

FREGE *

M. Dummett

All Souls College, Oxford

FREGE FUE EL FUNDADOR tanto de la lógica contemporánea como de la contemporánea filosofía del lenguaje. El uso de esta última frase, en conexión con él, tiene un extraño sonido, ya que él siempre expresó desprecio por el lenguaje; lo hizo así porque, para él, 'lenguaje' quería decir 'lenguaje natural', y creía que los lenguajes naturales eran instrumentos muy defectuosos para la expresión del pensamiento. En las oraciones del lenguaje natural, no solamente las apariencias superficiales son groseramente engañosas, sino que además, según Frege, los lenguajes naturales son incoherentes en el sentido de que no se podría construir una explicación sistemática completa del uso de las oraciones de tales lenguajes. En nuestro discurso cotidiano, no estamos meramente jugando a un juego cuyas reglas no hemos formulado por completo, sino a uno para el que no se podría extraer ningún conjunto consistente de reglas. Tal observación difiere de lo que otros filósofos han dicho acerca del lenguaje natural solamente en su fuerza: Frege se diferenció de sus predecesores en el remedio metodológico que adoptó. Otros han pensado que la tarea del filósofo es desnudar el pensamiento de su vestido lingüístico, introducirse en todas las formas de meras expresiones para llegar al puro pensamiento que las subyace: Frege fue el primero que dio la debida importancia al hecho de que no podemos tener un pensamiento que no expresamos, si no a nosotros mismos, a otros. Cualquier

*Si se le
con la
"Marquesado
Newman
↓*

* Comunicación entregada en inglés en el Departamento de Lógica de la Universidad de Valencia en 20 de febrero de 1975. La presente versión castellana ha sido realizada por Alfonso García Suárez y María Teresa Begueristain.

intento de escudriñar nuestros pensamientos tomados aparte de su expresión, acabará por lo tanto confundiendo la experiencia interna del pensar, o los simples acompañamientos contingentes mentales del pensar, con los pensamientos mismos. El pensamiento difiere de otras cosas de las que también se ha dicho que son objetos mentales, por ejemplo los dolores o las imágenes mentales, en que no es esencialmente privado. Yo puedo decirte cómo es mi dolor, o qué estoy visualizando, pero no puedo transferirte mi dolor o mi imagen mental. Es esencial al pensamiento, no obstante, el que sea transferible, el que yo pueda transmitirte exactamente lo que estoy pensando: como dijo Wittgenstein, en un pasaje en el que critica esta idea de Frege, por decirlo así, tomas el pensamiento en tu mente; yo hago algo más que decirte como es mi pensamiento —yo te comunico ese mismo pensamiento. Por lo tanto cualquier intento de investigar los pensamientos que culmine en un estudio de lo que en esencia es privado, esto es, de experiencias mentales internas, debe haber errado el tiro.

El remedio que Frege adoptó para el carácter insatisfactorio y engañoso de nuestros lenguajes cotidianos fue, por lo tanto, reemplazarlos, para el propósito de estudios filosóficos, por formas lingüísticas carentes de tales defectos. Las oraciones de un tal lenguaje idealmente construido expresarían pensamientos en virtud de los principios que gobiernan el uso de sus palabras constituyentes, principios susceptibles de formulación sistemática. Porque el pensamiento es comunicable sin residuo y porque lo que se comunica depende solamente de la común aprehensión de los principios que gobiernan el lenguaje, tales principios deben estar relacionados solamente con el uso efectivo de oraciones del lenguaje, esto es, con lo que en el uso del lenguaje está abierto a observación sin ayuda de ningún otro supuesto contacto entre mente y mente que vía el medio del lenguaje. Se sigue que, para estudiar el pensamiento, nuestra tarea no es describir algo mental en el sentido de algo que cae fuera de los límites del mundo físico; es, más bien, primero construir mejores medios para la expresión del pensamiento, y, segundo, formular exacta y explícitamente los principios que gobiernan

el empleo de tales formas de expresión —principios que generalmente dejamos meramente implícitos, y en el caso del lenguaje natural no hubieran podido en absoluto formularse coherentemente.

La primera idea de Frege fue que las oraciones juegan un papel primario en la teoría del significado. La oración es el complejo lingüístico más pequeño que uno puede utilizar para *decir* algo: por tanto el significado de una palabra ha de darse en términos de su contribución a la determinación de lo que puede decirse por medio de una oración que la contiene.

Su segunda idea es que la noción de *verdad* juega un papel crucial en la explicación que hemos de dar del significado de una oración, y por tanto, según el primer principio, del significado de cualquier expresión. Para entender una emisión asertórica, hemos de saber qué es aseverar algo, y cuál es el contenido de esta oración particular, esto es, qué condiciones debe cumplir lo dicho para que sea verdadero. Asimismo, para entender una emisión interrogativa hemos de saber qué es formular una pregunta, y cuál es el contenido de esta pregunta particular, esto es (para una pregunta que requiere la contestación 'Sí' o 'No'), cual es la condición bajo la que debiera recibir la contestación 'Sí'. Podemos asimilar ambas extendiendo la noción de verdad para que cubra también oraciones interrogativas, esto es, considerándolas como verdaderas cuando la contestación es 'Sí': si lo hacemos, entonces, en ambos casos, el hablante emite una oración que es verdadera solamente en el caso en el que se dé una cierta condición; la diferencia consiste totalmente en qué más está haciendo el hablante, esto es, en el primer caso, aseverando que la condición se cumple, y, en el segundo, preguntando si se cumple. Frege se refiere a esta diferencia como a una diferencia en la fuerza asignada a la oración en ambos casos (asertórica o interrogativa). A menudo alguna característica lingüística de la oración servirá para indicar la fuerza que se le asigne, y, si fracasamos en entender la convención que gobierna esta indicación de fuerza, no comprenderemos el significado de la oración. La fuerza es así un aspecto del significado, a distinguir de aquel ingrediente

FUERZA



de los significados de las palabras que va a determinar la condición bajo la que la oración es verdadera. A este último ingrediente Frege lo llama el sentido, y es de éste del que voy a hablar.

La noción de verdad es también, claramente, central para la lógica en el sentido tradicional de ciencia de la inferencia. La lógica comenzó con el descubrimiento de Aristóteles de que la validez de un argumento podía caracterizarse por ser una instancia de un esquema argumental válido, en donde un esquema argumental es como un argumento salvo que contiene letras esquemáticas en ciertos lugares en lugar de las palabras o expresiones reales, y es válido si cada instancia con premisas verdaderas tiene una conclusión verdadera, obteniéndose una instancia al reemplazar cada letra esquemática por una expresión real de la categoría lógica apropiada. Esta noción pre-semántica de la interpretación de un esquema por reemplazo era la única con la que la lógica podía operar hasta Frege. Frege nos proporcionó por primera vez una semántica, es decir, un análisis del modo en el que se determina una oración como verdadera o no de acuerdo con su composición a partir de sus palabras constituyentes. Para llegar a la semántica, debemos llevar a cabo una tarea preliminar —dar una explicación adecuada de la estructura sintáctica de nuestras oraciones, de cómo se componen de palabras. El doble logro de Frege, que le hace el fundador de la lógica moderna, está en haber sido el primero en ofrecer tal análisis sintáctico, y después establecer, en términos de él, una semántica. Para hablar más exactamente, no dio directamente ningún análisis sintáctico del lenguaje natural, pero, en su lugar, inventó un lenguaje formalizado cuyas oraciones tienen una estructura sintáctica precisa y que está a la vista, ya que su apariencia superficial la revela; y, habiendo hecho esto, nos proporcionó una semántica para este lenguaje. Es bastante general suponer que llegaremos a un análisis sintáctico satisfactorio del lenguaje natural solamente mostrando que sus oraciones tienen una estructura subyacente (o profunda) análoga a aquélla que tienen las oraciones del lenguaje formalizado de Frege, que, por supuesto, ha llegado a ser el tipo *standard* de lenguaje cuantificacional utilizado

por los lógicos. Frege no alegó tal cosa, suponiendo que el lenguaje natural se resiste a cualquier sistematización coherente completa; pero es en alguna versión de este alegato, al menos para una gran parte del lenguaje natural, en donde descansa el título de Frege de ser el fundador, no sólo de la lógica contemporánea, sino también de la filosofía del lenguaje contemporánea.

Al menos nadie ha puesto en duda esto: que un lenguaje que tenga la clase de sintaxis que Frege ideó para su lenguaje formalizado sea adecuado para formular cualquier teoría matemática o de ciencias naturales; las discusiones surgen solamente acerca de las formas de construcción oracional importantes para otros propósitos. La sintaxis de Frege está basada en la idea fundamental de que la construcción de una oración ocurre en dos etapas: primero construimos las formas más simples de oración, y luego, por medio de dispositivos reiterables de composición oracional, construimos a partir de ellas oraciones complejas. Las oraciones atómicas deben concebirse como compuestas de predicados simples, con un número fijo aunque arbitrario de lugares argumentales, y un número correspondiente de términos singulares (los términos singulares pueden ser complejos, pero no me interesaré por su estructura interna). Los modos de composición oracional son dobles: por medio de operadores oracionales, que pueden ser unarios, como la negación, o binarios, como las conectivas 'o' y 'si'; y por medio de cuantificación, el dispositivo utilizado en el lenguaje formalizado para expresar generalidad. La cuantificación, como generalmente se considera hoy, tiene dos formas principales, universal ('para todo x , ...') y existencial ('para algún x , ...'): la oración cuantificada se forma asignándole uno u otro cuantificador a un predicado complejo. Un predicado complejo es, esencialmente, el resultado de eliminar una o más ocurrencias de algún término singular de la oración; la oración cuantificada debe indicar dónde ocurren las lagunas resultantes en el predicado complejo. La oración de la que se formó el predicado complejo debe ser considerada como un constituyente de la oración cuantificada; pero, ya que no podemos determinar a partir de la oración cuantificada ningún cons-

tituyente único, tomaremos como constituyente de una oración cuantificada *cualquier* oración a partir de la cual podría haberse formado el predicado complejo, esto es, cualquier oración que resulte de poner uno y el mismo término singular en las lagunas que ocurren en el predicado complejo. (Aquí solamente he descrito la cuantificación de primer orden; no nos interesaremos por la cuantificación de orden superior.)

