ed.), *Antología semántica*, Nueva Visión, Buenos Aires, 1960. Agradecemos a esta editorial su amable permiso para reproducir la traducción en el presente volumen.

— «Verdad y significado» apareció por vez primera en *Synthèse*, 17 (3), 1967. Se incluye y traduce aquí con permiso de la editorial Reidel. En 1984 se recogió en Donald Davidson, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, pp. 17-36. El texto es el mismo que el de 1967, con el solo añadido de unas breves notas.

— «Significado y verdad» es una lección inaugural pronunciada en la Universidad de Oxford en 1969. Aquí recogemos la traducción publicada en P.F. Strawson, *Ensayos lógico-lingüísticos*, Tecnos, Madrid, 1983.

— «Interpretación radical» apareció en *Dialéctica*, 27 (1973), pp. 313-328, y se publica aquí con permiso del editor de dicha revista. En 1984 se recogió en Donald Davidson, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, El texto es el mismo de 1973, con el solo añadido de unas breves notas.

— «¿Qué es una teoría del significado?» fue publicado por vez primera en Samuel Guttenplan (ed.), *Mind and Language*, Clarendon Press, Oxford, 1975. Se reproduce aquí con el correspondiente permiso de Oxford University Press.

— «Emisiones realizativas» es una transcripción, con correcciones verbales menores, de una charla no escrita emitida por el tercer programa de la BBC. Esta charla se incluyó como parte de J. L. Austin, *Philosophical Papers*, Oxford University Press, 1961. La traducción reproducida aquí, con el correspondiente permiso, es la recogida en la versión castellana de dicha obra publicada por la *Revista de Occidente* en 1975.

— «¿Qué es un acto de habla?» fue originalmente publicada en Max Black (ed.), *Philosophy in América*, Allen & Unwin, 1965. El autor ha concedido el correspondiente permiso para su inclusión en este volumen. La traducción recogida apareció en *Cuadernos Teorema*, Valencia, 1977. La revista *Teorema* ha dado la correspondiente autorización.

— «Una taxonomía de los actos ilocucionarios» apareció originalmente en K. Gunderson (ed.), *Language, Mind and Knowledge*, University of Minnesota Press, 1975. El autor ha concedido su permiso para que este artículo se incluya aquí. La versión castellana apareció en *Teorema*, 1976. Los editores de dicha revista han concedido la correspondiente autorización.

— «Las intenciones y el significado del hablante» se publicó en *The Philosophical Review*, vol 78, 1969, pp. 147-177. Se reproduce aquí con el permiso de los editores de *The Philosophical Review*.

— «Lógica y conversación» apareció por vez primera en P. Cole y J.L. Morgan (eds.) *Syntax and Semantics. Speech Acts*, vol.3, Academic Press, 1975. Se reproduce aquí con permiso del autor.

— «Situaciones y actitudes» se publicó originalmente en *The Journal of Philosophy*, vol. 78 (1981). Se incluye con permiso de los autores y del editor de *The Journal of Philosophy*.

— «Semántica del rol conceptual» apareció en el *Notre Dame Journal of Formal Logic*, vol 23, 2, 1982. Se incluye aquí con permiso del editor del *Notre Dame Journal of Formal Logic*.

— «Sobre la definición de relevancia» forma parte de un libro colectivo en homenaje al Profesor H.P. Grice, *Philosophical Grounds of Rationality*, editado por Richard E. Grandy y Richard Warner, Clarendon Press, Oxford, 1986. Oxford University Press ha concedido la correspondiente autorización para incluir aquí este artículo.



## COLECCIÓN FILOSOFÍA Y ENSAYO

Dirigida por Manuel Garrido

- Austin, J. L.: *Sentido y percepción*.
- Boden, M. A.: *Inteligencia artificial y hombre natural*.
- Bottomore, T.; Harris, L.; Kiernan, V. G.; Miliband, R.; con la colaboración de Kolakowski, L.: *Diccionario del pensamiento marxista*.
- Brown, H. I.: *La nueva filosofía de la ciencia* (2.<sup>a</sup> ed.).
- Bunge, M.: *El problema mente-cerebro* (2.<sup>a</sup> ed.).
- Chisholm, R. M.: *Teoría del conocimiento*.
- Dampier, W. C.: *Historia de la ciencia y sus relaciones con la filosofía y la religión*.
- Díaz, E.: *Revisión de Unamuno*. Análisis crítico de su pensamiento político.
- Eccles, J. C.: *La psique humana*.
- Edelman, B.: *La práctica ideología del Derecho*.
- Fann, K. T.: *El concepto de filosofía en Wittgenstein*.
- Ferrater Mora, J., y otros: *Filosofía y ciencia en el pensamiento español contemporáneo (1960-1970)*.
- Feyerabend, P.: *Tratado contra el método*.
- García Suárez, A.: *La lógica de la experiencia*.
- Garrido, M.: *Lógica simbólica*.
- Gómez García, P.: *La antropología estructural de Claude Lévi-Strauss*.
- Habermas, J.: *La lógica de las ciencias sociales* (2.<sup>a</sup> ed.).
- Habermas, J.: *Teoría y praxis* (2.<sup>a</sup> ed.).
- Hierro, J. S.-P.: *Problemas del análisis del lenguaje moral*.
- Hintikka, J.: *Lógica, juegos de lenguaje e información*.
- Lakatos, I., y otros: *Historia de la ciencia y sus reconstrucciones racionales* (2.<sup>a</sup> ed.).
- Lindsay, P. H., y Norman, D. A.: *Introducción a la psicología cognitiva* (2.<sup>a</sup> ed.).
- Lorenzo, J. de: *El método axiomático y sus creencias*.
- Lorenzo, J. de: *Introducción al estilo matemático*.
- Mates, B.: *Lógica matemática elemental*.
- McCarthy, Th.: *La teoría crítica de Jürgen Habermas*.
- Popper, K. R.: *Búsqueda sin término*. Una autobiografía intelectual.
- Popper, K. R.: *Realismo y el objetivo de la ciencia*. Post Scriptum a la lógica de la investigación científica, vol. I.
- Popper, K. R.: *El universo abierto*. Post Scriptum a la lógica de la investigación científica, vol. II.
- Popper, K. R.: *Teoría cuántica y el cisma en física*. Post Scriptum a la lógica de la investigación científica, vol. III.

