



**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**

FACULTAD DE FILOSOFIA Y LETRAS

**NOMBRES PROPIOS
Y
REFERENCIA DIRECTA**



**FILOSOFIA
Y LETRAS**

**T E S I S
QUE PARA OBTENER EL GRADO DE:
DOCTORA EN FILOSOFIA
P R E S E N T A:**

Margarita María Valdés González Salas

MEXICO, D. F.

AGOSTO DE 1997





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

**NOMBRES PROPIOS
Y
REFERENCIA DIRECTA**

TESIS

**QUE PARA OBTENER EL GRADO DE
DOCTORA EN FILOSOFÍA**

PRESENTA

Margarita María Valdés González Salas

México, D.F., agosto de 1997

AGRADECIMIENTOS

Quiero expresar mi agradecimiento a todos quienes me alentaron para escribir esta disertación doctoral y, muy especialmente, a quienes me sugirieron reunir en este texto reflexiones sobre los nombres propios que había publicado en diversos artículos y expresado en varios escritos no publicados, pero discutidos en su oportunidad con muchos colegas y amigos. El ejercicio no pudo haber sido más iluminador y gozoso.

Entre quienes siempre estuvieron convencidos de que debía escribir una tesis doctoral quiero mencionar, antes que nadie, a mis hijos, Renata y Miguel, a quienes desde hace muchos años les parecía muy extraño que no tuviera yo el doctorado en filosofía y, a base de repetirlo, llegaron a convencerme de que, efectivamente, era muy extraño.

Quiero agradecer muy especialmente a mi amigo y asesor, Mark Platts, todo su interés, afecto e intenso trabajo a lo largo de la elaboración de esta tesis; él estaba más seguro que yo de que podría concluir la redacción de este texto en un plazo breve y, para ayudarme a lograrlo de la mejor manera, leyó con enorme cuidado y celeridad cada capítulo que le entregué; sus comentarios fueron siempre atinados, agudos, sugerentes, alentadores, y su actitud de confiada espera lo que más me estimuló para llegar al final de este trabajo.

No puedo dejar de mencionar aquí a Paulette Dieterlen, amiga y Jefa de la División de Estudios de Posgrado de la Facultad de Filosofía y Letras, quien, conociendo mi especial torpeza para llevar a cabo procedimientos burocráticos, supo evitarme papeleos y llegó a hacerme creer que era posible presentar el examen de grado sin necesidad de realizar casi ningún trámite.

Sin el apoyo, afecto y paciencia de quienes aquí menciono, este trabajo no hubiera sido posible.

INDICE

INTRODUCCIÓN 1

CAPÍTULO I. Dos concepciones de la referencia en el Cratilo.

Introducción	11
1. Dos tesis sobre los nombres	13
2. El "naturalismo" como una especie de descripcionismo	19
3. Relación referencial y naturaleza de los nombres en T1 y T2	26
4. La imposibilidad de la referencia equivocada	30
5. Dos concepciones del lenguaje y de la referencia	34
6. Dos intuiciones sobre los nombres	38

CAPÍTULO II. Frege: sentido y referencia indirecta.

Introducción	43
1. El argumento de Frege para introducir el sentido	47
2. Diferencias en valor cognoscitivo	52
3. La crítica de Millikan al argumento de Frege	54
4. Objeciones de Kripke a la noción fregeana de sentido de un nombre	65



**CAPÍTULO III. Referencia directa y conocimiento a priori de
verdades contingentes.**

Introducción	77
1. Contingencia, designadores rígidos y descripciones que fijan la referencia	79
2. La solución de Donnellan	90
3. Propositiones singulares <u>vs.</u> "contenidos cognoscitivos"	98
4. La solución de Evans	107

**CAPÍTULO IV. La objeción de Kripke a la teoría de la identidad
mente-cuerpo.**

Introducción	123
1. Identidad, designación rígida y modalidades	127
2. La objeción de Kripke	136
3. Concebibilidad y posibilidad	143
4. La falsa impresión de contingencia en el caso de las identidades teóricas	149
5. Referencia directa y rígida e identidades psicofísicas	153

CAPÍTULO V. Referencia directa y actitudes proposicionales.

Introducción	160
1. Enigmas kripkeanos	167
2. Sobre la posibilidad de creer proposiciones contradictorias	173
3. Predicados parcialmente definidos	186
CONCLUSIONES	203
BIBLIOGRAFÍA	208

INTRODUCCIÓN

El lenguaje natural contiene dos tipos de expresiones prima facie irreducibles: los términos singulares y los términos generales. Dentro de los primeros se encuentran los nombres propios ordinarios, los pronombres y los deícticos o demostrativos; dentro de los segundos, las expresiones predicativas. Ambos tipos de términos se combinan para formar la estructura lingüística fundamental: la oración simple, la "combinación básica", como la llamó Peter Strawson.¹ Gottlob Frege denominó "nombres propios" a los términos singulares y los consideró expresiones "saturadas", a los términos generales los llamó "palabras concepto" y los caracterizó como expresiones "no saturadas", sugiriendo, así, que requerían ser completadas por uno o más nombres propios; sostuvo que la distinción entre uno y otro tipo de expresión se debía, en último término, al hecho de que denotaban diferentes tipos de entidades: los nombres propios, según Frege, denotan objetos (entidades saturadas o completas) y los términos concepto denotan conceptos, entidades que buscan ser completadas por uno o más objetos.²

¹ Véase, Strawson, P.F., (1974), cap. 1.

² La distinción entre términos singulares y generales no coincide exactamente con la clasificación que hace Frege; éste coloca a las descripciones definidas del lado de los nombres. Véase, por ejemplo, Frege, G., (1892) y su artículo "Sobre concepto y objeto" en Frege, G., (1971).

Podemos trazar la distinción entre los términos singulares y los generales desde diferentes puntos de vista: (i) desde una perspectiva semántica, decimos que los términos singulares designan individuos, en tanto que los términos generales se aplican a individuos, son verdaderos o falsos de ellos y, en caso de designar algo, sus designata son propiedades, no individuos; (ii) desde un punto de vista sintáctico, los términos singulares se distinguen por poder ocupar la posición de sujeto en la oración y por poder ceder su lugar a una variable individual cuantificable; las expresiones predicativas, en cambio, no pueden ocupar la posición de sujeto gramatical, se forman sustrayendo de una oración simple uno o más términos singulares y nunca pueden ceder su lugar a una variable individual; (iii) desde un punto de vista pragmático, los términos singulares y los generales se distinguen porque permiten al hablante realizar diferentes tipos de actos de habla: los términos singulares le permiten al hablante referirse a individuos, hablar acerca de ellos, en tanto que los términos generales le permiten describirlos, decir algo sobre ellos.³ Expresiones tales como " __ es calvo", " __ es mayor que __", son términos generales, en tanto que "Voltaire", "ella" "Cicerón" son términos singulares. Los predicados o términos generales, como su nombre lo indica, son tales que por principio pueden aplicarse a muchos individuos, y cuando efectivamente lo hacen decimos que el individuo al que se aplican satisface el predicado en cuestión o que el predicado es verdadero de él. Los nombres, en cambio, están en lugar de un individuo y no son ni verdaderos ni falsos de su portador.

La distinción anterior alude, pues, a tres aspectos: el sintáctico o gramatical,

³ Para la distinción entre los dos tipos de expresiones (términos singulares y generales, sujetos y predicados), véase, Strawson, P., (1959), Parte II, cap. 5, Evans, G., (1982), Introducción, y Geach, P., (1962), cap. 2.

el semántico y el pragmático. Los términos singulares pueden ser sujetos de la oración y ceder su lugar a una variable, designan (o están en lugar de) individuos y son usados por los hablantes para referirse a individuos, para hablar sobre ellos; los términos generales, en cambio, no pueden ser sujetos gramaticales ni ceder su lugar a variables individuales, están en lugar de propiedades, y son usados por los hablantes para describir a los individuos, de manera que pueden ser verdaderos o falsos de ellos. La filosofía del lenguaje y la semántica filosófica suelen dedicar apartados especiales para tratar cada uno de estos tipos de expresiones. El dedicado a explorar las características de los términos singulares, su funcionamiento y naturaleza semántica, constituye la teoría de la referencia. La presente disertación se inscribe dentro de la teoría de la referencia y tiene por objeto investigar la manera como los nombres propios ordinarios alcanzan a sus objetos, es decir, estudiar el modo como se conecta el nombre propio y el individuo por él nombrado. La explicación que demos de esa conexión semántica tendrá, como veremos, importantes repercusiones en diversas discusiones en varios campos de la filosofía.

Históricamente, la relación referencial entre los nombres propios ordinarios y sus designata se ha concebido de dos maneras diferentes: (i) de acuerdo con la primera, los nombres se conectan directamente con sus referentes por la intención referencial del hablante que introduce el nombre, y dicha relación se conserva en virtud de algún acuerdo o convención social; (ii) según la segunda, en cambio, los nombres se conectan sólo indirectamente con sus referentes, a través de algún contenido descriptivo, representacional o conceptual, el cual sólo es verdadero del individuo designado por el nombre o sólo él satisface. Uno de los objetivos centrales del presente estudio es ofrecer una defensa de la que considero ser hasta la fecha la teoría más satisfactoria de la referencia directa de los nombres propios,

a saber, la concepción de los nombres propuesta por Saul Kripke en su trabajo El nombrar y la necesidad.⁴ La defensa que ofreceré procederá en dos tiempos: primero, examinaré críticamente los argumentos que pretenden establecer la verdad de la teoría de la referencia indirecta y trataré de mostrar su insuficiencia; luego, presentaré los rasgos centrales de la concepción de los nombres propios ordinarios sostenida por Kripke, consideraré tres posibles objeciones a dicha concepción e intentaré responder a cada una de ellas.

El estudio que aquí presento parte de la consideración de dos intuiciones sobre los nombres que se hallan claramente expresadas en un texto tan antiguo como lo es el Cratilo de Platón. Esas dos intuiciones -a las que en el diálogo platónico dan voz Cratilo y Hermógenes respectivamente-, se hallan a la base de sendas teorías sobre la naturaleza semántica de los nombres. La teoría de la referencia indirecta se halla representada en su forma más extrema por la postura del viejo Cratilo, según el cual los nombres tienen en sí mismos un contenido descriptivo en virtud del cual se aplican "naturalmente" a sus referentes. Según Cratilo, no es, pues, necesaria ninguna convención que conecte al nombre propio con su referente, ya que ambos están desde siempre "naturalmente" conectados. Aunque la posición de Cratilo resulta insostenible, argumentaré que anuncia de alguna manera lo que en la filosofía del lenguaje de hoy se conoce como la teoría descriptivista de los nombres.⁵ Hermógenes, en cambio, centra su atención en la "arbitrariedad" del nombrar y defiende una teoría de la referencia directa de los nombres: lo que hace que un nombre N refiera a un individuo x es simplemente -

⁴ Kripke, S. (1980)

⁵ Véase Kripke, S., (1980). Al final de la Primera Conferencia y principio de la Segunda, Kripke caracteriza magistralmente las teorías descriptivistas de los nombres propios.

dice Hermógenes- que un hablante h use N con la intención de referirse a x y, luego, que esa especie de bautismo se acepte como la norma que haya de gobernar el uso correcto del nombre en una comunidad lingüística. Para Hermógenes no hay más relación semántica que la relación que se establece convencionalmente. Para Cratilo, en cambio, la relación semántica por excelencia -la única a fin de cuentas- es la relación de satisfacción, el nombre sólo puede referir a un individuo a través de la satisfacción de un conjunto de predicados.⁶

Las teorías que consideran que la relación referencial entre el nombre y su nominatum es una relación indirecta parecen sostener que, malgré las apariencias, los nombres propios no constituyen estrictamente hablando una categoría semántica diferente de la de los predicados o términos generales; es decir, que no hay realmente términos singulares, pues la manera como los nombres se relacionan con sus designata es a través de la satisfacción de predicados o, más generalmente -en la sugerente terminología de Genoveva Martí- en virtud de que "cumplen con un determinado perfil cualitativo o representacional"⁷. Tal vez sea Bertrand Russell quien con mayor claridad haya sostenido en épocas recientes que los nombres propios ordinarios no llevan a cabo funciones semánticas diferentes de las que

⁶ Tarski introduce la noción semántica de satisfacción para poder definir la verdad de todas las oraciones complejas en términos de alguna propiedad de sus componentes. Como las oraciones cuantificadas tienen como componentes funciones oracionales (o predicados) que no pueden ser ni verdaderas ni falsas, la propiedad buscada por Tarski no puede ser la verdad. La satisfacción es una relación entre funciones oracionales (o predicados) y secuencias de objetos. Para el caso de las oraciones simples en la que figuran un nombre "N" y un predicado Φ se dirá que el objeto designado por "N" satisface el predicado Φ si y sólo si N es Φ . Para una excelente explicación de la estrategia de Tarski y de la noción semántica de satisfacción, véase Platts, M., (1979), cap. I, sec. 4.

⁷ Véase, Martí, G., (1995).

llevan a cabo los predicados o términos generales al sostener que los nombres propios ordinarios no son más que descripciones definidas disfrazadas.⁸ Según este modo de ver las cosas no habría una diferencia semántica sustancial entre la referencia, propia de los nombres ordinarios, y la satisfacción, propia de los predicados, pues la referencia de todo nombre propio se reduciría, a final de cuentas, a la satisfacción de un conjunto de predicados.

Luego de contrastar y discutir las dos concepciones sobre los nombres que Platón explora en el Cratilo, pasaré a revisar, en el capítulo II, ciertos aspectos de una versión moderna, infinitamente más precisa y sofisticada que la propuesta por Cratilo, de la teoría de la referencia indirecta de los nombres. Me refiero, desde luego, a la teoría de los nombres que en diferentes escritos propone y defiende Gottlob Frege. En las primeras tres secciones de ese capítulo, centraré mi atención en los argumentos que ofrece Frege para introducir la noción de "sentido" de un nombre y con ella su semántica de dos niveles.⁹ Frege parece cometer una petición de principio en su argumentación; en efecto, parece dar por supuesto algo que tenía que demostrar: que dos nombres correferenciales diferentes hacen una contribución diferente al significado de las oraciones en las que figuran (o tienen

⁸ Véase, especialmente, Russell, B., (1905) y (1910-11). Russell, desde luego, acepta la existencia de expresiones genuinamente referenciales, pero estas no son los nombres propios ordinarios, sino los nombres propios lógicos, esto es, partículas egocéntricas tales como "ahora", "aquí", deícticos que refieren a los propios datos sensoriales y, tal vez, el pronombre "yo". Los nombres propios ordinarios, en cambio, son semánticamente equivalentes a descripciones definidas. "Lo que significa una proposición sobre Apolo -dice Russell- es lo que obtenemos cuando sustituimos [el nombre por] lo que, según los diccionarios de historia antigua, quiere decir Apolo, digamos, "el dios solar"', Russell, B., (1905).

⁹ La expresión "semántica de dos niveles" es de Gareth Evans, véase Evans, G., (1982), cap. I, I.1.

diferente "contenido"). Rechazaré este supuesto fregeano, y argumentaré que hay una manera mucho más económica de explicar los fenómenos lingüísticos que Frege pretendía explicar con su introducción de "sentidos". La última sección del capítulo II está dedicada a examinar las objeciones devastadoras de Saul Kripke a cualquier teoría que sostenga -como la de Frege- que los nombres refieren sólo a través de cierto contenido descriptivo o representacional que permite identificar una única cosa, que el referente del nombre es la única cosa que satisface dicho contenido, y que si hay varias cosas que satisfagan ese contenido o no hay nada que lo satisfaga, el nombre no refiere.

Una vez que haya dado razones para considerar insuficientes los argumentos de Frege para introducir los "sentidos" (en el sentido que da a ese término Frege), y que haya mostrado, además, que de la aceptación de la hipótesis fregeana se seguirían consecuencias falsas, pasaré a explorar una versión particular de la concepción de los nombres propios según la cual éstos se conectan directamente con sus referentes. La versión de esa teoría que elijo examinar, como mencioné antes, es la propuesta por Saul Kripke en su libro El nombrar y la necesidad.

En los capítulos III, IV y V de esta tesis, presentaré diferentes aspectos de la concepción kripkeana de los nombres propios ordinarios; a la vez, en esos mismos capítulos discutiré las tres grandes dificultades que dicha concepción tiene que enfrentar: (i) la aparente consecuencia de que es posible conocer a priori proposiciones (o hechos) contingentes, (ii) la aparente incompatibilidad de las tesis de Kripke sobre los nombres propios con la teoría materialista de la mente y (iii) la necesidad de explicar qué pasa en las atribuciones de actitudes proposicionales en las que aparentemente falla el Principio de Sustitución de Idénticos. Como es bien sabido, según este principio si tenemos dos nombres propios correferenciales, digamos "N" y "N'", entonces, la sustitución de "N" por "N'" en una oración

verdadera dará como resultado otra oración verdadera; sin embargo, -se alega-, los contextos de creencia son una excepción a tal principio, pues al parecer es posible que un sujeto crea, por ejemplo, que Voltaire es un filósofo y no crea que François Marie Arouet es un filósofo, siendo que "Voltaire" y "François Marie Arouet" designan al mismo individuo. Esa supuesta falla de del principio se ha solido considera como un elemento de prueba en contra de la teoría de la referencia directa.

Las tres dificultades anunciadas serán el tema de los capítulos III, IV y V respectivamente. En el capítulo III expongo primero la tesis de Kripke según la cual los nombres propios son designadores rígidos, así como su distinción entre categorías epistémicas y metafísicas y su noción de "fijar" la referencia mediante una descripción; luego examino su argumento para sostener que de sus propias tesis se sigue la posibilidad de conocer a priori verdades contingentes. En las secciones subsiguientes examino la discusión de Keith Donnellan y de Gareth Evans acerca de la existencia de supuestos conocimientos a priori de verdades contingentes. En la parte final del capítulo ofrezco mi propia respuesta a la dificultad y concluyo que en el tipo de situación en el que Kripke cree ver la posibilidad de que un sujeto conozca a priori un hecho contingente, hay, a veces, un mero conocimiento empírico y, a veces, ningún conocimiento en absoluto.

En el capítulo IV parto del examen de ciertas tesis de Kripke sobre las modalidades y la relación de identidad, así como de su interpretación de las oraciones de identidad en las que figuran nombres propios, para discutir su objeción a las teorías materialistas de la mente. Kripke equipara las identidades que típicamente aseveran los materialistas, tales como "todo estado mental es un estado cerebral" o "el dolor es la estimulación de las fibras C", con las identidades teóricas típicas de la ciencia, tales como "la luz es un haz de fotones" o "el calor

es la energía molecular cinética media", las cuales expresan identidades necesarias. Kripke argumenta -y esa es su objeción-, que no podemos eliminar el aire de contingencia que tienen las identidades de los materialistas de la misma manera como podemos explicar y eliminar el aire de contingencia que tienen las identidades teóricas de la ciencia. De modo que concluye que la aparente contingencia en el caso de las identidades sostenidas por los materialistas no es sólo una apariencia; esto es, que realmente son contingentes y, dado que no puede haber identidades contingentes, que las supuestas oraciones de identidad materialistas deben ser falsas. Dedico las tres últimas secciones de ese capítulo a examinar y discutir la objeción de Kripke y descubro una omisión en su argumento antes esbozado. Si mi argumentación es correcta, la teoría de la referencia rígida y directa de los nombres no es incompatible con una postura materialista en filosofía de la mente.

El capítulo V está dedicado a examinar el comportamiento de los nombres propios en contextos de actitudes proposicionales y a tratar de encontrar una explicación de por qué aparentemente falla el Principio de Sustitución de Idénticos en ese tipo de contextos. Si la teoría de la referencia directa es correcta y la única contribución semántica de un nombre al significado de la oración en la que aparece es su referente, entonces, parece que tiene que valer de manera general el Principio de Sustitución de Idénticos. En su famoso artículo "Un enigma sobre la creencia",¹⁰ Kripke ofrece una defensa indirecta del Principio de Sustitución, pero nos deja ante un dilema: o bien aceptamos que es posible en situaciones como las que él describe en su mentado artículo atribuir a un sujeto aparentemente racional la creencia en una contradicción, o bien tenemos que aceptar que hay situaciones

¹⁰ Véase, Kripke, (1979a).

en las que, a pesar de que un sujeto racional asienta sinceramente a ciertas oraciones, no es posible atribuirle ninguna creencia. En ese último capítulo investigo ambas posibilidades y ofrezco razones para preferir el segundo cuerno del dilema. Eso, a su vez, me obliga a desarrollar en la parte final la tesis de que el predicado de creencia es un predicado parcialmente definido. Un predicado parcialmente definido, como veremos, es tal que puede haber individuos respecto de los cuales el predicado ni se aplica ni no se aplica. En caso de ser el predicado de creencia un predicado parcialmente definido podrá dar lugar a "vacíos" de valor de verdad, es decir, a oraciones "indefinidas" que no son ni verdaderas ni falsas.

Los tres capítulos en los que exploro las dificultades antes señaladas constituyen el meollo de esta disertación. Sólo después de mostrar que las supuestas consecuencias problemáticas no son en realidad tales, podemos permitirnos ver la teoría de la referencia directa como un instrumento valioso que nos permite comprender el funcionamiento de los nombres propios en el lenguaje natural.

CAPÍTULO I

Dos concepciones de la referencia en el Cratilo¹

Introducción.

Las ideas de Platón sobre los nombres propios, sobre su naturaleza y funcionamiento, se encuentran expuestas en el Cratilo. En este diálogo Platón no se plantea el problema de la referencia de los nombres propios como algo distinto del problema de la significatividad de los nombres comunes o términos generales; es decir, no se propone investigar el tema de la referencia de los nombres como un problema específico concerniente sólo a los términos singulares, sino como parte de un problema mucho más general. El tema del Cratilo, podríamos decir, es el de la significación lingüística en su forma más general: ¿qué es lo que hace que algunos sonidos que emitimos sean significativos? ¿qué es lo que determina que una expresión lingüística se aplique o no a un individuo? ¿en virtud de qué refiere un nombre? En este sentido, podríamos considerar este diálogo como uno de los primeros intentos de ofrecer un tratado general de semántica filosófica en

¹ Una versión anterior de este capítulo fue publicada con el título "Dos teorías de la referencia en el Cratilo" en *Análisis Filosófico*, VII (1987), nº 2.

la historia de la filosofía.² Sin embargo, el hecho de que Platón haga uso en su discusión de innumerables nombres propios para ilustrar las tesis que analiza, nos autoriza a sostener que, además del problema de la significación lingüística en general, le interesa de manera muy especial el problema de la relación entre los nombres propios y sus nominata, es decir, el de la referencia de los nombres y el modo como ésta se determina.³ La cuestión central a la que Platón pretende responder en este diálogo es la siguiente: ¿qué condiciones han de satisfacerse para poder afirmar con verdad que un nombre N refiere efectivamente a un individuo particular x? En la terminología de Sócrates, la pregunta se plantearía así: "¿Qué hace falta para que un nombre sea correcto?" o "¿En que consiste la corrección (orthótes) de un nombre?"

El rasgo que me interesa destacar en relación con el Cratilo es que en él encontramos prefiguradas por primera vez dos maneras de concibir la relación entre nombres y objetos nombrados: (i) una según la cual el nombre refiere

² Como antecedentes del Cratilo podemos señalar, entre otros, la teoría de Demócrito, quien fue el primero en sostener una tesis convencionalista del significado, la tesis de Antístenes sobre la imposibilidad del discurso falso y las doctrinas de otros sofistas, como Gorgias y Protágoras, que se ocuparon también de los nombres. Véase Goldschmidt, V., (1982), esp. "Introduction Historique", pp.5-35. Véase también la "Notice" en la edición del Cratyle preparada por Louis Méridier, Societé d'edition "Les Belles Lettres" (Budé), Paris, 1931, esp. sección IV.

³ Cabe pensar también que Platón no ve como dos problemas diferentes el de explicar la relación referencial, propia de los términos singulares, y el de explicar la significatividad de los términos generales. Su teoría de las Formas, en tanto que referentes de los términos generales del lenguaje, es una prueba de ello. Robinson, R., (1969), da buenas razones para traducir "ónoma" por "nombre" y considerar que en su sentido más común la palabra significa "nombre propio"; "al oír un griego la palabra "ónoma" -dice Robinson- lo primero que venía a su mente era un nombre propio", p. 101.

directamente a su nominatum, es decir, según la cual no se requiere ningún tipo de intermediario descriptivo, ni representacional, ni conceptual, para que se establezca la relación referencial entre un nombre y el individuo nombrado; y (ii) otra que concibe los nombres como términos provistos de un contenido descriptivo o representacional y ve la relación referencial como una relación indirecta o mediada entre el nombre y su denotatum. Según esta segunda manera de ver las cosas, la referencia se establece entre un nombre N y un individuo x cuando el contenido descriptivo de N es satisfecho únicamente por x, o cuando la representación que encierra N se ajusta exactamente a x. Podríamos decir que esta segunda manera de concebir la referencia supone que no hay una diferencia sustantiva entre el funcionamiento semántico de los nombres y el de los términos generales del lenguaje natural;⁴ los nombres serían abreviaturas de conjuntos de predicados y la referencia no sería más que la satisfacción de esos predicados por parte de un solo individuo.

1. Dos tesis sobre los nombres.

Sócrates persigue y analiza en el Cratilo⁵ dos tesis opuestas sobre la

⁴ B. Williams expresa el contraste entre las tesis de Cratilo y Hermógenes de manera sumamente escueta y exacta: "Cratilo asimila nombrar a describir [Hermógenes] asimila describir a dar un nombre", Williams. B., (1981), p. 90.

⁵ Me he basado para las citas en la edición francesa publicada en la colección Budé del Cratyle mencionada en la nota 2. Existen dos traducciones recientes al castellano del Cratilo: la de J.L. Calvo publicada en 1983 en la Biblioteca Clásica Gredos, Madrid y la de Ute Schmidt O. publicada en 1988 en la Colección Bibliotheca Scriptorum Graecorum et Romanorum Mexicana que publica la Universidad Nacional Autónoma de México. La traducción que nos ofrece Francisco de P. Samaranch en las Obras Completas de Platón publicada por Aguilar Editores, Madrid, 1966, es una traducción literal de la francesa antes

referencia de los nombres: una tesis convencionalista y otra a la que suele llamársele "naturalista".⁶ Veamos en qué consiste cada una de ellas.

(A). La tesis convencionalista, sostenida por Hermógenes, es aquella según la cual "la corrección (orthótes) del nombre [no es] otra cosa que un acuerdo y una convención ... el nombre que uno pone a un objeto es el nombre correcto" (384c-d). Este acuerdo o convención podría cambiarse a voluntad del hablante "sin que el nombre sustituido sea menos correcto que el nombre precedente. Pues la naturaleza --dice Hermógenes-- no asigna propiamente ningún nombre a ningún objeto: es cuestión de uso y de costumbre entre aquellos que han adquirido el hábito de dar nombres" (384d). La relación referencial entre el nombre y el objeto nombrado se establece, según Hermógenes, por la intención del locutor de utilizar en su discurso un fonema determinado para designar a un individuo directamente identificado como tal. Esto es, la referencia semántica de un nombre queda determinada única y exclusivamente por la referencia que hace el hablante, es decir, por su intención de referirse a un individuo que él distingue directamente

citada, de la cual copia, además, la introducción y las notas explicativas, sin dar el crédito correspondiente a Louis Méridier. Samaranch presenta su traducción como si fuera directa del griego. En algunos pasajes he seguido la traducción francesa del Cratyle de Léon Robin publicada en la colección Pléiade, Editions Gallimard, París, 1950. Agradezco al Dr. Bernabé Navarro la ayuda inapreciable que me prestó verificando la exactitud de mis citas con el texto griego.

⁶ Como se verá en lo que sigue, se trata del convencionalismo lingüístico, es decir, la idea de que todo significado lingüístico tiene su origen en una convención y del naturalismo lingüístico que rechaza la idea de que el significado lingüístico requiera una convención. Una brillante dilucidación de la noción de convención lingüística puede encontrarse en Lewis, D., (1969). Para una discusión de la imposibilidad del convencionalismo generalizado, véase, Strawson, P.F., (1964).

de todos los demás en el momento de hacer uso de tal nombre.⁷ No hay, además, ningún intermediario entre el sonido y la cosa nombrada por el locutor.

La relación entre nombre y objeto nombrado es, según esta concepción, una relación "arbitraria" en tanto que uno puede dar cualquier nombre a cualquier cosa. En la primera etapa de la discusión, Hermógenes parece rechazar cualquier limitación al acto de dar un nombre: "yo puedo llamar a cada objeto con tal nombre establecido por mí, tú, con tal otro establecido por tí" (395 d). Es sólo después de que Sócrates le muestra que de ese convencionalismo individual se sigue la imposibilidad del discurso falso y, en último término, de la comunicación (385a-386d),⁸ que Hermógenes recurre a la segunda parte de su tesis, según la cual el "acuerdo o la convención" social, "la costumbre", es indispensable para poder decir que un nombre refiere a un objeto y hacer posible la comunicación.

La tesis de Hermógenes contiene, pues, dos subtesis. (i) La primera, que

⁷ Para la distinción entre referencia semántica y referencia del hablante o locutor, véase Saul Kripke, (1979b). Kripke argumenta, en contra de lo sostenido por Hermógenes, que la referencia del hablante no determina el valor semántico de un nombre en una lengua, aunque en el origen de todo nombre se encuentre la intención de un hablante de referirse a un objeto al usarlo. Véase, en contraste, Donnellan, K., (1979).

⁸ Si tenemos en cuenta los argumentos de Wittgenstein contra la posibilidad del lenguaje privado, podríamos inclusive objetarle a Hermógenes que su convencionalismo individualista hace imposible el lenguaje mismo. Para que haya significado lingüístico es necesario que haya regularidades, que los hablantes sigan reglas que permitan distinguir los usos correctos de los usos incorrectos de las expresiones; esto sería imposible en el convencionalismo individual, ya que si hubiera "reglas" individuales no se podrían distinguir los casos en los que un locutor sigue efectivamente una regla de otros en los que simplemente crea seguirla. Obedecer una regla es una práctica que no puede explicarse sin hacer referencia a una comunidad. Véase, Wittgenstein, L. (1953), Parte I, par. 143-242.; véase también, Kripke, S., (1982) y Peacocke, C., (1981).

podríamos llamar "convencionalismo individual", es la siguiente: todo hablante dispone, por un lado, de un conjunto potencialmente infinito de fonemas (o de grafías) y, por otro, es capaz de distinguir una multitud de individuos o de objetos en la experiencia. El hablante puede utilizar cualquier fonema con la intención de señalar cualquier objeto y esto basta para convertir al fonema en el nombre del objeto en cuestión. Más aún, el hablante puede cambiar en cualquier momento los nombres de los objetos y todos serán igualmente correctos. Según (i), cada hablante podría tener en principio un "lenguaje individual" e incorregible; sin embargo, ese supuesto lenguaje no serviría para la comunicación. (ii) La segunda subtesis, el "convencionalismo social", añade que una vez que un hablante particular ha establecido una relación nominativa entre un fonema y un objeto, es necesario el acuerdo o la convención social para consolidar esa relación referencial y hacer posible la comunicación entre los diversos miembros "de la ciudad". El acuerdo tendría dos partes: una, la intención individual del hablante que introduce el nombre de utilizar consistentemente el mismo fonema para referirse siempre a la misma cosa y, otra, el establecimiento de la convención social propiamente dicha consistente en el acuerdo tácito o explícito por parte de los miembros de una comunidad lingüística de usar todos ellos los mismos sonidos para referirse a las mismas cosas (cfr. 384d-385a).⁹

Al proponer la primera subtesis sobre los nombres, Hermógenes parece atender a la situación en la que se introduce un nuevo nombre en un lenguaje preexistente. En efecto, los nombres, cuando se introducen, no suelen tener nada en sí mismos que obligue a los locutores a usarlos para referirse a unas cosas y no

⁹ Para una clara exposición de las dos subtesis defendidas por Hermógenes, véase Kretzmann, N., (1971), esp. pp. 127-129.

a otras.¹⁰ Ese parece ser el grano de verdad que hay en la idea de que el nombrar es "arbitrario". Cabe insistir, sin embargo, que este enfoque parece presuponer la existencia de un lenguaje y el dominio del mismo por parte de quien impone el nombre -el dador de leyes o "nomothetes"-, al momento de nombrar la cosa.¹¹ La segunda subtesis, en cambio, se propone teniendo en cuenta que el nombre introducido por un hablante particular tiene que ser aceptado por la comunidad lingüística para poder servir a los fines de la comunicación. Como se verá más adelante, dar cuenta de la introducción de nuevos nombres en el lenguaje plantea problemas diferentes de los que plantea dar cuenta de la utilización correcta de un nombre previamente existente o de lo que supone el dominio de un lenguaje que contiene nombres propios.

(B). La tesis defendida por Cratilo, sostiene, por el contrario, que "un nombre no es la apelación que algunos dan al objeto después de un acuerdo...[sino que] existe de manera natural (physei) una justa denominación para cada uno de los seres que es la misma para griegos y bárbaros" (383a-b) y que la corrección o la justeza de los nombres consiste precisamente en su capacidad de "hacer ver la naturaleza de cada ser" (428e). Así, según Cratilo, hay un nombre naturalmente

¹⁰ Lo anterior no implica que no pueda haber casos en los que el hablante elija un nombre (por ejemplo, "Amadeo", "Felicidad", "Monterrey", "Wonderland") precisamente porque tiene una etimología adecuada a los deseos o expectativas de quien introduce el nombre respecto de la cosa nombrada. Esto, sin embargo, no relaciona necesariamente al nombre con la cosa que responde a su "contenido descriptivo", ya que es posible que "Felicidad", por ejemplo, nombre a una persona desdichada o que "Amadeo" nombre a un enemigo de Dios.

¹¹ Siguiendo a Wittgenstein, dar nombres es uno entre muchos otros juegos de lenguaje y sólo se es capaz de dar nombres una vez que se saben hacer otras cosas, que se saben jugar otros juegos lingüísticos, que se tiene un cierto dominio del lenguaje. Véase, Wittgenstein, L., (1957) Parte I, parags. 20-42

correcto para cada individuo y para que un nombre refiera efectivamente al individuo que pretende nombrar tiene que mostrararnos de alguna manera su naturaleza. "La esencia (ousía) del objeto --según esta tesis "naturalista"--se manifiesta en el nombre"(393d); es por esto que "el nombre de Zeus es propiamente dicho como una definición (lógos) ...[nos hace] ver claramente la naturaleza del dios, lo cual es precisamente el efecto que un nombre debe de producir" (396a).

De acuerdo con la teoría propuesta por Cratilo, el locutor no puede elegir a voluntad el nombre con el que ha de designar a un individuo. Si ha de referirse a un objeto tendrá que utilizar precisamente el nombre que muestra la naturaleza del objeto, el nombre que le corresponde. El nombre y su referente están necesariamente relacionados en virtud de cierto contenido semántico inherente al nombre y no hace falta ni un acto de habla, ni una intención referencial, ni acuerdo o convención alguna, para establecer la relación nominativa ni para conservarla.¹² Ésta, según Cratilo, existe independientemente de cualquier comportamiento lingüístico individual o colectivo; es una relación que puede ser descubierta por el nomothetes, pero no creada.¹³

¹² Cuando dice Cratilo que el nombre está necesariamente relacionado con el objeto que nombra, quiere decir el nombre en cuanto dotado de un contenido descriptivo. Es evidente que un mero sonido desprovisto de todo significado no está necesariamente relacionado con nada.

¹³ La función del legislador (nomothetes) que impone nombres sería precisamente la de descubrir, no crear, esa relación (cfr. 389a-390a). Como señala Bernard Williams, "no hay ningún acto que el nomothetes, ni nadie más, pueda realizar para convertir a 'N' en el nombre de Y, o bien 'N' tiene la relación exigida con Y o bien no la tiene"; véase Williams, B., (1981), p. 89.

Las teorías semánticas sostenidas por Hermógenes y Cratilo pretenden ser teorías generales que, como señalé al inicio de este capítulo, se aplican a todos los términos del lenguaje, singulares y generales. Haré caso omiso de los problemas que generan ambas teorías para el caso de la explicación de la significatividad de los términos generales. Intentaré precisarlas cuando se aplican a los nombres propios y ver algunas dificultades que plantean. Como veremos, la manera de conceptualizar la relación referencial propia de cada una de estas teorías es radicalmente diferente. En lo que sigue llamaré "T1" a la tesis convencionalista de Hermógenes y "T2" a la tesis "naturalista" atribuida a Cratilo. Pero antes de precisar T1 y T2, intentaré justificar la interpretación que hago de T2 como una teoría "descripcionista" de los nombres; en la bibliografía filosófica reciente sobre el Cratilo¹⁴ no se ha subrayado suficientemente este rasgo de la teoría.

2. El "naturalismo" como una especie de descripcionismo.

Al enunciar Cratilo que existe para cada cosa una denominación "naturalmente" correcta (383a) y repetir Sócrates en variados pasajes (387d, 389d, 390c, 391a, 391e) que los nombres propios convienen o no convienen "naturalmente" a sus nominata, uno no puede evitar preguntarse qué quiere decir

¹⁴ Véase Goldsmith, V., (1982), Gosling, J.B.C., (1973), Kretzman, N., (1971), Robinson, R., (1955) y (1956), Ryle, G., (1960), Schofield, M. (1981), Williams, B., (1981), y el no tan reciente pero clásico Gomperz, T., (1953), vol. III, cap. xvi, sección 4. En Ketchum, R. J., (1979) hay el intento de defender la tesis naturalista de Cratilo apelando a la idea de que los nombres tienen significado, pero la interpretación que ofrece Ketchum se aplica sólo para el caso de los nombres comunes que "pertenecen a un particular espacio-temporal si y sólo si son verdaderos de él." (p. 133). He sostenido anteriormente la interpretación que aquí sostengo en Valdés, M., 1987.

que un nombre "naturalmente" refiera a una cosa o que un nombre sea "naturalmente" correcto para un individuo. La significación lingüística, así como la referencia, parecen ser cuestiones esencialmente no naturales;¹⁵ es decir, los signos lingüísticos no son signos naturales en el sentido en el que lo es el humo del fuego; la relación entre un sonido usado para designar a un individuo y el individuo designado es de una índole diferente de las relaciones "físicas" o naturales. Más aún, el hecho de que podamos utilizar diversos sonidos en una misma lengua o en distintas lenguas para referirnos a una misma cosa es prueba de que no hay una relación "natural" entre los sonidos utilizados y las cosas nombradas.

Interpretar a Cratilo como si afirmara que existe una correspondencia natural, y necesaria, entre la grafía "Homero" y Homero parece a primera vista disparatado, pues es claro que Homero bien pudo haberse llamado de otra manera. Cratilo no puede, pues, estar considerando los nombres como meras grafías o sonidos desprovistos de significado;¹⁶ si afirma que los nombres corresponden "naturalmente" a los objetos nombrados, es porque los considera de alguna manera

¹⁵ Para una brillante dilucidación del concepto de significado no natural, véase, Grice, H. P., (1957). Grice distingue la noción de significado natural en, por ejemplo, "Las ronchas significan que tiene viruela", del significado no natural, propio de las expresiones lingüísticas (aunque no exclusivo de ellas) en, por ejemplo, "'Suavemente" significa delicadamente". Véase, también, Grice, H. P., (1968).

¹⁶ En Ketchum R., (1979), se argumenta que la tesis naturalista de Cratilo vale para el caso de los términos generales apoyándose en la idea de que éstos poseen significado en sí mismos y por lo tanto no se aplican en virtud de convenciones. Ketchum considera que los nombres propios "no son tema del diálogo". Tendría que explicar, entonces, por qué los ejemplos que ofrece Platón contienen en su mayoría nombres propios.

como cargados de contenido descriptivo. Sólo si concebimos de esta manera los nombres resulta comprensible tanto la tesis de Cratilo según la cual hay una correlación "natural" entre los nombres y sus nominata que es la misma "para griegos y bárbaros" (383b)¹⁷, como la tesis de que la función propia de los nombres es instruirnos o darnos a conocer la naturaleza de las cosas (cfr. 388c, 422d). Si suponemos que para Cratilo los nombres tienen algún contenido significativo o que de alguna manera encierran un conjunto de descripciones, entonces podemos entender la corrección "natural" del nombre como la relación de "aplicación" de ese contenido significativo a su nominatum o, dicho en otros términos, como la satisfacción por parte del objeto designado de las descripciones que constituyen el contenido del nombre. Podemos también comprender que los nombres sirvan para instruirnos, pues los nombres reflejan la naturaleza de las cosas: el mero hecho de conocer su contenido significativo nos hace conocer las propiedades de los objetos por ellos nombrados. Los nombres, sostiene Cratilo, nos muestran la esencia de sus nominata (393d) y esto sólo es posible si de alguna manera encierran la contraparte inteligible de esa esencia, esto es, un concepto individual, una descripción definida esencial o alguna suerte de definición.¹⁸

Al examinar Sócrates la tesis de Cratilo, explora, dentro de esa perspectiva, dos maneras como los nombres podrían cumplir su función referencial según que se trate de nombres derivados (o compuestos) o de nombres primitivos (simples).

¹⁷ Decir que la corrección de un nombre es la misma para griegos y bárbaros equivale a considerar que si hay varios nombres correctos para un individuo en distintas lenguas el contenido significativo de esos nombres tiene que ser el mismo. Así, si "Zeus" y "Júpiter" son nombres correctos del dios, los dos nombres tienen que tener el mismo contenido significativo. Cfr. Williams, B., (1981), p. 83.

¹⁸ Cfr. Schofield, M., (1981), p. 62.

(i) Los nombres compuestos están constituidos a partir de otros nombres, nos dice; por ejemplo, "Hermógenes" contiene el nombre de Hermes y el de origen o generación, su contenido significativo será, pues, "el del linaje de Hermes" o "el que se origina en Hermes" y su referencia está determinada precisamente por ese contenido significativo; es decir, "Hermógenes" corresponde al individuo que satisface la descripción que encierra el nombre. La parte media del Cratilo (391d-428c) está dedicada al intento de hacer plausible esta tesis ofreciéndonos para ello, no sin un permanente tono humorístico, la etimología de multitud de voces griegas:¹⁹ la etimología del nombre nos da su significado y éste determina el individuo al que habrá de referir.

(ii) La larga sección etimológica desemboca, como era de esperarse, en el planteamiento de la cuestión crucial: ¿qué es lo que determina la referencia de los nombres primitivos o simples? Cuando un nombre no puede analizarse mediante otros, ¿en virtud de qué es significativo? y ¿cómo se determina su referencia? Esto es, si la función general de los nombres es referir a las cosas en virtud de que muestran la esencia de sus referentes, ¿de qué modo muestra la esencia de su referente un nombre simple? Sócrates propone la siguiente hipótesis "onomatopéyica", aparentemente aceptada por Cratilo, para explicar cómo adquieren el contenido significativo que de hecho tienen los nombres simples y, así, cómo logran referir: "el nombre [primitivo] --dice Sócrates-- es, según parece,

¹⁹ Las interpretaciones sobre esta sección etimológica del Cratilo varían enormemente. Hay quienes no ven en ella sino un divertimento ingenioso cuyo propósito es ridiculizar la tesis defendida por Cratilo, y quienes, por el contrario, consideran que esa es la parte medular del diálogo. Compárese, por ejemplo, Goldschmidt, V., (1982), p. 8; Schofield, M., (1981), pp. 62-63; Kretzman, M., (1971), pp. 134-136.

una manera de imitar por la voz lo que uno imita y nombra" (423b).²⁰ Si esto es así, los nombres primitivos logran referir al "imitar" la esencia de los objetos nombrados y los medios para llevar a cabo tal imitación son las sílabas y las letras de que están compuestos (cfr. 423b-427c). Cada letra, sugiere Cratilo, es un instrumento apto para reproducir un concepto simple (por ejemplo, la "r" el movimiento, la "l" lo liso, la "i" lo ligero, etcétera. Cfr. 426d-427c), de manera que la unión de letras en sílabas y de éstas en nombres nos permite dotar de un contenido conceptual al nombre simple y, así, determinar su referencia.

Si no se hace intervenir el convencionalismo para explicar por qué la "r" representa el movimiento, la "l" lo liso, etcétera, la falsedad de esta última tesis resulta patente.²¹ Cabe observar, además, que Cratilo no parece estar obligado a considerar que así como hay ingredientes significativos más elementales en los nombres complejos, los hay también en los nombres simples o "primitivos" (entre otras cosas porque, en ese caso, dejarían de ser "primitivos"). Cratilo pudo haber sostenido que los nombres simples significan directamente conceptos simples y los

²⁰ Gomperz interpreta a Cratilo como si afirmara que el origen del significado se encuentra en la "asociación natural de ciertos sonidos con ciertos sentimientos", véase Gomperz, T., (1936), p. 498. En Ryle, G., (1960), encontramos también esta interpretación puramente onomatopéyica: "las letras suenan como cosas o sucesos y por ello califican para funcionar como sus nombres", p. 441. Sin embargo, la teoría naturalista de Cratilo puede entenderse de una manera más abstracta e interesante; Kretzmann y Williams, en sus trabajos citados, se inclinan por una interpretación estructural de la misma.