Fue al planear esta sintaxis, tan familiar en lógica contemporánea, y, en particular, al construir la notación de los cuantificadores, como se reveló principalmente el genio de Frege. Una vez que se tiene un lenguaje cuya sintaxis es de esta forma, no es tan difícil continuar proporcionándole una semántica. No pretendo que la semántica de Frege, conocida ahora como semántica clásica o bivalente, sea indisputable; por el contrario, algunos lógicos, por ejemplo los intuicionistas, la rechazan: sólo que no es tan difícil concebirla una vez que se tiene la sintaxis. La primera idea fundamental es que la condición para que una oración compleja sea verdadera depende solamente de su composición a partir de oraciones atómicas; aquí es esencial la concepción bajo la cual cada instancia de una oración cuantificada universal o existencialmente ha de tomarse como un constituyente suyo. *Esta* idea no es seriamente puesta en duda, al menos por cualquiera que acepte una sintaxis fregeana. Definamos ahora el *valor semántico* de una oración atómica como cualquier rasgo suyo que es a la vez necesario y suficiente que lo posea para que toda oración compleja sea determinada como verdadera o no, de acuerdo con su composición a partir de oraciones atómicas. Entonces el segundo principio fundamental de la semántica fregeana, *i.e.* clásica, es que el valor semántico de una oración atómica (y por tanto de cualquier oración) es precisamente su valor de verdad —el que sea verdadera o no. Dada la notación cuantificacional, no es difícil llegar a la conclusión de que el valor de verdad de una oración cuantificada depende justamente de los valores de verdad de los constituyentes. Es menos obvio que *éste sea el caso para una oración formada por medio de la conectiva 'si'*; pero ya no está en discusión que, funcione o no de este modo el 'si' del lenguaje natural, es adecuado para la formulación de

teorías matemáticas y científicas emplear un 'si' para el cual sea correcta la tabla de verdad bivalente. Es, no obstante, al segundo principio al que se oponen los que rechazan la semántica clásica.

Una vez que tenemos estos dos principios, podemos extender la noción de valor semántico a expresiones distintas de las oraciones; y no hay ninguna seria elección acerca de cómo ha de hacerse esta extensión. El valor semántico de cualquier expresión es, de nuevo, aquel rasgo suyo que debe adscribirse para que toda oración en que ocurra sea determinada como verdadera o no verdadera. Por el segundo principio fundamental, el valor semántico de un término singular o de un predicado depende solamente de cómo haya de determinarse como verdadera o falsa una oración atómica que lo contenga. Establecemos primero que el valor semántico de un término singular es el objeto particular al que se refiere: entonces se sigue que el valor semántico de un predicado unario es dado fijando de qué objetos es verdadero, de un predicado binario de qué pares de objetos es verdadero, etc.

La explicación que acabo de dar es defectuosa en un importante respecto: depende de suponer que, para cada objeto, el lenguaje contiene un término singular que se refiere a él. Muchos lenguajes no satisfarán esta condición. Para tal lenguaje, debemos primero especificar el dominio de la cuantificación, y luego considerar una expansión del lenguaje que contenga un nombre de cada objeto de ese dominio. Si el dominio es muy amplio (no-denumerable), tal lenguaje expandido será sólo una construcción abstracta, no un lenguaje que podamos hablar efectivamente; pero esto no importa, puesto que estamos sólo usándolo como un recurso auxiliar para dar la semántica del lenguaje original.

Una vez que tenemos tal semántica, podemos substituir nuestra noción de interpretación por reemplazo por la de una interpretación semántica bajo la cual hacemos una asignación directa a las letras esquemáticas de los valores semánticos de expresiones de las categorías apropiadas, rodeando las expresiones mismas.

→ El sentido de una expresión ha sido explicado como aquel ingrediente de su significado que es relevante para la deter-

minación como verdadera o falsa de una oración en la que ocurre: podemos ahora decir que es aquel ingrediente de su significado que determina el valor semántico de la expresión. Es importante aquí evitar el confundirse sobre un punto acerca del cual muchos filósofos se expresan ambiguamente. Se podría decir que el significado de una oración no puede, por sí mismo, determinar su verdad o falsedad, al menos en el caso general; de ahí que el sentido de las palabras pueda determinar sólo la condición de su verdad, no su valor de verdad; éste dependerá también de cómo sea el mundo, *i.e.*, de la realidad extralingüística. Pero el valor semántico de una expresión fue explicado de manera que los valores semánticos de las palabras que componen una oración la determinen conjuntamente como verdadera o falsa: cuando se dice que el sentido de una palabra determina su valor semántico, esto depende del supuesto de que la contribución de la realidad extralingüística se está tomando en cuenta. La posibilidad de explicar así la noción de valor semántico depende del supuesto, incorporado en el segundo principio fundamental de la semántica clásica, de que la condición de la verdad de cada oración es, determinadamente, o cumplida o no cumplida. Podemos considerarlo un supuesto metafísico —un supuesto de la existencia de una realidad objetiva independiente de nuestro conocimiento. Podemos, igualmente, considerarlo como un supuesto de la teoría del significado, a saber, que logramos conferir a nuestras oraciones un sentido que las convierte determinadamente en verdaderas o falsas. Una vez que se rechaza este supuesto, como hacen los intuicionistas para los enunciados matemáticos, ya no podemos explicar legítimamente el valor semántico de una expresión como contribuyendo a determinar cada oración en la que figura como verdadera o no, puesto que no podemos suponer que el valor de verdad de cada una de tales oraciones sea determinado. Según tal concepción, el valor semántico de una oración puede, a lo sumo, ser su condición de verdad, y la concepción de los valores semánticos de otras expresiones debe ser modificada consecuentemente.

La semántica clásica puede ser rechazada por los que piensan que hay formas de composición oracional, no inclui-

das en el lenguaje formalizado de Frege, en cuya presencia no puede ser mantenida; un ejemplo es la defensa de una semántica de mundos posibles por parte de quienes creen necesario dar cuenta de operadores modales como 'necesariamente'. O, más radicalmente, puede ser rechazada incluso para un lenguaje que use sólo formas fregeanas de composición oracional, por quienes niegan que toda oración así formada deba ser determinadamente verdadera o falsa.

La noción fregeana de la *referencia* de una expresión es, esencialmente, la de valor semántico, tal como he usado este término, entendido en el contexto de la semántica clásica. Hay, en realidad, otro rasgo de la noción de Frege del que no me ocuparé, a saber, su creencia en que el valor semántico de toda expresión puede darse asociando con ella alguna entidad extralingüística de un tipo lógico que depende de la categoría lógica de la expresión; así se supone que hay alguna cosa cuya asociación con un predicado determina, para cada objeto, el que el predicado sea verdadero o no de este objeto. No prestaré casi ninguna atención a este punto.

El sentido es un ingrediente del significado: dar una explicación del sentido de una expresión es, por tanto, dar una explicación parcial de lo que un hablante sabe cuando entiende esa expresión. En vista de lo que ya dije, que, al considerar que el sentido determina la referencia, estamos suponiendo que se está tomando con ello en cuenta la contribución de la realidad extralingüística, lo que Frege consideró como uno de sus descubrimientos fundamentales, que hay una distinción entre sentido y referencia, esto es, que el sentido de una expresión no puede consistir sólo en el hecho de que tenga la referencia que tiene, debe ser perfectamente obvio: es por tanto a primera vista sorprendente que quede como una de sus tesis más persistentemente controvertidas.

¿Cómo es siquiera posible que alguien rechace la distinción entre referencia y sentido? Nadie puede negar que un hablante capacitado del lenguaje debe conocer algo más que la referencia de una expresión compleja: debe al menos saber cómo se determina su referencia de acuerdo con su composición a partir de sus palabras componentes. De ahí que lo más que puede mantenerse es que la distinción falla para

palabras simples, *i.e.* expresiones lógicamente simples; y el camino queda entonces despejado, quizá, para la admisión de que un hablante pueda entender una expresión compleja sin conocer su referencia —conoce la referencia de las palabras componentes, y sabe cómo determinan conjuntamente la referencia del todo, pero no sabe efectivamente cuál es ésta. En realidad, debe concederse esto a no ser que haya de sostenerse que cualquiera que entienda una oración sabe si es verdadera o falsa. La dificultad es que una expresión que sea lingüísticamente simple puede tener un sentido complejo: esto es, su sentido puede darse como si fuera el de una expresión compleja, como cuando explicamos que 'es primo' significa 'tiene exactamente dos divisores' o que 'mentiroso' significa 'hombre que afirma lo que sabe que no es verdadero'. De ahí que o bien debe restringirse aún más la tesis de que no hay ninguna distinción entre sentido y referencia, o debe negarse que, cuando introducimos una palabra por tales medios, le transfiramos el sentido, así como la referencia, de la expresión compleja.