Putnam, H.: *Razón, verdad e historia*.  
Quine, W. V.: *la relatividad ontológica y otros ensayos*.  
Rescher, N.: *La primacía en la práctica*.  
Rivadulla, S.: *Filosofía actual de la ciencia*.  
Robinet, A.: *Mitología, filosofía y cibernética. El autómeta y el pensamiento*.  
Rodríguez Paniagua, J. M.<sup>a</sup>: *¿Derecho natural o axiología jurídica?*  
Sahakian, W. S.: *Historia y sistemas de la psicología*.  
Strawson, P. F.: *Ensayos lógico-lingüísticos*.  
Valdés Villanueva, L. M. (ed.): *La búsqueda del significado. Lecturas de filosofía del lenguaje*.  
Vargas Machuca, R.: *El poder moral de la razón. La filosofía de Gramsci*.  
Veldman, D. J.: *Programación de computadoras en ciencias de la conducta*.  
Villacañas, J. L.: *Racionalidad crítica. Introducción a la filosofía de Kant*.  
Wellman, C.: *Morales y éticas*.



Colección  
CUADERNOS DE FILOSOFÍA Y ENSAYO

Director: MANUEL GARRIDO

- Javier Aracil: *Máquinas, sistemas y modelos*. Un ensayo sobre sistémica.
- José Luis L. Aranguren: *Propuestas morales* (3.<sup>a</sup> ed.).
- Y. Bar-Hillel y otros: *El pensamiento científico*.
- Mario Bunge: *Controversias en física*.
- Mario Bunge: *Economía y filosofía* (2.<sup>a</sup> ed.).
- Mario Bunge: *Intuición y razón*.
- J. N. Crossley y otros: *¿Qué es la lógica matemática?*
- Manuel Cruz: *Del pensar y sus objetos*. Sobre filosofía y filosofía contemporánea.
- Charles Darwin: *Ensayos sobre el instinto*.
- Félix Duque: *Filosofía de la técnica de la naturaleza*.
- Javier Esquivel y otros: *La polémica del materialismo*.
- Andrew Feenberg: *Más allá de la supervivencia: el debate ecológico*.
- Paul Feyerabend: *Adiós a la razón*.
- Paul Feyerabend: *¿Por qué no Platón?*
- Gottlob Frege: *Investigaciones lógicas*.
- Sigmund Freud: *Compendio del psicoanálisis*.
- Sigmund Freud: *Los sueños*.
- Manuel Garrido (ed.) y otros: *Lógica y lenguaje*.
- Jürgen Habermas: *Ciencia y técnica como «ideología»*.
- Jürgen Habermas: *Sobre Nietzsche y otros ensayos*.
- Jürgen Habermas: *Identidades nacionales y postnacionales*.
- Hans Hermes: *Introducción a la teoría de la computabilidad*.
- José Jiménez: *La estética como utopía antropológica*. Bloch y Marcuse.
- Leszek Kolakowski: *Si Dios no existe...* Sobre Dios, el diablo, el pecado y otras preocupaciones de la llamada filosofía de la religión.
- L. Kolakowski: *«Horror metaphysicus»*.
- Ramiro Ledesma Ramos: *La filosofía, disciplina imperial*.
- Benson Mates: *Lógica de los estoicos*.
- H. O. Mounce: *Introducción al «Tractatus» de Wittgenstein*.
- Carlos P. Otero: *La revolución de Chomsky: ciencia y sociedad*.
- Karl R. Popper: *Sociedad abierta, universo abierto*.
- José Sanmartín: *Una introducción constructiva a la teoría de modelos* (2.<sup>a</sup> ed.).
- A. M. Turing, H. Putnam y D. Davidson: *Mentes y máquinas*.
- A. N. Whitehead: *La función de la razón*.