²¹ Sócrates refuta la versión puramente onomatopéyica del "naturalismo" mediante un ejemplo: las palabras "sklérôtês" y "sklérôtêr" significan lo mismo en Atenas y en Eritrea respectivamente, aun cuando difieren en la última letra (cfr. 434c-435b); de modo que "nos servimos de lo desigual para representar la misma cosa [...] las letras iguales o desiguales son igualmente expresivas una vez admitidas por el uso y la convención" (435b).

nombres complejos conceptos complejos.²² Quedaría, desde luego, pendiente la explicación de cómo es posible que los nombres simples signifiquen "naturalmente" un concepto simple y, por lo tanto, refieran, sin apelar a alguna variedad de convencionalismo.²³ Pero no es esto lo que me interesa destacar aquí. Lo que deseo subrayar es la idea de Cratilo de que todos los nombres, primitivos o derivados, encierran una representación o un significado que determina su referencia; esto es, la idea de que los nombres genuinos sólo logran referir a través de algún tipo de representación o contenido descriptivo que sólo conviene al objeto nombrado, es decir, a través de algo similar a lo que más tarde Frege denominará un "sentido".²⁴

Aunque es cierto que en la mayoría de los ejemplos elaborados por Sócrates para poner a prueba la pertinencia de T2, el sentido sugerido para el nombre examinado no es una descripción satisfecha únicamente por el nominatum (sino por

²² De hecho esta afirmación sería más coherente con su tesis de que la corrección de los nombres es la misma "para griegos y bárbaros". Sin embargo, esta forma de explicitar su teoría tampoco la salvaría del tipo de objeción planteado por Sócrates mencionada en la nota 13. Cabe notar que en algunos pasajes Sócrates considera que la unidad semántica mínima es la palabra, no las letras, véase, por ejemplo, 385c.

²³ Cratilo parece apuntar a una teoría según la cual a cada "Forma" le correspondería "por naturaleza" un nombre primitivo determinado. Esto equivale a sostener que los términos generales entran al lenguaje ya cargados de contenido significativo y que es función del nomothetes descubrirlo. Queda, desde luego, la pregunta de cómo "entra" ese contenido en el término.

²⁴ No entraré aquí en la discusión acerca de cuál es la manera correcta de interpretar la noción de "sentido" en Frege, sólo me interesa subrayar que la propuesta de Cratilo comparte con la de Frege la idea de que la referencia tiene que estar mediada por algo que desempeña un papel epistemológico, algo provisto de "valor cognoscitivo". En el capítulo II volveré sobre este asunto.

una pluralidad de individuos),²⁵ considero que la insistencia de Cratilo en que el nombre debe mostrarnos la naturaleza única de la cosa nombrada nos permite atribuirle una teoría descriptivista de los nombres semejante en aspectos importantes a la de Frege.²⁶ En efecto, si el nombre ha de referir a una única cosa, el "contenido" del nombre tendrá que ser tal que determine a un sólo individuo. En concordancia con la interpretación que propongo, Sócrates sugiere como contenidos de muchos nombres examinados ("Zeus", "Tántalo", "Pelopes", "Hades", "Perséfone") descripciones claramente definidas en virtud de las cuales el nombre refiere a un único individuo. Esto nos permite considerar la tesis naturalista de Cratilo como la primera versión históricamente ofrecida de una teoría descriptivista de los nombres y como un antecedente de otras teorías descriptivistas de los nombres propios ordinarios que se discuten en la actualidad.²⁷ Se trata, además, de una teoría descriptivista-esencialista, pues

²⁵ Agradezco esta observación a Raúl Orayen que atinadamente me señaló que no todos los nombres propios examinados en el Cratilo encierran sentidos equivalentes a descripciones definidas. Por ejemplo, "Agamemnon" significa admirable por su perseverancia (395b), "Apolo" significa algo que lava (405c).

²⁶ Desde luego, la tesis de Cratilo diferiría sustancialmente de la de Frege en tanto que este último está muy lejos de sostener que los nombres tienen como una cuestión de orden "natural" los sentidos que tienen. Tampoco aceptaría Frege (al menos para el caso de los lenguajes imperfectos, como son todas las lenguas) la idea de que un nombre sólo puede tener un sentido; menos aún, la idea de que el sentido de un nombre esté constituido por los caracteres esenciales de su referente. Véase la introducción al Capítulo II.

²⁷ En las primeras páginas de Schofield, M., (1981) encontramos algunas afirmaciones que concuerdan con la presente interpretación. Sin embargo, Schofield no se detiene a explorar en detalle el carácter descriptivista de la teoría de Cratilo sobre los nombres propios. La interpretación que ofrezco se apoya, entre otros pasajes del Cratilo, en 393d-e, 394e, 396a, 430d, 435c.

nombrar a un individuo equivaldría, según Cratilo, a describirlo en sus caracteres esenciales. El nombre nos entrega los caracteres esenciales del nominatum; por esta razón, afirma Cratilo, "cuando uno conoce los nombres, conoce las cosas" (431d), "cuando uno conoce de qué naturaleza es el nombre –y es de la misma naturaleza que el objeto– al mismo tiempo uno conocerá también al objeto, pues resulta ser semejante al nombre" (435e). A la tesis semántica de los nombres que defiende Cratilo va aunada una sorprendente tesis epistemológica: Para cualesquiera dos individuos x y y , si x posee el nombre correcto de y , entonces x conoce la esencia de y .

3. Relación referencial y naturaleza de los nombres en T1 y T2.

(A). La relación referencial.

(i) De acuerdo con T1 en su versión "individualista", para que se establezca una relación referencial entre un nombre y su nominatum se requiere que haya una relación triádica entre un locutor l , un nombre N y un objeto x . Un signo N se convierte en el nombre de x sólo cuando es usado por l con la intención de señalar a x o de referirse a él. El nombre no tiene en sí mismo una "dirección", por así decir, que apunte a un objeto particular y no a otros. Es sólo la intención del locutor de asociar un sonido (o una grafía) a un objeto particular la que lo convierte en el nombre de ese objeto. Según esta versión individualista de T1, si deseamos especificar la referencia de un nombre tendremos siempre que relativizarla a un locutor y sus intenciones referenciales en un tiempo t , de modo que tendremos que usar oraciones de la forma " N utilizado por l en t , refiere a x ", pues el nombre N podría referir a otros objetos diferentes de x si fuera usado por otros locutores o en otro momento por el mismo locutor.

En la versión social de T1, la relación referencial, propiamente dicha, entre un nombre N y un objeto x, aunque requiera en su inicio de una intención referencial individual no se reduce a ella, pues es el producto del uso o de la "costumbre". Es decir, si bien a la base de toda relación nominativa está una intención referencial individual, la relación semántica entre un nombre N y su referente x es el resultado del uso de N, extendido a lo largo del tiempo, por parte de los hablantes de una comunidad para referirse a x y está determinada por ese uso.

Según T2, por el contrario; los locutores y sus intenciones no son necesarios para que se establezca la relación entre un nombre N y su referente. La referencia de N depende sólo de que N mismo se ajuste a la cosa nombrada; que se dé tal relación es independiente de los usos o las intenciones referenciales de los locutores. Recordemos que para T2 el nombre es un sonido provisto de contenido significativo y es en virtud de ese contenido que se establece la conexión con su nominatum (si es que éste existe). La referencia de un nombre no está relativizada a tiempos, ya que el nombre contiene desde un principio los caracteres esenciales del objeto nombrado, ni a costumbres, pues la costumbre no puede cambiar el contenido del nombre. Así, la relación referencial podría especificarse mediante oraciones de la forma "N refiere a x".

(ii) Según T1, al utilizar un hablante un sonido para referirse a un objeto, no hay nada que medie entre el sonido y el objeto. La relación referencial se establece directamente entre el sonido elegido por el hablante para nombrar una cosa y la cosa nombrada. No hay ningún intermediario representativo ni conceptual. Esto explica que podamos dar cualquier nombre a cualquier objeto. Nombrar es análogo a etiquetar arbitrariamente. En la versión social del convencionalismo la referencia entre el nombre y su referente sigue siendo

inmediata, directa, sólo que es aprendida de otros.

Según T2, por el contrario, la relación entre un nombre (en tanto que sonido o grafía) y su nominatum es una relación mediata o indirecta. Para que el sonido "Zeus" pueda ser usado por un hablante para referirse a Zeus es indispensable que algún intermediario conecte el sonido con el individuo nombrado. Cratilo piensa en dos tipos de intermediarios posibles: el concepto esencial, los "caracteres esenciales" del objeto (cfr. 424a y ss.) o una imagen o "representación" de la cosa nombrada (cfr. 432a-435a). Sin tal intermediario conceptual o representacional no habría nombres; los sonidos carentes de contenido significativo usados por un hablante con intenciones referenciales no referirán a nada.

(iii) T1 presenta la relación referencial como una relación contingente. Resulta perfectamente concebible una situación en la que el nombre "Sócrates" no refiera a Sócrates. Basta imaginar que el nombre se introdujo de una manera diferente de como en realidad se introdujo. El nombre de Sócrates no es necesariamente su nombre. Este es el hecho que Hermógenes pretende asentar cuando afirma que podemos cambiar los nombres "sin que el nombre sustituido sea menos exacto que el precedente" (385d).

T2, por el contrario, considera la relación entre un nombre y el objeto designado como una relación necesaria.²⁸ El nombre encierra un contenido descriptivo o una representación que conviene única y necesariamente al objeto

²⁸ Cuando digo que según Cratilo la relación entre el nombre y el objeto nombrado es una relación necesaria quiero decir que, debido a que el nombre tiene un contenido descriptivo, en todo mundo posible en donde existan ese nombre con ese contenido, el nombre denotará al mismo objeto al que de hecho denota en el mundo real. Dicho en otras palabras, un nombre N con contenido C, si refiere en un mundo posible, no podría referir a un objeto distinto de aquel al que de hecho refiere.

nombrado. El nombre, nos dice Cratilo, "es propiamente como una definición" (396a), de modo que resulta inconcebible una situación en la que un nombre N que contenga el conjunto de caracteres esenciales que definen a un objeto x (o la representación de x) no refiera a x. Si "Sócrates" expresa la esencia de Sócrates o contiene una representación que sólo conviene a Sócrates, entonces es imposible que "Sócrates" no refiera a Sócrates.

(B) La naturaleza de los nombres.

De acuerdo con T1, los fonemas que un hablante utiliza para designar objetos pueden carecer absolutamente de significado previo. Los nombres se conciben como marcas que no quieren decir nada en sí mismas. Sólo en la medida en que el nombre se "conecte" con un objeto mediante la intención referencial de un hablante, el nombre lo representa, se refiere a él. En aquellos casos en los que un locutor usa como nombre de un individuo una palabra que, debido a convenciones previamente aceptadas por la comunidad lingüística a la que pertenece, tiene un significado general, digamos, "Salvador", "Dolores", "Blanca", no es de ninguna manera necesario para que el nombre refiera a su nominatum que éste tenga las propiedades mencionadas por la palabra utilizada. Así, de acuerdo con T1, en su versión social, incluso si admitimos que una palabra pudiera tener un significado convencional previo a su introducción como nombre, dicho significado no jugaría por sí mismo ningún papel en la determinación de la referencia del nombre. Sólo la intención del locutor la determina. "Teófilo" puede ser el nombre de un ateo fanático siempre y cuando exista por parte del locutor la intención referencial pertinente.

T2, por el contrario, concibe los nombres como paquetes de significaciones.

Para que el nombre pueda cumplir con su función referencial es necesario que tenga un contenido descriptivo o representacional, éste es la única condición necesaria para que el nombre refiera.²⁹ La referencia del nombre no es más que una función (en el sentido técnico del término) de su contenido descriptivo, el cual recoge los caracteres esenciales de la cosa nombrada. De esta suerte, los nombres no pueden jamás errar sus referentes; "Todos los nombres son correctos --dice Cratilo-- [...] todos aquellos al menos que son nombres" (429d). La referencia es independiente tanto de las convenciones como de las intenciones y las creencias de los usuarios del nombre. En ese talante, afirma Cratilo: "'Hermógenes' no es el nombre de Hermógenes, aunque todo el mundo le dé ese nombre" (383b), pues Hermógenes -- dice Cratilo-- no tiene nada en común con la raza de Hermes y esto basta para decir que no ha recibido ese nombre; aun cuando "parezca haberlo recibido, ese nombre pertenece a otro, a aquel que tiene también la naturaleza" (429c).³⁰

4. La imposibilidad de la referencia equivocada.

Por vías diferentes tanto la versión individualista de T1 como T2 llegan a

²⁹ Para que ésta condición se convirtiera en necesaria y suficiente, se requeriría, además, que el supuesto contenido descriptivo del nombre fuera satisfecho por un solo individuo.

³⁰ La inconsistencia práctica en la que incurre Cratilo es usar en el texto citado el nombre "Hermógenes" para referirse a Hermógenes a la vez que niega que ése sea su nombre.

la consecuencia paradójica de que no hay error posible en la referencia de los nombres; es decir, a la consecuencia falsa de que si un nombre refiere no puede referir a algo que no sea su portador.³¹

Según T1, en su versión individualista, el error no puede tener lugar debido a que los nombres por sí mismos no refieren a nada; su referencia se determina sólo por la intención del locutor en el momento de utilizar un nombre. Tener la intención de referirse a x mediante N es suficiente para que N refiera a x , de modo que siempre que un locutor haga uso de N con la intención de referirse a x , N referirá a x . De acuerdo con T1 no parece haber manera de cometer una equivocación al usar un nombre. Hermógenes afirma: "el nombre que uno asigna a un objeto es el nombre correcto" (384d) y si uno cambia el nombre "el nombre sustituido no es menos correcto que el nombre precedente" (384d). Lo que normalmente llamaríamos un caso de uso incorrecto de un nombre, por ejemplo, usar el nombre "Miguel" para referirse a Juan, sería interpretado por T1 como un caso en el que simplemente varía la referencia de "Miguel" en virtud de la intención referencial del locutor. Según esta versión de T1, todo nombre parece tener un ingrediente "deíctico" que lo hace que refiera precisamente a ese individuo al que se dirige el usuario del nombre con intención de referirse. Basta que exista el individuo al que el locutor dirige su intención referencial para que el

³¹ En el cap. xii de Gosling, J.C.B., (1973), encontramos una interpretación del Cratilo según la cual el diálogo en su conjunto "constituye un ataque contra las posiciones que sostienen que todos los enunciados del lenguaje tienen que ser verdaderos y que la falsedad es imposible." (p. 267 en la versión española). Esto me parece no ser la cuestión central; Platón parece asumir la falsedad de la tesis de la imposibilidad del discurso falso, de modo que al mostrar que tal tesis se sigue tanto de T1 como de T2, nos ofrece una reductio ad absurdum de las posiciones de Hermógenes y de Cratilo.

nombre refiera a ese individuo.

Sócrates refuta esta tesis mostrando (recordemos que se extiende la tesis a todos los términos del lenguaje) no sólo que hace imposible el discurso falso, sino que hace también imposible la comunicación. En efecto, los actos de habla que sirven para comunicar (para "instruir", dice Sócrates) no pueden ser tales que contengan nombres que el hablante use arbitrariamente; si el lenguaje ha de servir a los fines de la comunicación, los nombres usados tienen que estar de alguna manera sometidos a reglas de uso correcto. (Cfr. 386e-388c).

En la versión convencionalista social de T1, resulta claro que puede hablarse de un posible tipo de error: el consistente en que el usuario del nombre N no respete las convenciones que rigen el uso de N en su propia comunidad lingüística. Cabe señalar, sin embargo, que el locutor, en ese caso, estaría en error frente a su comunidad lingüística, pero no frente a sí mismo. Es decir, su error sería relativo, no sustantivo.³² Los nombres que usa podrían ser incorrectos con relación a las convenciones de su comunidad, pero no en sí mismos.

De acuerdo con T2, es también imposible el error referencial, pero por razones completamente diferentes. Como hemos señalado, según Cratilo, el nombre contiene en sí mismo la determinación de su propia referencia. Las

³² Bernard Williams distingue entre "corrección interna" y "corrección externa" en el uso de un nombre. La interna se decide por referencia a las prácticas lingüísticas de nuestra comunidad lingüística, a nuestras convenciones. Usamos ese sentido de "corrección", por ejemplo, cuando afirmamos que el nombre correcto de Hermógenes es "Hermógenes". La cuestión acerca de la corrección externa, en cambio, es una cuestión acerca de la corrección de nuestras propias prácticas lingüísticas. Sócrates concuerda con T1 en que el acuerdo y la costumbre son la única cosa que determina la corrección interna de un nombre. T1 niega que haya una cuestión relativa a su corrección externa, Sócrates piensa que sí la hay y en esto concuerda con T2. Cfr. Williams, B., (1981), pp. 90-91.

creencias y las intenciones de los locutores no pueden modificar la relación que necesariamente se da entre el nombre y la cosa nombrada; esto es, si un locutor \downarrow utiliza un nombre \underline{N} (que convencionalmente se usa en su comunidad para designar a \underline{x}) con la intención de referirse a \underline{x} , y resulta que \underline{x} no es el referente "natural" de \underline{N} , \downarrow no se referirá a \underline{x} con su uso de \underline{N} aunque así lo crea. En un pasaje notable en el que Sócrates acorrala a Cratilo, le pregunta: "si alguien, encontrándote en el extranjero, te toma la mano y te dice 'Hola, Hermógenes, extranjero ateniense, hijo de Esmicrion'³³ [...] ¿es a tí a quien se refiere o a Hermógenes aquí presente? ¿o a nadie?" (429e). Cratilo responde: "En mi opinión, Sócrates, no emitiría más que sonidos vanos. [...] Diría yo de tal hombre que sólo hace ruido agitándose inútilmente" (429e-430a). Cratilo no puede decir que "Hermógenes" refiera a Hermógenes porque, como vimos antes, considera que ese nombre "es realmente el nombre de otro, el cual tiene también la naturaleza" (429c); la costumbre entre los atenienses de llamar "Hermógenes" al hijo de Hipónico³⁴ no es decisiva cuando se trata de determinar la referencia del nombre. Tampoco refiere a Cratilo, porque, aunque el usuario del nombre tenga la intención de referirse a Cratilo, "Hermógenes" no describe correctamente la naturaleza de Cratilo. Si no existe ningún individuo que tenga la naturaleza que muestra el nombre "Hermógenes", el nombre no refiere a nada. Ni las convenciones vigentes en una comunidad lingüística ni las intenciones del hablante

³³ El ejemplo parece suponer que Hermógenes es efectivamente hijo de Esmicrion; sin embargo, en Schmidt O., U., (1988), p. cxxxviii, la traductora anota que Hermógenes es en realidad hijo de Hipónico, lo cual hace doblemente difícil determinar a quién refiere "Hermógenes" en el caso hipotético planteado por Sócrates.

³⁴ Véase la nota anterior y la nota 26.

pueden cambiar esta situación. Por otra parte, si existe un individuo x cuya naturaleza corresponda a la que encierra un nombre N , x será necesariamente el referente de N , independientemente de las intenciones referenciales de los hablantes y de las convenciones vigentes en su comunidad lingüística. "La atribución inexacta –afirma Cratilo– no es posible en los nombres, la exactitud reina ahí necesariamente siempre"(430d).³⁵

5. Dos concepciones del lenguaje y de la referencia.

Es enteramente correcto decir que el nombre "Saturno" refiere a Saturno; pero puede también decirse que un locutor l se refiere con "Saturno" a un amigo o a su perro. Es un hecho que los nombres refieren y que las personas se refieren.³⁶ T1 sólo toma en cuenta la segunda parte de esta afirmación; T2 sólo la primera. Según T1, la referencia es una relación cuyo primer término es

³⁵ En Williams, B. (1981), hay una elegante versión de la refutación de Sócrates a la tesis de Cratilo en la que muestra que "entender la teoría de Cratilo, exige que uno entienda las posibilidades que niega", véase esp. p. 87. Dar un nombre N a un individuo x supone el acto de dirigirse verbalmente a x usando N (en donde N , según Cratilo, tiene un contenido significativo) y ese acto, en efecto, puede realizarse incorrectamente (es decir, podemos dar el nombre a un individuo al que no le corresponda el contenido significativo del nombre en cuestión).

³⁶ Cfr. Donnellan, K., (1979), pp.86-89. En este pasaje Donnellan alude a una diferencia advertida (aunque, en mi opinión, no suficientemente explotada) por Peter Geach entre "referencia semántica" y "referencia personal" en su libro Reference and Generality, Cornell University Press, 1962, p. 31. En castellano se señala esta diferencia pronominalizando el verbo "referir" para lo que Geach llama "referencia personal". La distinción está emparentada con la que traza Grice entre "querer decir" (meaning) predicado de emisores (personas) y predicado de expresiones lingüísticas en Grice, H.P., (1968).

siempre una persona y tiene lugar cuando en una situación concreta tal persona usa un nombre con la intención de designar algún particular. Según T2, en cambio, la referencia es una relación cuyo primer término es siempre un término singular y cuyo segundo término es un particular o un individuo; las intenciones de los hablantes no pueden alterar la relación referencial entre el nombre propio y el individuo designado. A esta diferencia en la manera de concebir la relación referencial van aparejadas concepciones del lenguaje enteramente diferentes.

T1 considera que la significación lingüística, en general, y la referencia, en particular, son fenómenos que ocurren sólo en el ámbito de las intenciones de los locutores o de los actos de habla. Los nombres en sí mismos son meros trazos o ruidos que no significan nada, no refieren por sí mismos a nada. La referencia del nombre se establece en cada ocasión de uso mediante la intención referencial del usuario del nombre. El lenguaje no es considerado como una estructura fija con elementos significativos en sí mismos y de cuya significación dependan las condiciones de verdad de los enunciados en los que eventualmente aparecen. La significación y la referencia se encuentran del lado de los locutores, no de las palabras, ni de las oraciones, consideradas en sí mismas. La semántica se confunde con la pragmática. Dentro de la concepción de Hermógenes, la relación entre lenguaje y mundo es totalmente contingente.

T2, por el contrario, supone una concepción del lenguaje como una estructura con elementos constitutivos que son en sí mismos significativos, guardan relaciones necesarias entre sí y, en el caso de los nombres propios, refieren sólo si hay un único individuo que satisfaga su contenido significativo. Los elementos de esta estructura describen, "imitan" o "representan" las cosas tal y como ellas son en la realidad y las relaciones entre los elementos de la estructura modelan las relaciones que se dan entre las cosas reales. (Recordemos que para Cratilo "quien conoce los

nombres conoce las cosas" y quien puede ver las relaciones entre los nombres será capaz de ver las relaciones entre las cosas nombradas por ellos.) De esta manera, la estructura del lenguaje (del lenguaje verdadero) representa la estructura del mundo y la relación entre el lenguaje y el mundo es necesaria.³⁷ T2 coloca la significación y la referencia del lado de las frases y los nombres de un lenguaje, no del lado de las intenciones de los locutores. Es sólo porque las palabras tienen un contenido significativo en sí mismas que un hablante puede usarlas para aplicarlas a las cosas o para referirse a los individuos. Como indicamos antes, según T2 un nombre refiere a un individuo sólo en la medida en que su contenido significativo se aplique a ese individuo y sólo a él.

Desde la perspectiva de T2, la referencia que hace un hablante está determinada por el significado objetivo de los nombres que usa, de modo que no hay distinción posible entre la referencia del hablante y la referencia semántica de un nombre. La pragmática queda totalmente absorbida por la semántica.

En relación con el asunto acerca de la naturaleza "social" o "individual" del lenguaje, podemos constatar también algunas diferencias obvias entre T1 y T2.

La concepción de los nombres presentada por Hermógenes (Cfr. 384d-385a, 385d-e) presupone una idea del lenguaje como un fenómeno fundamentalmente individual, como una creación absolutamente libre, arbitraria, por parte del sujeto.

³⁷ La estructura del lenguaje no puede no representar la estructura del mundo debido a que, como he indicado, cada término tiene un contenido que determina su referencia o aplicación y los contenidos de los términos bastan para determinar las oraciones verdaderas; el lenguaje verdadero refleja, así, la estructura del mundo. Por otra parte, la relación entre lenguaje y mundo es necesaria debido a que todos los términos (singulares y generales) tienen un contenido que hace imposible que se apliquen a algo diferente de aquello a lo que de hecho se aplican. La frase en el texto tiene una resonancia claramente "tractatusiana".

La significación y la referencia no se constituyen al nivel comunitario o social, sino al nivel individual y, como señalamos antes, no habría límites que pudieran restringir por principio la posibilidad del locutor de dar cualquier nombre a cualquier cosa. De no ser por las exigencias que impone la comunicación, T1 podría aceptar la posibilidad de tantos lenguajes como hablantes hay. La dimensión social del lenguaje, las convenciones, los acuerdos entre los miembros de una comunidad lingüística de dar los mismos nombres a las mismas cosas, se construye a partir de la significación y la referencia individual. Hermógenes no advierte las dificultades que acarrea admitir la posibilidad de lenguajes individuales o privados.

T2, por el contrario, concibe el lenguaje como un instrumento de uso común que hace posible la comunicación y nos permite avanzar en el conocimiento de las cosas ("los nombres están hechos para instruir" 428e). Aunque, según Cratilo, el lenguaje no es propiamente una creación social, sí es un "bien común" al que tienen acceso los miembros de la comunidad para intercambiar información. El hecho de que la significación y la referencia sean inherentes a las palabras hace innecesarias las convenciones y garantiza la posibilidad de que unos y otros se comprendan al usar el lenguaje, así como la objetividad de los significados. Nada más alejado de T2 que admitir la posibilidad de lenguajes individuales o privados. El problema que Cratilo tiene que enfrentar, y que Platón no aborda en este diálogo, es el de explicar cómo acceden los hablantes al conocimiento de los significados de los nombres, ya que su significado no tiene nada que ver con usos establecidos, ni con intenciones referenciales que podamos descubrir a partir de la conducta de los hablantes.

6. Dos intuiciones sobre los nombres.

La importancia del Cratilo radica en que recoge con enorme lucidez dos intuiciones sobre el funcionamiento de los nombres en el lenguaje y presenta las concepciones de la referencia que se desprenden de cada una de ellas. Ambas concepciones, tal como son expuestas por Hermógenes y Cratilo, conducen a conclusiones inaceptables o a callejones sin salida, de modo que el Cratilo puede considerarse como una larga demostración por reducción al absurdo de la falsedad de las dos tesis sobre la naturaleza de los nombres que allí se examinan. Sin embargo, las intuiciones quedan ahí para ser rescatadas y sistematizadas por una teoría adecuada de los nombres propios. Veamos, pues, cuáles son esas intuiciones.

(A). La intuición que está a la base de T1 es la que pudiéramos llamar "de la arbitrariedad" del nombrar. Es efectivamente en algún sentido arbitrario que México se llame "México", Napoleón se llame "Napoleón", etcétera; nada hay en el individuo denotado, ni en el nombre que lo denota, que exija que la relación nominativa se establezca del modo como se estableció. Si alguien alega que no es arbitrario que México se llame "México", pues, después de todo, hay algunas razones por las que se eligió ese nombre y no otro (por ejemplo, ser la tierra de los mexicas), se nos concederá, al menos, que es enteramente contingente que así se llame; pudo haberse llamado "Imperio Azteca" o "Nueva España" o de cualquier otra manera. La elección de un nombre depende en último término de los intereses o el gusto de quienes lo confieren. La intuición de que el nombrar es arbitrario se origina por el modo como de hecho introducimos un nuevo nombre en el lenguaje: de un conjunto potencialmente infinito de fonemas elegimos sin

ninguna limitación (excepto, tal vez, de orden práctico) aquel que hemos de utilizar para designar a un individuo. Decimos apuntando al individuo: "se llamará "Diego"". Es la elección libre del nombre y la intención de usarlo para referirnos a ese niño lo que convierte a "Diego" en el nombre de Diego. No es necesario revisar el significado del nombre (si es que lo tiene), ni buscar "descripciones o representaciones" asociadas con el nombre (si es que las hay), para determinar si le conviene o no al niño.

Dada la manera como se introduce un nombre en el lenguaje, resulta también enteramente natural pensar que la relación referencial es una relación directa; que no hay ningún tipo de intermediario descriptivo ni representacional entre el nombre y la cosa nombrada. Los nombres son marcas carentes en sí mismas de significado a las que podemos asignarles individuos. Cuando queremos hablar acerca de un individuo x usamos en la oración el nombre N que previamente le hemos asignado y decimos que N está en lugar de x .

Es cierto que al introducir un nombre para designar a un individuo se requiere identificarlo de alguna manera, también es cierto que el buen uso posterior del nombre depende en gran medida de que seamos capaces de reconocer al mismo individuo en ocasiones posteriores, y es cierto también que en ocasiones reconocemos a los individuos que hemos conocido en el pasado (directamente o mediante descripciones) a través de propiedades de esos individuos que somos capaces de detectar. Esa capacidad de reconocer a través de sus propiedades a los individuos a los que podemos dar un nombre que ya habíamos aprendido en el pasado, explica que, en general, seamos capaces de dar alguna descripción, por vaga que ésta sea, de esos mismos individuos. Estas parecen ser cuestiones que no toma en cuenta la primera intuición sobre los nombres, y que nos acercan, por el contrario, a la intuición (B).

(B). La intuición que parece estar a la base de T2 es la de que cuando conocemos un hecho singular expresado por una oración en la que aparece un nombre previamente introducido en la lengua, podemos parafrasear lo que conocemos mediante frases en las que puede ya no aparecer el nombre original, sino, en su lugar, ciertas descripciones (identificadoras o no) del portador del nombre. En efecto, cuando entendemos una oración en la que aparece un nombre y tratamos de explicar lo que entendemos, solemos ser capaces de utilizar en la explicación, en lugar del nombre, otras expresiones que con mayor o menos precisión describen al portador del nombre. Esta intuición parece subyacer en la idea de que los nombres son términos provistos de contenido; si en el contexto de explicitar nuestro conocimiento de un hecho singular el nombre es sustituible por ciertas descripciones, es porque --se piensa-- el nombre de alguna manera encierra tales descripciones. Si los nombres estuvieran desprovistos de todo contenido descriptivo -se piensa-, no podríamos parafrasear lo que dice una oración en la que figura un nombre mediante otra oración en la que se sustituye el nombre por expresiones descriptivas.

En concordancia con esta segunda intuición, parece natural pensar que identificamos a los individuos a través de sus propiedades. Y de allí, se pasa fácilmente a pensar que cuando conocemos el referente de un nombre asociamos en nuestro pensamiento con el nombre un conjunto de conceptos que no son más que las contrapartes mentales de las propiedades que nos sirvieron para identificarlo. De esta manera, a partir de la idea de que identificamos a los individuos por sus propiedades, se llega a pensar que los nombres están cargados de un contenido descriptivo que es tal que nos permite reidentificar en ocasiones posteriores al referente original y, de esta manera, utilizar consistente y

correctamente el nombre.³⁸

Por otro lado, si el nombre tiene un cierto contenido descriptivo o representacional y es a través de ese contenido que reconocemos al portador del nombre, resultará natural pensar que conocemos directamente las propiedades o representaciones asociadas con un nombre y sólo indirectamente al portador del nombre y, de manera correspondiente, en el terreno semántico, que el nombre se relaciona directamente con su contenido descriptivo y que sólo a través de éste refiere al individuo nombrado.

Las dos intuiciones descritas captan, en algún sentido, fenómenos lingüísticos reales. Una de ellas, como hemos visto conduce a una teoría de la referencia directa de los nombres, la otra, por el contrario, conduce a concebir la referencia como una relación mediada o indirecta. Una teoría aceptable de la referencia de los nombres propios tiene que combinar la economía semántica con la capacidad de explicar la situación epistemológica de los locutores que aparentemente son capaces de usar nombres propios, es decir, tiene que atender a las intuiciones que guían tanto a Hermógenes como a Cratilo sin llegar a los callejones sin salida a los que éstos parecen llegar. Si bien en el Cratilo no encontramos una teoría aceptable de la referencia de los nombres propios, si encontramos la definición de una polémica que, como veremos en capítulos posteriores, llega hasta nuestros días.

³⁸ En concordancia con lo anterior encontramos en el Cratilo la idea de Sócrates, expresada en 385c, de que para que el discurso pueda ser falso o verdadero tiene que estar constituido por elementos que sean ellos mismos falsos o verdaderos. Esto parecería requerir que los nombres que figuran en un discurso verdadero sean verdaderos ellos mismos de sus referentes, es decir, que sus contenidos descriptivos sean satisfechos por sus referentes, lo cual supone considerar a los nombres como equivalentes a paquetes de predicados o a abreviaturas de descripciones.

Me refiero a la polémica sobre si los nombres refieren directa o indirectamente a sus portadores.

Como señalé en la introducción, mi propósito en esta tesis es defender una versión contemporánea de la teoría de la referencia directa, es decir, de la tesis según la cual los nombres carecen de contenido descriptivo. Pero antes de proceder a exponer y defender dicha versión, consideraré los mejores argumentos que se hayan ofrecido para considerar que los nombres refieren indirectamente, me refiero, desde luego, a los argumentos de Frege. Sólo si podemos mostrar que tales argumentos no son concluyentes, podremos considerar viable una concepción de los nombres que prescindan de intermediarios. En el capítulo que sigue examinaré, pues, los argumentos que ofrece Gottlob Frege para introducir, como un elemento semántico diferente del referente, el sentido de los nombres propios.

CAPÍTULO II

Frege: sentido y referencia indirecta

Introducción.

Existen indudablemente diferencias notables entre la postura sostenida por Cratilo con respecto a los nombres, tal como la caractericé en el capítulo anterior, y la de Gottlob Frege¹ a la que me referiré en el presente capítulo. Sin embargo, comparten por lo menos tres tesis: (i) que asociado con cada nombre hay una descripción identificadora o algún otro criterio de identificación del portador del nombre,² (ii) que la referencia de los nombres es indirecta, en el sentido de que

¹ Me refiero, desde luego, a la postura que defiende Frege después de la introducción de la distinción entre el sentido (*Sinn*) y la referencia (*Bedeutung*) del nombre, alrededor del año 1890, y que da lugar a lo que Gareth Evans denomina "una teoría semántica de dos niveles". Según ésta toda expresión lingüística tiene dos tipos de propiedades semánticas: sentido y referencia, relacionadas de tal manera que el sentido determina la referencia. Véase, Frege G., (1892) y (1892-1895), y Evans, G., (1982), cap. 1.

² Algunos intérpretes juzgan que Frege realmente no se compromete con la tesis de que todo nombre contiene un criterio de identificación del referente; véase especialmente Evans, G., (1982) cap. 1. Michael Dummett, por su parte, juzga que Frege no se compromete con la tesis de que el criterio identificador del referente sea una descripción definida, véase Dummett, M., (1982), cap. 5, pp. 97-98. Sin embargo, son muchos los pasajes en los que Frege caracteriza los sentidos de los nombres como descripciones identificadoras (por ejemplo, "el discípulo de Platón y maestro de Alejandro Magno" en Frege, G., (1892) p. 5), o como "modos de presentación" de un objeto que nos permiten reconocerlo. Saul Kripke, Mark Platts y Manuel García-Carpintero se inclinan, como yo, a considerar que no se comete una injusticia si decimos que para Frege el sentido de

refieren sólo a través del criterio identificador asociado con cada uno de ellos, y (iii) que el criterio identificador asociado con un nombre no es alguna representación subjetiva, algo que sólo acontece en la conciencia de los usuarios, sino un componente semántico objetivamente existente que conocen todos los que son capaces de entender lo que dice una oración en la que figura el nombre.³

La diferencia más sobresaliente entre Frege y Cratilo es que, según este último, el nombre contiene "naturalmente" la descripción esencial de su referente, en tanto que para Frege esto no es el caso. Es decir, según Cratilo, el nombre posee un contenido descriptivo aún antes de haber sido usado (por ejemplo, la etimología de un nombre determina su contenido descriptivo); y esto hace que la referencia sea siempre indirecta y que la relación entre el nombre (con el contenido que tiene por sí mismo) y el objeto designado sea necesaria. Para Frege, en

(1979), cap. 6 y García-Carpintero M., (1996), caps. VI y VII.

³ La objetividad del contenido descriptivo del nombre es lo que conduce a Cratilo a sostener que el nombre "Hermógenes" no le pertenece a Hermógenes. El argumento de Frege a favor de la objetividad de los sentidos asociados con los nombres se encuentra en Frege, G., (1918-19) y (1892) y en una carta que dirigió a Philip Jourdain en 1914, recogida en Frege G. (1980) y citada en Evans, G., (1982) pp. 14-15. La tesis de que los sentidos son objetivos es diferente de la tesis de que todos los usuarios de un mismo nombre tengan que compartir un mismo sentido. Platts insiste en mantener esas dos tesis bien separadas y considera una confusión sostener que el sentido ("la ruta a la referencia") forme parte del significado del nombre y, por ende, que tenga que ser compartido por todos los que usan el nombre. Véase, Platts, M., (1979), p. 140. Gareth Evans, en cambio, piensa que la objetividad del sentido implica que los usuarios de un mismo nombre tienen que compartir un mismo sentido y que esta tesis es crucial para la semántica de Frege; véase Evans, G., (1982), pp. 16 y 40. Cabe aclarar que, a diferencia de Platts, Evans interpreta el sentido de un nombre como algo esencialmente no epistemológico, es decir, no como un criterio de identificación del referente (lo cual efectivamente puede variar de un usuario del nombre a otro), sino como "una manera particular de pensar acerca de [él]".

cambio, la relación entre un nombre N y la cosa nombrada, al momento de introducir N , es contingente,⁴ pues N no contiene en sí mismo antes de ser usado ninguna descripción, ni representación, de nada; es sólo una vez que N es usado para referirse a un objeto x , que queda asociado con N "el modo m como x se presenta" al (a los) usuario(s) del nombre. Una vez que se ha efectuado esta asociación, N refiere a x a través de m ; esto es, la relación semántica entre un nombre y su referente está mediada por un modo de presentación m que es caracterizado como "una ruta hacia el referente". Pero, según Frege, la intermediación de m es sólo una consecuencia de la manera como se introduce el nombre, no una propiedad "natural" del nombre. Una diferencia más relacionada con lo anterior es que, según Cratilo, el criterio identificador conectado con el nombre es una descripción esencial del referente, en tanto que para Frege, el modo de presentación m asociado con un nombre N no necesariamente nos entrega la "esencia" de la cosa nombrada por N .

Otra diferencia filosófica importante entre la postura de Cratilo y la de Frege es que, para Cratilo, la idea de que cada nombre tiene al menos un criterio identificador asociado con él, parece ser más la expresión de una intuición cuyas consecuencias está dispuesto a afrontar que la conclusión de algún razonamiento claramente estructurado; en tanto que para Frege se trata de la conclusión de un argumento o, mejor dicho, de diversos argumentos perfectamente articulados que pretenden concluir que las expresiones lingüísticas, en general, y los nombres, en

⁴ Frege se refiere al carácter contingente de esta relación diciendo que es "arbitraria". En Frege, G., (1892), p.4, leemos: "esta conexión [entre el nombre y la cosa designada] es arbitraria. A nadie se le puede prohibir que use como signo de algo cualquier objeto o proceso creados arbitrariamente.". Cabe observar que dado que Frege considera que las descripciones definidas son "nombres", tendría que matizar su afirmación para evitar que sea falsa.

particular, tienen, además de un referente, un componente semántico adicional. A ese componente semántico de los nombres al que, siguiendo a Michael Dummett,⁵ me he referido como "criterio identificador" del referente, Frege lo denominó "sentido" (Sinn).

En este capítulo me propongo examinar el argumento que usa Frege para introducir la noción de sentido de los nombres y señalar las serias dificultades que presenta. El capítulo tiene cuatro partes. Primeramente, analizo el argumento de Frege que parte de la llamada "paradoja de la identidad". Luego, examino los criterios de Frege para introducir la noción de "diferencia en valor cognoscitivo" entre oraciones referencialmente equivalentes; me interesa esa noción por dos razones: primero, por el papel crucial que juega en el argumento fregeano para introducir los "sentidos" y, luego, porque me permitirá intervenir, en la tercera parte, en el debate en relación a si las diferencias notacionales entre oraciones referencialmente equivalentes pueden o no explicar los fenómenos que la noción fregeana de sentido pretende explicar. Si logramos mostrar que hay maneras semánticamente más económicas de resolver las dificultades que motivaron la introducción del sentido, la noción de sentido resultará dispensable; es decir, no estaremos racionalmente obligados a aceptar la existencia de intermediarios entre el nombre y su referente. Por último, en la parte final de este capítulo me referiré a las famosas objeciones que Saul Kripke ha presentado en contra de la idea fregeana de que los nombres poseen un sentido y que éste determina su referencia.

⁵ Véase, Dummett, M., (1981), caps. 5 y 6, esp. pp. 95-101 y 179-180. Dummett, sin embargo, rechaza que el sentido sea un componente semántico. Ofrece una interpretación epistemológica del sentido de un nombre, caracterizándolo como un criterio de identificación del referente del nombre: "conocer el sentido de un nombre implica conectar el nombre con una manera particular de identificar un objeto como el referente del nombre" (p.95).

1. El argumento de Frege para introducir el sentido.

La extensa bibliografía sobre la noción fregeana de "sentido"⁶ nos da una idea de las dificultades que supone comprender y evaluar adecuadamente los argumentos de Frege para introducir el sentido como un componente semántico de las expresiones lingüísticas.⁷ Frege introduce su noción de sentido como la única manera de resolver ciertas dificultades a las que se enfrenta al desarrollar su teoría del lenguaje. Frege, en efecto, vió la dificultad que hay en sostener conjuntamente las siguientes dos tesis: (i) que la identidad no es una relación entre signos, sino una relación de cada objeto consigo mismo,⁸ y (ii) que dos oraciones

⁶ Véanse, entre los más importantes, Dummett, M. (1981), Evans, G., (1982), McDowell, J. (1977), Platts, M. (1980), Salmon, N., (1986), Simpson, T. M., (1975), Sluga, H. (1980) y (1988), Valdivia, L., (1989).

⁷ Michael Dummett, piensa que el sentido no es propiamente un componente semántico. Distingue entre el significado (meaning) y el valor propiamente semántico de una expresión lingüística. El primero -dice- es lo que el usuario de la expresión entiende al oírla o al usarla y que le permite decidir el valor de verdad de la oración en la que figura la expresión; el valor semántico, en cambio, es el designatum de la expresión. Entender una expresión -afirma- "nunca consiste, ni siquiera parcialmente, en asociar con la expresión o la palabra algo en el mundo" (p. 93), por esa razón el valor semántico no forma parte del significado ni el sentido forma parte del valor semántico de una expresión. Véase Dummett, M. (1981) esp. cap. 5. Los límites de este estudio me impiden evaluar aquí la propuesta de Dummett; baste señalar que su interpretación puramente epistemológica de la noción fregeana de sentido lo acerca demasiado al verificacionismo, por lo menos en su versión "ideal"; véase Evans, G., (1982), cap.1, n. 17. Entre quienes sí consideran que el sentido es una noción semántica están Evans, G., (1982), Ezcuardia, M., (1994), Millikan, R., G., (1991) Platts, M., (1979) y Sluga, H. (1980) p. 154.

⁸ Frege había supuesto en su Begriffsschrift (1879) que la identidad era una relación entre signos, no entre objetos. Véase esp. Frege, G., (1879), parágrafo 8. Frege considera allí que los enunciados de identidad son acerca de los nombres

de identidad verdaderas de la forma "a=a" y "a=b" pueden tener diferente "valor cognoscitivo" (Erkenntniswerte). Veamos esas dos tesis.

En su Begriffsschrift, Frege había supuesto que la información que aportaban las oraciones de identidad era de índole metalingüística y que esto explicaba la aparente "diferencia en valor cognoscitivo" entre dos oraciones de identidad verdaderas de la forma "a=a" y "a=b". Más tarde rechazó esta explicación, pues -además de que implicaba la extraña idea de que los nombres "de pronto están en lugar de sus contenidos... de pronto están en lugar de sí mismos" (Begriffsschrift, párrafo 8)- se percató de que no funcionaba para el caso de oraciones de identidad cuantificadas, tales como "Todo es idéntico a sí mismo", $(x)(x=x)$, o "Algo es idéntico a cada individuo", $(\exists x)(y)(x=y)$.⁹ En 1892 Frege corrige su posición anterior y sostiene explícitamente que la identidad es una relación entre objetos, y no entre nombres de objetos y, por ello mismo, que dos oraciones de identidad verdaderas de la forma "a=a" y "a=b" relacionan de la misma manera los mismos objetos. Sin embargo, -y esta es la segunda tesis-, Frege se percató de que, a pesar de ser referencialmente equivalentes, "a=a" y "a=b" tienen propiedades lógicas y epistemológicas diferentes: una es analítica y aparentemente

que figuran en ellos, no de sus contenidos. En general -dice- los nombres "son meramente representantes de sus contenidos, de manera que toda combinación en la que aparecen expresa sólo una relación de sus contenidos", pero en los enunciados de identidad de pronto se representan a sí mismos, "se muestran ellos mismos cuando se combinan mediante el símbolo de igualdad [...] Por tanto, al introducir un símbolo para la igualdad de contenido, necesariamente se produce una ambigüedad en el significado de todos los símbolos, ya que de pronto están en lugar de su contenido, de pronto están en lugar de sí mismos". Frege, G.,(1879), p. 26.

⁹ Véase Dummett, M., (1981), cap. 16, "Identity", esp. pp. 544 y ss. y Sluga, H., (1980), p. 152

a priori,¹⁰ en tanto que la otra constituye una genuina ampliación del conocimiento y es a posteriori; y no sólo eso, sino que también se percata de que es posible que una persona competente en el lenguaje entienda ambas oraciones y, no obstante, adopte actitudes proposicionales contrarias respecto de lo que ellas expresan.¹¹

La interpretación objetual de los enunciados de identidad y los fenómenos epistemológicos recién mencionados constituyen conjuntamente la motivación principal para que Frege introduzca la noción de "sentido" de un nombre en su ensayo "Sobre el sentido y la denotación".¹² En ese ensayo, escrito en 1892, se aprecia la estrecha relación que hay entre la noción fregeana de "sentido" y la de "valor cognoscitivo" de una expresión lingüística: la primera, como veremos, se introduce precisamente con el fin de explicar "diferencias en valor cognoscitivo" entre oraciones extensionalmente equivalentes.