El ataque a la distinción sentido/referencia viene de tres direcciones: de la escuela de Davidson, según la cual una teoría del significado para un lenguaje debe tomar la forma de una teoría de la verdad al estilo de Tarski; de los proponentes de una concepción holista del lenguaje, capitaneados por Quine; y de los partidarios de la llamada teoría causal de la referencia, conducidos por Kripke. Tarski se ocupó de la cuestión de cómo podríamos convertir la semántica informal clásica, introducida por Frege para un lenguaje formalizado, en una definición formal de verdad para las oraciones de ese lenguaje: nos ocupamos por tanto ahora de *dos* lenguajes formalizados, en vez de sólo uno —el lenguaje-objeto para el que estamos definiendo la verdad, y el metalenguaje en el que la estamos definiendo. Para Frege, la referencia de una expresión es una entidad extralingüística, y, en la semántica informal, o teoría de modelos, que se ha desarrollado a partir de sus ideas, una interpretación asocia a cada constante individual, predicado, etc., del lenguaje una entidad no-lingüística de un tipo adecuado. Sin embargo, cuando tratamos de construir una definición de verdad se

traslada la atención desde las entidades asociadas con las expresiones del lenguaje-objeto como referencias de éstas hacia los medios de que se dispone en el metalenguaje para especificar dicha asociación. Además Tarski, al igual que Frege, pero a diferencia de los teóricos de modelos en general, se ocupa, no de la noción de verdad bajo una interpretación arbitraria, sino de la verdad *simpliciter*, esto es, de la verdad bajo alguna interpretación determinada (la pretendida). Con el fin de disponer de un criterio para el éxito de una definición de verdad para un lenguaje-objeto dado, Tarski establece un requisito que se puede formular muy simplemente en el caso de que el metalenguaje sea una extensión del lenguaje-objeto; a saber, que la definición sea tal que seamos capaces de probar cada instancia del esquema

S es verdadera si y sólo si P

que se obtenga poniendo en lugar de 'S' un término que denote una oración del lenguaje-objeto, y en lugar de 'P', dicha oración. Con el fin de definir 'verdadero', es necesario definir algunas nociones auxiliares, como 'satisface' y 'denota'; y, con el fin de poder probar cada instancia del anterior esquema, es muy simple dar a las cláusulas de sus definiciones la forma más directa, esto es, una en la que también se use la expresión del lenguaje-objeto que se menciona, como en

'O' denota O

y

(u, v) satisface ' x es mayor que y ' si y sólo si u es mayor que v .

Cuánta luz se aporta al análisis del concepto de verdad mediante tal definición de 'verdadero' es algo que no nos interesa; Davidson ha propuesto que, si la reformulamos como teoría axiomática y no como definición, podemos, tomando la definición de verdad como ya dada, reinterpretarla como constituyendo una teoría del significado del lenguaje-objeto. Esto es, tanto si se les puede atribuir a los hablantes reales del lenguaje un conocimiento de tal teoría como si no,

el conocimiento de la misma por parte de alguien que posea la comprensión necesaria del concepto de verdad le proporcionará una intelección del lenguaje. Por supuesto, si estamos interesados en el tipo de conocimiento ejemplificado por el dominio de una persona de su lengua materna, dicho conocimiento no puede consistir en la capacidad de dar una formulación verbal de la teoría de la verdad; lo que nos interesa es *qué* conoce, no su manifestación de dicho conocimiento. Para evaluar dicha afirmación deberemos indagar, como es natural, qué lleva implícita la atribución de un conocimiento de tal teoría a un hablante, un conocimiento que, en general, no será conocimiento explícito. Yo creo que los argumentos de Frege para distinguir entre sentido y referencia ya encierran una demostración de lo inadecuado de tal planteamiento.

Frege presentó dos de esos argumentos, de los que ofreceré una reconstrucción en vez de una cita directa. Lo que intentamos mostrar es que captar el sentido de una expresión simple no puede hacerse igual a conocer su referencia; de modo que es natural que deseemos saber cómo entender una instancia del esquema (C1) del diagrama. El diagrama está construido de modo que, por ejemplo, (B5) sea un caso especial de (A5), y (C5) un caso especial de (B5), y así sucesivamente siempre que exista más de una entrada en cualquier fila. Ahora bien, es natural decir que (C1) es equivalente a (C2); y por lo tanto nos gustaría saber cómo entender las instancias de (B2). Diré que una instancia de (B2) adscribe 'conocimiento-qué' al sujeto, en oposición al 'conocimiento-que' que les es adscrito por una instancia de (A5) que es obtenida mediante reemplazo de 'P' por una oración completa que pueda usarse independientemente. El conocimiento-qué comprende también el conocimiento-cuál, el conocimiento-dónde, el conocimiento-cuándo y el conocimiento-quién; no comprende el conocimiento-cómo en el sentido de Ryle —por ejemplo, saber curar una picadura— pero incluye cosas como saber cómo fue inculado el veneno. Una instancia de (B2) es, por tanto, 'La policía sabe quién asesinó a Bexley'. Una sugerencia plausible es que (B2) es equivalente a (B3) [y, en consecuencia, (C1) y (C2) lo son a (C3)]. Una

DIAGRAMA

- | | | |
|---|--|---|
| (A5) X sabe que P. | | |
| (A6) X sabe que S es verdadera. | (B2) X sabe qué es F. | (C1) X conoce la referencia de w . |
| (B6) X sabe que \bar{b} es verdadera. | (B3) Para algún u , X sabe, de u , que es F. | (C2) X sabe a qué se refiere w . |
| (B5) X sabe que b es F. | (B4) X sabe, de a , que es F. | (C3) Para algún u , X sabe, de u , que w se refiere a él. |
| (B6) X sabe que \bar{b} es verdadera. | (B5) X sabe que b es F. | (C4) X sabe, de a , que w se refiere a él. |
| (B6) X sabe que \bar{b} es verdadera. | (B6) X sabe que \bar{b} es verdadera. | (C5) X sabe que w se refiere a b . |
| (B6) X sabe que \bar{b} es verdadera. | (B6) X sabe que \bar{b} es verdadera. | (C6) X sabe que \bar{w} se refiere a \bar{b} es verdadera. |

instancia de (B3) es una cuantificación existencial de una instancia de (B4), y por tanto será verdadera sólo si alguna instancia correspondiente de (B4) es verdadera: por ejemplo, 'La policía sabe, de Redmayne, que asesinó a Bexley'. También diré, de la instancia de (A5), que adscribe *conocimiento proposicional* al sujeto, y, en particular, que le adscriben conocimiento de la proposición expresada por la oración que reemplaza a 'P'; análogamente, diré que una instancia de (B4) adscribe *conocimiento predicativo* al sujeto.

Interpolaría aquí la observación de que la igualdad de (C1) y (C2) será estrictamente correcta sólo cuando exista una entidad que haga de referencia de la palabra en cuestión, cosa que Frege pensaba que sucedía siempre. Cuando, por ejemplo, se reemplaza 'w' por un término que se refiere a un predicado unario, podemos, sin invocar la doctrina especial de Frege, decir solamente que alguien conoce la referencia del predicado cuando sabe de qué objetos es verdadero, y este enunciado no es exactamente de la forma (B2); obtendríamos una instancia de (B2) sólo diciendo algo como 'sabe a qué concepto se refiere'. No obstante, aun cuando no aceptemos la doctrina especial de Frege, las subsiguientes observaciones pueden transplantarse a casos distintos de aquél en que la palabra en cuestión es un término singular.

Necesitamos ahora saber bajo qué condiciones será verdadera una adscripción de conocimiento predicativo (B4). Es natural aventurar que la verdad de una instancia de (B5) entrañará a veces la verdad de una instancia correspondiente de (B4), donde 'a' y 'b' se reemplazan por expresiones de la misma referencia, a veces por la propia expresión misma. Por ejemplo, podría suceder que la verdad de 'La policía sabe, de Redmayne, que asesinó a Bexley' se sigue de la verdad de 'La policía sabe que Redmayne asesinó a Bexley'. Cuando esto suceda, diré que la adscripción de conocimiento predicativo *se basa en* una adscripción de conocimiento proposicional. Así, pues, la verdad de determinadas instancias de (B5) garantizará la verdad de la correspondiente instancia de (B3), y, en consecuencia, de (B2). Por supuesto, no con toda instancia de (B5) sucederá lo mismo: la policía puede saber que el hombre que se ocultó en el tejado asesinó

a Bexley, mas ello no será fundamento para decir que sabe quién asesinó a Bexley a menos que podamos decir independientemente que sabe quién se ocultó en el tejado.

Puede ahora surgir la sospecha de que el conocimiento-qué no es una noción precisa, o de que, en todo caso, no es una noción que admita caracterización uniforme: porque, podría decirse, dependerá del contexto qué piezas de conocimiento proposicional validarán una adscripción de conocimiento-qué. A la policía, por ejemplo, le gustaría saber del asesino lo suficiente para poder arrestarlo y acusarlo, aunque pueda existir algún sentido en que no sepa quién es (en realidad). Es ésta una razón plausible; el problema es si se considera que vale igualmente para el conocimiento predicativo, o si se considera que constituye una base para poner en entredicho mi igualación de (B2) y (B3). Supondré, con vistas a la discusión, que el conocimiento predicativo es una noción precisa y libre de contexto; que, en tanto en cuanto el conocimiento-qué no es una noción tal, ello sólo muestra que la igualación de (B2) y (B3) no es totalmente correcta; pero que, sin embargo, la igualación de (C2) y (C3) es correcta.

Adscribir a alguien el conocimiento de la referencia de una palabra es, pues, sobre esta base, afirmar la verdad de alguna oración de la forma (C4); y a veces se puede justificar esto mediante apelación a la verdad de una oración de la forma (C5), si bien hasta ahora no hemos procurado circunscribir aquellas oraciones de esta forma que puedan servir a este propósito. ¿Qué es, por tanto, adscribir a alguien un conocimiento *neto* de la referencia de una palabra? Es, evidentemente, adscribir una pieza de conocimiento predicativo (C4), y añadir que ello constituye una caracterización *completa* de dicha pieza de conocimiento por su parte. Si la adscripción es correcta, se sigue que la posesión por parte del mismo de dicha pieza de conocimiento predicativo no puede basarse en una pieza correspondiente de conocimiento proposicional (C5). Pues una oración de la forma (B4) no determina individualmente una oración correspondiente de la forma (B5) que la implique, ya que la ' α ', en (B4), se encuentra en una llamada posición 'transparente', esto es, pode-

mos poner en el lugar de 'a' expresiones diferentes, con la misma referencia, sin alterar el valor de verdad. Por lo tanto, determinar cuál sea la pieza de conocimiento proposicional sobre la que se basa una pieza dada de conocimiento predicativo deberá ser caracterizar dicha pieza de conocimiento más de lo que se hace con la mera adscripción del conocimiento predicativo: en consecuencia, adscribir a alguien un conocimiento *neto* de la referencia de una palabra es adscribirle una pieza de conocimiento predicativo que no se basa en ningún conocimiento proposicional.