Al inicio del texto, Frege presenta lo que se suele conocer como "la paradoja

¹⁰ Frege escribe que las identidades de la forma "a=a" "valen a priori y han de ser consideradas analíticas, de acuerdo con Kant", ya que no contienen ampliaciones de nuestro conocimiento y son verdaderas independientemente de toda experiencia, cfr. Frege, G., (1892), p. 3. Cabe observar, sin embargo, que, estrictamente hablando, no es el caso que conozcamos la verdad de una identidad de la forma "a=a" independientemente de toda experiencia, ya que saber que a=a requiere saber que "a" refiere, y no podemos saber eso independientemente de toda experiencia.

¹¹ La idea de que es posible adoptar diferentes actitudes proposicionales frente a enunciados referencialmente equivalentes en los que figuran nombres diferentes se encuentra en Frege, G., (1892), (1918-1919) y en la carta a Philip Jourdain fechada en 1914 y recogida en Frege, G., (1980).

¹² Frege, Gottlob, (1892). Traducción al castellano de Eugenio Bulygin y Eduardo Rabossi, "Sobre el sentido y la denotación", en T.M. Simpson (comp.) (1973); las páginas en las citas se refieren a esta última edición.

de la identidad". Primeramente, señala que "a=a" y "a=b" son, "obviamente, oraciones de diferente valor cognoscitivo" (p.3), ya que "a=a" es analítica y vale a priori en tanto que "a=b" puede contener una ampliación valiosa de nuestro conocimiento y puede no valer a priori; luego, observa que la identidad es una relación entre lo que los nombres "a" y "b" designan y que por lo tanto "a=a" y "a=b" relacionan de la misma manera los mismos objetos y, por lo tanto, "no pueden diferir" (p.4). Frege acepta la ley de Leibniz de la Indiscernibilidad de los Idénticos,¹³ según la cual para cualesquiera dos objetos x y y , si $x=y$, entonces, para toda propiedad F , $Fx \leftrightarrow Fy$; es decir, Frege piensa que lo que es idéntico no puede diferir en ninguna propiedad. El problema al que se enfrenta, entonces, es el de explicar cómo es posible que oraciones que predicen lo mismo de los mismos objetos "difieran" en cuanto a sus propiedades lógicas y epistemológicas.

Si la identidad es una relación entre objetos, entonces, -piensa Frege correctamente- la diferencia en "valor cognoscitivo" entre dos oraciones de identidad de la forma "a=a" y "a=b" no puede explicarse apelando sólo a la referencia de los nombres involucrados, pues, suponiendo que son verdaderas, ambas oraciones son indistinguibles desde una perspectiva meramente referencial: no hallamos en la referencia de las oraciones que nos ocupan ninguna diferencia que pueda servirnos para explicar el fenómeno que nos interesa. Y, si no podemos dar cuenta de la diferencia en "valor cognoscitivo" entre oraciones tales como "Venus es Venus" y "Venus es la Estrella Matutina" en términos puramente referenciales, entonces, piensa Frege, las oraciones tienen que tener semánticamente hablando- algo más que sus referencias. Un poco más adelante en el mismo ensayo, Frege sostiene que la diferencia en valor cognoscitivo entre

¹³ Véase, Frege, G., (1884), parágrafo 65; y Dummett, M., (1981), cap. 16.

oraciones de identidad referencialmente equivalentes sólo puede explicarse si "a la diferencia entre signos corresponde una diferencia en el modo de presentación de lo que es designado" (p.4); es decir, sólo si admitimos que, además de referir, los nombres cumplen otra función semántica que es la de presentar de algún modo particular aquello a lo que refieren. A este modo de presentar el referente Frege lo llama el "sentido" del nombre; en el sentido del nombre -dice Frege- "está contenido el modo de presentación" (p.5) de lo designado.

No entraré aquí en la discusión acerca de si el sentido de una expresión lingüística es un componente "semántico" de la expresión o no. Baste señalar que Michael Dummett piensa que no lo es,¹⁴ en tanto que otros consideran obscura esta idea de Dummett y toman la noción de sentido directamente como una noción semántica. Sostener que el sentido es un ingrediente semántico del nombre parece ser independiente de considerar que sea o no una guía epistemológica al referente e independiente también de considerar que a cada nombre le corresponda por principio uno o varios sentidos. Mark Platts, por ejemplo, quien ofrece una interpretación epistemológica del sentido señala que el sentido de un nombre es uno de sus componentes semánticos por ser "por lo menos parte de lo que un hablante con un dominio suficiente del nombre entiende o sabe"; señala, además, que Frege se compromete con la tesis de que a cada nombre le pueden corresponder por principio diversos sentidos.¹⁵ Por otra parte, Gareth Evans, quien rechaza toda interpretación epistemológica del sentido, otorga también un carácter semántico al sentido de un nombre en virtud de que es parte de lo que un

¹⁴ Véanse las notas 5 y 7 más arriba y más en general Dummett, M., (1981), caps. 5 y 6.

¹⁵ Platts, M., (1979), p. 135.

hablante competente tiene que entender al escuchar y entender la emisión de una oración en la que figura el nombre; sostiene, además, que a cada nombre le corresponde un solo sentido.¹⁶

He indicado cómo llega Frege a postular los sentidos de los nombres. Su argumento parte de considerar que puede haber diferencias en valor cognoscitivo entre oraciones de identidad extensionalmente equivalentes pero en las que figuran nombres codesignativos diferentes. Quisiera preguntar en la siguiente sección, primeramente, qué es el "valor cognoscitivo" de una oración; luego, abordar la cuestión de si el sentido es indispensable para explicar las "diferencias en valor cognoscitivo".

2. Diferencias en valor cognoscitivo.

El "valor cognoscitivo" es algo que tienen las oraciones usadas en contextos determinados; sin embargo, Frege no explica ni dilucida directamente en qué consiste el valor cognoscitivo de una oración relativamente a un contexto; es decir, no nos dice directamente qué es ni cómo se determina éste. La dilucidación de esta noción es siempre indirecta, a través de una noción comparativa o relacional: la de tener dos oraciones Q y Q' diferente (o el mismo) valor cognoscitivo. Esta noción de tener diferente valor cognoscitivo se introduce, a su vez, mediante tres

¹⁶ Como indiqué antes, Evans concibe el sentido de un nombre como una manera particular de pensar sobre el referente: "La idea de Frege es que puede ser propiedad de un término singular, en cuanto elemento de un lenguaje público, que para entender las emisiones que lo contienen haya que concebir no sólo un objeto particular -su Significado-, sino que haya que concebirlo de una manera particular: o sea que todo usuario del lenguaje competente que entienda la emisión concebirá al objeto de la misma manera." Evans, G., (1981) p.16.

criterios diferentes:¹⁷

(i) Para cualesquiera dos oraciones Q y Q' referencialmente equivalentes, Q y Q' tienen diferente valor cognoscitivo si y sólo si saber que Q puede representar una genuina ampliación del conocimiento en tanto que saber que Q' no.

(ii) Para cualesquiera dos oraciones Q y Q' referencialmente equivalentes, Q y Q' difieren en valor cognoscitivo si y sólo si podemos conocer a priori la verdad de Q, en tanto que la verdad de Q' sólo la podemos conocer a posteriori.

(iii) Una oración Q y otra Q' tienen diferente valor cognoscitivo si y sólo si es posible que un sujeto racional y competente en el lenguaje sostenga actitudes proposicionales diferentes, incluso contrarias, frente a Q y Q'.

La manera como la noción de "diferencia en valor cognoscitivo" se conecta con la noción de "sentido" en la semántica de Frege es la siguiente. Dadas dos oraciones Q y Q' extensionalmente equivalentes, digamos, "Venus es Venus" y "Venus es la Estrella Matutina", (1) si aprender la verdad de Q no constituye una genuina ampliación de nuestro conocimiento mientras que aprender la verdad de Q' sí, o (2) si la verdad de Q se conoce a priori y la de Q' a posteriori, o (3) si un sujeto racional puede adoptar diferentes actitudes frente a Q y Q', digamos, puede creer que Venus es Venus, pero no que Venus es la Estrella Matutina, entonces, Q y Q' difieren en valor cognoscitivo y, ya que esa diferencia no puede explicarse apelando sólo a la referencia de Q y Q', tenemos que suponer que hay un ingrediente semántico adicional -el sentido- que la explica.

¹⁷ Los primeros dos criterios a continuación están formulados en Frege, G., (1982) p.3; el tercero aparece, entre otros lugares, en Frege, G., (1980), p. 80 y p. 153. La diferencia en "valor cognoscitivo" entre dos oraciones extensionalmente equivalentes equivale, como se verá, a una diferencia en el pensamiento expresado por ellas.

Una vez introducida la noción de sentido, Frege sostiene una teoría semántica de dos niveles: las expresiones lingüísticas en general y, en particular, los nombres propios, tienen la propiedad semántica de referir a un objeto, pero no logran sólo en la medida en que tienen otra propiedad semántica, la de expresar un sentido. Es decir, todo nombre tiene un referente (por lo menos en un "lenguaje ideal"), pero sólo porque contiene en sí mismo un modo de apuntar a su referente, y ese modo de apuntar a su referente es su sentido.¹⁸ El sentido del nombre determina, así, su referente. Es sólo a través de su sentido que "Venus" refiere al planeta Venus, "Cratilo" refiere a Cratilo, etcétera. La relación entre sentidos y referente es de muchos a uno; es decir, un mismo objeto puede ser el referente de diversos nombres con diferentes sentidos. Y no puede haber, según Frege, ningún término singular, ningún nombre, que refiera directamente sin la mediación de un sentido: la referencia de los nombres está siempre mediatizada por sentidos. El sentido de un nombre determina su referencia o, dicho de otro modo, la referencia de un nombre es una función de su sentido, de modo que si no hay sentido no hay referencia.

3. La crítica de Millikan al argumento de Frege.

Para que el argumento de Frege logre probar la existencia del sentido, tiene que ser el caso que las llamadas diferencias en valor cognoscitivo no puedan explicarse de ninguna otra manera que resulte semánticamente más económica.

En la bibliografía filosófica reciente, entre los diversos estudios críticos del argumento de Frege, destaca el trabajo Ruth Garrett Millikan, destacada filósofa

¹⁸ Véase, Church, A., (1956), p.6, n.14.

británica. Millikan¹⁹ objeta el tercer criterio usado por Frege para introducir sentidos y ofrece una explicación alternativa de cómo es posible tener actitudes distintas frente a oraciones de la forma "a=a" y "a=b" cuando "a" y "b" refieren al mismo individuo. El error que comete Frege, de acuerdo con Millikan, es, en pocas palabras, el de "confundir la igualdad [o la diferencia] en el vehículo de representación con una representación de la igualdad [o de la diferencia]".²⁰ Es decir, cuando Frege concluye a partir de la posibilidad de tener actitudes proposicionales diferentes frente a lo que expresan oraciones referencialmente equivalentes pero notacionalmente diferentes, que los pensamientos expresados por ellas tienen que ser diferentes, Frege parece confundir -señala Millikan-, la diferencia en "el vehículo" de representación de un contenido semántico con la diferencia en contenido semántico. Ante el dato de un sujeto X que adopta distintas actitudes proposicionales en relación con dos oraciones P y Q referencialmente equivalentes, pero notacionalmente diferentes, no hay razón para postular componentes semánticos SP y SQ diferentes entre sí y distintos de las referencias de P y de Q, que hubieran de ser los objetos de las actitudes proposicionales de X. Hacerlo equivaldría a "internalizar" la diferencia oracional, es decir, a trasladar al ámbito de los contenidos de pensamiento una diferencia que pertenece sólo a las oraciones que los expresan. Podemos aceptar la igualdad a través de diferentes vehículos de representación y la diferencia a través de la igualdad en vehículos. Es decir, un sujeto puede creer que "Tulio es Cicerón" tiene exactamente el mismo contenido que "Tulio es Tulio" aun cuando los "vehículos de representación" sean diferentes; igualmente puede darse en el pensamiento, por

¹⁹ Véase, Millikan, R. G. (1991) y (1993b).

²⁰ Millikan, R. G., (1991), p. 439.

principio, el fenómeno de la ambigüedad: alguien puede captar dos posibles contenidos mediante un solo vehículo de representación, digamos, al pensar "Juárez no es hermoso".²¹ Si esto es así, Frege comete una petición de principio al concluir, a partir de la posibilidad de adoptar racionalmente actitudes distintas frente a oraciones de la forma " $a=a$ " y " $a=b$ ", que las oraciones en cuestión tienen que diferir en cuanto a su contenido semántico.²² Lo único que necesitamos para explicar la posibilidad de adoptar diferentes actitudes proposicionales -sostiene Millikan- es un ingrediente "psicológico", no semántico: el contenido perceptual cuando se le presentan al sujeto las oraciones mismas. Esto es, el mero hecho de contener notaciones diferentes las oraciones "Tulio es Tulio" y "Tulio es Cicerón" basta para que se dé la posibilidad de mover la mente de diferente manera. La mente, después de todo, es sensible a la diferencia en notación, es decir, reacciona de diferentes maneras según que un mismo contenido semántico se presente

²¹ Me refiero a la posibilidad de entender la misma oración de dos maneras: una, como si dijera que el presidente Benito Juárez no era hermoso y, otra, como diciendo que el lugar Juárez no es hermoso.

²² Millikan, en Millikan, R., (1991), argumenta que no hay ningún razonamiento que nos permita concluir, a partir de la posibilidad de tener actitudes iguales (o diferentes) frente a dos oraciones emitidas en un mismo contexto, que sus respectivos contenidos semánticos sean iguales (o diferentes). Es decir, la igualdad/diferencia que podemos captar entre dos oraciones no garantiza la igualdad/diferencia de sus contenidos semánticos, pues la igualdad/diferencia de contenido no es transparente para la mente. El error de Frege, según Millikan, es "externalizar" la igualdad/diferencia percibida, es decir, considerar que si un sujeto cree que dos pensamientos son iguales/diferentes, entonces hay algo que es objeto de la creencia y que es efectivamente igual/diferente. Esto lo lleva a introducir intermediarios semánticos (los Sinne) que, por definición, tienen las características que parecen tener; es decir, a pensar que el sujeto no puede equivocarse respecto a sus propios contenidos. La "externalización" de la diferencia está, según Millikan, a la base de la petición de principio que comete Frege.

codificado de diferentes maneras, y eso basta para explicar que un sujeto racional pueda adoptar actitudes distintas (es decir, que su mente pueda moverse de diferentes maneras) cuando escucha en un contexto dado oraciones extensionalmente equivalentes pero en las que figuran nombres diferentes. En suma, Millikan argumenta que la diferencia en "inputs" perceptuales -una cuestión meramente psicológica- es suficiente para mover a la mente de diferentes maneras y, si esto es así, que la diferencia en notación basta para explicar que un sujeto pueda tener diferentes actitudes proposicionales frente a oraciones extensionalmente equivalentes formuladas con distintas notaciones.

El argumento de Millikan equivale exactamente al rechazo de lo que Stephen Schiffer ha llamado "el constreñimiento de Frege"²³, según el cual, basta que un sujeto racional crea, por ejemplo, que Marco Tulio es calvo y que Cicerón no lo es para que tengamos que admitir que "Marco Tulio" y "Cicerón" tienen asociados diferentes sentidos, diferentes modos de presentación o contenidos. En esta sección he tratado de mostrar que Millikan rechaza la noción de sentido por considerar que hay un paso inválido en el argumento de Frege. En la siguiente sección veremos otra manera de rechazar el supuesto fregeano relativo a la existencia de sentidos; es decir, veremos cómo también puede rechazarse al reflexionar sobre sus consecuencias.

Pero vayamos a los otros criterios de Frege para detectar diferencias en "valor cognoscitivo" e introducir sentidos, y veamos si la hipótesis de Millikan puede también responder a las diferencias que ellos registran.

Con respecto al criterio (i), notemos que la diferencia en notación entre identidades verdaderas de la forma " $a=a$ " y " $a=b$ " puede explicar que pensemos

²³ Véase, Schiffer, S., (1987) cap. 3, y (1992).

que la verdad de una de ellas pueda conocerse a priori -sin necesidad de llevar a cabo una investigación empírica-, en tanto que la de la otra no. En efecto, si conocemos el uso del signo de identidad, sabemos a priori que, en general, cuando el signo está flanqueado por nombres notacionalmente iguales, la oración es verdadera.²⁴ Esto es, frente a una identidad tal como "Tulio es Tulio", generalmente sabemos que es verdadera sin necesidad de hacer una investigación sobre la historia de Roma. Por otra parte, cabe notar que el hecho de que la verdad de una oración P, sea conocida a priori y la de otra oración Q referencialmente equivalente sea conocida a posteriori no parece ser un buen indicio de que P y Q difieren en contenido semántico, pues es perfectamente concebible una situación en la que la verdad de una misma oración, tal como "Los ángulos de un triángulo trazado sobre una esfera suman más de 180°", sea conocida a priori por algún sujeto (digamos, Riemann) y a posteriori por otro (un alumno que mide los ángulos de un triángulo trazado sobre una esfera), y no diríamos que esto indicaría que la oración en cuestión tuviese diferente valor semántico "para cada uno de ellos".

Con respecto al criterio (ii) propuesto por Frege para determinar si dos

²⁴ Digo "en general", en primer lugar, por algo que ya había mencionado antes: para que una identidad de la forma "a=a" sea verdadera es necesario que "a" refiera y esto no lo sabemos a priori, salvo, tal vez, en el caso en que nosotros mismos fijamos la referencia de "a" o, si creemos en la realidad de los números, en el caso en que "a" sea un numeral. En segundo lugar, es posible formular en ciertos contextos enunciados de identidad de la forma "a=a" que sí sean informativos, por ejemplo, "Aristóteles (el filósofo) es Aristóteles (el maestro de Alejandro)" cuya verdad no se sabe a priori. Sin embargo, en circunstancias en las que sabemos que "a" refiere y que tiene el mismo referente en ambas figuraciones, en general sabemos a priori que "a=a" es verdadera. Nótese que concluir a partir de la forma de "a=a" que es verdadera, es "internalizar", en la terminología de Millikan, una igualdad.

oraciones tienen o no diferente valor cognoscitivo, alguien podría pensar que las diferencias en notación a las que apela Millikan para explicar que un sujeto pueda adoptar diferentes actitudes proposicionales frente a oraciones que tienen un mismo valor referencial, no bastan para dar cuenta del hecho de que conocer algunas identidades constituya una ampliación del conocimiento, en tanto que conocer otras no; es decir, no explican lo que el sentido pretende explicar. Y si esto es así, si las diferencias en notación no pueden dar cuenta del hecho en el que se basa el segundo criterio para introducir diferencias en valor cognoscitivo, entonces Frege parecería estar justificado en introducir los sentidos.

Basándose en una idea del propio Frege, Maite Ezcurdia ha desarrollado un argumento,²⁵ que pretende mostrar la insuficiencia de la explicación notacional de Millikan para distinguir aquellos casos en los que conocer una identidad constituye una ampliación de nuestro conocimiento de aquellos otros en los que no. Ezcurdia señala primeramente que hay identidades de la forma "a=b" que sí amplían nuestro conocimiento y otras de la misma forma que, por ser introducidas mediante una estipulación hecha por el propio sujeto que conoce la identidad, no lo amplían. Si esto es así, señala Ezcurdia, aún cuando resulte verosímil la explicación de Millikan de la posibilidad de adoptar diferentes actitudes frente a oraciones referencialmente equivalentes codificadas en signos perceptuales diferentes, su explicación notacional no puede dar cuenta de por qué algunas identidades de la forma "a=b" constituyen genuinas ampliaciones de nuestro conocimiento en tanto que otras (cuya función es estipular simplemente una

²⁵ Véase Ezcurdia, M. (1994), cap. 1. Ezcurdia se basa en la idea expresada en Frege, G., (1892) p. 4 relativa a que la conexión entre un nombre y su referente es "arbitraria" y que a nadie se le puede prohibir usar cualquier cosa como signo de algo.

identidad) no, cosa que sí se puede explicar si aceptamos la existencia de sentidos.²⁶ Si el argumento de Ezcurdia es correcto, entonces la explicación de Millikan no habrá refutado la existencia de sentidos como genuinos componentes semánticos de las expresiones lingüísticas.

Podemos resumir de la siguiente manera el argumento de Maite Ezcurdia para mostrar las limitaciones del enfoque de Millikan:

Premisa 1. Para cualquier hablante, aprender una identidad tal como "Héspero es Fósforo" constituye una valiosa extensión de su conocimiento, en tanto que aprender la identidad referencialmente equivalente "Héspero es Héspero" no constituye una extensión tal.

Premisa 2. Un hablante *h* puede introducir un nombre arbitrario, digamos "*", estipulando que refiere al mismo individuo al que refiere "Héspero"; una vez así introducido, *h* sabe que "Héspero es *", pero no diríamos que con esto ha ampliado su conocimiento.

La dificultad que estas dos premisas plantean a Millikan es la siguiente: ella,

²⁶ Para reiterar, Ruth Gareth Millikan sostiene que es una petición de principio concluir, a partir del hecho de que "Cicerón es Tulio" es informativa (amplía nuestro conocimiento) en tanto que "Cicerón es Cicerón" no lo es, que la primera oración tiene que tener un contenido (semántico) diferente del de la segunda. Véase, Millikan, R.G., (1991) p. 455. El problema que plantea Maite Ezcurdia, como veremos en seguida, depende de interpretar a Millikan como si sostuviera, además, que todos los enunciados de identidad de la forma "a=b" son informativos (o amplían nuestro conocimiento), cosa que Millikan está lejos de sostener, ya que, en sus propios términos, equivaldría a "internalizar" una diferencia.

dice Ezcurdia, no puede explicar en términos notacionales por qué "Héspero es Fósforo" sí expresa una genuina ampliación del conocimiento de h, en tanto que "Héspero es *" no. Ya que ambas oraciones de identidad comparten la característica notacional de usar nombres diferentes a los lados del signo de identidad; si Millikan explica con base en esto el hecho de que aprender la primera constituye una genuina extensión del conocimiento de h, entonces -dice Ezcurdia- tendrá que sostener que aprender la segunda también constituye una extensión tal.²⁷

Conclusión. Si Millikan no puede distinguir con su explicación notacional el hecho de que "Héspero es Fósforo" amplíe el conocimiento de h, en tanto que "Héspero es *" no, la explicación notacional de Millikan no puede explicar todo lo que la noción de "sentido" sí puede explicar.

El argumento anterior es sugerente; sin embargo, creo que presenta algunas dificultades. Véamoslas.

Primeramente, hay que señalar que no siempre aprender la verdad de la oración "Héspero es Fósforo" constituye una extensión sustancial del conocimiento; un hablante, h, puede aprender ambos nombres que figuran en la identidad por referencia a una mismísima situación epistémica. Alguien, un astrónomo, por ejemplo, le dice a h: "Mira, aquel cuerpo celeste brillante es

²⁷ Cabe señalar en este punto que Millikan no parece estar obligada a sostener que "Héspero es *" constituye siempre una extensión del conocimiento; si, en la mente de una persona, "*" no tiene asociada ninguna información diferente de la que tiene asociada "Héspero", la identidad en cuestión no será informativa. Véase más adelante.

Héspero, o Fósforo, como quieras llamarlo". En ese momento, h, que por hipótesis no tiene ninguna creencia adicional sobre Venus, sabe que es verdadera la identidad "Héspero es Fósforo", pero esto no parece ampliar sustancialmente su conocimiento, esto es, su conocimiento en esa situación es muy similar al que tiene el hablante hipotético de Ezcurdia cuando sabe que Héspero es *.

La cuestión general que quiero señalar y que, coincide con un señalamiento de Millikan,²⁸ es la siguiente: tal parece que una identidad verdadera de la forma "a=b" amplía nuestro conocimiento (o modifica la red de nuestras creencias) si y sólo si tenemos conocimientos (o creencias) previos, obtenidos por vías diferentes, acerca del referente de "a" y de "b", y expresamos tales conocimientos mediante oraciones en algunas de las cuales figura el nombre "a" y en otras de las cuales figura el nombre "b". Así, conocer la identidad "Cicerón es Tulio" amplía nuestro conocimiento sólo en la medida en que tengamos conocimientos previos acerca de Cicerón que expresamos mediante oraciones en las que figuran a veces uno de sus nombres y a veces el otro; en ese caso, al aprender la verdad de la identidad "Cicerón es Tulio" juntamos en nuestro sistema de creencias en un solo archivo, por así decirlo, las creencias que antes manteníamos en archivos separados y ponemos al archivo la etiqueta "Cicerón o Tulio".²⁹ Esto, además, nos permitirá

²⁸ Véase especialmente, Millikan, R., (1991), sección 13.

²⁹ Peter Strawson, en su libro Subject and Predicate in Logic and Grammar, ofrece una explicación similar del carácter informativo de ciertos enunciados de identidad. Mi explicación, sin embargo, difiere de la de Strawson en un respecto importante: Strawson considera que el sujeto que aprende la verdad de una identidad en la que figuran nombres propios tiene que tener, asociado con cada uno de los nombres, un conjunto de creencias que identifican (o, al menos, así lo cree el sujeto en cuestión) el referente del nombre; yo, siguiendo a Kripke, creo que esta es una exigencia excesiva. Véase, Strawson, P.F., (1974), pp. 51-56.

llevar a cabo inferencias que pueden aportarnos conocimientos novedosos. Por ejemplo, si no sabíamos que Tulio tuviera dificultades para hablar y sí sabíamos que Cicerón fue tartamudo, ahora sabremos que Tulio tenía dificultades para hablar. Pero, si no tenemos conocimientos previos asociados con cada uno de los nombres que figuran en una identidad, saber que la oración de identidad es verdadera no constituye una "ampliación del conocimiento".

En segundo lugar, y de acuerdo con lo anterior, llegar a conocer la verdad de "Héspero es *" puede constituir una ampliación sustancial del conocimiento, no para quien introduce el nombre, desde luego, sino para alguien que haya aprendido algo acerca de Venus bajo el nombre de "*". Nada impide que si una persona introduce el nombre "*" para referirse a Venus, alguien diferente de ella pueda utilizarlo de manera que el enunciado "Héspero es *" exprese una genuina ampliación de su conocimiento, de acuerdo con lo sugerido en el párrafo anterior.

Si lo anterior es correcto, me parece que la primera premisa del argumento de Ezcurdia es falsa y la segunda es también falsa si no la restringimos al caso de su hipotético hablante, es decir, si la generalizamos para cualquier hablante, esto es, si tomamos en serio la propuesta de que la semántica de "Héspero es *", según ella la introduce, es tal que aprender su verdad no constituye una genuina ampliación del conocimiento. Sin embargo, aun cuando el argumento presente estas dificultades, Ezcurdia nos conmina a explicar en términos generales y consistentes con la postura de Millikan la diferencia entre identidades que no amplían nuestro conocimiento e identidades que sí lo amplían. Es decir, tenemos que dar cuenta de lo que sucede cuando conocer una identidad supone una "ampliación genuina de conocimiento", sin apelar a los sentidos de los nombres involucrados. En lo que resta de esta sección intentaré responder a esa cuestión.

En general, podemos decir que el aprendizaje de nombres no es

independiente de la adquisición de creencias, de modo que cuando aprendemos algún nombre nuevo simultáneamente adquirimos alguna creencia, alguna información, por vaga que sea, sobre su portador.³⁰ Si "N" y "N'" son nombres notacionalmente diferentes³¹ y cada uno de ellos figura como sujeto de oraciones que expresan creencias de un sujeto s, entonces si s aprende que $\underline{N}=\underline{N}'$, su conocimiento de esa identidad parece ampliar sus conocimientos (o creencias) previos en tanto que le permite añadir a las creencias que tenía bajo "N" las que tiene bajo "N'" y, así, reunir, como dije antes, en un solo archivo creencias que antes tenía repartidas en dos. Si sólo uno de los nombres que figuran en el enunciado de identidad " $\underline{N}=\underline{N}'$ " figura como sujeto de oraciones que expresan algunas de las creencias de s, en tanto que el otro nombre no, entonces, conocer la verdad de " $\underline{N}=\underline{N}'$ " no amplía las creencias de s. Nótese que, a diferencia de la explicación de Frege en su *Begriffsschrift*, en esta explicación de la diferencia entre identidades que amplían nuestro conocimiento e identidades que no lo amplían, no hemos atribuido a quienes conocen las identidades un conocimiento puramente metalingüístico. Y nótese también que no hemos tenido que apelar a ninguna

³⁰ John McDowell señala paralelamente que a alguien que haya adquirido un nombre y sea capaz de usarlo normalmente se le pueden atribuir actitudes proposicionales en cuya especificación del contenido se mencione el nombre en cuestión, y que si es posible atribuir a una persona cualesquiera actitudes proposicionales acerca de un objeto, tiene que podersele atribuir alguna creencia sobre dicho objeto. Véase, McDowell (1977), sección 6.

³¹ Como señalé antes, incluso conocer la verdad de oraciones de identidad en la que figuran nombres notacionalmente iguales, puede ampliar nuestro conocimiento en un contexto dado. Sólo tenemos que suponer que antes de adquirir el conocimiento en cuestión, tenemos en nuestro sistema de creencias dos archivos separados que nuestro nuevo conocimiento nos permite reunir en uno solo. Piénsese en el caso de Paderewski presentado por Kripke en Kripke, S., (1979a).

noción de "sentido", nos ha bastado apelar a ciertas creencias previas en el sistema cognoscitivo del hablante.

4. Objeciones de Kripke a la noción fregeana de sentido de un nombre.

Saul Kripke dedica, en buena medida, las conferencias I y II que integran El nombrar y la necesidad a examinar la teoría de Frege³² según la cual el sentido de un nombre es una descripción definida, o a un cúmulo de descripciones definidas que determinan la referencia del nombre, es decir, tales que el individuo que las satisface es el referente del nombre. La conferencia I está dedicada al examen y la refutación de la tesis general según la cual el sentido de un nombre es una descripción que identifica al referente distinguiéndolo del resto de los individuos. La conferencia II inicia con una lista de tesis semánticas, epistemológicas y metafísicas que conforman esa que Kripke denomina "la teoría descriptonista" de los nombres propios (la cual, según afirma, puede atribuirse no sólo a Frege sino también a Bertrand Russell),³³ y luego procede a refutarlas

³² O, si se prefiere, una interpretación de la teoría de Frege, según la cual el sentido de un nombre propio equivale a un cúmulo de descripciones definidas y el referente del nombre es el objeto que satisface todas o la mayoría de tales descripciones. Véase, Kripke, S., (1980). Cabe observar que las críticas de Kripke a esta teoría se seguirían aplicando aun cuando se especificara que el sentido de un nombre no siempre es una descripción definida, sino que puede ser algún otro tipo de criterio identificador del referente, tal como, por ejemplo, una imagen perceptual.

³³ La teoría de Russell sobre los nombres propios ordinarios es en efecto descriptonista, no así su teoría sobre los llamados por él "nombres propios lógicos". Los nombres propios lógicos refieren directamente, sin ningún tipo de intermediación, y sus referentes son datos sensoriales, universales y, tal vez, el propio yo; los nombres propios ordinarios, en cambio, no son más que descripciones definidas encubiertas y denotan al único individuo que las satisface.

mediante contraejemplos que revelan las consecuencias falsas que tienen o las conclusiones contraintuitivas a las que conducen.

En esta sección me interesa: (A) revisar las objeciones de Kripke a dos tesis semánticas, a las que llamaré TRE 1 y TRE 2,³⁴ que forman parte de la teoría descriptonista de los nombres propios; y (B), siguiendo también al propio Kripke,³⁵ plantear una dificultad para cualquier teoría que -como la de Frege- tenga como una de sus consecuencias que el valor cognoscitivo de una oración relativamente a un mismo contexto puede variar de un hablante a otro. Las dificultades que aquí mencionaré, así como las que se formularon en la sección anterior de este capítulo, abrirán el camino para presentar en el capítulo III una manera diferente de concebir la referencia de los nombres propios, similar en algunos aspectos a la teoría convencionalista sostenida por Hermógenes: la idea de Kripke según la cual la referencia es una relación directa entre el nombre y su nominatum.

(A). Veamos, pues, las dos tesis dentro de la teoría descriptonista de Frege que tienen que ver con la manera como se determina la referencia de un nombre y que al aplicarse a casos particulares pueden conducir a conclusiones falsas, es decir, pueden llevar a asignar al nombre una referencia equivocada:

TRE 1. Si el sentido de un nombre es un cúmulo de descripciones definidas y el sentido determina la referencia del nombre, entonces, para todo

Véase, Russell, B.,(1905) y (1910-11).

³⁴ La tesis TRE 1 corresponde a la tesis (3) y la tesis TRE 2 a la tesis (4) en la lista de Kripke; véase, Kripke, S., (1980) p. 71.

³⁵ Véase, Keipke, S., (1979a).

nombre propio N el referente de N es el individuo que satisface todas o la mayoría de las descripciones que constituyen el sentido de N.

TRE 2. Si el sentido de un nombre es un cúmulo de descripciones definidas y el sentido determina la referencia del nombre, entonces para todo nombre N, si no hay ningún individuo, o si hay más de uno, que satisfaga las descripciones definidas que constituyen el sentido de N, entonces N no refiere.

Nótese primeramente que las tesis **TRE 1** y **TRE 2** parecen resumir de manera inmejorable las ideas de Cratilo sobre los nombres propios; recordemos que, según Cratilo, si el contenido descriptivo de "Hermógenes", por ejemplo, es el de la descripción "El del linaje de Hermes", entonces el nombre "Hermógenes" refiere al individuo que satisface esa descripción; y si no hay nadie que sea del linaje de Hermes, el nombre no refiere. La teoría descripcionista de Frege no es muy diferente de la de Cratilo en este respecto. En efecto, "Aristóteles" refiere a Aristóteles, piensa Frege, sólo porque éste satisface (todas o la mayoría de) las descripciones que constituyen su sentido, digamos, "el filósofo que fue maestro de Alejandro Magno y discípulo de Platón". Si ningún filósofo hubiera sido maestro de Alejandro o si el filósofo que fue maestro de Alejandro no hubiera estudiado con Platón, o si varios individuos satisficieran esas condiciones, "Aristóteles", según Frege, no referiría. Así, "Odiseo" no refiere porque no existe nadie que cumpla con todas o la mayoría de propiedades que se le atribuyen a Odiseo en la Odisea, y si alguien asocia con el nombre "Carlos de Windsor" la descripción "el hijo de Isabel II", el nombre no refiere porque Isabel II tiene, de hecho, más de un hijo.

Tal parece que, por principio, TRE 1 y TRE 2 pueden interpretarse de dos maneras, según que se entienda el sentido: (a) a la manera de Cratilo, como algo contenido en el nombre mismo o (b) como los contenidos de ciertas creencias del hablante que, según él, bastan para identificar al referente del nombre. Podríamos pensar que el hecho de que Frege insista en la objetividad de los sentidos, nos obliga a interpretarlos como ingredientes semánticos contenidos en los nombre mismos, de manera que no varíen de un hablante a otro, y no como los contenidos de estados mentales de las personas que usan un nombre, los cuales por principio pueden variar de una persona a otra.

Hay, sin embargo, un problema con la primera interpretación sugerida: dado lo que Frege llama la "arbitrariedad" del nombrar, difícilmente entendemos la idea de que el sentido esté contenido en el nombre mismo,³⁶ pues, si lo estuviera parecería que -como sostiene Cratilo- los nombres no podrían asignarse

³⁶ Un interesante intento de explicar esa idea dentro de una teoría semántica en la que el sentido es despojado de toda implicación epistemológica es el de Gareth Evans que he mencionado anteriormente. Según Evans en el nombre mismo está su sentido, el nombre -dice Evans- "muestra o despliega" su sentido. Como indiqué antes, Evans interpreta los sentidos como "maneras particulares de pensar (o concebir) el referente" y sostiene, por ejemplo, que pensar a Cicerón como Cicerón es diferente de pensarlo como Tulio, no porque los nombres "Cicerón" y "Tulio" estén conectados con diferentes criterios identificadores o descripciones de Cicerón, sino simplemente porque el uso de los dos nombres supone dos maneras diferentes de pensar al famoso orador romano. Véase, Evans, G., (1982), cap. 1, esp. 1.4 y 1.6. Sin embargo, el interés de Evans por la semántica de Frege es más teórico que histórico, de modo que, más que explicar lo que Frege sostuvo, le interesa explicar lo que, en su opinión, debió de haber sostenido. Así, su interpretación de Frege, aunque brillante y original, (i) difiere demasiado de las tesis del propio Frege (por ejemplo, al sostener que a cada nombre le corresponde un solo sentido y que un nombre vacío no puede tener sentido) y (ii) está animada por un espíritu "nominalista" que la aleja demasiado del espíritu de Frege.

arbitrariamente. Por otra parte, tenemos buenas razones para creer que el sentido de un nombre esté constituido por los contenidos de ciertas creencias del hablante. Si nos fijamos en lo que dice Frege acerca de la objetividad, en los pasajes en los que argumenta que el sentido de una expresión tiene que ser objetivo,³⁷ vemos que, para él, la objetividad del sentido consiste fundamentalmente en el hecho de poder ser compartido por diversos hablantes, es decir, de no ser algo privado, sino público, comunicable. A la objetividad del sentido, Frege opone la subjetividad de las representaciones, las cuales son efectivamente incompatibles. Ahora bien, si ser compartible y comunicable es suficiente para ser objetivo, entonces, de acuerdo con Frege, podemos sostener que los contenidos de las creencias son entidades objetivas, ya que distintas personas pueden tener una misma creencia, pueden captar un mismo contenido, un mismo "pensamiento".³⁸ Y, si esto es así, entonces, interpretar a Frege como diciendo que lo que determina la referencia de un nombre N al ser usado por un hablante h es el contenido de ciertas creencias C de h -el cual tiene la peculiaridad de expresarse mediante oraciones en las que figuran descripciones definidas-, no es infiel al filósofo y satisface la condición de objetividad que Frege impone a todo sentido. Hay que subrayar, sin embargo, que la objetividad del sentido, así entendida, no garantiza en absoluto que el sentido asociado con un nombre N sea el mismo para todos los usuarios de N, pues como mencioné antes, las creencias asociadas con un nombre pueden variar de un hablante a otro.

³⁷ Véase esp. Frege, G., (1892) pp. 7-10, y Frege, G., (1918- 1919), pp. 34-38.

³⁸ "Por 'pensamiento' no entiendo el proceso subjetivo del pensar, sino su contenido objetivo que puede ser propiedad común de muchos." Frege, G., (1892), n. 5.

En su libro El Nombrar y la necesidad Kripke interpreta los sentidos fregeanos de la manera que acabo de mencionar, esto es, como creencias del hablante de que N es el tal y cual; es decir, como creencias que se expresan mediante identidades en las que figuran como términos el nombre N y una o más descripciones definidas.³⁹ A esta teoría del sentido Kripke la llama "teoría descriptonista de los nombres". De acuerdo con ella, si Juan cree, por ejemplo, que Giorgione es el autor de La Tempestad, y este es el rasgo mediante el cual Juan identifica al pintor, entonces el sentido de "Giorgione" para Juan será el sentido de "el autor de La Tempestad", y quien sea que satisfaga esa descripción será el referente del nombre "Giorgione" para Juan. (Lo cual da un resultado correcto en este caso, pues quien de hecho satisface la descripción que Juan asocia con "Giorgione" es Giorgione.). Una vez interpretado el sentido de esta manera, Kripke pasa a revisar las tesis que he llamado TRE 1 y TRE 2.

El problema con TRE 1 y TRE 2, dice Kripke, es que son falsas:⁴⁰ la referencia de un nombre propio ordinario no se determina a la manera como Frege y Russell lo pensaron. Kripke muestra mediante ejemplos que si aplicamos TRE 1 y TRE 2 para determinar la referencia de un nombre propio N usado por un hablante h, en muchas ocasiones obtendríamos una respuesta equivocada. Compliquemos el ejemplo de Juan y el nombre "Giorgione" con el fin de mostrar

³⁹ Mark Platts también sostiene en su interpretación de Frege que son ciertas creencias del usuario de un nombre las que determinan el referente del nombre; dice Platts: "Frege concibió el sentido de un nombre propio como si estuviera representado por algún conjunto de descripciones, descripciones que a su vez se combinan con el nombre propio para representar el contenido de diversas creencias aceptadas por los hablantes". Véase, Platts, M., (1979), p. 141.

⁴⁰ Los argumentos de Kripke en contra de las tesis que he llamado TRE 1 y TRE 2 aparecen en Kripke, S., (1985), esp. pp. 78-87.

lo anterior. Supongamos que Juan ha aprendido el nombre "Giorgione" asociándole, como dije antes, el sentido de la descripción "el autor de La Tempestad". Pero supongamos también que, sin saberlo nadie, el autor de esa pintura generalmente atribuida a Giorgione, es en realidad Tiziano; es decir, Tiziano es quien satisface la descripción que para Juan constituye el sentido de "Giorgione". Si este es el caso, entonces, de acuerdo con TER 1, al usar Juan el nombre "Giorgione" y afirmar "Giorgione pintó La Tempestad" se refiere, sin saberlo él, y sin que lo sepa nadie, a Tiziano y, peor aún, afirma algo verdadero, aun cuando en el ejemplo suponemos que es falso que Giorgione haya pintado La Tempestad.⁴¹

Alguien podría alegar que, en el ejemplo modificado, Juan en efecto se refiere a Tiziano cuando usa "Giorgione", pues él decide dar el nombre a quien sea que haya pintado La Tempestad. Admitamos que esto podría ser el caso, siempre y cuando Juan no tenga otras creencias ni sobre Giorgione, ni sobre Tiziano. Supongamos que Juan sólo ha visto La Tempestad y ha estipulado: sea "Giorgione" el nombre de quien sea que haya pintado este cuadro, y luego afirma "Giorgione pintó La Tempestad".⁴² Sin embargo, fijémonos que en ese caso sólo

⁴¹ En lo anterior parece estar implicada una especie de reducción al absurdo de la teoría descriptivista de los nombres (TDN): Si TDN es verdadera, entonces, los objetos de las actitudes proposicionales son transparentes para el hablante y éste sabe siempre a quién o qué se refiere (P); pero, como acabamos de ver, si TDN es verdadera entonces no es el caso que el hablante sepa siempre a quién o qué se refiere (-P), ya que como dijimos en el ejemplo, Juan se refiere sin saberlo a Tiziano. Si TDN -> P&-P, y puesto que -(P&-P), entonces -TDN.

⁴² "Giorgione" sería en la terminología de Gareth Evans un "nombre descriptivo", es decir, un nombre cuya referencia se estipula que es el individuo que satisfaga una determinada descripción. Véase, Evans, G., (1978 y véase, también, el Cap.III de esta tesis.

puede ser correcto decir que al afirmar "Giorgione pintó La Tempestad" Juan se refiere a Tiziano con el nombre "Giorgione", si Juan no tomó ese nombre de nadie más, es decir, si no lo leyó en algún libro, ni lo oyó de boca de otra persona de su comunidad, sino que lo inventó y lo introdujo él mismo. Si, por el contrario, Juan hubiera adquirido el nombre "Giorgione" a partir, por ejemplo, de la lectura del libro de Vasari o por haberlo escuchado de boca de algún conocedor de pintura veneciana renacentista, Juan se referiría con ese nombre a Giorgione, aun cuando creyera falsamente (en nuestro ejemplo) que Giorgione pintó La Tempestad; en particular, Juan no se referiría en ese caso a Tiziano, aunque éste fuera quien satisficiera la descripción que Juan asocia con el nombre.⁴³ Sin embargo, según TRE 1, como he señalado, bastaría que Juan diera a "Giorgione" el sentido de la descripción "el autor de La Tempestad" para decir -aun en el caso de que Juan adquiriera el nombre a partir de su lectura de Vasari- que Juan se refiere con "Giorgione" a Tiziano.

Variemos ahora el ejemplo un poco y supongamos que La Tempestad fue pintada por varios pintores y que todos ellos contribuyeron igualmente para producir esa misteriosa tela. En esa caso, según TRE 2, Juan, para quien el nombre "Giorgione" significa el autor de La Tempestad, no refiere. Igualmente, si por improbable que parezca, nadie hubiera pintado La Tempestad, sino que, como dice Kripke en uno de sus notables ejemplos, la pintura "se hubiera materializado debido a una distribución fortuita de átomos" sobre la tela, si Juan

⁴³ Digo que Juan no se referiría a Tiziano en esa situación porque tenemos la firme intuición de que, aun si fuera el caso que Giorgione no hubiera pintado La Tempestad, si Juan de hecho aprendió el nombre "Giorgione" al leer a Vasari, lo que haría verdadera su afirmación de la oración "Giorgione pintó La Tempestad" sería el hecho de que Giorgione, y nadie más, hubiera pintado La Tempestad. Cfr., Kripke, S., (1985), "Prefacio".

afirmara en ese caso: "Giorgione fue un famoso pintor veneciano", su afirmación, según TRE 2, sería falsa (o carecería de valor de verdad), ya que nadie satisfaría la descripción que según Juan constituye el sentido del nombre y, por lo tanto, Juan no se referiría a nadie al usar el nombre "Giorgione".

Los anteriores parecen ser contraejemplos que plantean dificultades insuperables para cualquier teoría que sostenga que los nombres tienen sentidos descriptivos y que esos sentidos determinan la referencia de los nombres, es decir, que refieren a los individuos que satisfacen las descripciones definidas asociadas por los hablantes con el nombre. Cabe observar que los mismos contraejemplos pueden construirse si se cambian las descripciones definidas por conceptos individuales o por algún otro tipo de representación del individuo al que intenta referirse un hablante al usar un nombre propio. Es decir, se puede mostrar mediante contraejemplos que hay casos en los que un individuo se ajusta a la representación o al concepto individual que el hablante asocia con un nombre N, y sin embargo ese individuo no es el referente de N.

Veamos ahora un segundo tipo de objeción que Kripke plantea a la teoría según la cual los nombres expresan sentidos y dos nombres que refieren al mismo individuo pueden expresar diferentes sentidos.