Con este instrumental, el primer argumento de Frege para la distinción de sentido y referencia, se puede establecer muy simplemente. Su premisa es que no existe conocimiento predicativo alguno que no se base en conocimiento proposicional. De ello se sigue que no puede existir algo como el conocimiento neto de la referencia de una palabra; la ulterior caracterización del conocimiento que un hablante tiene al conocer la referencia de una palabra muestra el sentido que dicha palabra tiene para él.

Una posible pega a este argumento sería negar la premisa: yo creo que la premisa es verdadera, pero no me ocuparé de su defensa. Pero otra réplica podría ser que el argumento, tal como yo lo he formulado, es formalmente inválido. Si pudiera existir conocimiento predicativo que no se basase en conocimiento proposicional, entonces ciertamente una adscripción de conocimiento proposicional siempre caracterizaría dicha pieza de conocimiento más de lo que lo hace la correspondiente adscripción de conocimiento predicativo. Pero si suponemos que no existe tal conocimiento predicativo, y, además, que una oración de la forma (B4) determina siempre una única oración de la forma (B5) que la implica, entonces la primera oración *es* una caracterización completa. El argumento que di contra este último supuesto no se sostiene: sólo muestra que la verdad de una instancia de (B4) no puede depender de la verdad de la correspondiente instancia de 'X sabe que a es F'.

El argumento requiere por tanto una complementación cuya forma dependerá de cómo el objetor especifique la instancia de (C5) de cuya verdad dependerá la verdad de una

instancia de (C2). No conozco ninguna sugerencia salvo la derivada de tomar una definición de verdad tarskiana como patrón de una teoría del significado: a saber, que donde, por ejemplo, reemplazamos '*w*' por un término que denote un nombre propio, debemos reemplazar '*b*' por ese mismo nombre propio. La sugerencia equivale a decir que conocer la referencia del nombre 'Valencia' o del nombre 'Goliat' es saber que 'Valencia' se refiere a Valencia o que 'Goliat' se refiere a Goliat. Ahora bien, tal afirmación parece a primera vista absurda, dado que parece que cualquiera que sepa que 'Valencia' es un nombre propio y que hay algo a lo que se refiere debe saber que 'Valencia' se refiere a Valencia, aun cuando no sepa lo más mínimo de qué es el nombre 'Valencia': la propuesta parece que confunde conocer la referencia de un objeto con saber que tiene una referencia. Esta obvia objeción puede salvarse, sin embargo, invocando todavía otra distinción: la existente entre conocer la verdad de una oración, en cuanto expresada por una instancia de (A6), y conocer la proposición expresada por esa oración, en cuanto expresada por una instancia de (A5). Si alguien sabe que una cierta oración es verdadera, una condición necesaria para que este conocimiento resulte en conocimiento de la proposición expresada por ella es que la entienda completamente (y con ello sepa qué proposición expresa). Por ejemplo, si alguien que nunca ha oído hablar del Profesor Geach o de semántica, y que sabe de la palabra 'Alemania' sólo que es el nombre de un lugar de Europa, oye afirmar autorizadamente que el Profesor Geach toma parte en una conferencia sobre semántica a celebrar en Alemania el 28 de mayo, nadie diría que sabe ahora que el Profesor Geach toma parte en una conferencia sobre semántica en Alemania el 28 de mayo. Sólo sabría que una cierta oración es verdadera; más exactamente, podría decir que sabe que alguien llamado 'Profesor Geach' toma parte en una conferencia sobre algo llamado 'semántica' a celebrar el 28 de mayo en Alemania, sea lo que esto fuere, indicando con ello de qué palabras de la oración tiene una imperfecta comprensión.

Si se invoca así la distinción entre (C6) y (C5) para defender la tesis de que conocer la referencia de 'Valencia' es

saber que 'Valencia' se refiere a Valencia, se nos debe decir qué se requiere para que alguien conozca la proposición expresada por una oración. Ciertamente no es necesario que entienda esa oración; podría entender y a la vez saber que es verdadera una oración equivalente en otro lenguaje. Si, por ejemplo, la oración contiene la palabra 'oveja', no necesita entender esa particular palabra castellana, pero es necesario que capte el concepto que expresa; similarmente, si la oración contiene el nombre 'Alemania', no necesita conocer el uso de ese nombre, pero debe también conocer el uso de un nombre, tal como 'Deutschland' o 'Allemagne', que tenga el mismo uso, o al menos tener la concepción de ese uso, *i.e.*, del modo de distinguir una específica área geográfica, aun cuando no lo asocie con ningún nombre efectivo de su vocabulario. Nuestro hipotético oponente al argumento de Frege no puede negar que hay algo asociado con un nombre que el sujeto debe captar para admitirle conocimiento de la proposición expresada por una oración en la que se usa el nombre, sin perder la distinción, para tal oración, entre conocer esa proposición y saber que la oración es verdadera; si la pierde, está indefenso contra la acusación de que ha confundido conocer la referencia de un nombre con saber que tiene una referencia. Lo que el sujeto debe captar es, precisamente, el sentido del nombre.

El repudio de la distinción entre sentido y referencia, para los nombres propios, es una manifestación de una tendencia a retornar a una doctrina del tipo propugnado por Mill de que la función lingüística de un nombre propio consiste en una asociación directa entre el nombre y el objeto de la cual no puede darse ninguna explicación ulterior. Una expresión natural de tal idea es concebir la captación del nombre por parte del hablante como conocimiento predicativo irreducible a conocimiento proposicional, *i.e.*, como saber, del objeto, que el nombre se refiere a él; pero, dada tal concepción de los nombres propios, motiva igualmente la idea de que, cuando 'b' se reemplaza por un nombre propio y 'a' por cualquier término singular con la misma referencia, la verdad de las resultantes instancias de (B4) y (B5) puede ser igualada. Pero, como hemos visto, esta ecuación, aunque legítima

según esta concepción de los nombres propios, no es explicativa, dado que, para establecer una distinción entre (B5) y (B6), tenemos que apelar a alguna noción del *uso* de un nombre, del modo en que el portador es distinguido por medio de él. Para contrastar efectivamente la captación del uso de un *nombre* con la de un término singular complejo tal como 'la ciudad sitiada por el Cid', el único recurso parece ser refugiarse en explicar una oración de la forma (B5) en términos de una de la forma (B4), es decir, apelar una vez más a la noción de conocimiento predicativo: y así la explicación se muerde su propia cola.

El único modo de romper el círculo es admitir —cosa que los miembros de la escuela davidsoniana son reacios a hacer— que discutir el significado solamente en términos de *qué* conoce el hablante cuando capta el uso de una palabra es insatisfactorio. Una teoría del significado, en el presente sentido, da una representación teórica de una habilidad práctica; en términos de Ryle, representa el conocimiento-cómo mediante el conocimiento-que. En muchos contextos, no surge ningún gran problema acerca de lo que está involucrado en el atribuirle a alguien una pieza de conocimiento proposicional, puesto que puede suponerse que él conoce algún lenguaje en el que pueda expresarse la proposición, y no habrá ninguna diferencia substantiva entre su conocimiento de la proposición y su conocimiento de la verdad de la oración; igualmente, pocas dificultades habrá en representar una pieza de conocimiento práctico como conocimiento de una proposición, puesto que la transición entre los dos será mediada por el dominio del lenguaje por parte del sujeto. Pero, cuando el conocimiento del que nos ocupamos es él mismo el requerido para el dominio de un lenguaje, estas cuestiones ya no se cuidan de sí mismas. Si representamos la comprensión de una palabra por parte de un hablante como su conocimiento de una cierta proposición, nuestra explicación queda incompleta a no ser que expliquemos también en qué consiste un conocimiento implícito de esa proposición, es decir, qué ha de contar como una manifestación de ese conocimiento.

En estos términos, lo que posee alguien que conoce la proposición expresada por una oración que contiene un nombre propio, no es un modo particular de distinguir o identificar el objeto, sino simplemente la capacidad de reconocerlo cuando se encuentra ante él. No debemos preguntar, según esta explicación, *cómo* o *por medio de qué* es reconocido el objeto; aun cuando haya una respuesta, el sujeto no tiene que conocerla. Nuestra comprensión de la palabra 'rojo' depende de nuestra capacidad de reconocer algo como rojo; no podemos responder a la pregunta de *cómo* reconocerlo.

Kripke ha etiquetado la tesis de Frege de que los nombres propios tienen sentido como 'la teoría descriptiva', identificándola con la concepción de que todo nombre propio tiene el mismo sentido que una descripción definida. Es, en realidad, esencial para la concepción de Frege el que un nombre *pueda* tener el mismo sentido que una descripción definida; pero pensar que un nombre no puede tener ningún otro tipo de sentido es malinterpretar seriamente a Frege. La idea de que alguien pueda tener una capacidad de reconocer un objeto que no puede explicar ulteriormente no es en modo alguno absurda, y sería enteramente incorrecto suponer que Frege tuvo algún motivo para negar que una captación de un nombre podría, en determinada ocasión, consistir en su asociación con esa capacidad precisamente; aunque conversamente, la sugerencia de que nuestra comprensión de todos, o incluso de muchos, nombres es de este tipo no es más que ridícula. Pero, naturalmente, aunque un sujeto pueda ser incapaz de dar *explicación* alguna de tal capacidad, requiere ulterior *descripción*. Ante todo, hemos de preguntar qué cuenta como el hecho de que el objeto se 'encuentre ante' él. Supongamos que el nombre es el de una ciudad, pongamos por caso 'Valencia': ¿se está diciendo que él puede distinguirla en un mapa, o que puede reconocerla a distancia, desde tierra o desde un avión, o solamente que puede identificarla siempre que esté en la ciudad? Si lo último, ¿bajo qué condiciones? —¿solamente cuando pasea por las calles, o incluso cuando está bajo techo?, y si está bajo techo, ¿solamente cuando puede ver por la ventana, o incluso cuando está a ciegas? No hay una noción uniforme de *encontrarse*

ante un objeto: la capacidad de reconocerlo es relativa a las condiciones bajo las cuales toma lugar el reconocimiento. En segundo lugar —un punto de inmensa importancia, debido a Frege— su capacidad de reconocer el objeto debe estar sujeta al conocimiento de que cae bajo el concepto que determina el criterio apropiado de su identidad (en nuestro ejemplo, el concepto de una ciudad); en particular, debe estar dispuesto a admitir error cuando puede mostrarse que sus pretensiones de haber reconocido el objeto contravienen el criterio de su identidad. Cuando la asignación al sujeto de tal capacidad de reconocimiento ha sido completada de estos dos modos, ya no hay ninguna base en principio para negar que su comprensión del uso del nombre no puede consistir en su habilidad para identificar así a su portador, aunque esto ya no es la simple cuestión que al principio pareciera ser. Eso justamente sería, en tal caso, lo que constituye el sentido que asignó al nombre; no habría fuerza para la sugerencia de que, dando una explicación de este sentido, no estuviéramos diciendo más de lo que se comunica por medio del neto enunciado de que él sabe, de la ciudad, que un cierto nombre se refiere a ella.