(B). Si (i) lo que determina la referencia de un nombre N es su sentido y éste es equivalente a ciertas creencias del hablante que se expresan mediante identidades en las que figura N y alguna descripción definida (por ejemplo, la creencia de Juan de que Giorgione es el autor de *La Tempestad*), y (ii) el sentido de una oración es una función del sentido de sus partes; entonces, una misma oración, en boca de diferentes hablantes, puede expresar diferentes sentidos completos, distintos pensamientos fregeanos. En efecto, si Juan cree que Giorgione

es el autor de La Tempestad, y esa creencia de Juan constituye el sentido de "Giorgione" para Juan, y Pedro cree que Giorgione es el más notable pintor veneciano y esa creencia de Pedro constituye el sentido del nombre "Giorgione" para Pedro, entonces Juan y Pedro asocian sentidos diferentes con un mismo nombre. Y, si el sentido de una oración es una función del sentido de sus partes, al afirmar Juan y Pedro: "Giorgione nació en Castelfranco", cada uno de ellos expresará un sentido diferente, Juan expresará el pensamiento de que el autor de La Tempestad nació en Castelfranco,⁴⁴ en tanto que Pedro expresará el pensamiento de que el más notable pintor veneciano nació en Castelfranco. Si conjuntamos con lo anterior otra tesis semántica de Frege, según la cual una oración que figura como subordinada en una cita indirecta o en una atribución de actitud proposicional deja de denotar su denotación normal y pasa a denotar lo que en contextos directos era su sentido, tendremos que concluir que si ambos, Juan y Pedro, afirman sinceramente: "Creo que Giorgione nació en Castelfranco", el objeto de la creencia de Juan será diferente del objeto de la creencia de Pedro.

Lo anterior plantea una dificultad insuperable a la teoría del sentido propuesta por Frege: si dos sujetos, como Juan y Pedro en nuestro ejemplo, asocian sentidos diferentes con un mismo nombre N y ambos afirman la misma oración "N es F", ninguno de ellos puede citar con verdad lo que el otro afirmó, ni puede, con base en la afirmación hecha por el otro, atribuirle con verdad una creencia, usando en la cita o en la atribución de creencia la misma oración que usó el otro al hacer su afirmación. Así, si, por ejemplo, Juan, en nuestro ejemplo, afirma: "Giorgione nació en Castelfranco" y Pedro quiere informar lo que

⁴⁴ No tomo aquí en cuenta el hecho de que los nombres propios "La Tempestad" y "Castelfranco" expresan a su vez sentidos que pueden también diferir de un hablante a otro.

literalmente afirmó Juan y dice: "Juan afirma que Giorgione nació en Castelfranco", lo que dice Pedro es falso. O si, después de oír la afirmación de Juan, Pedro quiere atribuirle la creencia correspondiente y dice "Juan cree que Giorgione nació en Castelfranco", su atribución será falsa también. Pues, como señalé antes, si Juan y Pedro asocian sentidos diferentes con el nombre "Giorgione" y, por lo tanto, expresan diferentes sentidos con la oración "Giorgione nació en Castelfranco", no pueden usar esa oración para citar indirectamente lo que el otro dijo ni para hacerle una atribución de creencia; y esto contraviene absolutamente nuestras prácticas de cita y de atribución de creencias. En particular, contraviene el principio de desentrecomillación para la atribución de creencias.⁴⁵

Dadas las dificultades que hemos expuesto, tenemos buenas razones para examinar en los siguientes capítulos de esta tesis una concepción alternativa de los nombres y la relación referencial que ha tenido gran resonancia y fecundidad y una aceptación cada vez mayor desde hace veinticinco años cuando se publicó la obra más señera en la que se expone tal concepción; me refiero a la teoría de la referencia de los nombres propios sostenida por Saul Kripke en El nombrar y la necesidad.

Además de exponer diversos aspectos de la concepción kripkeana de la referencia y los nombres, los tres siguientes capítulos abordarán las dificultades más importantes a las que tiene que enfrentarse tal concepción. El capítulo III

⁴⁵ Dicho brevemente, el principio de desentrecomillación establece que si un hablante *h* afirma sincera y reflexivamente "p", entonces *h* cree que p. Véase Kripke, S., (1979a). Ruth Barcan Marcus propone restringir el principio de manera que sólo puedan figurar en lugar de "p" oraciones que expresen estados de cosas posibles; véase Barcan-Marcus, R., (1981) y (1983). Véase también la discusión sobre este tema en el capítulo V más adelante.

abordará la cuestión de si efectivamente Kripke está comprometido a sostener que podemos conocer a priori hechos contingentes; el capítulo IV examinará la cuestión de si la concepción de los nombres como designadores rígidos nos compromete, como sostiene el propio Kripke, con el dualismo mente-cuerpo. Y, por último, en el capítulo V de esta tesis se abordará la difícil cuestión de cómo puede explicar una teoría de la referencia directa, como la propuesta por Saul Kripke, el comportamiento de los nombres en contextos de actitudes proposicionales.

CAPÍTULO III

Referencia directa y conocimiento a priori de verdades contingentes¹

Introducción.

En El nombrar y la necesidad,² Saul Kripke presenta una manera de ver la referencia de los nombres propios ordinarios radicalmente diferente de la de Frege. Kripke regresa a la tradición milliana³ y, en contraste con Frege, sostiene que la referencia de un nombre no se determina mediante la satisfacción de ningún contenido descriptivo asociado con él y que la referencia es, entonces, una relación directa, no mediada por "sentidos", entre el nombre y la cosa nombrada.

Cualquier teoría se evalúa por sus consecuencias. Consideramos como una

¹ Una versión anterior de este capítulo intitulada "Contingencia a priori" fue publicada en Crítica, vol.xx, nº 59, agosto de 1988.

² Kripke, S. (1980) Naming and Necessity (en adelante se citará N & N); los números de página en las citas corresponden a la versión revisada y aumentada de esta obra publicada en Basil Blackwell, Oxford.

³ John Stuart Mill, en su libro System of Logic, sostiene que los nombres propios se relacionan directamente con sus referentes y que, a diferencia de otra clase de términos, no tienen "connotación". "Los únicos nombres que no connotan nada son los nombres propios"; los nombres propios -dice Mill- no indican ninguna cualidad del individuo que denotan, nada que puedan tener varios individuos en común, un nombre simplemente "está en lugar de la cosa que nombra". Mill, J.S., (1843), Libro I, cap.2.

prueba a favor de una teoría científica el hecho de tener consecuencias que concuerden con otras teorías previamente aceptadas o con nuestras observaciones; en el caso de una teoría filosófica, el hecho de concordar con nuestras intuiciones es indicio de su verdad. Si, por el contrario, una teoría tiene consecuencias que chocan con proposiciones de otras teorías aceptadas, con la experiencia o con nuestras intuiciones más generales, muy probablemente la rechazaremos, como también rechazaremos aquellas tesis de las que se deriven por considerar que tienen consecuencias falsas. En el presente capítulo examinaré una tesis que Kripke presenta como una consecuencia de ciertas ideas centrales de su concepción de los nombres propios que parece ir en contra de algunas de nuestras intuiciones sobre nuestras capacidades epistemológicas. Me refiero a la tesis kripkeana según la cual es posible conocer a priori⁴ verdades contingentes, esto es, proposiciones que resultan ser verdaderas del mundo real, al que denominaremos m*, pero que podrían ser falsas respecto de algún otro mundo posible m_{N & N}, pp. 54-55).⁵

Kripke presenta esa tesis como una consecuencia de otras tres tesis defendidas en su obra: (1) la idea de que hay una distinción tajante entre las categorías epistemológicas y las metafísicas, (2) la idea de que los nombres son designadores rígidos, y (3) la idea de que si se usa una descripción para introducir

⁴ En la nota 26 de N & N, Kripke dice que llamará "a priori" a los enunciados cuya verdad se siga de una 'definición' que fija la referencia de un nombre. En otros pasajes, en cambio, utiliza la expresión "a priori" en la acepción más común que se origina en Kant y según la cual son a priori aquellos conocimientos que obtenemos "independientemente de la experiencia", esto es, sin necesidad de mirar cómo es el mundo. Es en este segundo sentido, bastante vago por otra parte, que entenderemos la expresión "a priori" en este capítulo. De otra manera la verdad de la tesis de Kripke se seguiría trivialmente de su propia estipulación.

⁵ Para la noción de mundo posible, véase, Kripke (1980), "Prefacio". Véase también, Bradley, R., y Swartz, N., (1979).

un nuevo nombre o para explicar a qué cosa refiere un nombre previamente introducido, con ello no se convierte al nombre en "sinónimo" de la descripción. En la sección 1 examinaré brevemente cada una de estas tesis; en esa misma sección trataré de mostrar cómo parece derivarse de ellas la consecuencia que advierte Kripke. En las secciones subsiguientes examinaré la discusión de esta dificultad por parte de Keith Donnellan, Michael Dummett y Gareth Evans. Concluiré haciendo un balance que nos permita ver lo que realmente se sigue de las tesis semánticas propuestas por Kripke, lo cual no parece chocar realmente con nuestras intuiciones.

1. Contingencia, designadores rígidos y descripciones que fijan la referencia.

Veamos, pues, las tesis (1), (2) y (3) de las que Kripke deriva la conclusión de que podemos conocer a priori verdades contingentes.

(1). Kripke distingue la necesidad, en tanto que categoría metafísica, de la aprioridad en tanto que categoría epistemológica.⁶ Lo que expresa cualquier enunciado verdadero es tal que o bien podría haber sido de otra manera de como de hecho es (como sucede con "Giorgione pintó La Tempestad") o bien no podría haber sido de otra manera de como de hecho es (como sucede con " $2+2=4$ ", "Giorgione es Giorgione" o "El agua es H_2O "); en el primer caso decimos que se trata de una verdad contingente, en el segundo caso decimos que se trata de una verdad necesaria. La necesidad y la contingencia son categorías metafísicas, no epistemológicas, en tanto que conciernen a lo que podría o no podría haber sido

⁶ Véase Kripke, (1980), esp. pp. 34-40.



el caso, no a nuestro conocimiento de ello. Por otra parte, hay por lo menos dos maneras distintas de conocer verdades: conocemos algunas verdades sin necesidad de investigar cómo es el mundo real, en tanto que conocemos otras verdades sólo mediante la investigación empírica. Si un sujeto conoce una verdad de la primera manera decimos que la conoce a priori, si la conoce de la segunda manera decimos que la conoce empíricamente o a posteriori; estas categorías son epistemológicas, no metafísicas, pues conciernen a nuestras maneras de conocer, y no al estatus metafísico que tienen las cosas que conocemos.⁷ Ahora bien, Kripke argumenta que las nociones de necesidad y aprioridad, lejos de ser equivalentes (como han supuesto muchos filósofos), ni siquiera están lógicamente conectadas. En contra de lo que supone Kant en su Crítica de la Razón Pura,⁸ Kripke sostiene que no todas las verdades necesarias se conocen a priori; muestra mediante ejemplos que puede haber verdades necesarias conocidas a priori (como la que expresa " $2+2 = 4$ "), otras conocidas a posteriori (como la que expresa "El agua es H_2O "), y otras, tal vez, que nadie conoce ni a priori ni a posteriori (como la conjetura de Goldbach, según la cual todo número par mayor que 2 es la suma de dos primos); de modo que el carácter necesario o contingente de una proposición no exige que se le conozca de una manera y no de otra. Las categorías metafísicas y las epistemológicas -argumenta Kripke- son categorías lógicamente independientes y -

⁷ Nótese que nada impide que una verdad que es conocida a priori por un sujeto sea conocida a posteriori por otro; las categorías epistemológicas no son absolutas, sino relativas a los diferentes sujetos epistemológicos. Nótese también que, en concordancia con lo anterior, Kripke evita caracterizar una verdad a priori como una que en general puede conocerse a priori.

⁸ Véase, Kant, E., (1787), B3-B6. En ese pasaje Kant toma la necesidad de un "juicio" como único criterio de que la verdad que expresa proviene de una fuente a priori.

en contra de lo que Kant supuso- no hay argumentos que las conecten, esto es, que demuestren que una proposición p es necesaria si, y sólo si, p es conocida a priori.

(2). Kripke llama "designadores" tanto a los nombres como a las descripciones definidas.⁹ Un designador puede o bien designar al mismo individuo en toda situación real y contrafáctica o bien a diferentes individuos según que se hable de diferentes situaciones contrafácticas; en el primer caso decimos que el designador es rígido, en el segundo, que es no rígido. Los nombres propios - sostiene Kripke- son siempre designadores rígidos, es decir, si designan a un individuo x en el mundo real, m^* , designarán a ese mismo individuo en todo mundo posible en el que x exista.¹⁰ Las descripciones definidas accidentales, en cambio, son designadores no rígidos,¹¹ pues es posible que designen a un individuo x en m^* y a otros objetos diferentes de x en otros mundos posibles o situaciones contrafácticas.¹² De acuerdo con lo anterior, el nombre propio

⁹ Kripke, S., (1980), p. 24.

¹⁰ Kripke, S., (1980), p. 48.

¹¹ Una descripción definida designa al único individuo que la satisface. Kripke distingue entre descripciones definidas esenciales y accidentales. Las primeras sí son designadores rígidos, ya que son satisfechas por el mismo individuo en todo mundo posible en el que dicho individuo exista. Las descripciones definidas accidentales, en cambio, son designadores no rígidos, ya que seleccionan a un objeto x por ser el único que tiene en m^* una propiedad accidental. A lo largo de este capítulo, cuando se hable de "descripciones" ha de entenderse que se trata de descripciones definidas accidentales. Esta clase de descripciones son las comúnmente usadas para fijar la referencia de un nombre y, como veremos, las que se requieren para que surja la posibilidad de conocer a priori verdades contingentes.

¹² M. Dummentt interpreta la distinción entre designadores rígidos y no rígidos como equivalente, en último término, a la distinción entre términos singulares con alcance largo y corto en contextos modales, véase Dummett, M., (1981), pp. 112-

"Aristóteles" designa en todo mundo posible o situación contrafáctica al mismo individuo al que de hecho designa ese nombre en m^* , en tanto que la descripción "el discípulo más notable de Platón" designa a Aristóteles en m^* pero podrá designar a otros individuos diferentes de Aristóteles en otros mundos posibles, por ejemplo, en una situación contrafáctica en la que Aristóteles muere a los cinco años.

(3). Un hablante h puede introducir un nombre N estipulando que el referente de N será la única cosa que satisfaga la descripción "el Φ "¹³ o, si N es un nombre previamente introducido en el lenguaje, h puede usar una descripción definida para explicar a otros a qué refiere N . Ahora bien, una descripción definida usada de esas maneras, sea para introducir un nombre o sea para identificar al referente de un nombre, fija simplemente la referencia del nombre en cuestión, pero no nos proporciona un sentido necesariamente asociado con él, ni convierte al nombre en sinónimo de la descripción; la relación entre un nombre

115. Kripke rechaza esta interpretación argumentando: (i) que la noción de designador rígido se aplica a cualquier nombre, incluso a los que aparecen en oraciones simples, sin operadores modales ni de otro tipo, en las que, por lo tanto, no cabe hacer distinciones de alcance, y (ii) que la idea de rigidez tiene que ver con nuestras intuiciones acerca de qué situación tendría que darse para que una oración en la que figura un nombre fuera verdadera en una situación contrafáctica. Si "Aristóteles" refiere a Aristóteles en m^* la única situación que puede hacer verdadera la oración "Aristóteles ama a los perros" en cualquier otro mundo posible es una en la que Aristóteles mismo, y nadie más, ama a los perros; véase, N & N, esp. pp. 5-9 y 14.

¹³ Este tipo de estipulación se hizo, por ejemplo, cuando se introdujo el nombre "UNA Bomber" para designar a quien sea que haya cometido ciertos actos terroristas en universidades americanas entre 1993 y 1996. Kripke ofrece el ejemplo de "Jack el Destripador" y, como veremos más adelante, el de "Neptuno".

y una descripción accidental que fija su referencia no es una relación analítica.¹⁴ En consecuencia tampoco es analítico un enunciado de identidad de la forma "N es el F" en el que figuran un nombre y la descripción mediante la que se fijó su referencia.¹⁵ Esto explica el hecho de que cuando fijo la designación de un nombre, pongamos por caso "Kripke", mediante una descripción definida, digamos "el autor de N & N", puedo más tarde afirmar sin contradecirme que Kripke podría no haber sido el autor de N & N;¹⁶ "el autor de N & N" sólo se usó en m* para fijar la referencia de "Kripke" en todo mundo posible, incluso en aquellos mundos en los que Kripke no escribe N & N, y no para señalar ni establecer una identidad de "sentidos" entre la descripción y el nombre. La rigidez del nombre y la no rigidez de la descripción, por otra parte, hacen imposible sostener que el nombre N y la descripción que fija su referencia "signifiquen" lo mismo, o que necesariamente hayan de designar al mismo individuo en toda situación contrafáctica o mundo posible.

A las tres tesis anteriores podríamos añadir una cuarta tesis kripkeana: la de que el nombre no tiene, en contra de lo que pensó Frege, un sentido. Conocemos

¹⁴ Frege parece pensar, por el contrario, que sí lo es: la relación entre la descripción que usa un hablante h como sentido de un nombre y el nombre mismo, es analítica, al menos para h. Por ejemplo, si h introduce el nombre "Aristóteles" mediante la descripción "El maestro de Alejandro Magno y discípulo de Platón", entonces "Aristóteles fué discípulo de Platón" expresa una verdad analítica (y a priori) para h. Véase, Frege, G., (1892), esp. n.2.

¹⁵ Un enunciado es analítico es "verdadero en virtud de su significado y verdadero en todos los mundos posibles en virtud de su significado." Kripke, (1980), p. 39.

¹⁶ Incluso en casos como el del nombre "UNA Bomber" mencionado en la nota 13, no es contradictorio decir que UNA Bomber podría haber sido una persona inofensiva.

directamente a los individuos en nuestras inmediaciones (y no sólo indirectamente a través de sus propiedades observables) y les damos nombres que refieren directamente a ellos (y no indirectamente, a través de las propiedades que los distinguen de los demás individuos).¹⁷ Decir que los nombres no tienen un sentido fregeano quiere decir dos cosas: (a) que no tienen un contenido conceptual o descriptivo asociado necesariamente y (b) que en caso de que un hablante (o una comunidad de hablantes) asocie mentalmente alguna descripción a un nombre, no será esa descripción la que determine la referencia del nombre.¹⁸

No son, pues, las descripciones que usamos eventualmente para fijar la referencia de un nombre las que determinan su referencia. Ésta se determina de una manera diferente: (i) en el caso de introducir un nuevo nombre, la referencia se determina mediante un acto en el que un hablante impone un nombre a un individuo, es decir, mediante un bautismo; y (ii) en el caso en que un hablante h usa un nombre N previamente introducido en el lenguaje, por la cadena de comunicación que va desde el uso que hace h de N en esa ocasión particular hasta el bautismo originario en el que alguien más le asignó por primera vez un referente a N. De modo que lo que determina a qué individuo refiere un nombre N cuando es usado por un hablante h, no es nada que esté en la cabeza o en la mente de h,

¹⁷ Véase, Kripke, S., (1980), pp.49-54.

¹⁸ Véase Kripke, S., (1980), p. 59. Volviendo a la terminología del capítulo 1, podemos decir no sólo que un individuo puede cometer un error "interno" al asociar una descripción con un nombre, es decir, un error relativo a la descripción que asocia con un nombre su comunidad, sino que la comunidad misma puede asociar una descripción equivocada con un nombre. Kripke da el ejemplo del nombre bíblico "Jonás" que según los estudiosos de la Biblia refiere realmente a un profeta al que, sin embargo, se la suele identificar mediante una información falsa: haber sido tragado por una ballena.

sino algo externo a h.¹⁹ De acuerdo con lo anterior, para que un hablante h que aprende un nuevo nombre N se refiera al referente original de N, al igual que todos los hablantes que integran la cadena de comunicación de la que él forma parte, h tiene que tener la intención de usar el nombre con la misma referencia que la persona de quien lo aprendió, pero no tiene que contar con ningún criterio descriptivo o representación mental que identifique al referente de N y le permita distinguirlo del resto de los individuos. Así, para que Lucas se refiera, por ejemplo, a Hamurabi con "Hamurabi", Lucas no necesita conocer una descripción que sólo se aplique a Hamurabi, basta que use el nombre "Hamurabi" con la intención de hacer la misma referencia que el hablante de quien tomó el nombre y que exista, además, una cadena causal que vaya desde el uso que hace Lucas de "Hamurabi" hasta Hamurabi mismo.

Si aceptamos las tesis anteriores se sigue -dice Kripke- la consecuencia de que al menos en algunas ocasiones es posible conocer a priori verdades contingentes expresadas mediante identidades que hacen uso de un nombre y de una descripción. Veamos cómo se deriva dicha supuesta consecuencia elaborando para ello un ejemplo del propio Kripke.²⁰

Consideremos el caso en el que, antes de haber observado a Neptuno, Leverrier introduce el nombre "Neptuno" (X de aquí en adelante) para designar a cualquiera que sea el objeto que satisfaga la descripción "la causa de las

¹⁹ Véase, Kripke, S., (1980), pp. 91-97. La explicación de Kripke se distingue de la posibilidad que menciona P. Strawson, en Strawson, P., (1959), p. 182. Según este último, aún cuando h no cuente con una descripción que identifique al referente de un nombre N que usa, lo que determina la referencia de N es una descripción identificadora en la mente de otro hablante de quien h toma la referencia de N.

²⁰ Véase, Kripke, S., (1980), n. 33.

perturbaciones en la órbita de Urano" ("el Φ " de aquí en adelante). Parece claro que cuando Leverrier introduce X de esta manera, no está usando X como sinónimo de "el Φ ", ni proponiendo que se considere el nombre como una abreviatura de la descripción. Si así fuera, Leverrier no podría afirmar sin contradecirse que X podría no haber sido el Φ , y su afirmación de que X es el Φ sería trivial o analítica. El hecho de que Leverrier pueda pensar consistentemente en una situación contrafáctica en la que debido, por ejemplo, a algún accidente cósmico X no fuese el Φ , muestra por lo menos dos cosas: (1) que la proposición de que X es el Φ es contingente, pues su negación es posible, y (2) que el significado o sentido de "el Φ " no es el significado o sentido (si es que hay alguno) de X ; " X es el Φ " es una oración cuya verdad no se sigue solamente de los significados de los términos que aparecen en ella, no es una oración analítica. De lo anterior parece resultar que al afirmar Leverrier que X es el Φ , antes de haber observado a Neptuno por el telescopio, estaba expresando una proposición verdadera (puesto que X efectivamente es el Φ), contingente (dado que podría ser el caso que X no fuese el Φ) y, dado que Leverrier fijó la referencia de X mediante la descripción "el Φ ", una proposición cuya verdad Leverrier sabe aparentemente a priori, pues Leverrier sabe que X es el Φ sin necesidad de recurrir a ninguna experiencia u observación ulterior.²¹

La tesis de Kripke, como vemos, no se basa en casos aislados de enunciados

²¹ El ejemplo clásico de Kripke para argumentar que hay conocimiento a priori de verdades contingentes es el de la introducción del metro como unidad de medida. La verdad de "La barra B mide un metro en t_0 ", dice Kripke, es conocida a priori por quien haya fijado la referencia de "un metro" mediante la descripción "la longitud de la barra B en t_0 "; sin embargo, es contingente que B mida un metro, ya que esa misma barra podría haber tenido otra longitud. Véase, Kripke, S., (1980), pp. 56 y ss.

"raros" que serían contingentes si bien conocidos a priori, sino que es una tesis general²² de acuerdo con la cual contamos con un mecanismo en nuestro lenguaje que nos permite la introducción sistemática de enunciados cuya verdad tiene aquellas características. Veamos, pues, la estructura de dicho mecanismo.

Sean N un nombre cualquiera (un designador rígido) y "el Φ " una descripción que de hecho, en m*, satisface un objeto o y mediante la cual un hablante h fija la referencia de N. Supongamos que "el Φ " es una descripción accidental de o, es decir, una que satisface o en m* pero que podría no satisfacer en otro mundo posible m, y supongamos también que h no tiene ninguna otra forma de identificar a o independientemente de la descripción "el Φ ". Si en esas condiciones h afirma que N es el Φ , la afirmación de h parece expresar una verdad contingente cuyo conocimiento no depende de ninguna observación ulterior y, en ese sentido, es a priori para h. Cualquier nombre y cualquier descripción que se ajusten al esquema anterior podrían generar proposiciones contingentes de las que el hablante tendría un conocimiento a priori. Veamos este mecanismo en un caso más.

Tomemos, por ejemplo, el nombre "Acautli" y la descripción "el primer azteca sacrificado en honor a Huitzilopochtli"; estipulo entonces:

(E) Llamaré "Acautli" a quien sea que haya sido el primer azteca sacrificado en honor a Huitzilopochtli.

Con esto he fijado la referencia de "Acautli". A continuación afirmo:

²² Véase, Donnellan, K., (1979b).

- (C) Si existe (existió) el primer azteca sacrificado en honor a Huitzilopochtli, Acautli es el primer azteca sacrificado en honor a Huitzilopochtli; o simplemente
- (C') Acautli es el primer azteca sacrificado en honor a Huitzilopochtli.

(C) expresa una proposición contingente, pues hay un mundo posible en el que su antecedente es verdadero y su consecuente falso. (C') también expresa una proposición contingentemente verdadera (asumiendo, desde luego, que algún individuo satisfaga la descripción "el primer azteca sacrificado en honor a Huitzilopochtli"). Sin embargo, dada mi estipulación (E), parece que no tengo que apelar a ninguna observación ni experimento para establecer la verdad de (C) y (C') y, en este sentido, parece que sé a priori que (C) y (C').²³

La mera posibilidad de que existan enunciados contingentes cuya verdad sea conocida a priori parece ir en contra de nuestras intuiciones y de las de toda una tradición filosófica. En efecto, por un lado, se ha pensado que si hay algo que podamos conocer a priori, es decir, sin necesidad de mirar cómo es este mundo, lo conocido tiene que ser verdadero de todo mundo posible, pues no depende su verdad de ningún rasgo distintivo del mundo real. Es decir, si la manera como es este mundo no tiene que ser investigada para establecer la verdad de *p*, entonces *p* tiene que ser necesaria. Por otro lado, si *p* es contingente, esto es, si *p* es

²³ Un nombre introducido de la manera antes señalada es, en un sentido importante, un nuevo nombre. Si antes de introducir el nombre "Acautli" a la manera como se sugiere en el texto, ese mismo fonema fuese el nombre de un tlatoani azteca notable, al estipular que "Acautli" haya de ser el nombre del primer azteca sacrificado en honor a Huitzilopochtli se inicia una nueva cadena causal y, en ese sentido, se introduce un nuevo nombre.

verdadera de este mundo pero no de cualquier mundo posible, se piensa que es necesario investigar empíricamente cómo es este mundo para conocer la verdad de *p*, pues ¿cómo podríamos llegar a saber algo verdadero del mundo real, pero no cualquier mundo posible, si no usamos nuestros sentidos? Kripke mismo expresa este tipo de reflexiones en N & N (pp. 38-39); sin embargo, si de veras son lógicamente independientes las nociones epistemológicas de las metafísicas, no es trivial que algo conocido a priori tenga que ser necesario, como tampoco es trivial que las verdades contingentes tengan que conocerse a posteriori.

La tesis de Kripke ha dado lugar a una amplia discusión en la literatura filosófica reciente.²⁴ Dummett expresa directamente que si tal tesis es realmente una consecuencia de la concepción kripkeana de los nombres propios, entonces algo debe estar equivocado en dicha concepción.²⁵ Donnellan, como trataré de mostrar en la próxima sección, argumenta que lo que Kripke piensa que es una consecuencia de su tesis acerca de la rigidez de los nombres en realidad no lo es. Gareth Evans, a su vez, argumenta en contra de Donnellan y pretende encontrar en otro rasgo del lenguaje, no en la rigidez de los nombres, el origen de los

²⁴ Véase, Donnellan, K. (1979b); Evans, G. (1978); Stephen Schiffer, (1979); Dummett, M. (1981), cap. 5, y "Appendix"

²⁵ Véase M. Dummett, (1981), p. 121. No examinaré aquí, por falta de espacio, la discusión de Dummett. Sólo deseo señalar que según él, sería incorrecto decir que al afirmar Leverrier "Neptuno es el Φ " expresa su conocimiento de un hecho contingente. En las pp. 122 y 123 de su libro, leemos "si alguien no tiene ninguna idea de qué longitud es un metro, salvo por el hecho de saber que es la longitud de una barra en París que él nunca a visto, puede saber a priori que el metro patrón mide un metro, pero en un sentido intuitivo él no sabe cuánto mide el metro patrón y, por lo tanto, al afirmar lo que afirma no expresa ningún conocimiento de ningún hecho contingente."

enunciados contingentes a priori.²⁶ En la sección 4 de este capítulo, examinaré como Evans intenta disolver la aparente paradoja distinguiendo entre dos tipos de contingencia, una profunda y otra superficial y mostrando que donde Kripke creyó ver una contingencia profunda, en realidad no hay más que una contingencia superficial que no tiene por qué inquietarnos.

2. La solución de Donnellan

En su artículo "The Contingent A Priori and Rigid Designation",²⁷ Donnellan introduce una útil distinción en esta discusión, a saber, la distinción entre saber que p es una oración verdadera y conocer la verdad expresada por p.²⁸ En la mayoría de los casos en que sé que p expresa una verdad, conozco la verdad que expresa; pero bien puede darse el caso de que, aun sabiendo que p es verdadera, desconozca la verdad expresada por p. Donnellan ofrece algunos ejemplos que ilustran esta distinción. Supongamos que en un Congreso Internacional de Matemáticas escucho al matemático Y demostrar ante un público de especialistas el teorema p. Los matemáticos más capaces de juzgar sobre la

²⁶ Según Evans, la posibilidad de que haya oraciones contingentes conocidas a priori depende, como veremos, de que contemos con propiedades en cuya especificación se hace referencia al mundo real; por ejemplo, la propiedad de ser de hecho rojo o de ser realmente rojo o rojo en m*. Véase la sección 4 más adelante.

²⁷ Donnellan, K., (1979b).

²⁸ Kripke reconoce la distinción que hace Donnellan en N & N, p. 69: "por supuesto, cualquiera que sepa el uso de 'es llamado' en el idioma, sabe que si 'quarks' significa algo, entonces 'los quarks son llamados "quarks"' expresa una verdad, aun cuando no sepa lo que el enunciado significa".

corrección de la prueba se hallan presentes y celebran que se haya probado la verdad de p . Sin embargo, la prueba es larga, Y salta algunos pasos en su demostración y, por otra parte, debido a mi falta de entrenamiento en matemáticas no comprendo el significado de algunos símbolos usados por Y en la demostración. Meses más tarde aparece publicada la prueba de p en el Journal of Mathematics, lo cual confirma mi creencia de que la prueba es correcta pues sé que los árbitros de esa revista son sumamente meticulosos. Sobre la base de la aceptación generalizada por parte de la comunidad matemática del teorema p , estoy en la posición de saber que p es verdadero aun cuando -dice Donnellan- desconozca la verdad expresada por p , esto es, desconozca (parcialmente) el significado de p . Otro ejemplo sería el de una persona A quien escucha a otra persona B decir "Xirau se llama 'Xirau'" y A no sabe quién o qué es Xirau. A puede alegar que sabe que la oración dicha por B es verdadera, aunque no sepa la verdad que expresa, pues no sabe acerca de qué objeto es la oración, no sabe ni siquiera qué clase de individuo forma parte del hecho en el mundo que hace verdadera a p . (A no podría decir que la situación que hace verdadera la afirmación de B es el hecho de que Xirau se llame "Xirau", pues si no sabe quién o qué cosa es Xirau, no sabe cuál es ese hecho. Obsérvese que Xirau podría protestar y decir: " A no sabe que yo me llamo 'Xirau'".)

De acuerdo con Donnellan, lo único que se sigue de la tesis de Kripke de que hay designadores rígidos que pueden introducirse mediante el uso de descripciones definidas que fijan su referencia, es que podemos saber a priori que ciertas oraciones son verdaderas, pero no que conozcamos a priori la verdad expresada por ellas. Esto es, en relación con el ejemplo de Leverrier y Neptuno, es correcto sostener que Leverrier sabe a priori que "Neptuno es el Φ " es una oración verdadera, e incluso que puede saber a priori que es contingentemente

verdadera, pero no que sabe cuál es la verdad que expresa la oración, esto es, no conoce (ni a priori ni a posteriori en el ejemplo) la proposición expresada²⁹. Esto, de acuerdo con Donnellan, es tanto como afirmar que Leverrier, en el ejemplo aducido por Kripke, desconoce al menos parte del significado de la oración que profiere.³⁰

Donnellan ofrece ciertas razones intuitivas a favor de su interpretación. Veamos esto mediante otro ejemplo. Supongamos que otro astrónomo contemporáneo de Leverrier, digamos Johann Galle (quien de hecho observó por primera vez a Neptuno el 23 de septiembre de 1846) hubiese visto a través del telescopio el planeta en cuestión, lo hubiese bautizado como "Neptuno", hubiese estudiado su trayectoria y después de múltiples observaciones y cálculos hubiese concluido que Neptuno, ese planeta observado por él, es la causa de las perturbaciones en la órbita de Urano. ¿Podemos, en algún sentido, sostener que Leverrier (antes de observar a Neptuno) y Galle (habiendo observado a Neptuno) saben lo mismo cuando afirman que X es el Φ ? Parece que no. Es claro que Leverrier sabe que hay algo que es el Φ . Esto lo sabe a posteriori, esto es, sobre la base de una serie de observaciones y cálculos, los cuales lo llevan a postular como la única explicación posible de los datos observados, que existe un cuerpo celeste que sigue una determinada trayectoria y que es la causa de las desviaciones registradas en la órbita de Urano. Leverrier puede incluso predecir la posición

²⁹ Donnellan se refiere a la proposición singular expresada, es decir, al par ordenado constituido por Neptuno mismo y la propiedad de ser la causa de las perturbaciones en la órbita de Urano; es decir, el par ordenado $\langle \text{Neptuno}, \langle \text{el } \Phi \rangle \rangle$. Véase, más adelante, n.36 y n.37.

³⁰ Es decir, el significado de la oración es la proposición $\langle \text{Neptuno}, \langle \text{el } \Phi \rangle \rangle$ y Leverrier desconoce uno de sus componentes. Cfr. Donnellan, K., (1979b), p. 51.

exacta en la que el planeta postulado se encontrará en ciertas fechas precisas. Pero Leverrier no sabe cuál objeto en el mundo real es el Φ . Galle, en cambio, sabe que ese objeto, con el que él está en relación epistemológica directa mediante la observación, es el Φ .³¹

La diferencia entre el conocimiento de Leverrier y el de Galle es que parece que lo que el primero sabe es:^{32, 33}

$$(1) \text{ (Ex)(}\underline{\Phi x}\text{)}$$

en tanto que lo que sabe Galle es:

$$(2) \text{ (Ex)(}\underline{\Phi x}\text{) \& } x=\underline{a}\text{)}$$

(en donde a es un nombre con un componente deíctico: ese a). Otra manera de ver que Leverrier no sabe lo mismo que sabe Galle es la siguiente: si hubiese habitantes en Neptuno y si tuvieran un lenguaje, podrían decir con verdad que Galle sabe que el planeta en donde ellos viven es el Φ , pero no podrían decir lo

³¹ Una manera de destacar la diferencia es decir que Galle conoce a Neptuno directamente, que ha interactuado directamente con él, en tanto que Leverrier lo conoce sólo indirectamente o por descripción.

³² En N & N, p. 79, n. 33, Kripke afirma: "En esta etapa [cuando Leverrier no había observado a Neptuno] valía una equivalencia material a priori entre los enunciados 'Neptuno existe' y 'Existe algún planeta que perturba la órbita de tales otros planetas en tal y cual posición'."

³³ En la fórmula que sigue utilizo el subrayado, siguiendo a Evans, como un operador de "unicidad", es decir, para señalar que sólo un individuo satisface el predicado Φ .

mismo de Leverrier.

El conocimiento que tiene Leverrier de (1) es un conocimiento a todas luces empírico.³⁴ Lo mismo el que tiene Galle de (2). ¿Qué añade Leverrier a su conocimiento cuando decide darle el nombre "X" a aquello que satisface "el Φ "? Nada sustancial. Es evidente que no sabe un dato nuevo acerca del mundo real por el mero hecho de darle un nombre a cualquiera que sea el objeto que satisface "el Φ ", sin saber cuál es ese objeto. Sólo abre la posibilidad de decir que sabe que "X es el Φ " expresa una verdad, aunque, como señala Donnellan, en algún sentido no conozca cabalmente la verdad expresada, no sepa cuál es el hecho (entendido éste como el tener un individuo concreto una propiedad determinada) que hace verdadera su afirmación. El conocimiento que tiene Galle de (2) es también a posteriori, pero se distingue del de Leverrier en que además de saber que algo es el Φ , sabe qué objeto es el Φ , es decir, tiene una manera de identificar a Neptuno independiente de la descripción "el Φ " y de otras descripciones que eventualmente Leverrier pudiera usar para hablar del planeta que produce las perturbaciones en la órbita de Urano. Es decir, Galle puede caracterizar a Neptuno como ese objeto que observa en determinada posición en la bóveda celeste en un momento dado.

De acuerdo con la posición de Donnellan tenemos, pues, lo siguiente:

³⁴ Para que la afirmación de Leverrier de que N es el Φ tenga siquiera valor de verdad parece necesario que "el Φ " efectivamente seleccione un objeto en el mundo real. Leverrier sabe que hay algo que satisface la descripción "el Φ " y esto lo sabe a posteriori. Ese conocimiento a posteriori es la única base que tiene para afirmar que sabe que N es el Φ . Si el conocimiento de Leverrier de que N es el Φ se apoya en (o presupone) un conocimiento empírico, ¿en qué sentido podemos afirmar que sabe a priori que N es el Φ ? Para una interpretación similar, véase S. Schiffer, (1979), p. 62. Gareth Evans argumenta en contra de lo aquí sugerido; véase la sección 4 más adelante.

(i) Antes de observar Leverrier a Neptuno, Leverrier no sabe (ni a priori, ni a posteriori) que X es el Φ . Su conocimiento se reduce a saber que algo es la causa de las perturbaciones en la órbita de Urano y que la oración "X es el Φ " expresa una proposición verdadera.

(ii) Para poder siquiera conocer el contenido significativo de una oración en la que aparece un nombre propio no basta tener un conocimiento de dicto del referente del nombre en cuestión, es necesario tener un conocimiento de re.³⁵ Y lo que determina que ese conocimiento sea de re es la existencia de una conexión causal (directa o indirecta) apropiada entre el hablante que usa el nombre y el referente del nombre.³⁶ Ello hace que las oraciones en las que figura un nombre

³⁵ Donnellan, K., (1979b), p. 54. La distinción entre de dicto y de re es discutible. En general, se dice que un hablante h sabe de dicto que o es F si con base en la información que tiene h puede afirmar justificadamente la oración "o es F" y la oración es verdadera, pero, por no haber interactuado causalmente con o, el conocimiento de h no es propiamente acerca de o, y el nombre usado para referirse a o no parece estar abierto a la cuantificación existencial; según Donnellan, Leverrier sabe de dicto que Neptuno es el Φ . En cambio, decimos que h sabe de re que o es F, si h tiene conocimiento perceptual de o, y aun cuando h no sea capaz de ofrecer una descripción de o su conocimiento es acerca de o por haber interactuado causalmente con o y el nombre usado para referirse a o parece estar abierto a la cuantificación existencial.

³⁶ Gareth Evans señala, en Evans, G., (1978), p. 173, que esta condición que impone Donnellan hace realmente imposible que alguien tenga en su repertorio algún nombre vacío que pueda figurar en oraciones significativas, y que, si esto es así, ni siquiera puede plantearse el problema señalado por Kripke. Donnellan podría replicar que lo que se requiere para plantear la posibilidad de conocer a priori verdades contingentes, no es que haya o pueda haber nombres vacíos con significado -como sostiene Evans- sino sólo que sea posible que al escuchar h un nombre vacío se produzca en él un estado psicológico indistinguible para desde su punto de vista del que se produce cuando escucha y comprende un nombre que sí refiere.

N estén abiertas a la sustitución de N por cualquier otra expresión codesignativa y que se les pueda aplicar la regla de Generalización Existencial. (Lo anterior equivale a rechazar los "nombres fregeanos", es decir, a rechazar que pueda haber nombres vacíos que figuren en oraciones significativas.) Si N carece de referente, cualquier oración p en la que aparezca N carecerá, según Donnellan, de valor de verdad y parcialmente de significado. Si, por otro lado, N se introduce mediante la estipulación: "Sea N el Φ " y N resulta tener un referente del que el hablante no tiene conocimiento de re, entonces, aunque la oración "N es el Φ " exprese una verdad el hablante ignorará la verdad que expresa por ignorar parcialmente su significado.

(iii) De lo anterior parece desprenderse la consecuencia de que si un hablante h pretende introducir un nombre N mediante una descripción cuando carece de conocimiento de re del referente de N, el nombre no le servirá para nada sustantivo, pues al no comprender h las oraciones en las que figura N, no podrá usar N para expresar ninguna supuesta creencia o conocimiento singular; no se diga ya un conocimiento de un hecho singular contingente.

(iv) Si lo anterior es correcto, entonces la única manera como un hablante h puede introducir un nuevo nombre en el lenguaje de modo que realmente pueda usarlo en oraciones que expresen proposiciones singulares de las que h sí tenga conocimiento, es mediante un "bautismo" en el que h impone el nombre a ese objeto con el que se relaciona causalmente y al que conoce directamente o de re.

Las exigencias que impone Donnellan para poder decir con verdad que un hablante conoce una proposición singular excluyen, pues, la posibilidad de atribuirle tal conocimiento a h mediante oraciones subordinadas en las que ocurran nombres que h haya introducido mediante bautismos in absentia. Leverrier, en

nuestro ejemplo, no puede propiamente expresar su conocimiento respecto de Neptuno hasta no verlo. Introducir el nombre "X" mediante la descripción "el Φ " y afirmar a continuación que X es el Φ , sin haber observado a Neptuno, no lo coloca en la posición de conocer, no digamos ya un hecho contingente a priori, sino ni siquiera el contenido significativo de su propia afirmación.

La mera sugerencia de que Leverrier desconozca el significado de la oración que profiere hace reaccionar a Michael Dummett y a Gareth Evans.³⁷ ¿Cómo no va a saber Leverrier lo que significa la oración "X es el Φ "? ¿No hay acaso algo que Leverrier entiende cabalmente cuando hace su afirmación? Parece innegable que, dada la manera como Leverrier introdujo "Neptuno", entiende algo cuando afirma que X es el Φ y eso que entiende es, según Dummett y Evans, el significado de la oración que afirma. Si efectivamente resulta incorrecto decir que Leverrier conoce o sabe cuál es el hecho concreto en el mundo que hace verdadera la oración que profiere, o cuál es la proposición singular expresada por la oración que emite, entonces, en lugar de concluir -como lo hace Donnellan- que Leverrier no conoce el significado de su afirmación, tal vez resulte más adecuado sostener -como lo hacen Dummett y Evans- que lo que Leverrier entiende cuando dice que N es el Φ , es el significado que para él tiene la oración que afirma y que es algo diferente de la proposición singular expresada por la oración. ¿Qué es, pues, eso que entiende Leverrier? ¿Qué es lo que sabe y cómo lo sabe cuando afirma con verdad y justificación que X es el Φ ?

³⁷ Véase, Dummett, M. ((1981), Apéndice al cap. 5, y Evans, G., (1978), pp. 171-173.

3. Proposiciones singulares vs. "contenidos cognoscitivos".³⁸

Tanto Kripke como Donnellan parecen comprometerse con la idea de que lo que conoce quien asevera con verdad y justificación una oración singular de la forma "Fa" es justamente la proposición singular que expresa "Fa", esto es, el tuplo ordenado $\langle \underline{F} \langle \underline{a} \rangle \rangle$, en donde \underline{F} es una propiedad monádica y \underline{a} el individuo que tiene la propiedad \underline{F} .³⁹ La proposición singular que Fa es, a la vez, el contenido semántico de la oración "Fa", de modo que, si h sabe o cree que Fa, el contenido de su creencia es idéntico al contenido semántico de "Fa".

Ahora bien, "lo que conoce h " cuando asevera con verdad y justificación "Fa", puede entenderse, por lo menos, de tres maneras diferentes: (a) como el hecho concreto que, en el contexto, denota "Fa" en el mundo real y que hace verdadera la oración; (b) como la proposición expresada por "Fa",⁴⁰ esto es, una

³⁸ El uso de comillas no tiene más propósito que indicar una actitud neutral respecto de la existencia de tales entidades.

³⁹ Es decir, el individuo mismo, no una representación de él. Para la noción de proposición singular véase Schiffer, S., (1987), cap. 3. y (1992) sección I y García-Carpintero, M., (1996), cap. VII. Las proposiciones singulares suelen también llamarse "proposiciones russellianas". Bertrand Russell sostuvo que una oración singular expresa una proposición constituida por el referente del término sujeto y la propiedad denotada por el predicado; así, es el mismísimo Mont Blanc un constituyente de la proposición expresada por la oración "El Mont Blanc mide más de 4,000 metros de altura"; una consecuencia de lo anterior es que cuando un nombre no refiere las oraciones en las que aparece carecen de significado; véase la carta que Russell dirigió a Frege el 12 de diciembre de 1904 en Frege, G. (1980), p. 169, también en Salmon N. y Soames S., (1988).

⁴⁰ Al aseverar una oración hacemos una especie de partición entre aquellos mundos en los que la aseveración es verdadera y aquellos otros en los que es falsa. La proposición expresada por una oración emitida en un contexto se caracteriza

entidad abstracta, estructurada, constituida por la propiedad F y el individuo a (independientemente de la manera como se le identifique) y (c) como lo que hay en la mente de h cuando entiende y acepta la oración "Fa" y que por ello mismo h conoce inmediatamente y es transparente o diáfano para él; es decir, como lo que llamaré en este capítulo el "contenido cognoscitivo" de h cuando cree o sabe que Fa. Estas nociones no han de confundirse ni identificarse, pues, como vimos en el capítulo anterior,⁴¹ es posible que la misma proposición expresada por "Fa", es decir, el mismo contenido semántico, sea representado mediante diferentes vehículos o "contenidos cognoscitivos" en las mentes de diferentes hablantes (o del mismo hablante en diferentes momentos); también es posible -tal y como veremos que argumenta Evans- que un mismo "contenido" en la mente de h represente diversas proposiciones. Sin embargo, como señalé antes, y a pesar de las diferencias indicadas entre proposiciones y "contenidos cognoscitivos", la concepción de Kripke de la referencia como una relación puramente semántica (no epistemológica), su tesis sobre la rigidez de los nombres y su teoría causal del nombrar, parecen comprometerlo a sostener que el contenido en la mente de h cuando sabe que Fa es precisamente la proposición que Fa.

Nótese que las tres posibles maneras de responder a la pregunta acerca de qué es lo que h conoce cuando sabe que Fa, se relacionan con las tres condiciones del conocimiento. Según la ortodoxia en epistemología, si h sabe que Fa, tienen

como una función asociada a la oración que nos lleva de mundos posibles a valores de verdad. Véase Bradley, R., y Swartz, N., (1979).

⁴¹ Véase esp. secciones 2 y 3.

que darse (por lo menos) tres condiciones.⁴² La condición de verdad, que distingue al conocimiento de la mera creencia, destaca el componente "externo" del conocimiento al exigir que éste no se reduzca a un mero acto mental, sino que tenga un objeto extramental, es decir, que no se agote en el mero estado "interno" del sujeto que conoce.⁴³ Para que h sepa que Cicerón es romano, se requiere, en efecto, que se dé el hecho externo consistente en que Cicerón posea la propiedad de ser romano. Pero, el conocimiento también tiene un componente interno; en ese respecto, no se distingue de la creencia. La condición de creencia resalta justamente ese componente interno, así como la naturaleza proposicional de su objeto. El objeto del conocimiento no se distingue, en efecto, del de la creencia: es una proposición, es decir, una entidad susceptible de tener un valor veritativo, abstracta, objetiva, en el sentido de que no se agota en una mera representación privada, y tal que los diferentes sujetos epistémicos que comparten una misma creencia o conocimiento se relacionan con ella, independientemente de las diferentes modalidades como en sus mentes se la representen. Así, si h cree que Cicerón es romano y h' cree que Marco Tulio es romano, h y h' creen la misma proposición singular, el mismo par ordenado.⁴⁴ Pero, además de enfatizar el

⁴² Véase, por ejemplo, Chisholm, R., (1977), cap. 6; Sosa, E., (1992), cap. III; Villoro, L., (1982), caps. 4, 7 y 8.

⁴³ Cabe la posibilidad de que lo que conoce h sea algo sobre un estado interno propio. Saber uno mismo que piensa, es un caso tal. En ese tipo de situación, lo que haría verdadera la proposición creída no sería una realidad "extramental". Algunos ilustres filósofos han considerado ese tipo de situación epistémica constituye el fundamento de todo el conocimiento; otros, no menos ilustres, han negado que ese sea siquiera un caso de conocimiento.

⁴⁴ Decir que la proposición que creen es la misma significa que la función asociada a las oraciones que expresan sus respectivas creencias y cuyo dominio son

carácter proposicional y objetivo del conocimiento humano, la condición de creencia apunta a un tercer componente del mismo, un componente que pudiéramos llamar "subjetivo" en tanto que puede variar de un hablante a otro y que consiste en la manera distintiva como cada quien se representa la proposición creída. Frege llamó a este último componente "el pensamiento" en la mente de quien cree o sabe algo, el cual, en caso de tratarse de una creencia sobre un individuo al que nos referimos por su nombre, sería diferente tanto de la proposición singular creída (pues el "pensamiento" varía al variar las maneras como los distintos sujetos se representan los objetos y las propiedades que constituyen una misma proposición), como del hecho sabido (ya que los "pensamientos" no necesariamente son verdaderos). Si bien en el ejemplo anterior es correcto decir que h y h' creen la misma proposición, no parece correcto afirmar que la manera como h y h' se representan esa proposición sea la misma y, en ese sentido, que el "contenido en la mente" de h y h', el cual les resulta diáfano o directamente accesible a cada uno de ellos, sea el mismo. Nótese que ese componente "subjetivo" de la creencia es sumamente importante para la justificación epistémica.⁴⁵

Para los teóricos de la referencia directa, la manera distintiva como diferentes sujetos epistémicos pueden representarse mentalmente la proposición creída -es decir, el ingrediente "subjetivo" de la creencia o el "sentido" que cada

mundos posibles y su contradominio valores de verdad, es la misma. Véase Evans, G., (1978) p. 176.

⁴⁵ Si la relación de justificación entre unas creencias y otras es inferencial, el sujeto tendrá que percatarse de ella (la justificación es "internista"), por esa razón el modo como se represente el sujeto las proposiciones creídas resultará crucial para la justificación.

sujeto asocia con la oración que expresa lo que cree-, es irrelevante para determinar el contenido propiamente semántico de la oración y, por ende, el objeto de la creencia; el componente "subjetivo" de la creencia es considerado sólo como eso: como un ingrediente psicológico o, en el mejor de los casos, como un vehículo epistemológico conveniente para relacionarse con la proposición creída, pero no constitutivo de ella.⁴⁶ Nada, aparte de la proposición singular expresada por una oración de la forma "Fa", es un componente semántico de la oración. Así, al disociar las categorías semánticas y metafísicas de todo aspecto epistemológico, Kripke y los teóricos de la referencia directa se inclinan a sostener que las oraciones singulares expresan proposiciones singulares y que tales proposiciones constituyen el objeto del conocimiento de sujetos como Leverrier.⁴⁷ Si representamos a Neptuno mediante "a" y al predicado "ser la causa de las perturbaciones en la órbita de Urano" mediante "F", tenemos que la proposición expresada por "X es el Φ " es $\langle \underline{F}, \langle \underline{a} \rangle \rangle$ y que lo que Leverrier conoce, de acuerdo con una teoría de la referencia directa de los nombres, es precisamente esa proposición.

Gareth Evans y Michael Dummett rechazan la identificación entre proposición expresada y contenido cognoscitivo; según ellos, lo pertinente para determinar qué cree o sabe un hablante que afirma una oración singular de la forma "Fa" es lo que el hablante entiende cuando escucha esa oración o cuando

⁴⁶ Véase, por ejemplo, Millikan, R. G., (1991) y Salmon, N., (1986) y (1989).

⁴⁷ En el capítulo V de esta tesis examinaré con mucho mayor detalle la dificultad que representa para una teoría como la propuesta por Kripke encontrar objetos adecuados para las actitudes proposicionales. Aquí sólo me interesa señalar su compromiso con entidades extensionales como objetos de las actitudes.

la usa, lo cual es transparente para él y es diferente de la proposición expresada.⁴⁸ Evans y Dummett coinciden en señalar que lo que lleva a Kripke a concluir que es posible conocer a priori verdades contingentes es justamente la identificación que implícitamente acepta entre contenidos cognoscitivos y proposiciones.⁴⁹ Ambos filósofos argumentan que a una misma proposición pueden corresponderle distintos contenidos cognoscitivos en la mente de diferentes hablantes, y Evans, como veremos, argumenta además que a un mismo contenido cognoscitivo en la mente de diferentes hablantes pueden corresponderle proposiciones diferentes.⁵⁰

La diferencia entre los teóricos de la referencia directa y sus críticos parece

⁴⁸ Para la distinción entre proposición y contenido, véase esp. Evans, G., (1978), pp. 176 y ss. y Dummett, M. (1981) pp. 126 y ss. Ambos filósofos, a diferencia de Kripke, consideran en las obras citadas que el "contenido cognoscitivo" sí es tema de la semántica.

⁴⁹ Véase Evans, G., (1978), sección III. Michael Dummett señala esta cuestión al escribir: "un hecho comunicado por un enunciado, tal como lo entiende Kripke, no puede igualarse con su contenido cognoscitivo ... [eso precisamente] prepara la trampa en la que cae Kripke cuando habla de alguien que conoce a priori un hecho contingente". Véase, Dummett, M., (1981) p. 126; también las pp. 132 y ss. en donde discute la noción kripkeana de significado en tanto que noción no epistemológica.

⁵⁰ Para la noción de "contenido" o de "lo que dice" una oración tal y como la entiende Gareth Evans, véase Evans, G., (1978), pp. 176-177. Para el tratamiento de Dummett de estos asuntos, véase, Dummett, M., (1981) caps. 5 (más el Apéndice) y 6. Cualquier discusión sobre lo que haya de considerarse como el contenido de una creencia tiene que evitar lo que Millikan llama "internalizar" y "externalizar" igualdades y diferencias; es decir, tiene que evitar la petición de principio consistente en concluir que un sujeto s tiene creencias con contenidos diferentes (iguales) a partir del dato de que s juzga tener creencias con contenidos diferentes (iguales). Véase, cap. 2 sección 3.

radicar, en último término, en si hemos de considerar el significado de las oraciones de una lengua como algo objetivo, perteneciente a las oraciones mismas (como las proposiciones que expresan) o si ha de relativizarse el significado de una oración a los hablantes que la entienden y la usan. Pues, lo que entiende un hablante h cuando escucha una oración p , es decir, el contenido cognoscitivo que p tiene para él, puede ser diferente del que tiene para otros hablantes h' , h'' , etcétera; las diferencias en historia, cultura y conocimientos que de hecho se dan entre diferentes hablantes inciden en la manera como entiende una oración.⁵¹ En el capítulo anterior me referí a las dificultades que enfrentan quienes, como Frege, admiten la posibilidad de que una misma oración pueda tener diferentes "sentidos" para diferentes hablantes y, así, relativizan el valor semántico de las oraciones a sus diferentes usuarios. Sin embargo, no es ese el tema que aquí me interesa examinar y no debemos adelantar conclusiones antes de examinar la manera como Evans traza la distinción entre la proposición creída y lo que él llama el "contenido" de una creencia y, luego, en la siguiente sección, su propuesta de solución al problema planteado por Kripke.

De acuerdo con Evans, dos oraciones p y p' tienen el mismo contenido o "dicen lo mismo" para un hablante h , si y sólo si, lo que h cree cuando entiende y acepta como verdadera p es lo mismo que lo que h cree cuando entiende y acepta como verdadera p' , de manera que si h entiende p y p' es imposible que acepte como verdadera una de esas oraciones y rechace la otra.⁵² Evans señala,

⁵¹ Como me señaló Raúl Orayen en una ocasión, lo que Einstein y yo entendemos al escuchar " $E=mc^2$ " deben de ser cosas diferentes.

⁵² Véase Evans, G., (1978), p. 176. Nótese que el criterio que usa Evans para identificar "contenidos" es un criterio "interno". Obsérvese, también, la similitud entre lo que Evans llama el "contenido" de una oración y lo que Frege

sin embargo, que p y p' pueden expresar distintas proposiciones desde un punto de vista semántico modal, ya que, dicho en sus propias palabras, " p y p' pueden encajar de diferentes maneras dentro del alcance de un operador modal".⁵³ Así, si p es la oración "Acautli fue sacrificado" y p' la oración "El primer azteca sacrificado en honor a Huitzilopochtli fue sacrificado", para alguien (como yo) que sólo sabe de Acautli que fue el primer azteca sacrificado en honor a Huitzilopochtli, las oraciones p y p' tienen, dice Evans, el mismo contenido o son epistémicamente equivalentes: ambas oraciones son verificadas (hechas verdaderas) exactamente por el mismo estado de cosas "y lo que uno cree al entender una oración y aceptarla como verdadera es, precisamente, que un determinado estado de cosas es el caso" ⁵⁴. Sin embargo, señala, si los nombres son designadores rígidos (y, por lo tanto, "Acautli" designa al mismo individuo en todo mundo posible, es decir, a ese azteca al que de hecho designa en el mundo real) y las descripciones son no rígidas (es decir, "el primer azteca sacrificado en honor a Huitzilopochtli" designa a distintos individuos en diferentes mundos posibles o situaciones contrafácticas), p y p' expresan proposiciones con diferentes características modales: la primera es efectivamente contingente en tanto que la segunda es necesaria. Hay situaciones contrafácticas o mundos posibles en los que Acautli no es sacrificado; en tanto que no hay ningún mundo posible m_n en el que alguien satisfaga en m_n la descripción que figura en p' y no sea sacrificado. El hecho de tener diferentes condiciones de verdad en diferentes mundos posibles es

llamó su "sentido".

⁵³ Cfr. Evans, G., (1978), p. 177.

⁵⁴ Evans, G., (1978), p. 180.

el responsable de que p y p' tengan diferentes propiedades modales.⁵⁵

Que las proposiciones expresadas por p y p' son modalmente diferentes (o diferentes *tout court*) puede mostrarse de la siguiente forma: la proposición expresada por p es el par ordenado $\langle \underline{S} \langle \underline{a} \rangle \rangle$, en donde " \underline{S} " representa la propiedad de ser sacrificado y " \underline{a} " a Acautli, de manera que la única condición que la hace verdadera en cualquier mundo posible es que Acautli, ese individuo designado por "Acautli" en \underline{m}^* , tenga la propiedad \underline{S} ; ningún otro individuo que tuviese la propiedad \underline{S} en algún otro mundo posible \underline{m}_n haría verdadera la proposición en cuestión. En cambio, para que p' sea verdadera en alguna otra situación contrafáctica o mundo posible \underline{m}_n , la condición que tiene que darse es que el objeto que satisface en \underline{m}_n la descripción "el primer azteca sacrificado en honor a Huitzilopochtli" tenga la propiedad \underline{S} .⁵⁶ Es evidente que hay algún mundo posible en el que Acautli no satisface la descripción definida mediante la que se le introdujo (por ejemplo, el mundo posible en el que Acautli logra escapar antes de ser sacrificado), sino que la satisface otro individuo; lo que haría verdadera a p' en ese otro mundo posible sería el hecho de que ese otro individuo, y no Acautli, fuese sacrificado. Vemos, pues, que p y p' expresan proposiciones diferentes.

Ahora bien, se puede aceptar sin mayores cuestionamientos que cuando asevero "Acautli fue sacrificado" la oración que profiero exprese efectivamente

⁵⁵ En una semántica de mundos posibles dos proposiciones son diferentes si difieren sus valores de verdad en diferentes mundos posibles; según esto, p y p' no sólo tienen características modales diferentes, sino que son proposiciones diferentes.

⁵⁶ En eso radica la no rigidez de las descripciones por contraste con la rigidez de los nombres. Véase Kripke, S., (1980) Prefacio.

una proposición contingente, esto es, algo verdadero del mundo real, pero no verdadero de todo mundo posible o situación contrafáctica; según la concepción semántica de Kripke, el carácter contingente de lo que afirmo depende del hecho de que "Acautli" es un designador rígido. Sin embargo -señala Evans-, se pueden tener dudas acerca de que esa proposición constituya justamente el contenido de mi creencia. Si Acautli forma parte esencial de la proposición expresada por "Acautli fue sacrificado", y no sé cuál es el individuo designado por "Acautli", difícilmente pensaré que ese individuo pueda formar parte del contenido de mi creencia. Este y otro tipo de consideraciones conducen a Evans a distinguir entre la proposición que en una semántica formal se le asigna a una oración al momento de interpretarla y el contenido cognoscitivo en la mente de quien conoce lo que la oración expresa. La solución que ofrece Evans al problema planteado por Kripke se apoya precisamente en esa distinción.

4. La solución de Evans.

En su artículo "Reference and Contingency" Evans explora la posibilidad de contar en el lenguaje con "nombres fregeanos", esto es, nombres que sí tienen un sentido, de manera que pueden figurar en oraciones significativas aun cuando carezcan de referente.⁵⁷ Dichos nombres descriptivos o fregeanos, sostiene Evans, son los que se introducen mediante una estipulación del tipo siguiente:

⁵⁷ Evans parece haber modificado su opinión sobre los nombres vacíos más tarde, ya que en Evans, G., (1982), capítulo 9, sostiene que alguien que hace uso de un término singular carente de referente con el ánimo de comunicar o informar algo realmente no dice nada. Los nombres descriptivos, dice Evans, requerirían ser tratados mediante una "Lógica libre" en la que se restringen las reglas de Generalización Existencial y de Especificación Universal.

(E) Llamaré "N" al individuo que sea el Φ

en donde "N" es un nombre y "el Φ " una descripción definida, la cual, para simplificar, ha de ser tomada como el único criterio para la aplicación del nombre descriptivo. Notemos que, por ser un nombre propio, "N" es un designador rígido; esto es, si refiere en m*, referirá al mismo individuo al que refiere en m* en todo mundo posible. En el ejemplo de Leverrier resulta evidente que la manera como él introduce el nombre "Neptuno" se ajusta a lo anterior y, por lo tanto, podemos considerarlo como un nombre descriptivo en el sentido especificado por Evans.⁵⁸

Introducir nombres descriptivos no es lo mismo que introducir abreviaturas de descripciones definidas. Evans admite la rigidez de los nombres así como la no rigidez de las descripciones definidas, de modo que si un nombre descriptivo es efectivamente un nombre no podrá ser semánticamente equivalente a ninguna descripción. Sin embargo, Evans sostiene que el contenido en la mente de h cuando usa una oración cuyo sujeto es un nombre descriptivo es idéntico al contenido en la mente de h cuando sustituye en la oración el nombre descriptivo por la descripción mediante la cual introdujo el nombre. Si Leverrier usó la descripción definida "la causa de las perturbaciones en la órbita de Urano" para introducir el nombre "Neptuno", al afirmar "Neptuno es F" el contenido de su pensamiento es idéntico al contenido de su pensamiento cuando afirma "la causa de las perturbaciones en la órbita de Urano es F" aun cuando las proposiciones

⁵⁸ Debo aclarar que Evans no discute en el artículo citado especialmente el caso en el que Leverrier nombra a Neptuno, sino otros casos; en especial, usa el ejemplo del nombre descriptivo "Julius", cuya referencia se fija al estipular que "Julius" referirá a quienquiera que satisfaga la descripción "El inventor del zipper". En lo que sigue, aplico el análisis y las conclusiones de Evans al caso de Leverrier.

asociadas a las dos oraciones mencionadas sean diferentes en el sentido de ser diferenciables en una semántica de mundos posibles.

Ahora bien, si Φ es la propiedad de ser causa de las perturbaciones en la órbita de Urano, de lo anterior parecería seguirse que el contenido en la mente de Leverrier cuando afirma que Neptuno es el Φ , es analítico o tautológico; sin embargo, sabemos que esto no es así, pues, como indicamos en la sección 1 de este capítulo, Leverrier podría haber dicho sin contradecirse que Neptuno no es el Φ . A fin de evitar esa consecuencia indeseada Evans propone interpretar la descripción que usa Leverrier, no como "la causa de las perturbaciones en la órbita de Urano", sino como "la causa de las perturbaciones en la órbita de Urano en el mundo real". De esta manera, podemos decir que el contenido en la mente de Leverrier cuando escucha o profiere "Neptuno es el Φ " es que la causa de las perturbaciones en la órbita de Urano en \underline{m}^* es la causa de las perturbaciones en la órbita de Urano. Nótese que este no es un contenido que resulta verdadero en todo mundo posible, es decir, su verdad no es necesaria. Se trata de un contenido tal que si es verdadero es sólo contingentemente verdadero, ya que es claramente posible una situación contrafáctica o un mundo posible \underline{m}_n en el que la causa de las perturbaciones en la órbita de Urano en \underline{m}^* no sea la causa de las perturbaciones en la órbita de Urano en \underline{m}_n . Lo que sucede, piensa Evans, es que al afirmar Leverrier "Neptuno es el Φ " con respecto a \underline{m}^* , lo que dice, el contenido de su afirmación, se convierte en lo siguiente: la causa de las perturbaciones en la órbita de Urano en \underline{m}^* es la causa de las perturbaciones en la órbita de Urano en \underline{m}^* . Y, como señala Evans, parece poco sorprendente que

Leverrier sepa a priori que su afirmación es verdadera.⁵⁹

Nótese, sin embargo, que si se asume -como Evans lo hace en su texto- el análisis russelliano de las descripciones definidas,⁶⁰ no se puede sostener que alguien que, como Leverrier, afirma una oración en la que figura un nombre descriptivo, como "Neptuno", (cuyo "contenido" es equivalente al de la descripción definida mediante la que se le introdujo), pueda conocer a priori la verdad del contenido de su afirmación, ya que esto equivaldría a sostener que puede conocer a priori la verdad que expresa una oración existencialmente cuantificada. Si queremos identificar más exactamente lo que Leverrier sabe a priori -escribe Evans- tenemos que tener en cuenta, primero, que "Neptuno" es un nombre descriptivo en el sentido especificado y, segundo, que tales nombres requieren de una "Lógica libre" y, por lo tanto, no están abiertos a la cuantificación existencial. Así, al discutir la solución que propone Donnellan, Evans insiste en que "Neptuno" tiene que ser un nombre descriptivo (y tener "alcance corto") para que siquiera pueda surgir la paradoja,⁶¹ pues, de otra manera -dice- la oración "Neptuno es el Φ " podría ser tratada conforme a una

⁵⁹ La tesis de Evans, como veremos más adelante, es más general: "podemos saber a priori que el mundo real posee ciertas propiedades que no poseen todos los mundos en la medida en que la especificación de esas propiedades exija hacer referencia al mundo real o a lo que realmente es el caso". Véase, Evans, G., (1978), p. 183. En efecto, tal parece que si estipulo que F es la propiedad de ser azul en m^* , sé a priori que "Todo lo F es azul" es verdadera dicha respecto de m^* ; sin embargo, la misma oración podría ser falsa respecto de otros mundos posibles; basta un mundo en donde exista Peter O'Toole pero en el que sus ojos no sean azules. Véase, más adelante, p. 40.

⁶⁰ Véase, Evans, G., (1978), p. 170.

⁶¹ Véase, Evans, G., (1978), p. 172.

lógica clásica y estaría abierta a la cuantificación existencial.⁶²

¿Cómo podemos formular, entonces, en una lógica clásica el contenido cognoscitivo que, según Evans, tendría la oración afirmada por Leverrier? Evans piensa que lo que conoce Leverrier podría expresarse en una lógica clásica mediante una oración de forma condicional, a saber:

(S) Si existe el Φ , Neptuno es el Φ ,⁶³

en la cual -señala Evans- "Neptuno" tiene "alcance corto", de manera que la oración (S) podría formalizarse así:⁶⁴

$$(3) \text{ (Ex) } \underline{\Phi(x)} \rightarrow [a] \Phi(a)$$

Esta oración expresa una proposición contingente, pues es posible una situación en la que el antecedente sea verdadero y el consecuente falso. El carácter contingente de (3) depende esencialmente del hecho de ser "Neptuno" un nombre. Sin embargo, dada la manera como Leverrier fija la referencia de "Neptuno", el contenido de su pensamiento cuando afirma que Neptuno es el Φ parece ser

⁶² Es decir, a partir de "N es el Φ " podríamos inferir la verdad de la oración existencialmente cuantificada que resulta de aplicar la regla de Generalización Existencial a "N es el Φ "; y obviamente, dice Evans, nadie sostiene que se pueda conocer a priori la verdad de una oración existencialmente cuantificada.

⁶³ Véase, Evans, G., (1978) esp. p. 184.

⁶⁴ En la fórmula a continuación sigo a Evans al usar tanto el subrayado como un operador de "unicidad" (es decir, para indicar que hay sólo un individuo que satisface el predicado Φ), como "[a]" para indicar que el nombre "a" que en la fórmula representa a "Neptuno" tiene alcance corto.

equivalente a lo siguiente.⁶⁵

$$(4) \quad \underline{\exists x} \Phi(x) \rightarrow (Ix)(\Phi(x, \underline{m}^*); \Phi(x)).$$

(4) es también contingente, pues, de nuevo, es posible que su antecedente sea verdadero y su consecuente falso; es decir, es posible un mundo \underline{m}_n en el que exista el Φ , pero en el que el Φ en \underline{m}_n no sea el Φ en \underline{m}^* . La propiedad que tiene que tener un mundo posible para que (4) sea verdadera respecto de él es la siguiente:⁶⁶

$$\lambda \underline{m} [(\underline{\exists x})(\underline{\Phi}(x, \underline{m})) \rightarrow (Ix)(\Phi(x, \underline{m}^*); \Phi(x, \underline{m}))]$$

y esta es una propiedad, como señala Evans, que no todo mundo posible posee.⁶⁷

⁶⁵ Sigo la notación utilizada en Evans, G., (1978) p. 170, en donde, en el contexto de investigar la forma lógica de las descripciones y su comportamiento semántico, Evans introduce "(Ix)" como un cuantificador binario, "el", el cual a partir de dos oraciones abiertas " $\Phi(x)$ " y " $\psi(x)$ " da lugar a una oración cerrada, y en donde "(Ix)($\Phi(x); \psi(x)$)" se lee como "el Φ es ψ ". Evans asume que la oración con "(Ix)" es demostrablemente equivalente a la expansión russelliana correspondiente: $(\exists x)(\Phi(x) \ \& \ \psi(x))$.

⁶⁶ El operador λ que aparece en la fórmula a continuación es el que introduce Alonzo Church, en Church, A., (1956), §§ 03 y 06 y que Rudolph Carnap explica con gran claridad en Carnap, R., (1958), sección 33. Al prefijar con el operador $\lambda \underline{x}$ cualquier fórmula oracional $\underline{P}(\underline{x})$ obtenemos una expresión predicativa que denota la propiedad que $\underline{P}(\underline{x})$ atribuye a \underline{x} . λ

⁶⁷ Evans, G., (1978), p. 184. Nótese que la propiedad en cuestión se especifica por referencia al mundo real, \underline{m}^* . Según Evans, "X es el Φ ", dicha con respecto al mundo real, y "X es el Φ en \underline{m}^* ", tendrían un mismo contenido cognoscitivo,

Pero si aplicamos (4) al mundo real, como sucede con la afirmación que hace Leverrier, el contenido de (4) se convierte en:

$$(5) \quad \underline{(Ex)(\Phi(x, \underline{m}^*))} \rightarrow (Ix)(\Phi(x, \underline{m}^*); \Phi(x, \underline{m}^*))$$

y esto es de acuerdo con Evans lo que Leverrier conoce a priori cuando afirma que Neptuno es la causa de las perturbaciones en la órbita de Urano.⁶⁸ Obsérvese que (5) resultará verdadera tanto en el caso en que algo sea el Φ en \underline{m}^* como en el caso en que nada sea el Φ en \underline{m}^* , esto es, resultará verdadera en cualquier caso. Por eso no debe sorprendernos que Leverrier conozca a priori el contenido de (5).

Ahora bien, Evans explica el carácter a priori de (5), no como el resultado de tener Leverrier facultades epistémicas especiales, ni como el resultado de ser "Neptuno" un designador rígido, sino por el hecho de que en la especificación de la propiedad que tiene que tener un mundo posible \underline{m} para que en él sea verdadera (4), se hace referencia al mundo real. Es esa referencia implícita al mundo real lo que constituye la clave de la explicación de Evans de cómo es posible conocer a priori proposiciones contingentes. Siempre que se introduzca una propiedad en cuya especificación se haga referencia implícita o explícita al mundo real, a lo que realmente es el caso -sostiene Evans- podremos contar con oraciones que si bien expresan proposiciones contingentes (es decir, que pueden ser falsas respecto de algún otro mundo posible) son tales que su verdad respecto de este mundo puede conocerse a priori. Así, por ejemplo, si especificamos que \underline{G} es la propiedad de

es decir, serían epistémicamente equivalentes, a pesar de encajar de diferente manera bajo operadores modales.

⁶⁸ Véase, Evans, G., (1978) pp. 184-185.

ser de hecho rojo, o rojo en m^* , sabremos a priori que es verdadera la oración "Todo lo que es G es rojo" cuando es afirmada con respecto a m^* ; no obstante, la oración expresa una proposición contingente, ya que es posible una situación contrafáctica m_n en la que algo que es rojo en m^* no sea rojo en m_n . Sin embargo, argumenta Evans, las oraciones en las que figuran predicados como G son sólo superficialmente contingentes; pues son contingentes sólo en el sentido de que su falsedad es posible en algún mundo posible diferente de m^* ; no son profundamente contingentes porque no nos informan nada acerca del mundo real, su verdad no depende de la existencia de ningún rasgo contingente de m^* . Una oración superficialmente contingente conocida a priori es tal que no puede ser falsa con respecto al mundo real, pero no puede ser falsa, no porque diga algo extraordinario, sino por tener un cierto rasgo lingüístico, a saber, por contener un predicado que denota, no una propiedad a secas, sino una propiedad en el mundo real. Que la verdad de "Todo lo que es G es rojo" pueda conocerse a priori se explica por el hecho de que la propiedad G que se exige de un mundo para hacer verdadera la oración se especifica de tal manera que al afirmar la oración con respecto al mundo real queda garantizada su verdad. Podríamos decir que las oraciones superficialmente contingentes cuya verdad se conoce a priori son verdaderas en virtud del tipo de predicado que figura en ellas y no porque exista algún rasgo del mundo real que las haga verdaderas. La oración 4, cuya verdad conoce Leverrier a priori es de este tipo: sólo es superficialmente contingente, es decir, es contingente en el sentido de que, aun cuando carezca de todo valor informativo, es posible su falsedad en algún otro mundo posible diferente de m^* , pero no es profundamente contingente en el sentido de que no le informa a Leverrier nada que no supiese de antemano, nada nuevo sobre lo que de hecho hay. Que haya oraciones superficialmente contingentes y conocidas a priori, nos

dice Evans, no es algo que deba inquietarnos.⁶⁹

La brillante argumentación de Evans nos deja, sin embargo, con otra inquietud. ¿Es adecuado interpretar la oración que afirma Leverrier como una especie de condicional encubierto? De ser así, bastaría, como señalé antes, con que el antecedente del condicional fuese falso, esto es, con que no hubiera ninguna causa de las perturbaciones en la órbita de Urano, para que pudiéramos decir con verdad que Leverrier sabe que Neptuno es el Φ .⁷⁰ Esto resulta dudoso al menos por tres razones: (i) si se comprobara que Neptuno no existe (como sucedió con el éter y con Vulcano), resultaría muy extraño insistir en que, aún en ese caso, Leverrier sabría que Neptuno es el Φ ; (ii) el hecho de que Leverrier sepa que hay una causa de las perturbaciones en la órbita de Urano, es decir, el hecho de que sepa que el antecedente de (5) es verdadero, justifica nuestra impresión de que Leverrier hace una afirmación categórica, y no hipotética cuando dice que Neptuno es el Φ ; y (iii) Leverrier no reconocería el condicional (5) como el contenido de

⁶⁹ La razón por la que no debe inquietarnos la existencia de tales verdades - parece pensar Evans- es que son de naturaleza más lingüística que fáctica. La solución que propone Evans presenta una dificultad que sólo mencionaré aquí. Según Evans, "Neptuno es el Φ " tiene para Leverrier un contenido cognoscitivo especificable mediante la oración (5). Pero, (i) la proposición asociada con 5 no es la misma que la asociada con la oración que afirma Leverrier; y (ii) 5 es analítica, es verdadera en virtud de su significado, en tanto que la oración que afirma Leverrier no lo es, pues, como señalé antes, él pudo haber afirmado sin contradecirse "Neptuno podría no haber sido el Φ ". Esto puede arrojar dudas respecto de que (5) represente adecuadamente el contenido de la afirmación de Leverrier.

⁷⁰ Podríamos en ese caso decir que Leverrier sabe que Neptuno es el Φ , porque (i) cree lo que afirma, (ii) lo que afirma es verdadero (vacuamente) y (iii) Leverrier tiene buenas razones para afirmar lo que afirma.

su afirmación.⁷¹ Leverrier se apoya en el conocimiento empírico que tiene de que existe el Φ para introducir el nombre "Neptuno" y afirmar, después, que Neptuno es el Φ . "Neptuno" nombra, pues, algo que Leverrier sabe que existe (aunque no lo haya percibido), de modo que resulta extraño interpretar su afirmación como si no incluyera la afirmación de que existe el Φ . Por otra parte, no podemos negar que la tesis de Evans de que el contenido cognoscitivo de una oración en la que aparece como sujeto un nombre descriptivo es idéntico al contenido cognoscitivo de otra oración que resulta de sustituir dicho nombre descriptivo por la descripción mediante la que se introdujo el nombre resulta sumamente convincente. ¿Qué es, pues, lo que Leverrier sabe cuando sabe que N es el Φ ?

De acuerdo con lo dicho más arriba, parece que lo que sabe Leverrier es algo como lo siguiente:

$$(6) \text{ (Ex)}\Phi(x)$$

y

$$(7) \text{ (Ix)(}\Phi(xm^*); \Phi(x))$$

esto es, resulta más adecuado describir lo que afirma y lo que sabe como una conjunción de (6) y (7) que como un condicional. (6), como he señalado, es algo que Leverrier sabe empíricamente y su conocimiento de (7) supone a todas luces su conocimiento de (6). La oración (7) es contingente y no expresa una verdad que pueda conocerse a priori, pues es, como (6) un enunciado existencial (recordemos

⁷¹ Parece ser crucial para la argumentación de Evans que Leverrier reconozca que la forma de su afirmación es condicional, ya que se supone que 5 recoge el contenido de su afirmación, lo que él entiende al momento de hacerla, y eso que él entiende se supone que es transparente o diáfano para él.

que es equivalente a su correspondiente expansión russelliana que tiene la forma de una oración existencialmente cuantificada). (6) es una oración contingente (profundamente contingente), en el sentido de que requiere para su verdad que el mundo real sea de una determinada manera y no de otra, que se dé un cierto estado de cosas, el cual no podemos saber a priori si se da o no se da. (7) es una oración que Evans calificaría de "superficialmente contingente" por contener el predicado "ser el Φ en \underline{m}^* ", pero fijémonos que sólo podemos saber que la oración completa es verdadera si sabemos (a posteriori) que algo satisface la descripción "el Φ " en \underline{m}^* . Por lo anterior, no parece haber razón para sostener que el nombre "Neptuno" en la oración que afirma Leverrier deba ser tratado con una "lógica libre" y que el nombre no esté abierto a la cuantificación existencial. Podemos aceptar que "Neptuno" sea para Leverrier un "nombre fregeano" sólo en el sentido de que, dada la manera como lo introdujo, tenga en su mente asociado con el nombre un cierto contenido descriptivo, no en el sentido de que al afirmar "Neptuno es el Φ " Leverrier no se comprometa, desde un punto de vista semántico, con la existencia de Neptuno, es decir, de la causa de las perturbaciones en la órbita de Urano en \underline{m}^* . Si nada satisface la descripción mediante la que Leverrier introduce el nombre "Neptuno", al afirmar "Neptuno es el Φ " Leverrier entenderá, sin duda, algo, es decir, su afirmación tendrá un cierto "contenido" para él, pero su afirmación no podrá ser tomada como verdadera.⁷²

Detengámonos y hagamos un balance luego de revisar la polémica entre Donnellan y Evans: (a) Donnellan parece equivocarse al sostener que Leverrier desconoce parcialmente "el significado" de la oración "Neptuno es el Φ " (si con ello quiere decir que Leverrier no asocia con su afirmación ningún "contenido

⁷² Tal vez tampoco debemos decir que es falsa, sino "fallida".

cognoscitivo" determinado, que al proferirla no entiende algo preciso); la distinción de Evans entre la proposición expresada por una oración y el contenido que la oración tiene para el hablante que usa la oración, nos permite decir que al proferir Leverrier la oración "Neptuno es el Φ ", Leverrier entendía algo preciso.

(b) También parece equivocarse Donnellan al sostener que Leverrier sabe a priori que "Neptuno es el Φ " expresa una verdad, pues para que Leverrier sepa eso es necesario que sepa que el nombre "Neptuno" refiere, es decir, que existe una y sólo una causa de las perturbaciones en la órbita de Urano y eso sólo lo puede saber empíricamente. Si esto es correcto, la única manera como Leverrier puede justificar que sabe que "Neptuno es el Φ " expresa una verdad, es apelando a su conocimiento empírico de que existe el Φ en m^* y trayendo a colación su propia estipulación. Si no sabe que existe el Φ en m^* , o si no recuerda su propia estipulación, no puede saber, menos aun a priori, que la oración "Neptuno es el Φ " expresa una verdad.

(c) Evans parece tener razón al señalar que el contenido cognoscitivo que Leverrier asocia con la oración que afirma es epistémicamente equivalente al contenido de la oración "El Φ en m^* es el Φ ". Si le preguntásemos a Leverrier que entiende por "Neptuno" seguramente diría: "el Φ " o, si supiera algo de lógica modal, "el Φ en m^* ". Sin embargo, hemos señalado en el capítulo anterior las dificultades que implica sostener que sean los aquí llamados "contenidos cognoscitivos" que los hablantes de una lengua asocian con sus oraciones, y que pueden variar de un hablante a otro, los contenidos semánticos de las oraciones mismas. Una teoría semántica tiene buenas razones para querer asignarle a "Neptuno es el Φ " uno y el mismo significado cuando es usada por Leverrier y por Galle, aun cuando ellos entiendan cosas diferentes cuando usan esa misma oración, es decir, aun cuando la oración tenga para cada uno de ellos diferente "contenido cognoscitivo".

(d) Por otra parte, he dado algunas razones

para no atribuirle a la afirmación que hace Leverrier la forma lógica de un condicional cuando él sabe que es verdadero el antecedente de ese condicional.⁷³

(e) El análisis adecuado de la descripción "el Φ en \underline{m}^* " parece ser el russelliano. Esto es, para que la oración "el Φ en \underline{m}^* es el Φ " sea verdadera es necesario que exista algo en \underline{m}^* que sea el Φ . Por lo tanto no parece haber lugar para atribuirle a Leverrier ningún conocimiento a priori cuando sabe que Neptuno es el Φ . (f) Donnellan parece acertar cuando señala que el nombre "Neptuno" en la frase de Leverrier debe estar abierto a la cuantificación existencial, pero tal vez no por las razones que él aduce, sino más bien porque Leverrier asocia con "Neptuno" el sentido de la descripción definida mediante la cual introduce el nombre y, en ese sentido, se compromete con la existencia de algo que sea la causa de las perturbaciones en la órbita de Urano, es decir, con la existencia de Neptuno.

¿Por qué Evans insiste en la falta de compromiso existencial cuando se conoce una oración contingente en la que aparece un nombre descriptivo? Creo que la respuesta es doble: (i) En primer lugar, Evans quiere tratar de la misma manera aquellos casos, como el de Leverrier, en los que sabemos que hay algo que satisface la descripción mediante la que se introduce el nombre que otros casos en los que los nombres descriptivos pudieran ser introducidos mediante descripciones de las que ignoramos si hay individuos que las satisfacen. Evans piensa que no es relevante para determinar lo que sabe un hablante que afirma con justificación una oración en la que aparece un nombre descriptivo el hecho de saber si la

⁷³ Si yo asevero "El asesino de Colosio es un traidor" me parecería sumamente extraño que alguien interpretara mi afirmación como si dijera "Si existe uno y sólo un individuo que haya asesinado a Colosio, ese individuo es un traidor"; dado que sé que alguien asesinó a Colosio, lo que afirmo tiene, más bien, la siguiente forma: "Existe uno y sólo un individuo que asesinó a Colosio y él es un traidor".

descripción mediante la que introdujo el nombre tiene o no un referente semántico. Así, por ejemplo, si estipulo: "'Pepe' será el nombre del primer oso polar en m* que pise suelo africano", y a continuación digo que sé que Pepe es el primer oso polar en m* que pisa suelo africano, de acuerdo con Evans el conocimiento que tengo es a priori y del mismo tipo que el conocimiento que tiene Leverrier cuando sabe que Neptuno es la causa de las perturbaciones en la órbita de Urano, aun cuando ningún oso polar pise nunca suelo congelado en m*. La ventaja que pueda haber en tratar semánticamente de la misma manera a "Neptuno" y "Pepe" parece desvanecerse cuando nos percatamos de estas consecuencias epistemológicas. (ii) En segundo lugar, Evans asume correctamente que cuando usamos oraciones en las que aparecen nombres descriptivos, nuestra intención implícita es la de denotar al individuo que en el mundo real satisface la descripción mediante la cual introdujimos el nombre y, por lo mismo, que el "contenido cognoscitivo" de un nombre descriptivo es el de la descripción introductora subindexada con "m*";⁷⁴ pero, por otro lado, como nada nos garantiza que toda descripción mediante la que podemos introducir un nombre descriptivo tenga un referente semántico, y dado que Evans quiere tratar de la misma manera todos los posibles nombres descriptivos (sea que refieran o no), él se ve obligado a atribuir a nuestras pretendidas aseveraciones categóricas (tales como "Pepe es F") en las que figuran nombres descriptivos la forma de modestos condicionales (Si existe el primer oso polar que pise suelo africano en m*, el primer oso polar que pise suelo africano

⁷⁴ Nótese, sin embargo, que aceptar que el "contenido cognoscitivo" en la mente de un hablante h que usa un nombre descriptivo N pueda ser el de la descripción introductora "el Φ ", no equivale a aceptar que el contenido semántico de N sea el de la descripción "el Φ ", ya que h podría afirmar sin contradecirse: "N podría no haber sido el Φ ".

en \underline{m}^* es F). Sin embargo, he argumentado a favor de analizar de diferente manera la situación epistemológica de quien no sabe si se satisface o no la descripción mediante la que pretende introducir un nombre de la situación de quien sí sabe que existe el referente de la descripción "el Φ en \underline{m}^* " mediante la que introduce un nombre descriptivo \underline{N} (lo que, por otra parte, sucede en la mayoría de los casos). Si esto es así, resulta inadecuado tratar ambos casos de nombres descriptivos como si fueran un solo tipo. Como también resulta inadecuado insistir en la ausencia de compromiso existencial cuando un hablante afirma " \underline{N} es el Φ ".

Como señalé antes, hay varias maneras de entender la frase "lo que conoce \underline{S} cuando sabe que $\Phi_{\underline{x}}$ ". Donnellan y Kripke consideran que lo que conoce \underline{S} cuando sabe que $\Phi_{\underline{x}}$, es la proposición singular que una semántica formal asigna a la oración " $\Phi_{\underline{x}}$ ". Evans, en cambio, considera que lo relevante para determinar lo que conoce \underline{S} cuando sabe que $\Phi_{\underline{x}}$ es el "contenido cognoscitivo" en la mente de \underline{S} cuando entiende y acepta la oración en cuestión. El problema que se les presenta a los primeros es que, al identificar "lo que conoce" \underline{S} con la proposición asignada a la oración mediante la que expresa su creencia, quizás le conceden al sujeto "más" conocimiento del que en realidad posee; por ejemplo, a un sujeto sabe que Cicerón es calvo, Donnellan y Kripke parecen concederle el conocimiento de que Marco Tulio es calvo. El problema que se les presenta, en cambio, a quienes, como Evans, distinguen entre el "contenido cognoscitivo" y la proposición expresada, es que en ocasiones le conceden al sujeto "menos" conocimiento del que en realidad tiene, como pudimos ver en el caso de Leverrier.

Para concluir: si nuestro interés es determinar lo que conoce un sujeto \underline{S} que afirma con justificación una oración singular en la que figura un nombre descriptivo (en el sentido especificado por Evans), parece crucial distinguir entre dos tipos de situaciones epistemológicas: (i) aquella en la que \underline{S} sabe que hay algo

que satisface la descripción accidental "el Φ " mediante la que introdujo el nombre, y (ii) aquella en la que S no sabe tal cosa. El primer caso (que sería el de Leverrier), si decimos que S sabe que N es el Φ , parece inevitable, como he argumentado, atribuir a S un conocimiento empírico, no a priori, pues su conocimiento de que N es el Φ depende esencialmente de su conocimiento empírico de que algo satisface la descripción "el Φ "; en el segundo caso, resulta contrario a nuestras intuiciones atribuirle al sujeto cualquier tipo de conocimiento.

Si mi análisis en este capítulo es correcto, Kripke parece estar justificado en sostener que al afirmar Leverrier que Neptuno es el Φ , emite una oración verdadera que expresa una proposición contingente, pero no en sostener que el conocimiento que expresa sea de alguna manera a priori para Leverrier.

CAPÍTULO IV

La objeción de Kripke a la teoría de la identidad mente-cuerpo¹

Introducción.

Como señalé en el capítulo anterior, una teoría se evalúa por sus consecuencias, y éstas en ocasiones pueden presentarse en campos muy alejados de aquel en el que originalmente la teoría fue propuesta. Según Saul Kripke, la tesis de la referencia directa y rígida de los nombres que él propone tiene entre sus consecuencias una muy importante en el terreno de la filosofía de la mente: la de la falsedad del fisicalismo, esto es, de la teoría que sostiene, en general, que todo lo que hay (objetos, hechos o sucesos) es físico y, en particular, que los estados mentales son estados cerebrales o corporales y, en ese sentido, que son estados físicos de las personas.² Obsérvese que si el fisicalismo es falso, es decir, si los

¹ Una versión anterior del presente capítulo fue publicada con el mismo título en *Dianoia*, Anuario del Instituto de Investigaciones Filosóficas, 1980.