El primer argumento de Frege para la distinción entre sentido y referencia tiene, no obstante, un importante defecto: no tiende a mostrar que el sentido de una palabra sea una característica del *lenguaje*. Muestra, en el mejor de los casos, que cada hablante, si ha de asociar una referencia con una palabra, debe asignarle un sentido particular; no muestra ninguna necesidad de que diferentes hablantes asignen el mismo sentido a cualquier palabra, siempre que los sentidos que todos le asignan determinen la misma referencia. Por lo tanto deja abierta la posibilidad de que el sentido de una palabra no sea en absoluto parte de su significado, si el significado ha de ser algo objetivo y compartido por todos los hablantes, como mantiene Frege; que, lo mismo que dijo Frege de las imágenes mentales asociadas con una palabra, sea meramente un acompañamiento psicológico, o, en el mejor de los casos, parte del mecanismo psicológico por medio del cual un hablante asigna un significado a la palabra, y no un ingrediente genuino del significado. Esta

duda la contradice el segundo argumento de Frege para la distinción sentido/referencia.

El primer argumento de Frege dice que debemos adscribir *más* a un hablante que el mero conocimiento de la referencia de una palabra; el segundo dice que no podemos adscribirle *tanto*. Este es el argumento en donde Frege introdujo por primera vez la noción de sentido, en su famoso ensayo "Sobre sentido y referencia". El primer argumento estaba relacionado con el conocimiento que un hablante tiene de la referencia de una palabra, o, más generalmente, con lo que conoce cuando conoce su significado. El segundo argumento también tiene que ver con el conocimiento, esta vez el conocimiento transmitido a un oyente por medio de una aserción, o, más generalmente, adquirido por un hablante cuando llega a aceptar una oración como verdadera. Para que una oración sea capaz de utilizarse para transmitir información, debe ser posible entenderla antes de saber que es verdadera. Frege considera, específicamente, un enunciado de identidad, esto es, uno de la forma '*a* es igual a *b*', y argumenta que si, para entender el enunciado, se requiere que el oyente conozca la referencia de los términos '*a*' y '*b*', entonces, si es verdadero, debe, si lo entiende, saber de antemano que es verdadero; por lo tanto, con esta suposición, tal enunciado no podría transmitir información. Claramente, este argumento debe asumir que, si uno conoce la referencia de dos términos, debe saber si tienen la misma referencia o no. Esta suposición obviamente no está justificada bajo todos los modos posibles de interpretar la expresión 'conocer la referencia', por ejemplo, no bajo uno de acuerdo con el cual la verdad de una instancia de (C1) se sigue de la de una instancia de (C5), para la misma sustitución de '*w*' y *cualquier* sustitución de '*b*'. Puede justificarse reforzando en dos distintos modos la tesis que Frege estaba refutando. Primero, tomémosla como la tesis de que conocer el significado de un término consiste en parte en conocer su referencia (y no meramente lo entraña). Es una característica innegable de la noción de significado —oscura como lo es tal noción— que el significado es *transparente* en el sentido de que, si alguien asigna un significado a cada una de dos palabras, debe saber si estos

significados son el mismo. Se sigue por lo tanto de la tesis en esta forma que quien conoce los significados de dos términos sabe si tienen la misma referencia. Alternativamente podemos establecer la tesis a refutar como que, si un hablante entiende un término, debe tener un conocimiento *neto* de su referencia, en el sentido previamente explicado. Entonces se seguirá la conclusión siempre que podamos decir que, si alguien sabe de un objeto u que es F y que es G , debe saber, de aquel objeto, que es ambas cosas, F y G . De nuevo, es obvio que no se sostendrá bajo todo modo posible de asignar la condición de que alguien tenga una pieza de conocimiento predicativo; pero se seguirá si añadimos la premisa de que atribuirle estas dos piezas de conocimiento predicativo es una caracterización completa de los trozos relevantes de conocimiento por su parte, ya que de otro modo no habría nada que pudiéramos citar como la pieza adicional de conocimiento que tendría que adquirir para pasar de las dos piezas de conocimiento que él tenía al conocimiento, de u , de que era simultáneamente F y G .

El sentido fue definido como un ingrediente del significado, y consecuentemente como algo que es transparente en el modo en el que el significado es transparente, y como aquel ingrediente que va a determinar la referencia. Entonces, si es necesario, para que un enunciado de identidad sea informativo, que alguien que entiende los sentidos de dos términos pueda no conocer que tienen la misma referencia, se sigue que los sentidos pueden diferir aunque la referencia sea la misma. Ahora bien, el que prefiramos decir que alguien que conoce el sentido también conoce la referencia es una cuestión de cómo preferimos construir la expresión 'conocer la referencia', *i.e.* de las condiciones que nosotros imponemos para la verdad de una atribución de conocimiento predicativo. Parece menos engañoso *no* decirlo; pero, si lo decimos, lo podemos hacer sólo porque admitimos que el conocimiento de la referencia de cualquier término dado se basa en un conocimiento de cualquiera de varias proposiciones distintas, de modo que las diferencias entre las proposiciones reflejan la diferencia entre varios sentidos que pueden ser atribuidos al término sin alterar su referencia.

Este argumento puede parecer una mera variación del primero, empleando una vía más circular hacia la misma conclusión, y, por supuesto, está estrechamente conectado con él —yo no lo he desmenuzado detalladamente, ya que supondría seguir sobre la misma base que tendríamos en el primer argumento. El segundo argumento tiene, no obstante, una ventaja importante, ya que proporciona una razón para suponer que diferentes hablantes deben asignar el mismo sentido a una palabra. Este segundo argumento concierne, pues, al uso del lenguaje para la *comunicación*, la cual depende de que el contenido informativo de una oración sea constante de un hablante a otro. Si el lenguaje ha de servir como un medio de comunicación, no es suficiente que una oración sea de hecho verdadera bajo la interpretación que le da un hablante sólo en el caso en el que sea verdadera bajo aquella que le da otro; también es necesario que ambos hablantes sean conscientes de este hecho.

La noción de contenido informativo está claramente conectada, de una manera que aquí no intentaré precisar, con aquello que reconocemos que justifica un enunciado, que establece su verdad: un enunciado sin contenido informativo no requiere justificación, y es plausible decir que captamos la información que transmite una oración sólo en el caso en que conozcamos lo que es necesario para establecerla como verdadera. Está claro, por tanto, que puede construirse un argumento paralelo, para la distinción entre sentido y referencia, en términos de lo que se nos exige para reconocer como verdadero un enunciado de identidad. Tal argumento proporcionaría igualmente una razón para tomar como común a diferentes hablantes el sentido asignado a un término, ya que el uso exitoso del lenguaje depende del acuerdo acerca de lo que es necesario para mostrar que un enunciado es verdadero. Aquí, por supuesto, existe un problema para una teoría del significado, como la de Frege, que se resuelve en una semántica clásica, bivalente. No toda oración tiene un sentido tal que seamos capaces, incluso en principio, de reconocerla como verdadera, teniendo en cuenta solamente el que sea verdadera. Se sigue que los sentidos de ciertas expresiones no pueden darse solamente en términos de nuestros

medios de reconocer como verdaderas las oraciones en las que ocurren; más bien deben darse de tal modo que captemos qué condición debe satisfacerse para que tal oración sea verdadera, independientemente, al menos en parte, de si somos capaces de reconocer aquella condición como dada o no. Por tanto, para una oración de este tipo se necesita una explicación de cómo conocemos qué es lo que se considera indicador de que es verdadera en el caso de que seamos capaces de reconocerlo, ya que el sentido de la oración no estaba dado directamente por referencia a nuestros medios de conocimiento. Está claro que, si una teoría del significado del tipo fregeano ha de ser capaz de explicar satisfactoriamente el uso efectivo del lenguaje, tal explicación deberá darse ulteriormente; la explicación detallada de los sentidos de nuestras palabras debe ser tal que muestre los medios por los cuales somos capaces de derivar, de una captación de la condición que debe darse para que una oración sea verdadera, un conocimiento de cuándo podemos reconocer aquella condición como cumplida.

El segundo argumento de Frege por la distinción sentido/referencia puede generalizarse a otras formas de oraciones, de hecho a cualquier oración atómica. Conocer la referencia de un predicado es saber de qué objetos es verdadero; es decir, si el conocimiento-qué debe interpretarse como conocimiento predicativo, saber de cada objeto del dominio, si el predicado es o no verdadero de él. Así, si alguien conoce la referencia de un predicado, y conoce también, de algún objeto, que es a lo que se refiere un término dado, debe saber entonces si es o no verdadera la oración resultante de asignar ese predicado a ese término. Este argumento es exactamente paralelo al argumento de Frege sobre enunciados de identidad, y es válido bajo el correspondiente modo de interpretar la tesis que se ha demostrado que lleva al absurdo.