² Thomas Nagel caracteriza el fisicalismo así: "la tesis de que una persona, con todas sus características psicológicas, no es nada además o por encima de su cuerpo, con todas sus características físicas", Nagel, T., (1965). Según David Lewis, la teoría de la identidad psicofísica "es la hipótesis de que cada experiencia es idéntica a algún estado físico. Específicamente, a un estado neuroquímico.", véase, Lewis, D. K., (1966), p. 5 en la versión castellana. Kripke describe de la siguiente manera la teoría de la identidad mente-cuerpo: "De acuerdo con esta teoría, un estado mental, tal como que una persona tenga un dolor, es idéntico a [...] determinado estado neural de su cerebro o de su cuerpo.",

estados mentales no son estados físicos o corporales, entonces, -aparte de convertirse en un misterio el hecho de que tengan conexiones causales con ciertos estados físicos (corporales)-, tendrá que aceptarse que tienen una realidad diferente de la del resto del mundo natural y, por lo tanto, que por principio su explicación escapa al conjunto de teorías científicas que explican ese mundo natural. Y si esto fuera el caso, resultaría inútil buscar una explicación científica natural de ellos. En esta época en la que los desarrollos recientes de la física han acrecentado nuestra confianza en su capacidad de suministrar explicaciones adecuadas de todo fenómeno natural,³ y en la que, por otro lado, las ciencias cognitivas han emprendido (no sin un éxito considerable) la tarea de explicar -echando mano de las teorías más recientes en las neurociencias, la cibernética y la inteligencia artificial- diversos fenómenos mentales como otros de tantos fenómenos naturales, la posición de Kripke resulta, por decir lo menos, poco atractiva.

El objeto del presente capítulo es examinar y evaluar la objeción que presenta Saul Kripke⁴ a la teoría de la identidad entre lo mental y lo físico. Decir que un estado mental es idéntico a un estado físico implica materialmente afirmar que los términos mentales y físicos usados para formular la identidad en cuestión, denotan en realidad una y la misma cosa, estado o suceso. Así, afirmar, por ejemplo, que el dolor es idéntico a la estimulación de la fibra C, supone aceptar

Kripke, S., (1971), p. 17 en la versión castellana.

³ La segunda premisa en el argumento de David Lewis a favor de la teoría de la identidad mente-cuerpo es precisamente que resulta incompatible con nuestra confianza en la adecuación explicativa de la física no aceptar la teoría de la identidad. Véase, Lewis, D. K., (1966), p. 15 en la versión castellana.

⁴ La objeción de Kripke a la teoría de la identidad se encuentra formulada en Kripke, S., (1971) y (1980).

que las expresiones "dolor" y "estimulación de la fibra C" denotan lo mismo. Dicho de otra manera, para que sea verdadero el enunciado "el dolor es (idéntico a) la estimulación de la fibra C", es necesario que ambos términos que figuran en la oración de identidad refieran a lo mismo. Si un término mental M y un término físico F refieren a cosas o sucesos diferentes, el enunciado "M=F" será falso, aun cuando M y F puedan estar de alguna manera correlacionados. Que ocurran por accidente simultáneamente M y F, no basta para establecer el enunciado de identidad correspondiente; pues, como señala correctamente Kripke, afirmar que M=F equivale a afirmar que M y F necesariamente ocurren simultáneamente.⁵

Hay por lo menos dos versiones de la teoría de la identidad mente-cuerpo que conviene distinguir aquí, antes de iniciar la discusión:⁶

a) la versión que sostiene que la identidad se da entre tipos de estados mentales -digamos, dolores- y tipos de estados físicos o cerebrales -digamos, estimulación de las fibras C -, de modo que la verdad de la teoría implicaría la existencia de regularidades sin excepción y la posibilidad en principio de contar con leyes psicofísicas estrictas, y

b) la versión que afirma que cada suceso mental particular -digamos mi dolor de cabeza en este momento- es idéntico a un suceso cerebral particular describable en términos puramente físicos que ocurre en un determinado momento -digamos, la estimulación de mis fibras C ahora-, pero que niega la identidad entre tipos de estados mentales y tipos de estados físicos, con lo cual niega también la posibilidad de contar con leyes psicofísicas estrictas.

⁵ Véase, Kripke, S., (1980), p.151.

⁶ La distinción entre dos tipos de fisicalismo se encuentra expuesta con claridad en Davidson, D., (1970), en donde el autor ofrece una defensa ya clásica de la segunda versión del fisicalismo que caracterizaré a continuación.

Hay quienes piensan que la objeción que presenta Kripke al fisicalismo sólo se aplica a la primera versión de la teoría de la identidad mente-cuerpo, la que sostiene que los tipos de estados mentales son idénticos a tipos de estados cerebrales o corporales; sin embargo, aquí sostendré que el argumento de Kripke se aplica por principio a ambas versiones, aunque, según sostendré al final del capítulo, la objeción, a fin de cuentas, no parece concluyente para ninguna de las dos versiones del fisicalismo.

La objeción que Kripke opone a la teoría de la identidad entre lo mental y lo físico puede examinarse críticamente desde dos perspectivas. La primera consiste en revisar el aparato conceptual que Kripke utiliza en las dos obras en las que aborda el tema de la identidad psicofísica -a saber, El nombrar y la necesidad y su artículo "Identidad y necesidad"-, con el fin de cuestionar la verdad de algunas de sus tesis o la validez de algunas distinciones que introduce en esos trabajos y sobre las cuales basa sus objeciones; la segunda, en cambio, que es la que aquí adoptaré, consiste en aceptar provisionalmente el aparato conceptual kripkeano y discutir con él en su propio terreno. En este capítulo no discutiré, pues, las tesis de Kripke sobre la naturaleza del nombrar, la designación rígida o los mundos posibles, ni pondré en cuestión la existencia de las modalidades de re, ni la verdad del esencialismo. Tampoco discutiré la concepción que propone Kripke acerca de las verdades de la ciencia. Más bien, exploraré la cuestión de si efectivamente a partir de las tesis kripkeanas se sigue la falsedad de la teoría de la identidad mente-cuerpo. En la primera sección, expondré algunas tesis del propio Kripke indispensables para entender su objeción a la teoría de la identidad mente-cuerpo; en las secciones subsiguientes, plantearé y analizaré su objeción y trataré de mostrar que no es concluyente, esto es, que dejan aún una puerta abierta por donde el fisicalista puede escapar y sostener su posición como una alternativa

conceptual viable.

1. Identidad, designación rígida y modalidades.

De acuerdo con Kripke, un enunciado de identidad en el que figuran nombres, tal como "a=b", es un enunciado acerca del objeto designado por "a" y por "b"; esto es, no es un enunciado sobre los nombres "a" y "b", ni sobre los posibles contenidos descriptivos que un hablante o una comunidad de hablantes pudieran asociar con "a" y con "b". Y lo que afirma tal enunciado es que a y b son lo mismo, es decir, que entre el referente de "a" y el referente de "b" se da una relación reflexiva, simétrica y transitiva que sólo pueda tener un objeto consigo mismo. "La identidad -dice Kripke- debe tomarse simplemente como la relación de una cosa consigo misma".⁷ Lo que dice, pues, un enunciado de identidad es que un objeto es idéntico a sí mismo. Esto, desde luego, no implica que todos los enunciados de identidad sean triviales, en el sentido de ser analíticos, ni que su verdad pueda establecerse a priori. Puede constituir un genuino descubrimiento empírico llegar a saber, por ejemplo, que Cicerón es Marco Tulio. Obsérvese que lo que dice un enunciados de identidad tal como "Cicerón es Marco Tulio" no es lo mismo que lo que dice el enunciado del metalenguaje: "'Cicerón' y 'Marco Tulio' refieren al mismo individuo", pues este último habla acerca de los nombres "Cicerón" y "Marco Tulio" y una condición para que sea verdadero es que exista un lenguaje y que existan los nombres que figuran en el enunciado; en ese sentido su verdad depende de la existencia del lenguaje, en tanto que la verdad que expresa un enunciado de identidad tal como "Cicerón es Marco Tulio" es

⁷ Kripke, S., (1980), p. 108.

independiente de la existencia de todo lenguaje. Por otra parte, y como argumenté en el capítulo III, los enunciados de identidad en los que figuran un nombre y una descripción definida no establecen sinonimias entre esas expresiones; si afirmo que Cicerón es el autor de Las Catilinas, puedo, sin contradecirme, asegurar que Cicerón pudo no haber escrito Las Catilinas. Más aún, estos enunciados con forma superficial de identidad, a diferencia de las identidades en las que sólo figuran nombres, pueden no afirmar una relación de identidad de una cosa consigo misma, ni decir algo necesario. Si aceptamos, como lo hace Kripke, el análisis russelliano de las descripciones, lo que afirma "Cicerón es el autor de Las Catilinas" es que Cicerón es el único individuo que satisface la propiedad de ser autor de Las Catilinas y, a menos que pensemos que esa sea una propiedad necesaria de Cicerón, tendremos que considerar que la oración en cuestión expresa una proposición contingente.⁸

Kripke demuestra formalmente⁹ que si aceptamos: (i) la necesidad de la identidad de un objeto consigo mismo, esto es, la verdad de

(x) $\text{Nec}(x=x)$;

(ii) el principio de la indiscernibilidad de los idénticos de Leibniz -en el que se

⁸ La cuestión es un poco más complicada, ya que en el caso de las oraciones de identidad en las que figuran descripciones definidas, se tienen que hacer distinciones de alcance. Si la descripción tiene alcance corto, entonces lo dicho en el texto es correcto; sin embargo, si tuviera alcance amplio, la oración sí expresaría una proposición de identidad necesaria. Véase, Kripke, S.,(1971), n. 4.

⁹ Véase Kripke, S., (1971), p. 136 y ss; en la traducción al castellano, p. 6 y ss.

apoya la ley de sustituibilidad de los idénticos-, esto es:

$$(\underline{x})(\underline{y})[(\underline{x}=\underline{y}) \rightarrow (\underline{F}\underline{x} \leftrightarrow \underline{F}\underline{y})];$$

y (iii) la existencia de propiedades modales tales como "ser necesariamente idéntico a \underline{x} ", entonces, tenemos que aceptar la consecuencia de que toda identidad verdadera es necesaria, esto es, que

$$(\underline{x})(\underline{y}) [(\underline{x}=\underline{y}) \rightarrow \text{Nec } (\underline{x}=\underline{y})];$$

dicho de otra manera, que no hay identidades contingentes.

Decir que toda identidad verdadera es necesaria es equivalente a sostener que no sólo es verdadera en el mundo que de hecho se da, sino que sería verdadera en cualquier mundo posible o situación contrafáctica, esto es, que no podría darse una situación en la que el enunciado que la expresa fuese falso.¹⁰ Para poder sostener consistentemente la tesis de que todas las identidades verdaderas son necesarias junto con la tesis de que las identidades pueden ser descubiertas empíricamente, Kripke traza dos distinciones -que ya mencionamos en el capítulo anterior- en las que apoya su argumento: 1) la distinción entre designadores rígidos y no rígidos, y 2) la distinción entre la noción metafísica de necesidad y la noción epistémica de aprioridad. Recordémoslas aquí brevemente.

1. Un designador es una expresión del lenguaje cuya función semántica es referir a un individuo. Tanto los nombres como las descripciones definidas son

¹⁰ Véase, Kripke, S., (1971), pp. 150 y ss.; en la traducción al castellano pp. 18 y ss.

designadores.¹¹ Los primeros refieren sin describir, en tanto que los segundos sólo refieren porque describen.¹² Si el designador designa al mismo individuo en todo mundo posible en donde el objeto existe, se dirá que es rígido; en tanto que si designa a un individuo en el mundo que de hecho se da, pero pudiera designar a otros individuos en otros mundos posibles o situaciones contrafácticas, se dirá que es no rígido o accidental.¹³ Según Kripke, un nombre propio tal como "Cicerón" es un designador rígido porque designa en éste y en cualquier otro mundo posible a Cicerón, en tanto que la descripción "el autor de Las Catilinarías" es no rígido, porque aunque de hecho designa a Cicerón en este mundo (por haber sido él el autor de Las Catilinarías), podría haber designado a otro individuo en otro mundo posible. Es posible que Cicerón no fuese el autor de Las Catilinarías y que alguien en su lugar hubiese escrito discursos cualitativamente idénticos a los que de hecho escribió Cicerón; en ese caso ese individuo, y no Cicerón, sería el autor de Las Catilinarías. Esto es, hay un mundo posible en el que Cicerón no escribe Las Catilinarías, sino que las escribe alguien más, en tanto que no hay ningún mundo posible en el que Cicerón no es Cicerón. Cabe observar que dentro de la concepción kripkeana de la relación referencial no sólo los nombres propios ordinarios son designadores rígidos, sino que también lo son los nombres de

¹¹ Cabe hacer nuevamente una distinción de alcance. Sólo una descripción con alcance largo parece puede ser un designador, las descripciones con alcance corto cumplen más bien una función predicativa. Véase, Kripke, S., (1971), n. 4.

¹² Un nombre sólo está en lugar de la cosa nombrada, su único valor semántico es la cosa nombrada; el referente de una descripción definida es la única cosa que satisface la descripción. Véase, Kripke (1980), n. 3.

¹³ Véase, Kripke, S., (1971), p. 145 y (1980) p. 48.

fenómenos y especies naturales,¹⁴ tales como "calor", "tigre", "dolor", etcétera; pues si éstos designan un fenómeno o una especie natural en el mundo real, designarán ese mismo fenómeno o especie natural en todo mundo posible.

La distinción entre designadores rígidos y no rígidos no corresponde, de acuerdo con Kripke, a la distinción entre nombres y descripciones, ya que tanto los nombres (propios y de clases naturales) como ciertas descripciones definidas, designan rígidamente a sus objetos. Son descripciones rígidas las que designan a su referente mediante una propiedad esencial; por ejemplo, "el doble de 5" denota rígidamente al número 10, "la energía molecular cinética media" denota rígidamente al calor. Admitir que ciertas descripciones son designadores rígidos supone aceptar la existencia de propiedades esenciales y, en ese sentido, de modalidades de re¹⁵. Ahora bien, lo que Kripke sostiene es que si los términos que entran en un enunciado de identidad son ambos designadores rígidos, y si el enunciado en cuestión es verdadero, entonces será necesariamente verdadero.¹⁶

¹⁴ Traduzco "natural kind name" como "nombre de algún fenómeno o especie natural" con el fin de evitar una interpretación equivocada de la tesis de Kripke que parece estar más conectada con el esencialismo que con alguna fobia a lo no extensional.

¹⁵ Véase, Kripke, S., (1980) pp. 39 y ss. Si "el Φ " es una descripción rígida que designa a un mismo individuo x en todos los mundos posibles, entonces Φ es una propiedad esencial de x . Las modalidades de re, por contraposición a las modalidades de dicto, conllevan la aceptación de propiedades esenciales y la posibilidad de cuantificar en contextos modales, cfr. Kripke, S., (1971), p. 139. Aceptar las modalidades de dicto supone aceptar que hay entidades lingüísticas que pueden ser necesariamente verdaderas, pero no supone aceptar que haya hechos necesarios en la realidad extralingüística. Cfr., Kripke, S., (1980), p. 40.

¹⁶ Si en una oración de identidad alguno de los dos designadores utilizados es no rígido, la oración no expresará algo necesario; cfr. Kripke, S., (1971), pp. 139-140 y (1980), pp. 143-144. Véase también la nota 8 más arriba.

Así, pues, dado que "Marco Tulio" y "Cicerón" son designadores rígidos, y puesto que es verdadero que Marco Tulio es Cicerón, entonces será necesario el enunciado "Marco Tulio = Cicerón". El hecho de que la verdad de esta identidad no haya sido establecida mediante razonamientos a priori, sino mediante la investigación empírica, no obsta para atribuirle necesidad a la identidad en cuestión.

2. A partir de lo anterior, se hace patente que es indispensable trazar, también, la distinción entre las nociones de necesidad y de aprioridad si se quiere defender la tesis de que las identidades verdaderas son necesarias. Tradicionalmente, dice Kripke, estas nociones han sido confundidas o, por lo menos, usadas de manera intercambiable.¹⁷ Muchos filósofos han sostenido que todo conocimiento necesario es a priori y que todo lo conocido a priori es necesario. Si efectivamente esto fuera el caso, esto es, si la necesidad y la aprioridad fuesen inseparables, sería insostenible la tesis conforme a la cual toda identidad verdadera es necesaria, ya que, como acabamos de señalar, el modo como llegamos a conocer un buen número de identidades es mediante la investigación empírica. Piénsese en las identidades "Héspero es Fósforo" o "el calor es la energía molecular cinética media". No bastó reflexionar sobre "Héspero" para descubrir que refiere al mismo astro al que refiere "Fósforo", ni bastó reflexionar sobre el calor para percatarse de que es la energía molecular cinética media. Si ha de sostenerse que estas identidades son necesarias, habrá que deslindar la necesidad de la aprioridad.

Como explicamos en el capítulo anterior, las nociones de necesidad y de aprioridad pertenecen a dos campos distintos de la filosofía: La necesidad tiene su

¹⁷ Véase, Kripke, S., (1971), pp. 149-151, y (1980), pp. 34-39.

lugar en la metafísica, en tanto que la categoría de aprioridad lo tiene en la epistemología.¹⁸ Un enunciado es necesario si y sólo si su falsedad es imposible, esto es, si no hay ninguna situación contrafáctica en la que el enunciado sea falso. Por otro lado, decir que un enunciado es a priori, es referirse al modo como conocemos (o podemos conocer) su verdad. Un enunciado e es a priori para un sujeto s si s sabe que e es verdadero sin apelar a la experiencia.

Una vez trazada la distinción anterior, Kripke argumenta que, lejos de haber una correspondencia biunívoca entre la clase de las verdades a priori y la de las verdades necesarias, son posibles todos los tipos de entrecruzamientos entre esas dos categorías y las de contingencia y aposterioridad. Si admitimos las modalidades de re y aceptamos que las esencias pueden descubrirse empíricamente, esto es, que la ciencia, al investigar los rasgos estructurales básicos de una especie, puede descubrir las propiedades esenciales de la especie,¹⁹ tendremos que admitir que existen verdades necesarias a posteriori. Las identidades teóricas en las que los términos de la identidad son dos designadores rígidos, por ejemplo, "el calor es la energía cinética media", sería el prototipo de esta clase de verdades. Si efectivamente el calor es la energía cinética media entonces no hay ningún mundo posible o situación contrafáctica en el que no lo sea. Hay, también, verdades contingentes a priori -por ejemplo, la que expresa "El metro patrón que está en el Museo de Pesas y Medidas en París en t mide un metro"-, así como, desde luego, verdades contingentes a posteriori y necesarias a priori.²⁰

¹⁸ Kripke, S., (1971), p. 150 y (1980) p. 35.

¹⁹ Cfr. Kripke, S., (1980), p. 138.

²⁰ Kripke escribe dos cosas significativas sobre los enunciados necesarios a priori o analíticos: (i) que es una especie de estipulación decir que un enunciado

Dado lo anterior, podemos adivinar la dirección que toma la primera parte del ataque de Kripke a la teoría de la identidad mente-cuerpo tal como ha solido presentarse históricamente. Tradicionalmente, los filósofos materialistas han aceptado, sin percatarse de que la opinión que sostienen es problemática, que las identidades psicofísicas, en caso de ser verdaderas, serían verdades contingentes.²¹ La manera como parecen alcanzar esa conclusión es la siguiente: si se llega a establecer la verdad de la teoría de la identidad, esto no será sólo reflexionando sobre los conceptos que intervienen en las identidades psicofísicas, sino que deberá ser el resultado de ciertas observaciones y descubrimientos empíricos; esto es, las identidades psicofísicas, si son verdades, son verdades a posteriori. A partir del carácter empírico de las identidades psicofísicas pasan a sostener -apoyándose en la idea de que la experiencia no puede suministrar conocimiento de hechos necesarios- que las identidades psicofísicas, en caso de ser verdaderas, son contingentes.

Kripke objeta este tipo de materialismo argumentando que no hay tales identidades contingentes. La única salida que deja al filósofo materialista es la de aceptar la distinción entre necesidad y aprioridad y sostener que las identidades psicofísicas que supone verdaderas son necesarias, si bien a posteriori. Sin embargo, esto supone abandonar la idea de que sea siquiera posible la existencia de un determinado estado mental sin su correlato físico o viceversa, esto es, supone deshacerse de la llamada "intuición cartesiana". Esta intuición -que se

analítico es verdadero en virtud de su significado y (ii) que los enunciados analíticos son verdaderos en virtud de su significado, no sólo en el mundo real, sino en todo mundo posible. Véase Kripke, S., (1980), p. 39.

²¹ Véanse, por ejemplo, Smart, J.J.C. (1959), Place, U.T., (1962), y la primera frase de Lewis, D. K., (1966).

muestra en el hecho de que aparentemente podemos concebir que una planta esté triste, que un pájaro piense o que Dios haga planes- es la de que se pueden tener estados mentales sin tener una constitución física semejante a la que tenemos los humanos o sin tener ninguna constitución física en absoluto. El hecho de tener esta intuición, por una parte, resta plausibilidad a la tesis de que las identidades psicofísicas sean necesarias y, por otra, completa la explicación de por qué los materialistas piensan que las identidades psicofísicas son contingentes.

El defensor de la teoría de la identidad tiene, pues, una tarea difícil: mostrar que lo que creíamos que era concebible en realidad no lo es, que lo que creíamos que podíamos imaginar "esto es, que pudiera darse un cierto estado mental M sin su correlato físico F", en realidad no es algo que podamos imaginar. La dificultad radica en que la concebibilidad de tales situaciones parece ser un hecho inapelable; tanto es así, que está a la base de la idea de los propios materialistas de que las identidades psicofísicas son contingentes.

En suma, si el defensor de la teoría de la identidad sostiene que las identidades psicofísicas son contingentes, Kripke lo refuta mostrándole que no hay tal clase de identidades. Si pasa a sostener que son necesarias, Kripke le exige entonces que explique por qué la llamada "intuición cartesiana" no es una intuición genuina. Y es en ese momento cuando entra la segunda parte de la objeción de Kripke que es sin lugar a dudas la más poderosa. Kripke argumenta que no es posible mostrar que la intuición cartesiana sea una intuición ilusoria, es decir, que cuando creemos poder concebir que es posible la existencia de un estado mental M independientemente de cualquier estado físico F, nuestra capacidad de concebir tal cosa sea una mera ilusión. El caso de las identidades psicofísicas no es como el de las identidades teóricas de la ciencia en donde si es posible mostrar que la aparente posibilidad de concebir su falsedad no es más que una falsa apariencia y

que, por tanto, en esos casos cuando decimos tener la intuición de que lo que expresa una oración tal como "el calor es la energía cinética media" es contingente, se trata de una mera ilusión de contingencia.

2. La objeción de Kripke.

Kripke distingue tres formas distintas en las que se ha solido presentar la teoría de la identidad mente-cuerpo:²²

(i) La identidad de una persona con su cuerpo.

(ii) La identidad de dos sucesos singulares, esto es, un suceso mental irrepetible y particular (por ejemplo, la sensación de dolor de cabeza que tengo en este momento) y un suceso físico también irrepetible y particular (por ejemplo, mi tener en este momento determinada configuración neuronal en alguna región cerebral). Esta posición no admite que se puedan identificar tipos de estados mentales con tipos de estados físicos, y por tanto niega la posibilidad de que existan leyes psicofísicas.

(iii) La identidad de un tipo de estado mental (por ejemplo, el dolor) con un tipo de estado físico (por ejemplo, la estimulación de las fibras C). Esta versión de fisicalismo implica la existencia de leyes psicofísicas estrictas.

Cada una de estas modalidades de la teoría materialista de la mente presenta dificultades peculiares, pero todas ellas parecen sujetas a un mismo tipo de objeción que ya Descartes adelantó y otros, incluyendo a Kripke, han insistido en

²² Kripke, S., (1980), p. 144.

presentar.²³ La objeción cartesiana a la teoría de la identidad enunciada en la forma (i) consiste en sostener que, dado que, para cualquier individuo x , es concebible la existencia de un cuerpo idéntico al de x , con configuraciones neuronales idénticas a las de x , pero sin ningún tipo de estado mental (o de experiencias), entonces la mente (o los fenómenos mentales) de x y el cuerpo (o los fenómenos neuronales) de x no pueden ser idénticos, sino que son "sustancias" (o fenómenos que ocurren en dos sustancias) distintas. Algunos críticos de Descartes lo acusaron de cometer una falacia modal, esto es, de pasar injustificadamente de un enunciado de posibilidad a la afirmación categórica de que la mente (o la sustancia pensante) es distinta del cuerpo (o de la sustancia extensa). Sin embargo, Kripke sostiene que esa crítica es infundada,²⁴ y la objeción cartesiana válida, debido a que los términos "mente" y "cuerpo" usados en el argumento de Descartes son designadores rígidos.²⁵ Esto es, si los términos con los que se enuncia la identidad designan rigidamente a sus objetos, de la premisa cartesiana se sigue efectivamente su conclusión. Así, si se acepta como verdadera la premisa acerca de la posible existencia de estados físicos sin sus respectivos

²³ Colin McGinn ofrece un argumento para mostrar que la objeción de Kripke no se aplica a la versión (ii) de fisicalismo. Véase, McGinn, C., (1977). Véase, también, Hansberg, O., (1978), esp. pp. 160-162.

²⁴ Véase, Kripke, S., (1971), p. 163, y (1980), pp. 145-146

²⁵ Descartes, de hecho, no usa el término "mente" en su argumento; en la versión francesa de Las Meditaciones usa indistintamente "âme", "esprit" o "pensée". Véase, Descartes, René, Méditations touchant la Première Philosophie dans lesquelles l'existence de Dieu et la distinction réelle entre l'âme et le corps de l'homme sont démontrées, Paris, 1647. (La versión original en latín de Las Meditaciones fue publicada en París en 1641.) Sin embargo, está enteramente dentro del espíritu de Descartes la defensa que hace Kripke de su argumento.

correlatos mentales, o viceversa, se tiene que aceptar inevitablemente la conclusión dualista.²⁶ De manera que un defensor coherente de la teoría de la identidad, está obligado a mostrar que la aparente posibilidad de concebir que la mente no es idéntica al cuerpo, o que los estados mentales son distintos de los estados neurofisiológicos o físicos, no es más que una posibilidad ilusoria, es decir, que la intuición cartesiana es, a fin de cuentas, una intuición equivocada. Hay quienes sostienen que el argumento de Kripke no se aplica a la forma (ii) de la teoría de la identidad mente-cuerpo.²⁷ En cambio, se suele aceptar que la objeción cartesiana se aplica a la forma (iii) de la tesis materialista. En lo que sigue, la discusión se centrará en esta última forma por las siguientes razones; primera, es más clara y más precisa que la (i); segunda, no es mi interés aquí mostrar que (ii) no se salva de la objeción; y, tercera, por ser la forma (iii) de la teoría a la que Kripke opone con más fuerza sus argumentos.

Sean "M" y "F" designadores rígidos. Digamos que "M" designa un tipo de estado mental, digamos, el dolor, y "F" un tipo de estado físico o neurofisiológico de los individuos que tienen M, digamos, la estimulación de las fibras C. El

²⁶ Kripke, S., (1980), p. 145.

²⁷ Véase la nota 21. Olbeth Hansberg, en Hansberg, O., (1978), sostiene que una teoría de la identidad del tipo (ii) permite explicar por qué tenemos la intuición de que las identidades psicofísicas son contingentes. Según Hansberg, el hecho de que estados mentales de un mismo tipo puedan ser idénticos a estados neuronales de tipos diferentes explica tal intuición. Sin embargo, nótese que el argumento de Kripke se aplica a cada una de las identidades verdaderas que acepta una teoría de la forma (ii). Así, si se acepta que un dolor particular d es idéntico a un estado neuronal particular n -rechazando explícitamente la identidad entre tipos-, se tiene la intuición -diría Kripke- de que d podría existir sin n y esta impresión de que d y n están correlacionados de manera contingente no parece poder explicarse apelando a la ausencia de correlaciones estrictas entre tipos de dolores y tipos de estados neuronales, que por hipótesis se han rechazado.

materialista, defensor de la teoría de la identidad mente cuerpo, sostiene que $M = F$. Entonces, por el argumento de Kripke acerca de la necesidad de tales identidades, el materialista tiene que sostener que necesariamente M es F . Esto es, no puede congruentemente sostener que M es (idéntico a) F y a la vez que M podría ser diferente de F . Aceptar la posibilidad de que M fuese diferente de F equivale a autorrefutarse.

El argumento I a continuación pretende reconstruir la primera parte del ataque de Kripke en contra de la teoría de la identidad. Su objetivo es mostrar que el materialista cae en contradicción si acepta la intuición cartesiana:²⁸

Argumento I

{1}	(1) $(M = F)$	tesis materialista
{ Λ }	(2) $(M = F) \rightarrow \text{Nec}(M = F)$	Necesidad de la identidad.
{3}	(3) $\diamond \neg(M = F)$	intuición cartesiana
{ Λ }	(4) $\text{Nec}(M = F) \rightarrow \neg \diamond \neg(M = F)$	equivalencia definicional
{3}	(5) $\neg \text{Nec}(M = F)$	de 3 y 4, por M.T.
{1,3}	(6) $\neg(M = F)$	de 2 y 5, por M.T.
{1,3}	(7) $(M = F) \ \& \ \neg(M = F)$	conjunción de 1 y 6

Más aún, "M" y "F" son designadores rígidos, de modo que si aceptamos

²⁸ En los argumentos a continuación uso la notación que introduce Benson Mates en Mates, B., (1972), cap. 6 y que usa en los caps. 7, 8, y 9. A la izquierda del número de cada línea aparece el conjunto de premisas de las que se deriva (de la que se sigue) la línea en cuestión. Si el número de línea es el mismo que el que aparece a la izquierda entre llaves eso significa que la línea es una premisa; si la línea se deriva del conjunto vacío, Λ , entonces la línea es un teorema de la lógica (expresa un verdad lógica).

el principio según el cual, si dos designadores rígidos D y D' denotan distintos individuos en un mundo posible, entonces denotarán distintos individuos en todos los mundos posibles en los que de hecho denoten algo, tenemos que aceptar que si $\neg(M = F)$, entonces necesariamente $\neg(M = F)$. Por ejemplo, si "Venus" y "Marte" son designadores rígidos y denotan diferentes individuos en el mundo real, de manera que $\neg(\text{Venus} = \text{Marte})$, entonces, en todo mundo posible (en donde existan los denotata de los nombres) denotarán diferentes individuos y no habrá ningún mundo posible en el que Venus sea Marte. Si lo anterior es correcto, entonces,

$$\neg(\text{Venus} = \text{Marte}) \rightarrow \neg\blacklozenge(\text{Venus} = \text{Marte})$$

Y el consecuente de este condicional es equivalente al enunciado modal $\text{Nec}(\text{Venus} = \text{Marte})$, de modo que si Venus no es Marte, entonces necesariamente no lo es. El argumento puede generalizarse de modo que pruebe el principio de la necesidad de la no identidad:²⁹

$$\text{PNNI. } (\underline{x})(\underline{y})(\neg(\underline{x}=\underline{y}) \rightarrow \text{Nec}\neg(\underline{x}=\underline{y}))$$

Una vez armados con este principio, añadimos al Argumento I:

$$\{\Delta\} (8) \neg(M = F) \rightarrow \text{Nec}\neg(M = F) \quad \text{PNNI}$$

y podemos mostrar que el materialista que acepta la intuición cartesiana cae en la

²⁹ Kripke se refiere a este principio como el principio de la Necesidad de la Diferencia en Kripke, S., (1980), nota al pie 56, p. 114.

contradicción modal (11), aparentemente más fuerte que la contradicción de la línea (7):

{2,3}	(9) Nec-(M = F)	de 6 y 8, por M.P.
{1}	(10) Nec(M = F)	de 1 y 2 por M.P.
{1,2,3}	(11) Nec(M = F) & Nec-(M=F)	Conjunción de 9 y 10.

Mediante un argumento similar, a partir de la tesis de Kripke de la necesidad de la identidad y de la aceptación de la intuición cartesiana acerca de la concebibilidad de una situación en la que M. no sea F, y añadiendo lo que he llamado "el principio de la necesidad de la no identidad", llegamos no sólo a la conclusión dualista, sino a la necesidad de dicha conclusión. Y en ese punto, precisamente, lo que dice el argumento comienza a parecer contraintuitivo, ya que resulta difícilmente aceptable que no sea ni siquiera posible (y, por ende, realmente concebible) la falsedad de la tesis dualista.

Pensemos que el materialista puede construir el Argumento II, a continuación, cuya premisa 2' sólo afirma que la posición materialista es concebible, esto es, que es realmente pensable que $M = F$. Mediante ese argumento el materialista puede llegar a la conclusión contraria a la que llega el Argumento I.

Argumento II:

{ Δ }	(1') $\neg(M = F) \rightarrow \text{Nec}\neg(M = F)$	PNNI
{2'}	(2') $\diamond(M = F)$	Intuición materialista.
{2'}	(3') $\neg\text{Nec}(M = F)$	de 2', por equiv. definic.
{2'}	(4') $\neg(M = F)$	de 1' y 3', por M.T.
{2'}	(5') $(M = F)$	de 4', doble negación.
{2'}	(6) $\text{Nec}(M = F)$	de 5', por la necesidad de la identidad.

En el Argumento II, partiendo de las tesis kripkeanas sobre la identidad, conjuntamente con la "intuición materialista" conforme a la cual se acepta que es al menos concebible que $M = F$ (bajo peligro de reducir al absurdo la posición materialista), se llega a la conclusión no sólo de que es verdadera la tesis de la identidad sino que es necesariamente verdadera. Si bien el argumento de Kripke a favor del dualismo obliga al materialista a mostrar que la intuición cartesiana es falsa, parece que el argumento materialista obligaría al dualista a mostrar que la intuición materialista es equivocada o, como diría Kripke, a mostrar el "carácter ilusorio" de esa intuición.

Cabe notar que en ambos argumentos, I y II, basta partir de la concebilidad de la identidad (o de la no identidad) para establecer la correspondiente necesidad.

El problema parece estar en las premisas 3 del Argumento I y 2' del Argumento II. Esto es, parece que en algún sentido del término "concebible", son efectivamente concebibles tanto que el dolor sea la estimulación de las fibras C como que no lo sea. El hecho de que históricamente se hayan sostenido tanto posiciones dualistas como materialistas, obliga a aceptar, prima facie, que son al menos pensables o concebibles las tesis que caracterizan a cada una de esas posiciones. Sin embargo, al decir que ambas tesis son concebibles, ¿queremos simplemente informar de un hecho empírico, esto es, de que históricamente ha habido filósofos y científicos que han sostenido la posibilidad tanto de que M sea F como de que M no sea F? O, por el contrario, cuando decimos que ambas cosas son concebibles, ¿estamos aceptando que las oraciones "M es F" y "M no es F" expresan posibilidades reales, metafísicas? De esta cuestión me ocuparé en la siguiente sección. Por el momento basta observar que dado que la intuición

cartesiana y la materialista tomadas conjuntamente conducen a conclusiones contradictorias, no pueden aceptarse ambas, ni considerar que ambas expresan posibilidades reales. En alguno de los dos casos tiene que tratarse de una mera ilusión de posibilidad.

Alguien podría decir que lo que hemos llamado la "intuición cartesiana" y "la intuición materialista" en realidad no tienen el mismo estatus; esto es, que en tanto que la primera es una verdadera intuición con alcances metafísicos, la segunda no pasa de ser una mera hipótesis (falsa, por cierto) que han sostenido algunos filósofos y que sirve tal vez de guía en sus investigaciones a algún grupo de científicos. A esto podemos contestar (a) que no parece haber nada que las distinga formalmente, esto es, una afirma que es concebible la no identidad de los estados mentales con estados físicos, la otra que es concebible tal identidad; (b) decir simplemente que una tiene alcances metafísicos, en tanto que la otra no, equivale a cometer una petición de principio; (c) si la intuición materialista parece problemática o dudosa, tanto más dudosa resulta la intuición cartesiana, conforma a la cual un ser desprovisto de sistema nervioso central pudiera tener vida mental: experiencias, sensaciones y actitudes proposicionales. Los avances científicos no pueden menos que afectar nuestras intuiciones acerca de lo que es posible y de lo que no lo es. Y dados los descubrimientos que se han logrado en el campo de las neurociencias y de las ciencias cognoscitivas, resulta problemático, por contraintuitivo, aceptar la concebibilidad de una situación en la que, por ejemplo, un manequí sienta o piense o, como dice Kripke, la posibilidad (concebibilidad) de que alguien tenga experiencias y, sin embargo, descubramos que carece de cualquier estructura nerviosa.³⁰

³⁰ Cfr., por ejemplo, Kripke, S., (1971), p. 161 y (1980), pp. 147 y 148

He sostenido que parece haber un sentido legítimo en el que podemos decir que son concebibles ambas posiciones antagónicas, por otro lado tenemos que admitir que alguna de las dos intuiciones tiene que ser falsa, bajo peligro de caer en una contradicción. Antes de revisar los argumentos de Kripke para defender que la intuición cartesiana es verdadera, quiero explorar la cuestión de qué queremos decir cuando afirmamos que algo, una situación, es concebible.

3. Concebibilidad y posibilidad.

Hay por lo menos dos maneras de interpretar la expresión "es concebible que M no sea F".

(a) Una interpretación epistemológica, conforme a la cual diríamos que, dada la evidencia con la que contamos y las teorías dentro de las que nos movemos, nos parece posible que resulte ser el caso que M no sea F. Esto sería equivalente a sostener que, hasta donde sabemos, decir que M y F son diferentes no implica contradicción. Nótese que la evidencia con la que contamos puede ser insuficiente o la teoría dentro de la que nos movemos puede resultar equivocada. Parece ser que, en este sentido de "concebir", distintas personas en un mismo tiempo t o una misma persona en dos momentos distintos t_1 y t_2 pueden concebir dos posiciones contradictorias: puedo pensar hoy que los estados mentales son estados cerebrales y dentro de cinco años que no es el caso que los estados mentales sean estados cerebrales. (En un momento dado se pensó que la luz eran ondas, en otro, que tenía una naturaleza corpuscular). La evidencia con la que contamos en la actualidad parece hacer concebibles, en este sentido epistemológico, tanto que M es F como que M no es F. Esto es, estamos en una situación en la que nuestra

ignorancia no nos permite descartar ninguna de las dos posibilidades.

b) Una interpretación metafísica, conforme a la cual lo concebible es lo realmente posible, independientemente de que lo pensemos o no; en este sentido, lo concebible sería lo que puede ser el caso independientemente de que lo sepamos o lo imaginemos.³¹ No serían "concebibles", en este sentido, las situaciones descritas mediante enunciados necesariamente falsos. Si fuese el caso que M es F, sería metafísicamente imposible que M no fuese F, y en este sentido metafísico de "inconcebible" sería inconcebible que M no fuese F. Sin embargo, cuando no sabemos si M es F, tal vez no resulte epistemológicamente inconcebible que M no sea F.

Reservemos la expresión "concebible" para referirnos a lo que epistemológicamente nos parece posible y hablemos de "posibilidad" sólo para referirnos a lo metafísica o realmente posible.

Algo es concebible sólo relativamente a un conjunto de conocimientos o creencias previas; si las creencias son falsas, lo que es concebible en un momento t puede no ser posible, esto es, podemos concebir situaciones necesariamente falsas, imposibles (por ejemplo, podemos pensar que la luz es un fluido, que los átomos de oxígeno tienen valencia 3, que los ángulos internos del triángulo miden 100°). Puedo creer en estos momentos que Saturno es un planeta cuya composición química es Q, y en este sentido concebir que Saturno está hecho de Q. Pero si resulta que en realidad Saturno tiene una composición distinta de Q, digamos que tiene una composición Q', lo que concibo (epistemológicamente) es metafísicamente imposible. Diríamos entonces que aunque es concebible

³¹ Tal vez lo metafísicamente posible coincidiría con lo concebible sólo en el caso en que las teorías que aceptásemos fuesen correctas y completas y, por ende, contásemos con un conocimiento cabal de la realidad.

(epistemológicamente) que Saturno, ese mimísimo planeta, tenga una composición Q , es metafísicamente imposible que tenga tal composición, puesto que de hecho tiene una composición Q' y nada puede tener a la vez dos composiciones químicas diferentes. Resulta también sostenible que lo que es epistemológicamente concebible con respecto a un conjunto de creencias C en un momento dado t , no sea concebible en un momento posterior t' en el que contamos con otro conjunto C' de creencias. Antes de que los científicos establezcan la verdad de una identidad de la forma $\alpha=\beta$, es concebible tanto que $\alpha=\beta$ como que $\alpha\neq\beta$.

Ahora bien, cuando decimos que la evidencia hace razonable sostener, como parece ser el caso, que es concebible tanto que M es F como que $M \neq F$, lo único que -si somos congruentes- podemos querer decir es que ambas situaciones son concebibles en el sentido epistemológico del término, pues, como argumenté antes, si quisiéramos decir que ambas representan posibilidades reales o metafísicas caeríamos en contradicción. Pero si las intuiciones cartesiana y materialista sólo expresan la concebibilidad epistemológica tanto de M es F como de M no es F , de allí no puede legítimamente inferirse la posibilidad real o metafísica de las situaciones descritas mediante las oraciones " M es F " y " M no es F ". Esto es, como se desprende de variados ejemplos en la historia de la ciencia, no contamos con una regla o principio general que garantice el paso de lo epistemológicamente posible a lo real o metafísicamente posible.

Si esto es así, entonces, tenemos que el argumento cartesiano o bien parte de la concebibilidad (epistemológica) de que $M \neq F$, en cuyo caso la premisa no es suficiente para establecer la conclusión deseada -esto es, que sea metafísicamente posible que $M \neq F$ -, o bien introduce una premisa con pretensiones metafísicas con la única justificación de que su contraparte epistemológica ha sido aceptada, lo cual, como indiqué anteriormente, no es un

paso inferencial generalmente garantizado. Sólo si es metafísicamente posible que $M \neq F$, entonces será necesariamente falso que $M = F$. Para que el argumento cartesiano refute la posición materialista, es necesario que tome como premisa, no la concebibilidad epistemológica de que $M \neq F$, sino su posibilidad metafísica; pero tomar esto último como premisa, sin más justificación que el hecho de que es epistemológicamente concebible que $M \neq F$, equivale a asumir lo que se quiere probar. De la misma manera, si el argumento materialista partiese de la premisa que dice que es metafísicamente posible que $M = F$ para establecer la verdad de " $\text{Nec}(M = F)$ ", se le podrían aplicar el mismo tipo de objeción.

En suma, el hecho de que sea concebible que $M \neq F$, hace concebible (epistemológicamente) que sea metafísicamente posible que $M \neq F$, pero esto de ninguna manera basta para establecer la posibilidad real de que $M \neq F$.³²

Como he señalado, las situaciones que concebimos las concebimos siempre en el marco de un conjunto de creencias. No podemos hablar de que p sea concebible considerada como un átomo aislado, sin relaciones con otras proposiciones creídas más o menos generales. Podemos en ciertos casos, desde luego, mostrar que lo que era epistemológicamente concebible, digamos p , no es realmente posible; para ello basta mostrar, por ejemplo, que alguna de las creencias con relación a las que resultaba concebible p , es una creencia falsa o que tal creencia, junto con otras creencias verdaderas, implica que $\neg p$. Así, por ejemplo, puedo mostrar que es metafísicamente imposible que Héspero no sea

³² David Wiggins, en Wiggins, D., (1975), sostiene algo semejante: "La afirmación sincera de que uno ha imaginado [algo] difícilmente puede demostrar la concebibilidad real [de ese algo]. Tal vez el que un hombre imagine que se da p muestra al menos que es concebible que sea concebible que p . No lo sé." (p. 22). Lo que Wiggins llama "concebible" parece corresponder a lo que he llamado metafísicamente posible.

Fósforo, porque la creencia de que un mismo cuerpo celeste no puede estar en diferentes lugares de la bóveda celeste en la mañana y en la tarde -que podemos suponer era una prueba a favor de la verdad de "Héspero no es Fósforo"-, se mostró falsa y, en cambio, se vió que otras creencias verdaderas implican la verdad de "Héspero es Fósforo". En muchos casos saber que algo es metafísicamente imposible sólo puede descubrirse a posteriori. De la misma manera, si se quiere eliminar la intuición cartesiana o la intuición materialista, parece ser necesario mostrar que alguna de las creencias en la que se fundan dichas intuiciones es en realidad falsa o insuficiente, y eso, también, es algo que en apariencia sólo podrá hacerse a posteriori, es decir, sólo podrá llevarse a cabo cuando los descubrimientos científicos pertinentes lo permitan.

Resumiendo, la posición que sostengo es que ni el cartesiano ni el materialista pueden dar como justificación de sus premisas 3 y 2', respectivamente, el hecho de que $M \neq F$ o de que $M = F$ sean epistemológicamente concebibles, porque, como he argumentado, esto en general no constituye una garantía de la posibilidad metafísica correspondiente. En este sentido, podemos afirmar que el problema con los argumentos I y II no es que cometan una falacia modal, sino que no ofrecen una justificación adecuada de sus premisas. Por otro lado, si el cartesiano o el materialista introducen 3 y 2' con la intención de afirmar posibilidades metafísicas, sin más justificación, considero que asumen justamente lo que deberían probar.

El problema al que nos enfrentamos aquí es de índole general: ¿Cómo es posible pasar de una posibilidad epistemológica a una posibilidad (o necesidad) real? ¿Es justificable pasar de la concebibilidad de una identidad a la necesidad de la misma?

Kripke distingue claramente el campo de la epistemología del campo de la

metafísica y subraya en repetidas ocasiones que no es legítimo pasar, sin más argumentación, de una consideración de índole epistemológica (por ejemplo, el carácter a priori de una verdad) a otra de índole metafísica (por ejemplo, el carácter necesario de la misma). Si él considera que el argumento cartesiano puede formularse de manera tal que no se le pueda acusar de hacer este tipo de inferencia ilegítima, es porque puede ofrecer otros argumentos (no la simple concebibilidad de que $M \neq F$) que le permiten defender el punto de vista conforme al cual la intuición cartesiana expresa una genuina posibilidad metafísica, y no una mera posibilidad epistemológica; esto en contraposición con ciertas otras identidades teóricas cuya falsedad es en apariencia posible, pero en donde puede mostrarse que tal posibilidad es puramente ilusoria.

En la siguiente sección examinaré el argumento de Kripke a favor de considerar que la intuición cartesiana expresa una auténtica posibilidad metafísica.

4. La falsa impresión de contingencia en el caso de identidades teóricas

Las identidades teóricas tales como "el calor es la energía cinética media" o "la luz es un haz de fotones" contienen términos que son designadores rígidos y, por lo mismo, si las identidades son verdaderas, son necesariamente verdaderas.³³ Si esto es así -argumenta Kripke-, resultará realmente inconcebible la falsedad de una identidad teórica verdadera; aunque creamos concebir la

³³ Véase, Kripke, S., (1971), p. 160 y (1980), pp. 140 y ss. La idea de Kripke es que los fenómenos y especies naturales existen independientemente de toda taxonomía impuesta por los humanos, de modo que los términos que refieren a ellos designan a esos mismos fenómenos o especies naturales en todo mundo posible o situación contrafáctica en donde existen.

posibilidad de su falsedad, eso no es posible. Acepta que, a veces, frente a tales identidades de la ciencia tenemos la misma impresión de contingencia que la que tenemos frente a identidades psicofísicas. Pero, investiga el origen de esa impresión de contingencia en ambos casos y argumenta que en el caso de las identidades teóricas de la ciencia tal impresión tiene un carácter puramente ilusorio, cosa que no sucede con la impresión de contingencia que tenemos frente a las identificaciones psicofísicas. En suma, Kripke pretende mostrar que hay una falta de paralelismo importante entre las identidades teóricas de la ciencia y las identidades psicofísicas: la impresión de contingencia que tenemos frente a identidades tales como "el calor = energía molecular cinética media" es meramente ilusoria; en cambio, la impresión de contingencia que tenemos frente a identidades como "el dolor es la estimulación de las fibras C" es una impresión verídica. Su argumento es que si no podemos mostrar que la impresión de contingencia que acompaña a las identidades psicofísicas sea ilusoria, entonces tenemos que aceptar que expresan posibilidades genuinas y, en ese caso, que la intuición cartesiana tiene alcances metafísicos.

Para mostrar que la aparente contingencia de las identidades teóricas es ilusoria, Kripke introduce la noción de "descripción que fija la referencia de un nombre". Podemos usar una descripción para fijar la referencia de un nombre, esto es, para identificar o seleccionar el objeto denotado por el nombre. Así, por ejemplo, usamos la descripción "lo que nos produce la sensación de caliente" para fijar la designación del nombre "calor", o "lo que nos produce la impresión visual de claridad" para fijar la referencia de la palabra "luz". Si bien "calor" es un designador rígido, la descripción que usamos para fijar su referencia es no rígida; esto es, seleccionamos al calor mediante una propiedad accidental del calor, pues resulta concebible que ese fenómeno natural que es el calor pudiera no

producirnos la sensación de caliente, sino la de frío o ninguna. Si esto es así, entonces, es contingente que aquello que causa en nosotros la sensación de caliente sea el calor y, por ende, la identidad "lo que causa en los humanos la sensación de caliente es el calor" no expresa un hecho necesario.³⁴ La razón por la que el hecho expresado no es necesario es que la descripción "lo que causa en los humanos la sensación de caliente" no es un designador rígido, en tanto que el nombre "calor" sí lo es. Lo anterior no significa que sea contingente el hecho que expresa "el calor = la energía cinética media". En este último caso tenemos una identidad necesaria, dado que la oración es verdadera y en ella figuran dos designadores rígidos.³⁵ La ilusión de contingencia proviene, pues, del hecho de que seleccionamos el calor mediante una propiedad accidental (ser sentido por los humanos de tal o cual manera). Así, podemos explicar en términos generales la apariencia de contingencia de una identidad que utiliza los designadores rígidos R_1 y R_2 , señalando que las descripciones D_1 y D_2 , mediante las cuales fijamos las referencias de R_1 y R_2 , son descripciones accidentales. Si bien " $R_1 = R_2$ " es necesaria, no lo serán ni " $R_1 = D_2$ ", ni " $R_2 = D_1$ " (ni, desde luego, " $D_2 = D_1$ "), aun cuando todas estas expresiones sean igualmente verdaderas. Lo que en el enunciado " $R_1 = R_2$ " nos produce una falsa impresión de contingencia es que fijamos la referencia de R_1 mediante una descripción no rígida D_1 y que tácitamente suponemos que " $R_1 = R_2$ " expresa la misma proposición que " $D_1 =$

³⁴ Habría que hacer aquí la distinción de alcance pertinente. Sólo si el alcance de la descripción es corto la identidad no será necesaria.

³⁵ Recordemos que las descripciones esenciales son también designadores rígidos y la descripción "la energía cinética media" es una descripción esencial del calor.

R_2 ", lo cual, como señalé antes, no es el caso.³⁶

Pero ¿qué pasa en el caso de las identidades psicofísicas? Se trata, en opinión de Kripke, también de identidades en las que figuran designadores rígidos: "dolor" denota rígidamente la experiencia de dolor y "estimulación de las fibras C" denota rígidamente un determinado estado cerebral o una determinada configuración neuronal. Sin embargo, sostiene Kripke, ambos objetos designados son seleccionados mediante propiedades esenciales, de modo que no podemos alegar aquí que la apariencia de contingencia provenga -como en el caso de las identidades teóricas- del hecho de haber fijado la referencia de los términos "dolor" y "estimulación de las fibras C" mediante propiedades accidentales. Seleccionamos el dolor porque lo sentimos de un modo especialmente desagradable, y es esencial al dolor ser sentido de esa manera; y seleccionamos el estado cerebral por ser una configuración neuronal de tal y cual tipo, y -sostiene Kripke- es también esencial al estado cerebral ser dicha configuración neuronal.³⁷

De esta manera, piensa Kripke, no sólo queda establecido que las identidades teóricas son realmente necesarias (a pesar de que tenemos la falsa impresión de que su falsedad es concebible), sino que también ha mostrado que la intuición cartesiana -acerca de la posibilidad de que los estados mentales no sean estados físicos o cerebrales- resiste el tipo de explicación que se ofrece para disipar la apariencia de contingencia de las identidades teóricas, y permanece, por tanto, como la expresión de una genuina posibilidad metafísica.

³⁶ Véase, Kripke, S., (1971), p. 160 y (1980), pp. 143-144.

³⁷ Véase Kripke, S., (1971), pp. 161-162 y (1980), pp. 144 y ss.

5. Referencia directa y rígida, e identidades psicofísicas.

La argumentación de Kripke, sin embargo, no parece ser concluyente.

(i) En primer lugar cabe señalar, que si bien es claro que seleccionamos el dolor por el hecho de sentirlo de cierta manera, lo cual efectivamente constituye una propiedad esencial del dolor (esto es, no puede existir un dolor sin ser sentido de la manera como de hecho se siente),³⁸ no resulta claro sostener que seleccionamos al estado cerebral F (la estimulación de las fibras C) mediante la propiedad esencial de "ser el estado cerebral F". Esto es, con respecto al dolor, es claro que la "evidencia" que nos justifica (si es que puede decirse tal cosa) para poder afirmar con verdad que tenemos un dolor, es el hecho de que lo sentimos; de manera que al fijar la referencia del término "dolor" mediante la descripción, por ejemplo, "la sensación corporal que sentimos de cierta manera especialmente desagradable", efectivamente usamos una propiedad esencial del dolor para identificarlo. La "evidencia sensorial" es en este caso definitiva; hay una relación necesaria entre la evidencia y el objeto seleccionado y, en ese sentido, podemos decir que seleccionamos el dolor mediante una propiedad esencial. En el caso del estado cerebral F, en cambio, no lo seleccionamos, ni lo identificamos directamente como el estado cerebral F, sino que lo identificamos como el estado cerebral que, dentro del marco de una teoría, está correlacionado con una

³⁸ Esto no equivale a sostener que la manera como descubrimos en todos los casos que existe un dolor sea porque lo sentimos, pues en el caso de atribuir dolor a terceras personas (si es que éste es un caso de descubrir que existe un dolor) lo hacemos sobre la base de la conducta verbal y no verbal de las terceras personas. No se trata, pues, de una afirmación acerca de *cómo sabemos* que existe un dolor, es decir, de una afirmación de índole epistemológica, sino de una afirmación con respecto a lo que el dolor *es*.

determinada evidencia E (la cual obviamente no incluye que sea sentido de ninguna manera). Así, por ejemplo, podríamos usar la descripción "el estado cerebral que existe cuando el encefalograma describe una gráfica del tipo G" o "el estado que se presenta cuando se registra mediante aparatos de laboratorio que se libera una cantidad n de acetil-colina en determinadas sinapsis", para fijar la referencia de la expresión "estimulación de las Fibras C". Nótese que esa evidencia que nos permite seleccionar un determinado estado cerebral, al que designamos rígidamente mediante "estimulación de las fibras C", no es la propia estimulación de las fibras C, ni ninguna propiedad esencial de la estimulación de las fibras C; dicha evidencia tiene sólo una relación contingente con tal estimulación. Es decir, es perfectamente imaginable que se dé una situación en la que la evidencia sensorial sea la misma y sin embargo no se dé la estimulación de las fibras C. Dicho en otras palabras, es posible que nos hallemos exactamente en la misma situación epistemológica en la que nos de hecho nos hallamos cuando diagnosticamos la estimulación de las fibras C y que, sin embargo, no haya tal estimulación. Y, si esto es así, es posible, entonces, explicar nuestra intuición de que es contingente la identificación entre el dolor y la estimulación de las fibras C apelando precisamente a esta posibilidad.³⁹

Podemos también imaginar que se identificase a F mediante una evidencia totalmente diferente. En el caso de identificar un estado cerebral F lo que hacemos es interpretar cierta evidencia sensorial a la luz de una teoría científica. Pero no hay nada que sea identificar a F al margen de toda teoría; esto es, no es que existan por ahí, para que se les mire, un conjunto de estados cerebrales bien determinados y que nuestra función sea la de seleccionar uno de ellos como quien

³⁹Véase, Kripke, S., (1980), pp. 141-143.

selecciona, de entre un conjunto de objetos coloreados, aquél que es rojo. En el caso de la identificación del dolor no es necesaria la intervención de ninguna teoría científica; en el caso de la identificación del estado cerebral, es indispensable tal teoría. Y la conexión entre la evidencia sensorial y la existencia de un objeto teórico es generalmente contingente. Lo anterior hace razonable sostener que, si bien la referencia del término "dolor" se fija mediante una de sus propiedades esenciales, no sucede lo mismo con el designador "estado cerebral F" (la estimulación de las fibras C), en donde la referencia se fija mediante una propiedad contingente, a saber, la de estar contingentemente relacionado con una evidencia E. "Ser un estado cerebral de cierto tipo B es [...] esencial a B", escribe Kripke, y esto parece inapelable; lo que, en cambio, parece cuestionable es la afirmación de que seleccionamos o identificamos un estado cerebral B por la característica esencial de ser un estado cerebral B, y no por estar correlacionado contingentemente, dentro del marco de una teoría, con determinada evidencia E.⁴⁰

⁴⁰ Hay un pasaje en Kripke, S., (1980), pp. 150-151, en donde Kripke parece acercarse a la idea de que el hecho de sentirse como dolor interviene en la manera como fijamos la referencia de "la estimulación de las fibras C". Si esto fuese así, y la teoría de la identidad resultase verdadera, efectivamente estaríamos utilizando una propiedad esencial del estado cerebral F (= dolor) para fijar la designación de "la estimulación de las fibras C". Sin embargo, esto parece falso después de lo argumentado en líneas anteriores: no fijamos la referencia de F por el hecho de ser sentido como dolor, sino por estar correlacionado con cierta evidencia del laboratorio. En el mismo pasaje, después de sostener Kripke que es posible que existiera el calor sin que se diera la situación que de hecho nos permite fijar la referencia de "calor", esto es, el hecho de que lo experimentamos como caliente, Kripke pregunta: "¿es acaso posible que la estimulación de la fibra C pudiera haber existido sin haberse sentido como dolor?" (p. 151) Nótese que la analogía planteada no es correcta; para que lo fuese, la pregunta tendría que haber sido: ¿es acaso posible que la estimulación de la fibra C pudiera haber existido y que no se la identificase de la manera como de hecho la identificamos? La respuesta que

(ii) En segundo lugar, parece posible que, aun tratándose de identidades teóricas científicas, figuren en ellas descripciones esenciales y no accidentales y, así, que tales identidades estén flanqueadas por dos designadores rígidos. Esto es, dentro de la perspectiva de Kripke, parece perfectamente posible que la ciencia descubra dos o más propiedades esenciales de un mismo objeto, esto es, propiedades sin las cuales el objeto no sería lo que es.⁴¹ Concedamos, por mor del argumento, que es esencial al planeta Saturno, por ejemplo, tener una composición química Q y que también es esencial a Saturno estar a una determinada distancia del sol (pues de no ser así su constitución química no podría ser Q). Sin embargo, parece francamente contingente la identidad "el planeta cuya constitución química es Q = el planeta que está a una distancia d del sol", a pesar de figurar en ella dos designadores rígidos. El paradigma de explicación de Kripke para mostrar que la contingencia es sólo aparente en el caso de ciertas identificaciones teóricas no se aplica en este caso, ya que no parece haber otras descripciones que "fijen la referencia" de las descripciones que figuran en la oración de identidad. Ahora bien, si el rechazo de las identidades psicofísicas se apoya en el hecho de que no se puede explicar y eliminar nuestra intuición de que son contingentes, tendríamos que rechazar, también, algunas identidades teóricas como la mencionada anteriormente cuya aparente contingencia tampoco se puede explicar ni eliminar.

(iii) Kripke podría alegar que lo que sucede en casos como el anterior es que, asociados a las descripciones "el planeta cuya constitución química es Q " y "el planeta que está a una distancia d del sol", hay ciertos enunciados "puramente

⁴¹ Para la noción de propiedad esencial, véase, Kripke, S., (1971), pp. 151 y 152, notas 11 y 12, y (1980), pp. 39 y ss., 110-115.

qualitativos", los cuales sí nos permiten explicar la aparente contingencia de ese tipo de identidades. Esos enunciados puramente cualitativos son descripciones de la evidencia sensorial sobre cuya base establecemos la verdad de "Saturno tiene una constitución química Q" y "Saturno está a una distancia d del sol" y están sólo contingentemente relacionados con las proposiciones que expresan estas oraciones. Pero si esto es así, entonces, también podremos argumentar, en el caso de las identidades psicofísicas, que aun si las descripciones mediante las que fijamos la referencia de ambos términos de la identidad fuesen descripciones esenciales y, por lo tanto, fuesen designadores rígidos, el término "estimulación de las fibras C" que refiere al estado cerebral pertinente, es un término teórico asociado a un conjunto de enunciados puramente cualitativos que describen la evidencia sensorial que usamos para identificarlo y con los cuales está contingentemente relacionado.

He argumentado que la relación entre el término teórico "estimulación de la fibra C" y los enunciados observacionales mediante los cuales fijamos su referencia, es contingente, pues sería posible estar frente a la misma evidencia y no existir el estado cerebral en cuestión; también sería posible identificar el mismo estado cerebral mediante otro tipo de observaciones. Como dice Kripke, sería posible estar "cualitativamente en la misma situación epistemológica", esto es, tener la misma evidencia sensorial, y estar, sin embargo, frente a un tipo de objeto diferente de aquel frente al que de hecho estamos.⁴² Para mostrar que la pretendida contingencia de una identidad teórica es ilusoria, dice Kripke, "nuestro

⁴² Véase, Kripke, S., (1980), p. 142. Hector Neri Castañeda sugiere que no se trataría en este caso de identidades estrictas, sino de lo que en su sistema denomina "consustanciaciones". La relación de consustanciación, de acuerdo con Castañeda, es una relación contingente que se descubre a posteriori. Véase Castañeda, H.N., (1972) esp. pp 51-57.

paradigma general es redescibir cualitativamente (es decir, en términos de experiencias sensoriales) tanto la evidencia anterior como el enunciado mismo y sostener que están relacionados sólo contingentemente".⁴³ Esto sería lo que explicaría en último término la impresión de contingencia. El problema que he planteado a Kripke es, precisamente, que ese tipo de explicación nos permitiría eliminar también la apariencia de contingencia de las identidades psicofísicas.

Cuando afirmamos, por ejemplo, que el dolor es determinado patrón P de actividad neuronal, no queremos implicar que se pueda observar directamente tal patrón, ya que no es algo que esté ahí para que lo miremos; lo único que podemos observar son, tal vez, ciertos registros de actividad eléctrica cerebral, o tal vez ni siquiera esto, sino aparatos cuyos indicadores se mueven y agujas que se deslizan sobre un papel y dibujan líneas. El asunto central en mi argumento es que es totalmente contingente que esas gráficas estén relacionadas con P. Podríamos haber detectado de otras maneras esa misma actividad neuronal o podríamos haber detectado algo diferente mediante la misma evidencia empírica. Sostener esto es distinto de sostener que la relación entre el dolor y P sea contingente. Kripke señala correctamente que no es posible estar frente a la misma situación epistémica que estamos cuando tenemos un dolor (esto es, la situación en la que sentimos un dolor) y decir sin embargo que no hay dolor; sentir un dolor es suficiente para que exista. Sin embargo, no se puede sostener que la evidencia observacional que nos permite detectar un estado cerebral P y el estado P mismo estén necesaria o esencialmente relacionados; como he señalado, sería posible encontrarnos en una situación epistemológicamente idéntica (cualitativamente descrita) en dos mundos posibles m y m' y que en m esa evidencia fuera prueba de la existencia de un

⁴³ Kripke, S., (1980), p. 143; véase también p. 150.

estado cerebral \underline{C} , y en \underline{m} ' fuera evidencia de otro estado cerebral \underline{C}' o de nada. La relación entre tal evidencia y \underline{C} o \underline{C}' sería contingente.

Las identificaciones teóricas difieren de las psicofísicas en que las primeras sólo contienen designadores rígidos de fenómenos físicos (externos, objetivos), en tanto que las segundas afirman la identidad de una experiencia o tipo de experiencia (subjetiva) con un fenómeno físico o tipo de fenómeno físico. La impresión de contingencia de esta segunda clase de identidades proviene del hecho de que seleccionamos el fenómeno físico pertinente mediante evidencias que pueden estar relacionadas sólo contingentemente con dicho fenómeno, esto es, mediante descripciones accidentales. Si esto es así, puede explicarse entonces el carácter ilusorio de la intuición cartesiana: si " $M = F$ " es verdadero, entonces es necesario; la impresión de contingencia deriva del hecho de que la relación de F con la evidencia E (que nos permite fijar la referencia de F) es contingente y, por tanto, es también contingente la relación entre M y E .

Si mi argumento ha sido correcto, se puede concluir que la teoría de la referencia directa y de la designación rígida de los nombres no tiene entre sus consecuencias la falsedad de la teoría de la identidad mente-cuerpo. Tampoco tiene la consecuencia contraria; simplemente es neutral frente a tal asunto. Mi argumento no niega que los términos "dolor" y "estimulación de las fibras C " refieran directamente y designen rígidamente; sólo subraya que la relación entre la expresión "estimulación de las fibras C " y la evidencia en relación con la cual se le introduce en una teoría neurofisiológica, no es una relación necesaria.

CAPÍTULO V

Referencia directa y actitudes proposicionales

Introducción.

La teoría de la referencia directa de los nombres propios se distingue de otras teorías de la referencia por su economía; como su nombre lo indica, concibe la relación referencial como no mediatizada por sentidos.¹ Los nombres propios, según esta teoría, no tienen más función semántica que la de referir; a diferencia de las descripciones definidas, un nombre no caracteriza a su portador de ninguna manera, es decir, no lo describe como el poseedor de propiedades que lo identifiquen y lo distingan de todos los demás individuos, simplemente refiere, está en lugar de su referente. La única contribución semántica de un nombre al

¹ Kripke rechazaría el apelativo de "teoría" para calificar la concepción de los nombres que él ofrece en Kripke, S., (1971), (1980) y (1979a), pues, como él mismo escribe, su rechazo de los sentidos fregeanos y su propuesta de ver la referencia de los nombres de una manera "milliana" no equivale a dar un conjunto de condiciones necesarias y suficientes para la referencia de un nombre. En lo que sigue, no se tome, pues, el término "teoría" en ese sentido fuerte que Kripke rechazaría.

significado de una oración es el objeto designado por él.²

Si la teoría de la referencia directa es correcta, entonces, está comprometida, entre otras cosas,³ con la tesis de que los nombres propios correferenciales son sustituibles en todos los contextos salva veritate; es decir, con la validez irrestricta del Principio de Sustitución (PS).⁴ Pues, si efectivamente la única contribución de un nombre al significado de una oración en la que figura es su referente, y el significado de la oración es una función del significado de sus partes, entonces, si se sustituye en ella un nombre por otro nombre correferencial, el significado de la oración permanecerá idéntico; y si no cambia el significado de la oración (sus condiciones de verdad), no podrá cambiar su valor de verdad. En la versión de la

² Scott Soames, en Soames, S., (1995), cap. 6, caracteriza los términos directamente referenciales como aquellos cuyo "contenido proposicional en un determinado contexto es su referente en ese contexto"; los nombres propios son directamente referenciales en ese sentido. Los demostrativos son los términos directamente referenciales par excellence: en caso de carecer de referente no hacen ninguna contribución al significado de la oración en la que figuran, de modo que la oración en cuestión no expresa una proposición y carece de valor de verdad. Véase, esp. Kaplan, D., (1979).

³ Digo "entre otras cosas", porque también está comprometida con la validez irrestricta de la generalización existencial. Si la contribución de un nombre al significado de la oración en la que figura es sólo el individuo designado, entonces, aparentemente se podrá inferir a partir de, por ejemplo,

(i) María cree que Santaclós es barbón,

la oración

(ii) Hay algún x tal que María cree que x es barbón.

No trataré aquí el tema de la cuantificación en contextos de creencia y modalidad. Para una excelente discusión del tema, véase, Kaplan, D., (1968).

⁴ PS establece que para cualesquiera dos términos t y t' , si $t = t'$ y $O(t)$, entonces $O(t')$ que es una oración que se obtiene a partir de $O(t)$ sustituyendo en ella todas o algunas figuraciones de t por t' .

teoría de la referencia directa que he venido examinando en otros capítulos, y que es la propuesta por Saul Kripke,⁵ los nombres se caracterizan, además, como designadores rígidos, es decir, como términos singulares que, si designan a un individuo x en \underline{m}^* , designarán a x en todos los mundos posibles en donde x exista.⁶

La economía -la "elegancia" dirían los lógicos- es sin lugar a dudas uno de los rasgos que nos hacen preferir una de entre varias teorías. Pero otro rasgo, aún más importante, por el que la adoptamos es su capacidad para resolver problemas especialmente difíciles en relación con el comportamiento de los individuos que constituyen su dominio. En el caso de una teoría de la referencia de los nombres propios, el rasgo fundamental que nos hará inclinarnos a favor de una teoría en particular será su capacidad para dar cuenta del funcionamiento de los nombres en contextos especialmente problemáticos, es decir, su capacidad de explicar fenómenos lingüísticos que intuitivamente todos parecemos reconocer y que son de algún modo enigmáticos. El más notable de entre ellos es, sin lugar a dudas, el fenómeno de la aparente falla de sustituibilidad salva veritate de nombres correferenciales en los contextos llamados "intensionales", esto es, en contextos

⁵ Véase Kripke, S., (1971), (1979 a) y (1980).

⁶ Hay una precisión que hacer: un nombre, según esa concepción, no es sólo una grafía o un sonido, si así fuera sería desde luego posible que en otro mundo posible la misma grafía designara a otro individuo diferente del que designa en \underline{m}^* . No entraré aquí en el problema de dar un criterio de identidad para un nombre, sólo diré que la relación referencial que en \underline{m}^* tiene un nombre N con un individuo x , es decir, la cadena causal que los conecta, forma parte de tal criterio. Por otra parte, un individuo existente en \underline{m}^* puede no existir en \underline{m}_n y, en caso de no existir, no podrá ser designado en \underline{m}_n por ningún designador. Kripke, S., (1980), p. 48.

modales y en contextos de actitudes proposicionales.⁷

En Kripke, S., (1971) y (1980), después de elaborar poderosos argumentos para mostrar que la teoría descriptivista de los nombres nos da una visión equivocada de la manera como se determina la referencia de los nombres propios, Kripke explora las consecuencias que tiene la concepción de los nombres que él propone para la modalidad de las oraciones en las que figuran nombres propios. Argumenta que la rigidez de los nombres hace que la modalidad de una oración se preserve cuando sustituimos en ella un nombre por otro nombre correferencial. Si los nombres son efectivamente designadores rígidos y el único valor semántico de un nombre es su referente, entonces, los nombres correferenciales son sustituibles no sólo salva veritate, sino salva modalitate. Si "Voltaire es Voltaire" expresa una verdad necesaria, si es verdadera en todos los mundos posibles, también la será "Voltaire es François Marie Arouet". Como señala el propio Kripke, si el único valor semántico de un nombre es su portador, "el que una oración exprese una verdad necesaria o contingente depende solamente de la proposición expresada y no de las palabras que se usen para expresarla".⁸ En

⁷ Los contextos modales son oraciones gobernadas por un operador modal tales como "necesariamente" y "posiblemente"; por ejemplo, "necesariamente $2+2 = 4$ ", "posiblemente Voltaire es español". Un contexto de actitud proposicional es una oración compuesta en cuya principal figura un verbo psicológico -tal como, "creer", "desear", "pensar"-, que toma como objeto directo una cláusula subordinada conectada con la principal mediante la partícula "que"; por ejemplo, "María cree que Voltaire es calvo", "Juan desea que Marcos venga a México".

⁸ Kripke, S., (1979). Reimpreso en Salmon N., y Soames S., (comps.), (1988), p. 104. Los números de página en las citas corresponden a esta última edición.

consonancia con lo anterior, y en contra de lo que Quine y otros han afirmado,⁹ sostiene que los nombres propios sí son sustituibles salva veritate en contextos modales. Si la oración "Necesariamente Voltaire es Voltaire" es verdadera, también es verdadera la oración "Necesariamente Voltaire es François Marie Arouet".¹⁰

Sin embargo, como apuntamos antes, si la teoría de la referencia directa y rígida de los nombres es verdadera, los nombres codesignativos tienen que ser sustituibles salva veritate no sólo en los contextos modales, sino de manera irrestricta; es decir, PS tiene que valer en todos los contextos, incluyendo los contextos de actitudes proposicionales. Congruente con esta exigencia, en su artículo "Un enigma sobre la creencia" Kripke se propone, precisamente, hacer una defensa indirecta de PS en contextos de actitudes proposicionales.¹¹ Muestra, y ése es el enigma que plantea, que las mismas consecuencias paradójicas que se obtienen al usar el PS en atribuciones de actitudes proposicionales pueden obtenerse sin usar tal principio, sino usando sólo otros dos principios aparentemente inobjetable: el Principio Desentrecorillador (PD) y el Principio

⁹ El tema de la falla de sustitución de términos correferenciales ha ocupado a Quine desde sus inicios. Sus artículos pioneros sobre el tema son, Quine, W.V.O., (1943), (1947), (1953) y (1956). Eliminar la llamada "opacidad" de los nombres en contextos intensionales es una de las ideas dominantes en la filosofía de Quine, al punto de proponer la eliminación de términos singulares en favor de términos generales. Véase Quine, W.V.O., (1960), esp., cap. 4, 5 y 6.

¹⁰ Recuérdese la distinción que hace Kripke entre la necesidad y la aprioridad. "Voltaire es François Marie Arouet" expresa una verdad (metafísicamente) necesaria cuya verdad es (epistémicamente) a posteriori.

¹¹ Kripke expone claramente su estrategia indirecta al final de la sección II de Kripke, S., (1979a), pp. 117-118.

de Traducción (PT).¹² Si el argumento de Kripke es correcto, entonces, no son concluyentes los argumentos con los que se suele rechazar la validez de PS en contextos de actitudes proposicionales; es decir, PS no puede ser tenido como el responsable de las atribuciones paradójicas de creencias. La supuesta "falla de sustitución" de nombres correferenciales en contextos de creencias y otras actitudes proposicionales sería síntoma de un problema mucho más amplio y profundo, el cual exigiría una respuesta distinta de las que tradicionalmente se han ofrecido para resolver el problema de las atribuciones paradójicas.

Recordemos que lo que conduce a muchos filósofos que han estudiado los contextos de actitudes proposicionales a rechazar la validez de PS en tales contextos es el hecho de que si lo aceptamos podemos llegar a atribuir creencias abiertamente contradictorias a un sujeto aparentemente racional. Juan puede asentir, por ejemplo, a las oraciones "Voltaire es filósofo" y "François Marie Arouet no es filósofo" y esto no parece amenazar en absoluto su racionalidad. Con base en el asentimiento sincero de Juan a ambas oraciones podemos hacerle la siguiente atribución: cree que Voltaire es filósofo y que François Marie Arouet no es filósofo. Pero, dado que Voltaire es François Marie Arouet, si suponemos que PS vale en todos los contextos, nos vemos obligados a atribuirle a Juan la creencia en una franca contradicción; es decir, tendremos que decir que Juan cree que Voltaire es filósofo y que Voltaire no es filósofo. En vista de lo anterior, y dado que por hipótesis Juan es un individuo racional, se suele concluir que los nombres "Voltaire" y "François Marie Arouet" no son sustituibles salva veritate en

¹² Además de PD y PT, Kripke hace uso implícito de la equivalencia T que Tarski introduce en su teoría semántica de la verdad (y que es también un principio que nos permite quitar comillas): "p" es verdadera si y sólo si p. Véase la nota 26 en Kripke, S., (1979).

contextos de creencia; esto es, o bien los nombres no son correferenciales en todos los contextos o simplemente PS no es válido universalmente.¹³

El objetivo central de este último capítulo es examinar el comportamiento de los nombres propios en contextos de creencia. Para ello estudiaré el enigma que plantea Kripke en su artículo "Un enigma sobre la creencia" y exploraré dos posibles respuestas acordes ambas con su concepción de la referencia directa de los nombres. Las posibles respuestas que exploraré, lejos de cuestionar el enfoque sobre los nombres defendido por Kripke, pretenden apoyar su viabilidad. Entender las aparentes fallas de sustitución de nombres correferenciales en contextos de actitudes proposicionales es, en mi opinión, el reto mayor que tiene que enfrentar una teoría de la referencia directa de los nombres propios; por esa razón he reservado este último capítulo para discutir la cuestión. En la sección 1, presento el enigma que plantea Kripke sobre la creencia; en la sección 2 sugiero una primera respuesta a las preguntas que plantea Kripke en su artículo, la cual rechaza la imposibilidad de atribuir creencias contradictorias a sujetos que (como Pierre) asienten a oraciones incompatibles en las que figuran diversos nombres de un mismo individuo; esa primera respuesta, como veremos, encierra dificultades que

¹³ Como se puede ver, el razonamiento anterior no constituye un rechazo obligado de PS en contextos de actitudes proposicionales, pues se puede sostener que aunque los nombres "Voltaire" y "François Marie Arouet" no son sustituibles en tales contextos, eso no invalida PS. Frege razonó de esa manera y sostuvo que en las cláusulas subordinadas de atribuciones de actitudes proposicionales los nombres dejan de referir a sus referentes usuales y pasan a referir a lo que en "contextos directos" (no psicológicos) son sus "sentidos"; así "Voltaire" y "François Marie Arouet" en el contexto de nuestro ejemplo no son correferenciales. Frege es un defensor de PS: tomó la no sustituibilidad de supuestos nombres correferenciales en contextos de actitudes proposicionales como prueba de que en esos contextos los nombres no son realmente correferenciales. Véase, Frege, G., (1892) y Kaplan, D., (1968).

me llevarán a estudiar en la sección 3 ciertas peculiaridades lógicas del predicado "__ cree __". En esa última sección exploraré la propuesta de tratar el predicado de creencia como un predicado "parcialmente definido", abriendo así la posibilidad de considerar que las atribuciones paradójicas de creencias dan lugar a auténticos vacíos de valor de verdad (truth value gaps). La posibilidad de que haya atribuciones de creencia que carezcan de valor de verdad podría tal vez explicar, no sólo las aparentes fallas de sustitución de términos correferenciales en contextos de actitudes proposicionales, sino esas otras atribuciones paradójicas que registra Kripke en su planteamiento del enigma sobre la creencia y para llegar a las cuales no es necesario usar PS.

1. Enigmas kripkeanos.

En "Un enigma sobre la creencia" Kripke introduce dos principios que regulan las atribuciones de creencias y cuya validez, como veremos, parece estar fuera de cualquier duda: el Principio Desentremillador (al que abreviaré "PD") y el Principio de Traducción (abreviado "PT").¹⁴ PD es un principio que conecta el asentimiento sincero a una oración O por parte de un sujeto S con la creencia de S en lo que O expresa. El asentimiento que juega un papel determinante en la atribución de creencias a terceras personas es, desde luego, el asentimiento externo, público, pero PD no excluye -en el caso de autoatribuirse uno mismo una creencia- que el asentimiento pueda ser interno. En este sentido, PD parece aceptable tanto para quienes sostienen una concepción "mentalista" de la creencia

¹⁴ Para la introducción, formulación y justificación de estos principios véase, Kripke, S., (1979), esp. sección II, pp. 112-115.

como para quienes favorecen una concepción conductista de ella.

PD. Si un hablante normal del castellano asiente sincera y reflexivamente a "p", entonces cree que p,

en donde "p" es cualquier oración declarativa del castellano que no contiene pronombres ni deícticos, pero que puede contener nombres propios que sean moneda corriente entre los castellanoparlantes, tales como, "España", "Napoleón", "ONU", "Pánuco", "Londres", etcétera, y en donde la expresión "sincera y reflexivamente" pueden completarse con los adverbios que el lector juzgue necesarios. PD se justifica, primeramente, por el hecho de que aprendemos el concepto de creencia precisamente en conexión con el asentimiento sincero a oraciones declarativas y, luego, porque nuestra práctica normal de atribución de creencias se basa, precisamente, en la aceptación tácita de PD. Una aplicación sencilla de PD sería la siguiente: Juan asiente sincera y decididamente a la oración "Voltaire es filósofo", con base en ese asentimiento le atribuimos la creencia de que Voltaire es filósofo. Hay, desde luego, tantos principios desentrecomilladores como lenguas existen; cada lengua tiene su propio principio desentrecomillador.

Ahora bien, es evidente que, a veces, nosotros los hablantes del castellano hacemos atribuciones de creencias en castellano a personas que hablan otras lenguas, a personas que no asienten a oraciones en castellano, sino a oraciones en otras lenguas (que, desde luego, entendemos), por ejemplo, en francés o en inglés. Así, si veo que Luc asiente sinceramente a la oración francesa "Aix la Chapelle est une ville charmante", puedo, sin temor a equivocarme, hacerle la siguiente atribución:

A. Luc cree que Aquisgrán es una ciudad encantadora.

En efecto, la versión francesa de PD me autoriza a hacer a Luc la atribución en francés:

B. Luc croit que Aix la Chapelle est une ville charmante,

y luego, con el fin de hacerle a Luc una atribución verdadera en castellano, me apoyo en el Principio de Traducción (PT) que se formula de la siguiente manera:

PT. Si una oración en una lengua L expresa una verdad en L, entonces, cualquier traducción de ella a a cualquier otra lengua L' expresará una verdad (en L').

PT puede verse como un análisis parcial de la noción de traducción. Traducir a L' una oración O de L es encontrar una oración O' de L' que signifique lo mismo que O, y si ambas oraciones significan lo mismo (tienen las mismas condiciones de verdad) no podrán diferir en valor de verdad. Lo que me justifica para hacer la atribución A es, precisamente, que es una traducción correcta al castellano de la atribución en francés B.

Una vez armados con esos dos principios, recordemos el enigma que plantea Kripke sobre la creencia y que surge precisamente al aplicar PD y PT al caso hipotético de un individuo llamado Pierre. Pierre es un francés monolingüe que nunca ha salido de Francia, pero que ha mirado con cuidado un album con postales a colores de Londres que muestran bajo un sol radiante el Big Ben, el Parlamento sobre el Támesis, la Catedral de Westminster, el Palacio de Buckingham, Hyde

Park, etcétera. Una vez que ha visto esas fotografías de Londres, Pierre asiente a la oración francesa:

"Londres est jolie".

Con base en ese asentimiento y usando PD y PT le hacemos la siguiente atribución:

(1) Pierre cree que Londres es bonita.

Más tarde, Pierre consigue un empleo a través de amigos que, como él, son franceses monolingües y que al contratarlo lo trasladan a una ciudad que nunca había visitado antes. Trabaja y vive en un barrio pobre y desangelado habitado por gente que habla una lengua diferente del francés y nunca se aventura fuera de su barrio; aprende "de oídas" el idioma local y se percata de que los nativos del lugar llaman "London" a la ciudad en donde viven. Alguien en sus alrededores emite un día la oración inglesa:

"London is not pretty";

Pierre la comprende y asiente a ella sin titubeos. Con base en su asentimiento sincero, y usando nuevamente PD y PT, podemos hacerle la siguiente atribución:

(2) Pierre cree que Londres no es bonita.¹⁵

Podemos imaginar una situación en la que Pierre conversa simultáneamente con dos amigos monolingües, uno francoparlante y otro angloparlante, y que en ese contexto asiente conjuntamente a las oraciones "Londres est jolie" y "London is not pretty".

La pregunta que plantea Kripke es la siguiente: ¿qué cree Pierre sobre Londres?, ¿qué creencia(s) respecto de Londres podemos atribuir a Pierre en esa situación con base en su asentimiento a aquel par de oraciones?

Hay sólo cuatro posibles respuestas a esas interrogantes: (i) cree que Londres es bonita, (ii) cree que Londres no es bonita, (iii) no cree ni que Londres es bonita, ni que Londres no es bonita y (iv) cree que Londres es bonita y que Londres no es bonita. Y Kripke rechaza cada una de esas posibles respuestas mediante poderosos argumentos. Las tres primeras -sostiene- son inaceptables porque no parecen darles el valor que merecen a los hechos que sirven de base para hacer las atribuciones (1) y (2). (iv) es inaceptable por una razón diferente: porque supone que Pierre puede creer una contradicción lógica sin haber cometido ningún error lógico, es decir, sin haber faltado a ningún principio o regla lógicos. Si, a pesar de las razones que da Kripke para no hacer la atribución (iv), se quiere defender la idea de que es legítimo atribuirle a Pierre la creencia de que Londres es bonita y de que Londres no es bonita, se tiene que explicar cómo, sin cometer ningún error lógico, Pierre puede llegar a creer semejante contradicción; es decir, se tiene que explicar cómo es posible llegar a creer una contradicción sin haber

¹⁵ Pierre, desde luego, no ha dejado de aceptar la oración francesa "Londres est jolie" y, en ese sentido, parece no haber abandonado su antigua creencia. Si se le pidiera dar el nombre de una ciudad bonita, podría decir: "Londres".

cometido ninguna falla de racionalidad.

Notemos que cuando en el ejemplo anterior parecemos estar orillados a atribuir a Pierre dos creencias contradictorias no se ha utilizado PS. Sólo se han utilizado PD y PT. Esto, precisamente, es lo que Kripke pretendía mostrar: que PS no es responsable de las consecuencias paradójicas a las que se puede llegar cuando atribuimos creencias; es decir, que podemos llegar a las mismas consecuencias indeseables (vernos orillados a atribuir creencias contradictorias a sujetos racionales como Pierre) cuando nos apoyamos sólo en PD y PT. Si esto es así, entonces los casos de atribuciones paradójicas que generalmente se aducen como prueba contra la validez de PS en contextos de actitudes proposicionales en realidad no logran su propósito. Y si lo anterior es correcto, entonces, las atribuciones paradójicas no le restan verosimilitud a la teoría de la referencia directa de los nombres que -como dije al inicio de este capítulo- parece implicar la validez de PS.

El intento kripkeano de resucitar la validez de PS en contextos de actitudes proposicionales, nos deja frente a un dilema: (i) o bien aceptamos que no es posible contestar de ninguna manera la pregunta acerca de qué cree Pierre, es decir, que no es posible atribuirle ninguna creencia, (ii) o bien contraargumentamos que es posible atribuirle la creencia en una proposición contradictoria. Si tomamos el primer cuerno del dilema, tenemos que explicar por qué, si aparentemente eran correctas las atribuciones (1) y (2) hechas a Pierre, no podemos atribuirle con verdad la creencia en la conjunción problemática. La dificultad que supone dar esa explicación, ha conducido a algunos defensores de la teoría de la referencia directa a preferir el segundo cuerno del dilema y argumentar que, malgré las apariencias, podemos con verdad atribuir a Pierre la creencia en una contradicción.

Las dos siguientes secciones son sendas investigaciones de los dos cuernos del dilema antes expuesto. En la sección 3 exploro la posibilidad de atribuir a Pierre una creencia contradictoria y centro mi explicación en dos cuestiones: por una parte, ciertas similitudes entre las sustituciones de nombres gobernadas por PS y nuestra práctica de traducir gobernada por PT y, por la otra, la distinción entre lo necesariamente falso (lo metafísicamente imposible) y las contradicciones lógicas a priori. La sección 4 es un intento de explicar por qué no se puede atribuir con verdad a Pierre ninguna creencia.

2. Sobre la posibilidad de creer proposiciones contradictorias.

Primeramente, quisiera destacar ciertas similitudes entre PS y PT.

(i) Tanto PS como PT regulan operaciones sobre oraciones que al efectuarse conservan el valor de verdad de la oración original, es decir, operaciones salva veritate. Para que la sustitución de un nombre propio "a" por otro "b" sea salva veritate, PS exige que "b" y "a" sean correferenciales; es decir, exige que sea verdadera la oración de identidad "a = b". PT, por su parte, garantiza que si la traducción de una oración O de una lengua L a una oración O' de otra lengua L' es correcta, entonces es salva veritate; y O' es una traducción correcta de O si y sólo si O y O' significan lo mismo. Ahora bien, si la única contribución de un nombre al significado de una oración en la que figura es su referente y nos permitimos hablar de "lo que significan los nombres", vemos que PS y PT no son tan diferentes como parecen a primera vista: ambos garantizan que si se sustituye en una oración una expresión por otra "co-significativa" no se altera el valor de verdad de la oración original. Ambos principios, parecen seguirse de la tesis según la cual el significado de una oración está dado por sus condiciones de verdad, las

cuales, a su vez, son una función de las contribuciones semánticas que hacen al significado de una oración las palabras que la constituyen, así como de la idea bastante obvia de que el valor de verdad de una oración en un contexto depende (en parte) de esas condiciones.

(ii) En el mismo orden de ideas, cabe notar que la práctica de traducción es en cierto sentido una práctica de sustitución: cuando traducimos sustituimos unas oraciones por otras que significan lo mismo. Esto podría llevarnos a sospechar que Kripke realmente no ha demostrado lo que se proponía, a saber, que podemos llegar a atribuir creencias contradictorias a un sujeto racional aun cuando no usemos PS en el proceso de atribución; después de todo, traducir un nombre "a" de L por un nombre "b" de L' parece equivalente justamente a un caso de sustitución de idénticos. Nótese, también, que para poder traducir correctamente un nombre de L por un nombre de L' tenemos que conocer (de la misma manera que cuando sustituimos nombres correferenciales) la verdad de una identidad: puedo traducir "London" por "Londres" si y sólo si sé que "London = Londres" es verdadera, en otras palabras, si y sólo si sé que London es Londres.¹⁶ Sin embargo, podemos decir algo a favor de distinguir estos dos principios: PT, a

¹⁶ Adviértase que esta oración de identidad es un híbrido del inglés y el castellano. Parte de la conclusión del argumento de Kripke es que no puede darse una respuesta correcta en una sola lengua a la pregunta sobre lo que cree Pierre; es decir, que los recursos de una lengua son insuficientes para expresar su creencia. Desde luego, Kripke admite que se puede describir la confusión de Pierre atribuyéndole una creencia metalingüística, es decir, una creencia acerca de los términos del lenguaje que usa (por ejemplo, puede decirse que cree que "Londres" y "London" nombran ciudades diferentes y que una es bonita y la otra no), lo que niega es que, una vez que Pierre vive en Londres, se le pueda atribuir alguna creencia sobre Londres, es decir, una creencia para atribuir la cual usemos una oración subordinada en la que se use el nombre de la ciudad y no sólo se le mencione.

diferencia de PS, parece inobjetable en el sentido de limitarse a expresar una condición necesaria para la traducción; toda traducción, en efecto, tiene que ser salva veritate. Más aún, la aceptación de PT se halla a la base de una práctica universal cuya legitimidad nadie pone en duda: todos aceptamos que la verdad se preserva al traducir de una lengua a otra, incluso cuando atribuimos creencias (afirmaciones, deseos, intenciones, etcétera) a personas que hablan una lengua extranjera L' usando para ello oraciones de nuestra propia lengua L. Esta última práctica se considera correcta si la oración de L que figura como cláusula subordinada en la atribución de creencia tiene las mismas condiciones de verdad (el mismo significado) y, por ende, el mismo valor de verdad, que la oración de L' a la que el hablante al que se hace la atribución asintió previamente. PS, en cambio, es un principio que introducen los lógicos con base en consideraciones teóricas en relación con la noción de identidad y la naturaleza semántica de las constantes individuales y de los nombres.¹⁷

(iii) Es importante notar que en la formulación de PS y PT no se hace explícito que las prácticas que ellos gobiernan supongan la posesión de ciertos conocimientos empíricos por parte de quienes las llevan a cabo.¹⁸ Podemos

¹⁷ Quine, en Quine, W.V.O, (1953a), señala que PS y el principio leibniziano de la indiscernibilidad de los idénticos, esquemáticamente formulado $(x)(y) ((x=y \& Fx) \rightarrow Fy)$, son de hecho uno y el mismo, véase p.139. Sin embargo, cabe mencionar una diferencia importante: PS es un principio cuya validez presupone la existencia de algún lenguaje; el principio de indiscernibilidad de los idénticos, en cambio, es un principio metafísico cuya validez es independiente de la existencia de cualquier lenguaje.