Hay, sin embargo, una diferencia entre el argumento de Frege y esta extensión suya. Decir que, en general, el valor semántico de un término singular es un objeto, no restringe de ningún modo la clase de teoría semántica que adoptamos (a menos que se presuponga alguna ontología particular, algu-

na doctrina sobre los tipos de objetos que contiene el mundo): por tanto, el argumento original de Frege no depende de la adopción de una semántica clásica bivalente. Pero la extensión del argumento a otras oraciones atómicas depende de que asumamos que el valor semántico de un predicado es su extensión, *i.e.* de que sea determinadamente verdadero o falso para cada objeto del dominio; y esto es válido sólo en la semántica clásica.

Pero, así como el supuesto subyacente a la extensión del argumento es más fuerte, así la conclusión es más poderosa. El argumento de Frege acerca de los enunciados de identidad sería satisfecho suponiendo que el sentido de un término singular está relacionado con su referencia como un programa con su ejecución, es decir, si el sentido provee un procedimiento efectivo de operaciones físicas y mentales mediante las cuales se podría determinar la referencia. Ya que, sin apelar al supuesto de que cada oración significativa *tiene* un valor de verdad determinado, no podemos pretender que el valor semántico de cada oración sea su valor de verdad; y por tanto, aunque podemos decir que es posible entender una oración sin conocer su valor semántico, no tenemos base para argüir que se pueda entenderla sin incluso ser capaz de descubrir efectivamente su valor semántico. Así, concediendo que el valor semántico de un término singular es el objeto al cual se refiere, no podemos, sin apelar al principio de bivalencia, asumir que puedan ser admisibles términos singulares cuya referencia no puede efectivamente ser determinada, ni tampoco, por tanto, cualquier enunciado de identidad cuyo valor de verdad no pueda efectivamente decidirse. La extensión del argumento a todas las oraciones atómicas presupone bivalencia; por tanto, deja abierta la posibilidad de que el lenguaje pueda contener predicados primitivos cuya aplicación no puede efectivamente decidirse, y que podemos, por tanto, entender sin ser capaces de determinar su valor semántico, ya que el valor semántico de un predicado se supone que es su extensión. Como ya indiqué, debe admitirse alguna distinción entre sentido y valor semántico cualquiera que sea la teoría semántica que adoptemos; pero es sólo en el contexto de la semántica clásica donde exigimos una

noción de sentido que determine el valor semántico de una expresión, pero no de una manera no efectiva.

Mi discusión de la noción de sentido ha sido sólo programática, lo mismo que era la de Frege: no he dicho específicamente qué debemos tomar como constituyente de los sentidos de expresiones de diferentes categorías, sino que únicamente he discutido los argumentos para suponer que necesitamos una tal noción. Mi propia opinión es que es altamente dudoso que la semántica clásica sea correcta; si no lo es, necesitamos admitir entonces sólo una distinción comparativamente trivial entre sentido y valor semántico. Sin embargo, para investigar esta cuestión eminentemente profunda necesitamos tener una concepción exacta de la forma que tomaría una teoría del significado que incluya la semántica clásica; y mi convicción es que una tal teoría debe incorporar una distinción substancial entre referencia y sentido de la clase que deseó trazar Frege.

Los argumentos de Frege en favor de la distinción no proporcionan por sí mismos una defensa contra el ataque de la escuela holista. Como expresión característica del holismo, tomemos el siguiente pasaje de un reciente ensayo de Davidson:

Renunciar a la distinción analítico-sintético como básica para la comprensión del lenguaje es renunciar a la idea de que podemos distinguir claramente entre teoría y lenguaje. El significado, como laxamente podríamos usar la palabra, está contaminado por la teoría, por lo que es tomado como verdadero.

Expondré el pensamiento aquí expresado en términos de las ideas de Quine más que en las del mismo Davidson (un tratamiento adecuado requeriría, naturalmente, un examen de todas las versiones diferentes del holismo). Una exposición detallada de la forma en que funciona un lenguaje específico —lo que he denominado hasta ahora una 'teoría del significado' para el lenguaje— se juzga correcta o incorrecta sólo si es posible derivar de ella las disposiciones lingüísticas observables de los hablantes. Esas disposiciones pueden considerarse simplemente como disposiciones a asentir a o a

disentir de las oraciones del lenguaje bajo ciertos estímulos sensoriales, incondicional o condicionalmente sobre el previo asentir a o disentir de otras oraciones. Las disposiciones condicionales constituyen las conexiones inferenciales entre las oraciones que son reconocidas por los hablantes. Las disposiciones lingüísticas poseídas por los hablantes en un momento dado pueden incluir o involucrar una predisposición a aceptar ciertas oraciones como verdaderas independientemente del estímulo externo actual. Supongamos ahora que una nueva oración se llega a aceptar generalmente como verdadera. Esto puede suceder por una multitud de razones: sobre la base de la observación, como la conclusión de una cadena de razonamiento, o simplemente por estipulación. Sin embargo sucede, en virtud de las conexiones inferenciales entre la oración recientemente aceptada y las demás oraciones, que alterará muchas de las disposiciones condicionales e incondicionales de los hablantes para asentir o disentir. Hará exactamente las mismas alteraciones en ellas sea cual sea la razón por la que se aceptó originalmente como verdadera: de hecho, no habrá nada en la subsecuente conducta lingüística de los hablantes sobre cuya base podamos diferenciar esta oración de cualquier otra que acepten como verdadera.

He intentado formular esta doctrina tan fríamente como me ha sido posible, pero es una doctrina candente, y se les ha subido a la cabeza a muchos de sus proponentes. Pues parece como si pudiéramos describir el resultado de la doctrina de este modo: Frege concibió el lenguaje como un juego desarrollado con reglas fijadas, tal que hay toda la diferencia del mundo entre un movimiento del juego y una alteración de las reglas; pero, si el holismo es correcto, todo movimiento del juego cambia las reglas —la distinción no tiene ninguna base en realidad. Algunos holistas, realmente, han llegado a pretender haber demostrado que la propia noción de significado es superflua; todos han tratado la de cambio de significado como inútil. Una razón por la que los oponentes del holismo han arrojado hasta ahora tan poca luz sobre la cuestión es que no distinguen entre la doctrina fundamental del holismo, tal como acabo de esforzarme en

exponerla, y la consecuencia que sus proponentes han pretendido que tiene. La doctrina fundamental, en mi opinión, es a lo sumo no más que una verdad a medias; pero mucho más importante es el hecho de que, aun cuando fuese totalmente verdadera, no tiene ni puede tener las consecuencias que se le atribuyen.

Imaginémonos que tenemos, para un lenguaje dado, una teoría del significado de la que podemos derivar las actuales disposiciones lingüísticas, condicionales e incondicionales, de los hablantes. La teoría proveerá por tanto lo necesario para que un hablante haga ciertos juicios en ciertas circunstancias; además mostrará, a través de las conexiones inferenciales encapsuladas en las disposiciones condicionales, cómo el conjunto total de las disposiciones del hablante será modificado por la subsecuente aceptación o rechazo de cualesquiera oraciones dadas. De ahí que, según los principios de la teoría del significado que tenemos, no somos capaces de determinar las disposiciones lingüísticas efectivas de un hablante en fechas futuras apelando a la teoría del significado solamente: sólo podemos hacerlo tomándolo conjuntamente con un acumulativo *stock* de juicios que él acepta.

Hasta ahora no hemos supuesto ni que la doctrina fundamental del holismo es válida para este lenguaje, ni que no lo es; pero supongamos ahora que lo es, es decir, que no hay forma de discriminar, apelando a la conducta lingüística de los hablantes en cualquier momento dado, entre cualquier oración generalmente aceptada como verdadera y cualquier otra, en particular que no hay tal discriminación entre ellas correspondiente al tipo de razón que los hablantes tuvieron originalmente para aceptarlas. Entonces habrá algo arbitrario en la representación de las disposiciones lingüísticas de los hablantes por medio de la teoría del significado constante junto con el acumulativo *stock* de juicios: los detalles de la representación dependerían de dónde nos encontrásemos eventualmente. Si hubiésemos comenzado la empresa antes o después, el *stock* de juicios sería, respectivamente, mayor o menor, y la teoría del significado fijada sería diferente. De ahí que, aunque nuestra representación no sea incorrecta, podamos preferir una diferente que no sea arbitraria de esta

forma. Una teoría del significado para el lenguaje de este nuevo tipo no nos permitiría, en sí misma, derivar disposición lingüística específica alguna en absoluto: produciría, en cambio, una función que proyectase todo conjunto (o, quizá, todo conjunto finito) de juicios en cuanto a la verdad y falsedad de las oraciones sobre un conjunto total de disposiciones, condicionales y no condicionales, a asentir y disentir. Una teoría del significado de este tipo sería uniforme: su forma no dependería de la selección arbitraria de un origen; pero seguiría ocurriendo que, dada tal teoría del significado, podríamos determinar las disposiciones lingüísticas efectivas de un hablante en cualquier momento sólo sabiendo primero qué oraciones acepta como verdaderas y cuáles rechaza como falsas.

No tengo ninguna idea de cómo sería precisamente, en detalle, tal teoría del significado; yo no acepto la doctrina fundamental del holismo, así que no necesito saberlo. Lo que es claro es que, al igual que cualquier teoría del significado desarrollable, debe tener una base finita: el conocimiento que un hablante tiene, al conocer el lenguaje, es una cantidad finita de conocimiento, aunque resulta en la comprensión de infinitamente muchas oraciones. Al igual que cualquier teoría del significado, sus ingredientes atómicos serían las representaciones de los significados de las palabras individuales y de las formas de construcción oracional del lenguaje: éstas, tomadas conjuntamente, determinarían colectivamente una proyección efectiva de conjuntos (finitos) de juicios sobre rangos totales de disposiciones lingüísticas. Una confusión frecuentemente cometida por los holistas es afirmar que, en una teoría holista, una oración no tiene un significado propio; esto es tan absurdo como decir que, en una teoría fregeana, una palabra no tiene un sentido propio. Lo que se sigue del holismo es que la teoría del significado para el lenguaje no determina por sí misma nuestra disposición a asentir o a disentir de cualquier oración particular bajo diferentes condiciones; pero esto no entraña que, en tal teoría, las oraciones e incluso las palabras del lenguaje no tengan un significado.

Es, además, claro que no hay ningún sentido plausible de 'significado' —por muy laxamente que empleemos el término— bajo el que pueda afirmarse que la aceptación general de un juicio provocará un cambio de significado, ni, por tanto, ninguno bajo el que podamos decir que las teorías abrazadas por los hablantes infectarán los significados de sus palabras. Es bien cierto, en *cualquier* teoría, que el hacer un juicio modifica las disposiciones que un hablante tiene a asentir o a disentir de otras oraciones; pero no hay más fundamentos para considerar estas disposiciones como determinantes del significado en una teoría holista que en una fregeana. La modificación de las disposiciones lingüísticas que ocurre como resultado de hacer un juicio tiene lugar de acuerdo con la comprensión general del lenguaje que está incorporada en una adecuada teoría del significado para él, y así un cambio en estas disposiciones no es ninguna indicación de cambio alguno en los significados de las palabras. Los lenguajes, ciertamente, cambian, en vocabulario, en pronunciación, en sintaxis, y, no hay razón para dudarlo, en significado. Pero necesitamos reconocer como cambio de significado sólo un cambio que no suceda según una regla, un cambio que no pueda ser explicado por una teoría; quizá debamos decir: que no pueda ser explicado por una teoría de la cual pueda atribuirse una captación implícita a todo hablante, dado que, si los hablantes son incapaces de prever que un cambio dado ocurriría en hipotéticas circunstancias, no pueden ser los significados de las palabras los que determinen ese cambio, aun cuando sea predecible. Pero los cambios que un holista pretende que se siguen del hecho de hacer cualquier juicio, o al menos de su adopción por la totalidad de la comunidad lingüística, y que son cambios de significado 'tal como podríamos usar laxamente la palabra', no son de este tipo.

Así pues, es claro que una teoría satisfactoria del significado para un lenguaje que satisfaga la doctrina fundamental del holismo tendría que ser de muchísima *mayor* complejidad que una para un lenguaje que no la satisfaga: en la construcción de una teoría del significado para un lenguaje del primer tipo, la identificación de las disposiciones lin-

güísticas poseídas actualmente por los hablantes a partir de un conjunto finito de observaciones de su conducta lingüística sería sólo un primer paso; el segundo sería el paso mucho más difícil de erigir principios generales para determinar un rango total de tales disposiciones para cualquier conjunto (finito) de oraciones generalmente aceptadas como verdaderas. Ciertamente nadie ha llegado a acercarse ni siquiera a un esquema de lo que sería una tal teoría holista del significado (*Palabra y objeto* de Quine se ocupa de la traducción y no de la teoría del significado). Hay una tendencia entre los holistas a hablar como si hubiesen prescindido de la necesidad de una teoría del significado enteramente, como si, habiendo demostrado, según creen, que el significado está contaminado por la teoría (de los hablantes), hubiesen con ello demostrado que no necesitamos en absoluto tener ningún trato con la noción. Esto, si es correcto, implicaría que todo lo que necesitamos que se nos diga es la teoría que un hablante, o una comunidad lingüística, sostiene en un momento determinado —el conjunto de oraciones aceptadas como verdaderas— y con ello seremos capaces de entender esas oraciones y también cualesquiera otras del lenguaje, habiéndose desvanecido como insostenible la distinción entre una teoría interpretada y una teoría ininterpretada. Esto es una completa ilusión. Dado simplemente un puñado de oraciones aceptadas como verdaderas por todos los hablantes, o por cualquier hablante, no hay forma en absoluto de llegar a cualquier interpretación única de esas oraciones, o de otras oraciones del mismo lenguaje. A la impresión contraria se llega, supongo, de alguna manera similar a ésta: Supongamos que consideramos un conjunto de oraciones que corresponde a las que todos los hablantes contemporáneos de castellano considerarían verdaderas, salvo que en todo lugar ‘mesa’ reemplaza a ‘águila’ y viceversa; entonces la información de que éstas son las oraciones consideradas verdaderas por los hablantes de un lenguaje bastaría para decirnos qué significan (como podríamos decir laxamente), en ese lenguaje, las palabras ‘mesa’ y ‘águila’; pues el conjunto incluiría oraciones tales como ‘Las águilas, cuando están sobre sus patas, tienen superficies superiores

horizontales planas', 'Las mesas hembras ponen huevos', etc. Pero, naturalmente, estamos aquí apelando tácitamente al supuesto de que ya conocemos las disposiciones lingüísticas de los hablantes que se relacionan con oraciones que no contienen las palabras 'mesa' y 'águila': abusa de la suposición de que, por referencia a este trasfondo, podemos exhibir las conexiones inferenciales entre oraciones que contienen estas palabras y otras oraciones indicando simplemente qué oraciones que las contienen son interpretadas por todos como verdaderas, y de que, por referencia al mismo trasfondo, estas conexiones inferenciales bastarán para determinar las disposiciones no condicionales a asentir a o a disentir de las oraciones que las contienen. No tiene tendencia ninguna en absoluto a demostrar que haya algún medio de derivar, precisamente a partir del conjunto de oraciones generalmente aceptadas, las disposiciones lingüísticas de los hablantes. El holismo no es una doctrina que nos permita abandonar la idea de que las palabras individuales tienen sentidos; por el contrario, es una doctrina que, si es correcta, exige que consideremos que nuestras palabras tienen sentidos de un tipo mucho más complejo de lo que hemos imaginado, de un tipo, en realidad, del que hasta ahora no nos hemos formado ninguna imagen clara.

Podría objetarse que me he basado en una simple negación de la tesis, defendida por Davidson, de que, al construir una teoría del significado para un lenguaje, no tenemos nada con que avanzar, al comienzo, excepto el conocimiento de las oraciones que los hablantes creen verdaderas y de las condiciones que se dan cuando hacen tales juicios. Davidson piensa, sin embargo, en un lenguaje que, como todos los lenguajes naturales, contiene una gran cantidad de recursos indéxicos, y los juicios que proporcionan los datos incluyen muchos juicios acerca de oraciones que contienen tales recursos. La doctrina fundamental del holismo, tal como la formulé, no es válida para las oraciones que son, por cierto, juzgadas verdaderas sólo en cuanto emitidas en una ocasión particular; nadie podría mantener que ningún rasgo de nuestra conducta lingüística discriminaría entre esas oraciones y las que un partidario de una teoría fregeana del significado

pretendería que son analíticas. Más aún, los datos de Davidson incluyen información sobre las condiciones prevalentes cuando se hizo originalmente un juicio, de manera que, aunque el propio Davidson es holista, toma como parte de sus datos lo que, en el caso de una oración sin rasgos indéxicos que llega a ser generalmente aceptada, Quine declara que es una irrelevante pieza de información histórica. La tesis de Davidson es totalmente diferente de la que yo estoy negando. Lo que yo estoy negando es que se puedan derivar, del conocimiento de que un cierto conjunto de oraciones —necesariamente sin rasgos indéxicos— comprende todas las aceptadas como verdaderas por todos los hablantes de un lenguaje, sin más información ulterior acerca de las condiciones que llevan a su aceptación, las disposiciones lingüísticas de los hablantes, o cualquier cosa que sea posible tomar como una interpretación del lenguaje. No digo que algún holista haya jamás propuesto explícitamente esta tesis; pero es la tesis que tendría que sostener para que las doctrinas fundamentales del holismo constituyesen un fundamento para decir que podemos prescindir de la noción de significado o que esta noción ha sido engullida por la de la teoría de los hablantes, y es patentemente falsa.

Por ejemplo, descendamos de las alturas y preguntemos qué pasa, en el contexto de una visión holista del lenguaje, con la concepción del sentido de un nombre propio. Es corriente que los holistas respondan a esto diciendo que, puesto que ya no consideramos que ninguna oración generalmente aceptada tenga un *status* privilegiado, para retener la noción de sentido debe interpretarse como dada por la totalidad de todas las oraciones que contienen el nombre y que son generalmente aceptadas como verdaderas; no puede haber una selección, entre esas oraciones, de las que son constitutivas del sentido del nombre, *i.e.* contribuyen a determinar su referencia, en cuanto opuestas a las que meramente registran lo que creemos que es verdadero de su portador. Esto puede que parezca razonablemente plausible de los nombres de objetos no o ya no accesibles a observación, aunque implausible de otros nombres; pero no da el sentido tal como éste tendría que ser presentado en una

teoría del significado genuinamente holista del tipo que he estado discutiendo, pues éste sería un sentido que no determinaría la verdad de ninguna oración particular que contuviese el nombre, sino que proyectaría cualquier conjunto dado de tales oraciones sobre las disposiciones lingüísticas apropiadas. Por el contrario, la noción de sentido ofrecida por el holista ejemplifica la suplantación del significado por la teoría de los hablantes, y cambiará según cambie esta teoría. La sugerencia ejemplifica de hecho en forma muy notable la justicia de mi afirmación de que los holistas tienden a sufrir la ilusión de que han mostrado cómo podemos pasar sin ninguna teoría del significado, mientras que meramente han hecho mucho más difícil ver cómo construir una.

Pero, tal vez se diga, todo lo que esto demuestra es que una aplicación irrestricta de la doctrina fundamental del holismo llevaría a una teoría del significado de complejidad inmanejable: no demuestra que esta doctrina no haya de aplicarse correctamente a ciertas oraciones, p. ej., a las que podrían usarse al decirle a alguien la referencia de un nombre propio, no que, aplicada así, tuviera alguna de esas extrañas consecuencias. ¿No es plausible en sí mismo el que no podamos hacer ninguna distinción efectiva entre el sentido de un nombre propio y la totalidad de lo que es generalmente creído acerca de su portador? Y, si es así, ¿no es ilegítimo objetar a esta doctrina el que no podríamos aplicar simultáneamente esa tesis a palabras de todos los tipos?

Indudablemente los nombres propios proporcionan el caso más difícil para la distinción entre sentido y referencia, razón por la cual las discusiones de esa distinción tienden a concentrarse en ellos; pero tales discusiones disminuyen de valor según la dificultad de generalizarlas a otras palabras, y, en el alegato que acabo de anunciar, esa generalización ha sido prejuzgada. Hay, en realidad, cierta plausibilidad en la afirmación; pero, para comprobarla, debemos apelar de nuevo a nuestra distinción entre saber que una oración es verdadera y conocer la proposición expresada por ella. Si un niño que nunca ha oído de Milán ve un encabezamiento de un periódico 'Huelga de correos en Milán', no sabe que el personal de correos está en huelga en Milán, sino sólo

que la oración 'El personal de correos está en huelga en Milán' es verdadera, o que el personal de correos está en huelga en un lugar llamado 'Milán'. ¿Qué tendría que aprender él para que nosotros dijésemos que conoce la proposición? No, ciertamente, *todo* lo que es generalmente creído acerca de Milán, p. ej., que S. Ambrosio fue su obispo, que primero los Viscontis y luego los Sforzas fueron sus duques, etc., etc. Incluso en el caso de los nombres propios, hay lugar para la distinción entre la explicación *standard* de su referencia y la provisión de información *standard* acerca de su portador: sólo una obsesiva adhesión a una *teoría* podría darnos alguna razón para tratar de negar esto.

La llamada teoría causal de la referencia ataca la distinción sentido/referencia desde la dirección opuesta; para ella, lo que se diría ordinariamente al explicar un nombre propio no tiene papel alguno en la determinación de la referencia de ese nombre, puesto que podría resultar falso. Por ejemplo, la manera *standard* de explicarle el nombre 'Edmund Gibbon' a alguien no familiarizado con él sería decir que Gibbon escribió un libro llamado *La decadencia y caída del imperio romano*; podríamos, sin embargo, descubrir que este libro no fue después de todo escrito por Gibbon. Se considera por tanto que lo que determina la referencia de un nombre no es nada que sea mostrado por el modo en que el nombre podría explicarse, sino la existencia de una cadena causal que conecta un uso dado de un nombre con su introducción original. Para evitar dificultades acerca del caso en que un individuo es llamado por otro, se establece entonces que cada emisión particular que forma un eslabón en esta cadena causal debe hacerse con la intención de preservar la referencia del nombre; pero entonces han de hacerse ulteriores salvedades para tener en cuenta la posibilidad de que tal intención pueda a veces frustrarse, de que pueda haber una transferencia involuntaria del nombre de un portador a otro; esto, se explica, puede suceder porque, aunque hay una intención de preservar la referencia, es acompañada por alguna otra intención prevalente. Esta posibilidad ha sido admitida en el caso de nombres de lugares (como en el excelente ejemplo debido a Evans de 'Madagascar', originalmente el

nombre de parte del continente, transferido por error a la isla); es difícil ver sobre qué base se podrían entonces excluir casos de transferencia del nombre de una persona durante su vida (por ejemplo, del cambio de un niño por otro), o, realmente, después de su muerte (está lejos de ser evidente, en el propio ejemplo de Kripke, que si de hecho 'Goliat' fue originalmente el nombre de otro filisteo, y no del gigante que David mató, nosotros, al usar el nombre para referirnos al último, estemos diciendo algo que debemos retirar). Una vez que se hacen todas estas concesiones, y dado que se concuerda en que el criterio de lo que determina la referencia es lo que los hablantes aceptarían como determinándola en un caso de conflicto, la teoría causal de ser un rival de la concepción de Frege de que la referencia de un nombre propio es determinada por su sentido se contrae en una propuesta, para una clase restringida de nombres propios, acerca de lo que debiera tomarse como un importante ingrediente de sus sentidos.

La teoría causal, en una forma madura, hace imposible distinguir entre conocer el uso de un nombre propio y simplemente haber oído el nombre y reconocerlo como nombre, y por tanto entre conocer una proposición y conocer la verdad de la oración que la expresa, cuando esto depende de conocer un nombre que figura en esa oración. ¿Qué diferencia introduce el que alguien sea capaz de dar la explicación *standard* del nombre 'Edmund Gibbon' si esa explicación puede ser factualmente incorrecta sin afectar a su captación del nombre? Supongamos que oigo por casualidad a alguien decir, 'Edmund Gibbon fue acogido en la Iglesia Católica' y que esta es la primera vez que he oído el nombre y no sé nada del tema de la conversación: si la teoría causal es correcta, yo no obstante entiendo implícitamente qué determina la referencia del nombre y podría usarlo yo mismo con la misma referencia meramente pretendiendo hacerlo. ¿En qué forma, pues, no capto el uso del nombre? ¿En qué forma estoy peor que alguien que pueda dar la explicación *standard*, particularmente si, como sucede, esa explicación *standard* es incorrecta?

La diferencia es ciertamente que si uso el nombre 'Edmund Gibbon' no sabiendo más sobre su portador que lo que aprendí de ese fragmento de conversación, se necesitaría retroceder a los hablantes particulares a quienes lo oí casualmente para encontrar de quién se estaba hablando; pero, si alguien está en posesión de la explicación *standard* del nombre, entonces nunca consideraríamos relevante seguir el modo particular en que oyó por vez primera el nombre, aun cuando se descubriese que la explicación *standard* es incorrecta, *i.e.* que no fue Gibbon quien escribió *La decadencia y caída*. Es por esta razón por lo que diríamos que, en el primer caso, yo no sabía que Edmund Gibbon fue una vez acogido en la Iglesia Católica, sino que sólo sabía que esto era verdadero de alguien llamado 'Edmund Gibbon'; aun cuando se admite que, en muchos casos, la explicación *standard* de un nombre no ha de tomarse como definitivamente verdadera de su portador, so pena de que el nombre sea privado de referencia, todavía puede que un conocimiento de esa explicación *standard* sea un ingrediente esencial de la captación del uso del nombre como palabra del lenguaje común. En la medida en que sea correcto decir que la línea de descendencia causal de un nombre determina su referencia, es la línea de descendencia causal de ese nombre en su uso generalmente admitido lo que cuenta, y no la línea de descendencia causal de cualquier antigua emisión del nombre hecha con la intención de preservar la referencia, fuera la que fuese. Kripke ha aducido, correctamente, que el punto de que Gibbon podría resultar no haber sido el autor de *La decadencia y caída* no puede salvarse declarando que el sentido del nombre 'Edmund Gibbon' es 'el hombre del que generalmente se cree que ha escrito *La decadencia y caída*', puesto que la creencia general a la que así se apela es la creencia de que Edmund Gibbon escribió *La decadencia y caída*, y tener tal creencia presupone una captación del nombre. Pero esto no debe oscurecer el hecho de que aquello cuya línea de descendencia causal debemos seguir es la creencia, no el nombre mismo.

Lo que hace posible, incluso para alguien que no sabe de Gödel más que él ha probado la incompletud de la arit-

mética, contemplar el supuesto de que podría resultar que no fue él quien lo hizo, es su conciencia de que hay otros medios, no en disputa, de identificar el portador del nombre 'Gödel'. Es por esta razón por lo que la noción de saber, de Gödel, que probó la incompletud de la aritmética tiene más substancia que la de saber que 'Gödel' es el nombre del hombre que probó ese teorema. Lo mismo, realmente, vale para el nombre 'Gibbon'. Pero aun cuando un nombre retenga una asociación con sólo un acto, p. ej. 'Obadías' con la composición de la profecía del Viejo Testamento, podemos suponer que esta tradición nos ha llegado del tiempo en que vivió el autor; y, puesto que la atribución original de la que esa tradición surgió podría haber sido errónea, pero habría tomado substancia de los otros medios, que entonces estaban a mano, de referirse al individuo en cuestión, es inteligible, y plausible, sostener que, aun en tal caso, el nombre en nuestros labios se refiere al autor supuesto en vez de al real. Pero aquí es la *tradición* la que conecta nuestro uso del nombre con el hombre; de dónde provino el propio *nombre* efectivo por vez primera tiene poco que ver con ello. Por esto es por lo que un caso como el de Goliat es enteramente diferente. Aquí la tradición tiene que ver con el gigante filisteo; no ha de suponerse que, en el momento, alguien tomó a un hombre diferente por aquél a quien David había matado con su honda, sino sólo que, entonces o en una fecha posterior, tomaron su *nombre* incorrectamente: y así la referencia del nombre, tal como *nosotros* lo usamos, es al hombre a quien la tradición se refiere, y no al que llevó ese nombre cuando vivía.

La teoría causal de la referencia incorpora una genuina intuición del modo en que se determina la referencia de una clase especial de nombres propios: nombres de personas, animales, barcos, etc., que fueron objetos de conocimiento directo, *i.e.* objetos posibles de ostensión, cuando se introdujeron sus nombres, pero que ya no están en existencia. Incluso para éstos he aducido que la teoría no es enteramente correcta tal como está, y que es en cualquier caso sólo una teoría *acerca de* los sentidos de tales nombres en vez de una que *reemplace* el sentido por algo diferente. Lo

que ciertamente no hace es dar una explicación del funcionamiento de los nombres propios en general. Estamos por cierto especialmente interesados por los nombres propios personales, lo cual es una razón por la que los filósofos que discuten sobre nombres propios tienden a usar nombres de personas en la mayoría de sus ejemplos. Pero no debemos olvidar que los nombres de personas tienen varios rasgos peculiares suyos que no son compartidos por los nombres propios de objetos de otros tipos, y que éstos son muy diversos: nombres de lugares, nombres de meses y días de la semana, nombres de carreras (como la 'Grand National'), nombres de juegos, de salidas en ajedrez, de estrellas y constelaciones, de teoremas matemáticos, de poemas, tipos de baile, religiones, vientos, enfermedades, escrituras y lenguajes, huracanes, exposiciones, guerras y tratados, remolinos, teorías científicas. De muy pocos de éstos es siquiera inteligible decir que podrían resultar desmentir los medios que usaríamos para explicar sus nombres a alguien que no los conociese; y de muy pocos tipos de nombre es plausible sostener que su referencia, tal como las usamos, depende de para qué fue introducido originalmente el nombre. Creo que sería un grave error suponer que, con la teoría causal, han sido descubiertos al final los mecanismos de referencia; menos aún que ha sido demostrado que la noción del sentido de un nombre es redundante.