¹⁸ Lo anterior es correcto para la práctica de traducir de una lengua pública a otra y la de sustituir un nombre propio ordinario por otro codesignativo. En matemáticas, sin embargo, la sustitución de un término por otro correferencial puede no requerir ningún conocimiento empírico; es posible saber a priori que

formular PS de la siguiente manera:

PS. Para cualesquiera dos términos singulares t y t' , si " $t=t'$ " expresa una verdad, entonces, si $O(t)$ es una oración verdadera, también será verdadera la oración $O(t')$ que se obtiene sustituyendo en $O(t)$ todas las figuraciones de t por t' .¹⁹

Aun cuando se acepte la validez a priori de un principio lógico tal como PS, hay que admitir que la verdad de la oración de identidad que aparece en el antecedente y que es indispensable para poder sustituir nombres propios correferenciales en un razonamiento particular, puede no ser ella misma conocida a priori; es decir, la aplicación de PS en una inferencia particular puede requerir que quien lo aplique posea una información empírica concreta; en el caso de Pierre, por ejemplo, que sepa que "London" y "Londres" designan la misma ciudad, de modo que pueda hacer la sustitución que autoriza PS. Pero esto sólo puede saberlo si conoce un hecho empírico: que las cadenas causales mediante las que estos nombres han

$7+5 = 12$ y, así, llevar a cabo la sustitución en, por ejemplo, "Los meses del año son 12", sin necesidad de apoyarse en ningún conocimiento empírico. Pero, en aquellos casos en que existen dos nombres para un individuo -como "Voltaire" y "François Marie Arouet"-, si no somos los introductores de uno de esos nombres, sólo podemos saber empíricamente que la identidad en la que figuran ambos nombres es verdadera.

¹⁹ Carnap expresa de la siguiente manera el principio de sustituibilidad: "Supóngase que "...a..." es una oración que contiene una o más figuraciones de "a". Supóngase que "...b..." es una oración que se obtiene a partir de "...a..." reemplazando "a" por "b" en una o más figuraciones de "a". Entonces, "...b..." está L-implicada por "a = b" y "...a..."." Carnap, R., (1958), p. 70. Como se puede apreciar tanto esta formulación de PS como la que ofrezco en el texto suponen la existencia de un lenguaje.

llegado hasta el propio hablante remiten en último término a una misma ciudad, o que de acuerdo con ciertas convenciones lingüísticas del inglés y del francés ambos nombres refieren a lo mismo.²⁰ De la misma manera, aun cuando PT sea analítico y valga a priori, la práctica de traducción gobernada por PT requiere que el sujeto que traduce de L a L' conozca ambas lenguas, conocimiento que, también, sólo puede adquirirse empíricamente.²¹ En suma: aun cuando PS y PT sean lógicamente verdaderos y aceptemos su verdad a priori, ambos requieren para su aplicación en inferencias concretas que quien los use conozca ciertas convenciones lingüísticas o ciertas cadenas causales que sólo pueden aprenderse mediante la observación y la experiencia.²²

Si lo anterior es correcto, entonces, un sujeto como Pierre en el ejemplo de Kripke que no tiene la información empírica requerida, que no sabe que "Londres" y "London" refieren a la misma ciudad, obviamente no podrá llevar a cabo la sustitución que autoriza PS, y su carencia de información aparentemente lo llevará

²⁰ El hablante tiene que saber que ambos nombres tienen el mismo referente; mi señalamiento es que tal conocimiento sólo puede obtenerse empíricamente. Menciono aquí el conocimiento de las cadenas causales sólo para señalar un tipo de conocimiento empírico que podría bastar para saber que una identidad es verdadera.

²¹ Tal conocimiento empírico sólo es necesario, desde luego, cuando L y L' son ambas "lenguas públicas". Si el traductor traduce de una lengua L a una lengua L' inventada por él mismo, no requerirá, desde luego, ningún conocimiento empírico adicional.

²² En este sentido, PS y PT no son diferentes de otros principios lógicos válidos a priori, tales como las reglas de inferencia Modus Ponens y Modus Tolens; éstas también pueden requerir para su aplicación correcta en contextos de conocimiento que quien los aplique posea ciertos conocimientos empíricos que hagan las veces de premisas en su inferencia.

a sostener creencias incompatibles, contradictorias, en el sentido de que no pueden ser conjuntamente verdaderas. Es decir, Pierre, en el ejemplo de Kripke, asiente a un par de oraciones que tomadas conjuntamente expresan una imposibilidad, algo falso en todos los mundos posibles. Su ignorancia de la identidad pertinente le impide, a la vez, percatarse de la imposibilidad que cree, esto es, no le permite derivar a partir del par de oraciones a las que asiente una contradicción lógica flagrante de la forma Fa y -Fa cuya falsedad él descubriría a priori y que, desde luego, rechazaría.

En lo anterior está implícita la tesis de que no todo lo que es necesariamente falso (no toda proposición metafísicamente imposible) se expresa mediante oraciones cuya forma superficial haga "evidente" a cualquier hablante el carácter contradictorio de lo que expresa y, así, le permita reconocer a priori su necesaria falsedad. La posibilidad de que oraciones necesariamente falsas no tengan una forma superficial de contradicción lógica explica el hecho de que, a veces, asintamos a oraciones que describen estados de cosas imposibles o a combinaciones inconsistentes de oraciones debido a que no podemos descubrir a priori su inconsistencia. No todas las imposibilidades metafísicas pueden reconocerse a priori, es decir, a partir de la mera estructura superficial de las oraciones que las expresan. "Voltaire es filósofo y François Marie Arouet no lo es" (asignando a "Voltaire" y "François Marie Arouet", desde luego, el referente usual) no tiene la forma superficial de una contradicción lógica, y eso hace imposible descubrir a priori su falsedad;²³ sin embargo, es necesariamente falsa, expresa algo

²³ Cuando digo que es imposible descubrir su falsedad a priori, no quiero dar a entender que el valor semántico de "Voltaire" y "François Marie Arouet" sea diferente, sino que aunque tengan el mismo valor semántico esto sólo puede saberse empíricamente. Si quisiera ser más cuidadosa, tendría que decir dos cosas:

imposible: que uno y el mismo individuo tiene una propiedad y no la tiene. Lo mismo sucede con el par de oraciones, "Londres est jolie" y "London is not pretty", a las que asiente Pierre; su conjunción es falsa en todos los mundos posibles, pero su falsedad necesaria no es evidente a partir de la forma superficial de la conjunción de las oraciones a las que Pierre asiente; esto es, su necesaria falsedad no puede ser reconocida por un sujeto que, como Pierre, carezca de la información empírica relevante. Por otra parte, el hecho de que Pierre asienta a esas oraciones conjuntamente no hace posible la conjunción que expresan.²⁴ Distingamos, pues, dos tipos de oraciones: las que expresan proposiciones metafísicamente imposibles (necesariamente falsas) y las que expresan contradicciones lógicas que un hablante que las entienda pueda reconocer a priori:

(i) que ninguna proposición parece ser en sí misma a priori, pues sólo será a priori relativamente a una oración que la expresa y a un sujeto que entienda dicha oración; (ii) tal vez la mera forma superficial de una oración no baste en ningún caso para reconocer que la oración expresa una verdad lógica; adviértase que incluso una oración de la forma "a es a" ("Voltaire es Voltaire"), en ciertos contextos, puede no expresar una verdad, no se diga ya una verdad necesaria. Sin embargo, en la mayoría de los casos, la forma de una oración que expresa una verdad necesaria (o una contradicción) es un claro indicio de que la oración expresa tal clase de verdad.

²⁴ La idea de que tiene que ser posible todo aquello a lo que un sujeto asiente (o aquello que cree), está relacionada con el llamado "principio de caridad" que recomienda interpretar a cualquier sujeto como si no aceptara contradicciones. Ruth Barcan Marcus, en Marcus, R. B., (1981) y (1983), pone como una condición necesaria para que P sea objeto de creencia para un sujeto que P sea posible. Un problema con esa propuesta es que, si las proposiciones de la lógica y la matemática son necesarias, resultaría imposible creer proposiciones lógicas o matemáticas falsas, por ejemplo, que 729 es un número primo o que no es el caso que $((P \rightarrow Q) \vee (Q \rightarrow R))$.

(MI) Una oración O expresa una proposición metafísicamente imposible si y sólo si O es falsa en todos los mundos posibles, es decir, si y sólo si O es necesariamente falsa.

(CLAP) Una oración O expresa una contradicción lógica a priori si y sólo si un sujeto que la entienda correctamente puede descubrir a priori, sin ningún conocimiento empírico adicional, que O expresa una proposición falsa en todos los mundos posibles.

Como se desprende de las definiciones anteriores, todas las contradicciones lógicas que un sujeto reconoce a priori expresan proposiciones necesariamente falsas, pero no a la inversa. Esto es, el conjunto de oraciones que expresan contradicciones lógicas a priori es un subconjunto propio de las que expresan proposiciones metafísicamente imposibles (necesariamente falsas). La oración "Voltaire es filósofo y François Marie Arouet no lo es" es MI, pero no CLAP. En cambio, la oración "Voltaire es filósofo y Voltaire no es filósofo", es, al menos para mí, a priori y, por lo tanto es un caso de MI y de CLAP.²⁵ Decir que una oración es MI es asignarle una categoría metafísica a la proposición que expresa, decir que

²⁵ Como he señalado antes, alguien podría insistir en que podemos usar a veces una oración CLAP tal como "Juárez es hermoso y Juárez no es hermoso" de manera que no exprese una contradicción lógica. Lo anterior está fuera de toda duda, pero también parece estar fuera de duda que la mera forma de la oración es indicio de que expresa una contradicción: si escuchamos una oración como la anterior reaccionamos naturalmente rechazándola por falsa a menos de recibir una aclaración; si no hay tal aclaración, concluimos justificadamente que la proposición que expresa es necesariamente falsa. Véase la nota 23.

es CLAP es asignarle, además de una categoría metafísica, una característica epistemológica a dicha proposición.²⁶

Lo que hasta aquí he sostenido es que hay oraciones MI cuya falsedad no es posible descubrir por medios puramente a priori, es decir, que no son CLAP, y he sugerido que tales oraciones pueden expresar el contenido de la creencia de un sujeto. El par de oraciones que expresan las creencias de Pierre sobre Londres son conjuntamente MI. Nótese que Pierre sólo podría llegar a saber que lo que cree es una imposibilidad si fuera capaz de poner en la forma de una contradicción a priori las oraciones que expresan su par de creencias, es decir, si pudiera transformar la conjunción MI que expresa su creencia en una conjunción CLAP; en ese caso obviamente abandonaría uno de los conyuntos, o los dos. No reconoce que lo que cree es necesariamente falso porque, como señalé antes, la conjunción de las oraciones a las que asiente no revisten la forma superficial de una contradicción lógica, es decir, porque su falsedad necesaria no puede ser reconocida a priori por Pierre.

Comparemos la situación en la que se encuentra Pierre con la del intérprete hispanoparlante (y políglota) que, con base en el asentimiento de Pierre a las dos

²⁶ La aprioricidad, dice Nathan Salmon, es una propiedad de proposiciones y sólo derivativamente se predica de oraciones; véase Salmon, N., (1986) Apéndice B. Sin embargo, en concordancia con lo dicho antes, considero que una proposición es a priori sólo relativamente a una oración que la expresa y a un sujeto que entiende la oración. "Voltaire es Voltaire" y "Voltaire es François Marie Arouet" expresan la misma proposición, pero la proposición es verdadera a priori para mí sólo relativamente a la primera de esas dos oraciones y en virtud de que la entiendo con "Voltaire" en ambas figuraciones refiriéndose al mismo individuo. Nuestro acceso epistémico a las proposiciones son las oraciones que las expresan y el carácter a priori o a posteriori de una proposición parece ser relativo al acceso epistémico que tengamos a ella.

oraciones problemáticas, le atribuye el par de creencias incompatibles. El intérprete sabe empíricamente que la traducción correcta al español de "London" y "Londres" (pronunciado en francés 'Londr') es "Londres"; sabe, por lo tanto, que London = Londres y que Londres [Londr] = Londres. Al traducir correctamente al español ambas oraciones a las que asiente Pierre, el conocimiento empírico que tiene el intérprete le permite expresar la imposibilidad que cree Pierre como una contradicción lógica que él sabe a priori, a partir de su mera estructura superficial, que es necesariamente falsa.

Si la anterior argumentación es correcta, afirmar que Pierre cree una contradicción -es decir, una imposibilidad metafísica- no implica "condenarlo a la irracionalidad".²⁷ Cree una conjunción de proposiciones necesariamente falsa; en este sentido, cree una contradicción, pero bajo la forma de una MI, no de una CLAP.

El argumento anterior pretende defender la posibilidad de atribuir a Pierre, en contra de la conclusión de Kripke, una creencia en una proposición necesariamente falsa. Pero, ¿cuál es la atribución concreta que se le puede hacer a Pierre? Según lo antes expuesto podremos atribuir a Pierre algo así como lo siguiente: cree que Londres [Londr] es bonita y que London no es bonita.²⁸ ¿Hay algo inadecuado en esta atribución?

²⁷ Según lo anterior, la racionalidad de un sujeto parecería tener que ver, no sólo con el contenido de sus creencias, sino también con la forma como expresa lo que cree; es decir, de acuerdo con lo dicho en el texto, un sujeto sería irracional sólo si creyera una proposición necesariamente falsa expresada mediante una oración cuya forma fuera lógicamente contradictoria.

²⁸ Nótese que no podemos traducir al castellano los nombres extranjeros de Londres que figuran en esta atribución, pues, si lo hiciéramos, le atribuiríamos a Pierre una creencia en una contradicción flagrante.

Creo que encierra tres serios problemas.

(1) Primeramente, si abrazamos la teoría de la referencia directa de los nombres, la anterior atribución a Pierre nos regresa al problema original; ya que si el único valor semántico de un nombre es su referente, podremos atribuirle a Pierre su creencia necesariamente falsa de dos maneras distintas:

(i) cree que Londres es bonita y cree que London no es bonita.

(ii) cree que Londres es bonita y cree que Londres no es bonita,

pues, según la teoría de la referencia directa de los nombres, (i) y (ii) son equivalentes, dicen lo mismo, atribuyen exactamente el mismo par de creencias a Pierre.²⁹ De modo que si el defensor de la referencia directa de los nombres acepta la atribución (i) parece que tiene que aceptar también la atribución (ii). Dicho de otra manera, la estrategia sugerida en la respuesta al enigma de Kripke que he explorado y que consiste en defender que se le puede hacer la atribución a Pierre usando una oración MI, pero no CLAP, no exime al defensor de la referencia directa de los nombres de tener que atribuir a Pierre su creencia mediante una oración CLAP.³⁰

(2) Creo que un intérprete ordinario que entienda el francés y el inglés,

²⁹ Véase el argumento de Scott Soames a favor de atribuir la creencia en una franca contradicción a quienes están en situaciones epistémicas y lingüísticas semejantes a la de Pierre en Soames, S., (1989).

³⁰ Hay por lo menos dos defensores de la teoría de la referencia directa que argumentan que es correcto atribuir a Pierre una creencia en una contradicción flagrante. Véase Nathan Salmon, (1986), (1989a) y (1989b). Véase también Scott Soames, (1989). Kripke, sin embargo, mucho más cauteloso, se abstiene de extrarar semejante conclusión.

hable castellano y tenga el conocimiento empírico relevante, pero carezca de motivaciones filosóficas, no estará de ninguna manera inclinado a hacer la atribución (i)³¹, sería insólito que lo hiciera; pero, desde luego, tampoco estará dispuesto a usar (ii), dado que contiene como cláusula subordinada una CLAP (que por hipótesis Pierre, que es un sujeto racional, rechazaría). Es probable que el intérprete atribuya a Pierre sólo una creencia metalingüística, por ejemplo, que cree que los nombres "Londres" y "London" refieren a dos ciudades distintas y que una de ellas es bonita y la otra no. Dado lo confundente de la situación de Pierre, es probable que simplemente se abstenga de atribuirle cualquier creencia sobre Londres, lo cual, como podemos apreciar, coincide exactamente con el veredicto de Kripke.³²

(3) Las atribuciones (i) y (ii), presentan una dificultad central. La dificultad que percibimos al hacer a Pierre la atribución (ii) tiene que ver con el hecho de

³¹ Obsérvese que el hacer desde el español la atribución (i) supone usar de manera ineliminable una ensalada de tres lenguas. En efecto, si no se puede traducir alguno de los dos nombres, no se puede traducir ninguno. ¿Por qué habríamos de traducir "Londres", y no "London", como "Londres"? Haciendo a un lado el carácter extravagante de la atribución (i), nótese que si un hablante h sabe que los nombres usados son correferenciales y, por las razones que sea, usa (i) para hacer la atribución a Pierre, no podrá evitar atribuirle una creencia en una proposición que desde su propia perspectiva es una contradicción flagrante, pues si bien es posible que Pierre tenga dos "ideas" de Londres correspondientes a cada uno de los dos nombres que usa, el intérprete sólo tiene una. Véase más adelante la dificultad (3).

³² Recuérdesse que el enigma en opinión de Kripke radica en el hecho de que parece inadecuado atribuir a Pierre cualquier creencia sobre Londres, a pesar de que antes de estar en la situación enigmática en la que se encuentra contábamos con evidencia suficiente para atribuirle la creencia de que Londres es bonita, y a pesar de exhibir la misma conducta que nos basta en el caso de otros sujetos angloparlantes para atribuirles con verdad la creencia de que Londres no es bonita.

que la creencia es una actitud proposicional que resulta adecuada frente a lo tenido por verdadero, y ni Pierre ni ningún sujeto mínimamente racional tomaría por verdadera la oración subordinada usada en la atribución (ii). La dificultad que sentimos al hacer la atribución (i) tiene que ver con dos cuestiones: primero, como señalé antes, el hecho de que supone usar de manera ineliminable una oración híbrida del español, el francés y el inglés, y resulta sumamente extraño que no baste una sola lengua para atribuir una creencia a un sujeto; segundo, dado que el intérprete sí sabe que "Londres" y "London" denotan una misma ciudad, es decir, dado que sabe que ambos nombres tienen exactamente el mismo valor semántico, él no puede dejar de entender como una contradicción flagrante la oración que él mismo profiere al hacer la atribución (i). Resumiendo, el problema parece ser el siguiente: sea que usemos (i) o (ii), nuestra atribución tendrá un cierto aire paradójico: o atribuimos la creencia usando en nuestra atribución una oración falsa a priori o atribuimos la creencia usando una oración ineliminablemente híbrida,³³ que aunque no sea falsa a priori, en el sentido de no ser una CLAP cuya falsedad cualquier hablante podría reconocer a priori, es algo semejante: expresa una proposición que quien hace la atribución sabe a priori, por el mero hecho de comprenderla, que es necesariamente falsa. Ambas atribuciones entran en conflicto con lo que mencioné antes: que la creencia es la actitud adecuada frente a lo tenido por verdadero. ¿Cómo podemos atribuir a un sujeto una creencia en una proposición que sabemos a priori que es necesariamente falsa?

Las dificultades mencionadas no han solido ser destacadas por quienes han respondido al enigma de Kripke según los lineamientos antes señalados, es decir, por quienes defienden que es correcto atribuir a Pierre la creencia en una

³³ Híbrida en tanto que es una mezcla de varias lenguas.

contradicción. Creo, sin embargo, que la discrepancia entre ese tipo de respuesta y nuestras intuiciones acerca de lo que es la creencia, nos conducirán tarde o temprano a tratar de responder la pregunta acerca de qué creencias se le pueden atribuir a un sujeto como Pierre de una manera que resulte intelectualmente más satisfactoria. En la sección siguiente exploro la posibilidad de ofrecer una respuesta totalmente diferente al enigma planteado por Kripke.

3. Predicados parcialmente definidos.

En su trabajo sobre la verdad,³⁴ Kripke defiende la tesis de que el predicado de verdad es un predicado monádico, parcialmente definido que, como tal, permite vacíos de verdad. Un predicado $P(x)$ está parcialmente definido si, dado un dominio no vacío D , $P(x)$ se interpreta mediante un par de conjuntos disyuntos (S_1, S_2) , en donde S_1 constituye la extensión de $P(x)$ y S_2 es la antiextensión de $P(x)$ (es decir, el conjunto de individuos en D a los que el predicado no se aplica), pero en donde algunos elementos en D no están ni en la extensión del predicado ni en su antiextensión. Se dice que $P(x)$ es verdadero de los elementos en S_1 , falso de los elementos en S_2 , e indefinido³⁵ para aquellos que

³⁴ Véase Kripke, (1975).

³⁵ Decir que el predicado es indefinido para algunos individuos en D , no equivale a aceptar un tercer valor de verdad para las oraciones en las que el predicado se aplica a esos individuos; en este sentido la introducción de predicados parcialmente definidos no supone el rechazo de la bivalencia. Hay que distinguir entre carecer de cualquier valor de verdad y tener un valor de verdad diferente de la verdad y la falsedad.

están en \underline{D} pero que no están en $\underline{S}_1 \cup \underline{S}_2$.³⁶ Scott Soames caracteriza los predicados parcialmente definidos como aquellos predicados que, para ciertas cosas, tanto la afirmación de que el predicado se aplica a esas cosas como la afirmación de que no se aplica, tienen que rechazarse.³⁷ La idea clave del modelo que propone Soames para tratar este tipo de predicados, es la de que un predicado puede introducirse en el lenguaje dando sólo condiciones suficientes (no necesarias) para su aplicación correcta y condiciones suficientes para su no aplicación. Tales condiciones habrán de ser mutuamente excluyentes, pero no conjuntamente exhaustivas, de modo que habrá algunos objetos respecto de los cuales no se cumplen ni las condiciones suficientes para aplicar el predicado, ni las condiciones suficientes para no aplicarlo.

Las nociones de "aplicarse un predicado a un objeto" y de "ser verdadero" están esencialmente relacionadas. Si un predicado \underline{F} se aplica a un objeto \underline{x} , entonces, la oración $\underline{F}\underline{x}$ en la que se hace la predicación correspondiente es una oración verdadera. De manera similar, si \underline{F} no se aplica a \underline{x} , la oración $\underline{F}\underline{x}$ es falsa. Las oraciones en las que figuran predicados parcialmente definidos serán verdaderas cuando el predicado se aplique a cosas que satisfacen las condiciones suficientes para la aplicación, falsas cuando se apliquen a otras cosas que satisfacen las condiciones suficientes para la no aplicación; pero no serán ni verdaderas ni falsas, cuando el predicado se predique de cosas que no satisfagan las condiciones suficientes ni para su aplicación ni para su no aplicación. La aceptación de predicados parcialmente definidos no supone el rechazo de la bivalencia en el

³⁶ Véase, Kripke, (1975), esp. sección III. Véase también, Kleene, S.C., (1971), § 64.

³⁷ Véase el cap. 6 de su manuscrito Truth and Meaning, inédito (ms. de 1995 en la biblioteca del Instituto de Investigaciones Filosóficas, UNAM).

sentido de que no nos obliga a introducir un "tercer valor" veritativo, sólo abre la posibilidad de contar con genuinos vacíos de valor de verdad. Esto es, podremos tener oraciones (gramaticalmente bien formadas) en las que figure el predicado en cuestión tales que ni ellas ni sus negaciones tengan un valor de verdad.

Parece innegable que hay predicados parcialmente definidos en el lenguaje natural. La razón por la que contamos con este tipo de predicados parece ser básicamente la siguiente: al introducirlos, nuestro interés es hacer ciertas distinciones burdas entre dos grupos de individuos. Predicados tales como "ser calvo", "ser rojo", "ser generoso" "ser democrático" son parcialmente definidos. Hay casos claros en los que el predicado "es calvo" ("rojo", "generoso", "democrático") se aplica y en ese caso la predicación es verdadera; otros casos en los que claramente no se aplica y en esos casos la predicación es falsa; y otros más respecto de los cuales las reglas con las que contamos para aplicar o no aplicar el predicado simplemente no dicen nada, se quedan silenciosas. En esos casos, el predicado ni se aplica ni no se aplica; son casos en los que se produce un vacío de valor de verdad dado que la oración resultante no es verdadera, ni es falsa. Los predicados vagos son un subconjunto de los predicados parcialmente definidos.³⁸

Mi propuesta en esta sección es explorar la posibilidad de tratar el predicado

³⁸ No todos están de acuerdo en tratar a los predicados vagos como parcialmente definidos. Tim Williamson, en Williamson, T., (1994), sostiene que para todo predicado supuestamente vago F y toda cosa x o bien F se aplica a x o bien no se aplica a x , lo cual equivale a sostener que no hay oraciones con predicados vagos cuyo valor de verdad sea indefinido. La vaguedad que percibimos en ciertos predicados no es, dice Williamson, más que el producto de limitaciones epistémicas. Otros, en cambio, caracterizan los predicados vagos como un tipo bastante común de predicados parcialmente definidos, véase Soames, S., (1995), cap. 7.

"__cree que__" como un predicado parcialmente definido y usar nuestras conclusiones para responder al enigma planteado por Kripke. La motivación para considerar la posibilidad de que el predicado "__cree__" sea parcialmente definido no es un mero artificio ad hoc para tratar el enigma, más bien proviene de la intuición de que, como tantos otros predicados, hay casos en los que es claro que el predicado se aplica, otros en los que es claro que no se aplica y otros más en los que parece posible que las reglas mediante las que el predicado se introdujo en el lenguaje pudieran simplemente no arrojar ninguna respuesta: tal vez no nos permiten ni atribuir con verdad una creencia ni negar su atribución. El propio Kripke, al plantear el enigma, parece sugerir lo anterior cuando argumenta que realmente no podemos atribuirle a Pierre ninguna creencia (ni positiva, ni negativa) respecto de Londres y su hermosura una vez que Pierre llega a vivir a Londres.³⁹ Si no es legítimo en esa situación aplicar a Pierre el predicado "cree que Londres es bonita", ni negar su aplicación, tal vez sea porque ese predicado se halla parcialmente definido y cuando tratamos de aplicarlo a casos como el de Pierre se produce un vacío de valor veritativo.

Al final de la sección anterior señalé que hay algo extraño en las atribuciones que llamé (i) y (ii), en el siguiente sentido: cualquiera que sepa que la oración usada para atribuir una creencia es necesariamente falsa considerará inadecuado hacer tanto la atribución (i) como la (ii). El hecho de que la creencia sea la actitud adecuada frente a lo verdadero, nos plantea un problema cuando nos vemos en la situación de tener que atribuir a un sujeto una creencia en una proposición que nosotros sabemos a priori que no puede ser verdadera, es decir, que es necesariamente falsa.

³⁹ Véase, Kripke, (1979 a), sección III.

La dificultad anterior podría estar conectada con las reglas o convenciones que existen en el lenguaje para atribuir con verdad una creencia, es decir, con las condiciones suficientes para que el predicado " cree " se aplique a un par compuesto de una persona y una proposición y, así, dé lugar a una predicación verdadera. ¿Cuáles podrían ser esas condiciones? Parece que tendrán que ver con el asentimiento razonado a una oración que el sujeto comprende y que expresa la proposición creída y, desde luego, con la posibilidad de que la oración pueda ser tenida por verdadera en el contexto.⁴⁰ Consideremos las condiciones de aplicación C1 a continuación:

C1. Si (i) x entiende una oración O, la cual relativamente al contexto en el que se hace la atribución expresa la proposición P, (ii) O no tiene la forma de una contradicción lógica a priori ni contradice a priori ninguna otra O' aceptada por x en el contexto, y (iii) apoyándose en razones pertinentes, x asiente sinceramente a O, entonces, x cree P y la oración "x cree que O" es verdadera en el contexto.⁴¹

⁴⁰ Ruth Barcan Marcus en su artículo "A Proposed Solution to a Puzzle about Belief", en French, p., et alt., (comps.) Midwest Studies in Philosophy, vol. vi, 1981, señala atinadamente la conexión que hay entre la creencia y la posibilidad; sin embargo, yerra al exigir que la proposición expresada por la oración subordinada en una atribución de creencia sea posible; lo único que se requiere, en mi opinión, es que el sujeto que cree que P considere que P es posible de modo que pueda tener por verdadera la oración que la expresa. Véase, Valdés, M.M. (1993), sección III.

⁴¹ Donde se dice "asiente sinceramente a O" podríamos también decir "da muestras claras en su conducta verbal y no verbal de que acepta O". No hay aquí ninguna razón especial para considerar que haya que privilegiar sólo la respuesta verbal, en ocasiones la conducta no verbal puede ser tan buena base para atribuir una creencia como la verbal.

C1 tiene que tomarse como una condición conjunta. Cualquier situación en la que el par $\langle x, O \rangle$ satisfaga C1 es una situación en la que x cree P y en la que la predicación " x cree que O " es verdadera. La condición relativa al entendimiento y al asentimiento razonado no parece requerir de mayor justificación; pensemos, si se nos pregunta "¿en qué situación podemos decir con verdad que x cree P ?" no dudaríamos en responder que cuando x entiende una oración que expresa P y, aduciendo razones pertinentes, x asiente sinceramente a ella. La condición de que O no sea una contradicción cuya necesaria falsedad x pueda descubrir a priori, y no contradiga a priori ninguna otra O' que x acepte en el contexto, puede justificarse alegando que las actitudes epistémicas de x se rigen por un principio de racionalidad mínima y señalando que el respeto a tal principio es una condición necesaria para que pueda haber un asentimiento razonado. En efecto, si x entiende O , y O es contradictoria a priori, x no podrá tener por verdadera a O y, por lo tanto, no asentirá a ella de manera reflexiva. Sólo se asiente sincera y razonadamente a lo que se toma por verdadero y si O expresa una proposición que x reconoce a priori como necesariamente falsa, x no podrá tomarla por verdadera y no asentirá a O sinceramente una vez que reflexione.

Por otra parte, dado que la creencia es un estado mental -una actitud proposicional-, parece concebible que un sujeto pueda tener una creencia y no la manifieste en su conducta, esto es, que haya estados mentales que no se manifiesten públicamente, por lo menos por un tiempo. Aún cuando la conducta verbal y no verbal sea indispensable para introducir en el lenguaje un predicado de actitud proposicional tal como "__cree__", una vez introducido, resulta perfectamente concebible que alguien tenga una creencia que no manifieste en su

conducta durante un período.⁴² Si esto es así, entonces, podemos ver que C1 nos da condiciones suficientes para la aplicación del predicado "__cree__", mas no condiciones necesarias; esto es, la comprensión de x de una oración O que expresa una proposición P, el asentimiento sincero y razonado de x a O y el carácter no contradictorio a priori de ésta última, bastan para que sea el caso que x crea que P y, así, bastan para hacer verdadera la oración "x cree que O", pero no son tales que tengan que ser satisfechas por cualquier sujeto del que sea verdad que cree que P, es decir, no son condiciones necesarias para la aplicación del predicado.

¿Hay también un conjunto de condiciones suficientes para la no aplicación del predicado "__cree que __"? Tal parece que sí y es el siguiente:

C2. Si (i) x entiende O, la cual relativamente al contexto en el que se hace la atribución expresa la proposición P y, (ii) aduciendo razones, x disiente sinceramente de O, entonces x no cree P, y la oración "x no cree que O" es verdadera. Si O es contradictoria a priori o contradice a priori alguna otra O' aceptada por x en el contexto, x no asentirá a O, no creará P y la oración "x no cree que O" será verdadera en el contexto.

Tal y como hemos formulado C2, tenemos razones suficientes para no aplicar a x el predicado "__cree que O" cuando x entiende y rechaza explícitamente (disiente de) la oración O. Para que la condición C2 fuera necesaria, tendría que ser el caso que siempre que un sujeto x no creyera una proposición P, x disintiera de alguna oración que expresara P o lo mostrara en su conducta. Pero, en

⁴² Lo anterior no implica negar la necesaria conexión de la creencia con la conducta verbal y no verbal, sólo implica rechazar una concepción conductista burda de la creencia.

concordancia con lo que sostuvimos antes, es posible que x no crea que P y sin embargo no dé ninguna muestra de ello. Nuevamente, para determinar los casos en los que no se aplica el predicado, es necesario que su no aplicación esté regulada por la ocurrencia de ciertas manifestaciones conductuales verbales y no verbales; pero, una vez introducido el predicado " __no cree__ ", podemos también concebir el caso en que alguien no crea una proposición y, sin embargo, no de muestras de ello. De hecho hay proposiciones que no creemos que nunca hemos tenido oportunidad de manifestar conductualmente que no las creemos (y tal vez nunca tengamos tal oportunidad). Así como añadimos como parte de $C1$ una condición necesaria para el asentimiento (a saber, que O no sea contradictoria a priori), añadimos a $C2$ una condición suficiente para el disentimiento, a saber, que O sea una oración que x reconozca como contradictoria a priori. Ésta no es más que una exigencia mínima para la racionalidad del sujeto: un sujeto racional rechazará una oración necesariamente falsa a priori o una oración que contradiga a priori alguna otra oración aceptada por él en el contexto.

Si lo anterior es correcto, entonces el predicado " __cree__ " es un predicado parcialmente definido y por principio podrá haber casos a los que ni se aplique ni no se aplique. Es decir, será posible que haya atribuciones de creencias que no sean ni verdaderas, ni falsas, sino simplemente carentes de valor de verdad.

Una vez armados con la idea de que el predicado de creencia está parcialmente definido, volvamos a preguntarnos qué creencias le podemos atribuir con verdad a Pierre en los distintos momentos de su historia. Primeramente, parece innegable, según la manera como se ha caracterizado su caso, que antes de que Pierre visite Londres podemos aplicarle el predicado " __cree que Londres es bonita", ya que claramente en ese contexto c , la condición $C1$ es satisfecha, pues, Pierre entiende la oración francesa "Londres est jolie", ésta no es contradictoria

a priori ni contradice a priori ninguna otra oración de aquellas a las que él asiente, y con base en razones pertinentes Pierre asiente a dicha oración en c.

Pero, cuando Pierre llega a Londres y asiente convencidamente a la oración "London is not pretty", ¿podemos aplicarle el predicado "__cree que Londres no es bonita"? Aquí la respuesta no es clara como en el caso de la atribución anterior. Por una parte, en vista de que en ese contexto, c', Pierre entiende la oración en inglés, la cual no tiene una forma sintáctica contradictoria, ni contradice a priori ninguna otra oración que Pierre acepte en c', y dado que Pierre asiente reflexivamente a ella en c', parece correcto aplicarle el predicado en cuestión. Tenemos sin duda la impresión de que Pierre es víctima de una confusión. Pero recordemos que Pierre no ha asentido a ninguna oración sintácticamente contradictoria o que contradiga a priori alguna otra oración a la que él asienta; él sólo ha asentido a la oración francesa "Londres est jolie" (llamémosla OF), y a la oración inglesa "London is not pretty" (llamémosla OI) y, aunque estamos suponiendo que sabe que "jolie" y "pretty" significan lo mismo, no puede saber a priori que las oraciones son contradictorias. Por otra parte, hay algo desconcertante, inadecuado, en la atribución a Pierre de la creencia de que Londres no es bonita: si lo que dice OI efectivamente contradice lo que dice OF, ¿cómo puede Pierre asentir a ambas oraciones, cómo puede tener ambas creencias? Más aún, después de haberle atribuido en c la creencia de que Londres es bonita, ¿se le puede atribuir justificadamente la creencia contraria? Nótese que son la historia de Pierre (no sus habilidades lógicas) y el hecho empírico de que "London" y "Londres" sean correferenciales, lo que hace problemático atribuirle en c' la creencia de que Londres no es bonita. No hay ninguna propiedad intrínseca de la oración OI que Pierre pase por alto y que sea la que la haga contradictoria o inadecuada como objeto de su creencia; son ciertos hechos empíricos los que

determinan que OI contradiga lo que dice OF y, por lo tanto, los que hacen que resulte inadecuada como objeto de creencia para Pierre.⁴³ Es decir, los hechos empíricos parecen incidir en lo que Pierre cree, y hacen problemático para cualquiera que conozca la historia, entienda el concepto de creencia y tenga los conocimientos lingüísticos relevantes, atribuir a Pierre la pretendida creencia. Nótese que, una vez que Pierre llega a Londres y ocurre la confusión reseñada, tampoco es claro que podamos atribuirle la creencia de que Londres es bonita. Esto es, en *c'* no parece aplicársele a Pierre el predicado "cree que Londres no es bonita", pero tampoco parece aplicársele el predicado "cree que Londres es bonita" en virtud de que los hechos empíricos le resultan demasiado desfavorables. En lo anterior hay, desde luego, la sugerencia de que el predicado "cree que Londres es bonita" está parcialmente definido y la sugerencia también de que cuando se intenta aplicar en *c'* a Pierre produce un vacío de valor veritativo: tanto la atribución de la creencia de que Londres es bonita, como la de la creencia de que Londres no es bonita parecen inadecuadas.

Haciendo a un lado las observaciones recién hechas sobre la aparente dificultad de aplicar a Pierre, una vez que está en Londres, tanto el predicado

⁴³ Como se podrá observar, lo dicho antes es análogo a algo que Kripke sostiene para el caso del predicado de verdad. Según él, para que una oración pueda evaluarse como verdadera tiene que ser fundada [*grounded*], y "el que una oración sea, o no, fundada, no es en general una propiedad intrínseca (sintáctica o semántica) de la oración, sino que generalmente depende de los hechos empíricos." Las oraciones infundadas dan lugar a oraciones paradójicas, y éstas a vacíos de valor de verdad. Resumiendo, el carácter paradójico de ciertas oraciones en las que figura el predicado de verdad, depende de que sea o no fundada y esto, a su vez, depende de ciertos hechos empíricos. Véase, Kripke, S., (1975), p. 11 en la versión castellana.

"cree que Londres no es bonita" como el predicado "cree que Londres es bonita",⁴⁴ preguntemos ahora si podemos, de acuerdo con C1, atribuirle a Pierre la creencia en la conjunción expresada por las oraciones incompatibles a las que asiente. Consideremos primero cómo podríamos eventualmente atribuir a Pierre la creencia en tal conjunción y confrontemos las atribuciones sugeridas con las condiciones establecidas en C1 y C2. Las oraciones que llamé (i) y (ii) en la parte final de la sección anterior parecen ser los únicos candidatos posibles para hacer semejante atribución. Veamos, pues, si (i) y (ii) son atribuciones correctas de acuerdo con C1 y C2.

Recordemos que entre las condiciones suficientes C1 para aplicar a x el predicado "__cree que O" está que la oración O mediante la que le atribuimos la creencia a x no exprese una contradicción lógica a priori y no contradiga a priori ninguna oración aceptada por el sujeto de la atribución en el contexto. Notemos que C1 no dice nada para el caso en el que la oración que figure en nuestra atribución exprese una proposición necesariamente falsa, pero cuya falsedad no sea reconocible a priori, ni para el caso en el que O contradiga de una manera no reconocible a priori otra oración aceptada por x en el contexto de atribución. ¿Cómo se le podría atribuir a Pierre la supuesta creencia en la conjunción contradictoria? Parece que, si hemos de satisfacer las condiciones en C1, no podemos usar en la atribución dos oraciones contradictorias a priori. Como indiqué antes, si usáramos un par de oraciones cuya conjunción fuera falsa a priori para

⁴⁴ Digo que hagamos de lado por el momento la sugerencia de que el predicado de creencia está parcialmente definido, pues si efectivamente no pudiéramos atribuirle a Pierre la creencia de que Londres es bonita, ni la creencia de que Londres no es bonita, entonces, menos podríamos atribuirle la creencia en la conjunción.

atribuir la creencia a Pierre, tendríamos que decir que él satisface también las condiciones suficientes para no creer la proposición expresada. Y entonces, si llegaríamos a una auténtica paradoja: Pierre tendría la creencia en cuestión sólo si no la tuviera.⁴⁵ De modo que, en contra de lo sugerido en la sección anterior, no parece que podamos hacer a Pierre, de ninguna manera, la atribución (ii).⁴⁶

La atribución (i) tiene la ventaja, en relación con C1, de no contener como cláusula subordinada una oración falsa a priori y, así, de recoger la creencia de Pierre de una manera aparentemente más verosímil. Sin embargo, recordemos que (i) tiene la desventaja de usar esencialmente expresiones de tres lenguas; aunque las palabras usadas tengan traducciones canónicas, no podemos traducirlas a una sola lengua so pena de atribuirle a Pierre la creencia en una contradicción a priori, cosa que, según argumenté antes, nos conduciría a una situación paradójica. Por otro lado, dado que, por hipótesis, quien le atribuye la creencia a Pierre sabe que "London" y "Londres" refieren a la misma ciudad (es decir, que tienen exactamente el mismo valor semántico) al hacer la atribución (i) será consciente de que le atribuye a Pierre la creencia en una franca contradicción. Es decir, el intérprete sabe que la creencia contradictoria que le atribuye a Pierre mediante la oración subordinada en (i) es exactamente la misma que la que expresa la oración

⁴⁵ Es decir, si hiciéramos la siguiente atribución: "Pierre cree que Londres es bonita y que Londres no es bonita", entonces, dado el carácter contradictorio de la oración subordinada usada en la atribución, de acuerdo con la última cláusula de C2, podríamos concluir que Pierre no cree tal cosa.

⁴⁶ De acuerdo con C1 no se puede atribuir a Pierre la creencia contradictoria mediante (ii). Esto no impide que podamos usar (ii) para fines distintos al de atribuirle la creencia, por ejemplo, para llamar la atención de Pierre sobre el carácter contradictorio de las proposiciones que expresan las oraciones a las que asiente.

"Londres es bonita y Londres no es bonita". ¿Cuál es pues la ventaja de usar en la atribución la oración híbrida? ¿no es acaso tan evidentemente contradictoria (para el intérprete) como la contradicción lógica a priori?⁴⁷ Si efectivamente es así, al hacer la atribución (i) el intérprete difícilmente puede pensar que satisface C1, ya que la oración que usa en su atribución, aunque no sea ella misma contradictoria a priori, es tal que él sabe que dice exactamente lo mismo que una oración contradictoria a priori. De modo que un intérprete razonable se abstendrá de aplicar a Pierre el predicado "cree que Londres es bonita y que London no es bonita" en c'.

Hemos visto que las atribuciones (i) y (ii) son tales que los predicados de creencia que en ellas se pretender aplicar a Pierre no satisfacen C1. Veamos ahora si las condiciones C2 nos permiten negar la aplicación a Pierre del predicado "cree que Londres es bonita y que Londres no es bonita" y "cree que Londres es bonita y que London no es bonita". Ya hemos concedido antes que si usáramos en la atribución la oración conjuntiva subordinada "Londres es bonita y Londres no es bonita", entonces, se cumpliría una condición suficiente para que Pierre no tuviera la creencia en cuestión. De modo que (ii) puede ser negada sin problema: Pierre no cree que Londres es bonita y que Londres no es bonita. Recordemos que hay dos condiciones suficientes para atribuir a un sujeto el predicado "no cree que

⁴⁷ Este problema se plantea incluso para un intérprete que acepte los sentidos fregeanos: aunque la confusión de Pierre se explica, según tal intérprete, suponiendo que Pierre asocia dos sentidos o representaciones distintas de Londres con "Londres" y "London", el intérprete que sí sabe que ambos nombres designan la misma ciudad no tiene más remedio que asociar con ambos nombres un único "sentido" o representación de Londres. Si esto es así, al hacer la atribución (ii) el intérprete atribuirá una creencia en una franca contradicción. Véase, Kripke, (1979a), sección 1.

O", a saber, (a) que con base en razones pertinentes el sujeto disienta de O o (b) que O sea una oración contradictoria a priori. Pero, ¿que sucede con la atribución (i)? Ya que (i) no es contradictoria a priori, la única manera como pudiéramos decir que el predicado en (i) no se aplica a Pierre tendría que ser si Pierre no entiende al menos una de las oraciones OF y OI que sirven de base para hacer la atribución o si disiente de alguna de ellas. Pero, hemos asumido que Pierre entiende ambas oraciones y que asiente tanto a OF como a OI; más aún, en la conversación hipotética en la que Pierre usa las dos lenguas, Pierre asiente a la conjunción "OF and (et) OI". Por lo tanto, no podemos decir que disienta de la conjunción "OF and (et) OI" y, dado que la conjunción de OF y OI no es contradictoria a priori, no podemos decir tampoco sobre esa base que no crea que OF y OI. Por lo tanto, el predicado __no cree que Londres es bonita y que London no es bonita no se aplica a Pierre. Pero, dado que, como vimos antes, tampoco se aplica a Pierre el predicado contrario, la atribución "Pierre cree que Londres es bonita y que London no es bonita" no parece ser ni verdadera ni falsa. Esto no quiere decir que no podamos hacerla, sólo que si la hacemos nuestra atribución carecerá de valor de verdad.

Hemos visto que las condiciones suficientes que establecimos para aplicar a x el predicado "cree que O" y que hacen verdadero que x cree P (en donde P es la proposición expresada por O), y las condiciones suficientes para aplicar a x el predicado "no cree que O" y que hacen verdadero que x no cree P, no dicen nada, se quedan calladas, sobre el caso en el que el carácter contradictorio de P se expresa en la atribución mediante una oración O necesariamente falsa pero no obviamente falsa, no falsa a priori. Es decir, la atribución (i) parece ser un caso en el que la atribución no es ni verdadera ni falsa, sino "indefinida", carente de

valor de verdad. Nuestro interés al introducir el predicado " __cree__ " es distinguir un grupo de individuos, de los que claramente podemos decir con verdad que creen que algo es el caso, de otro grupo de individuos de quienes podemos decir claramente que no creen que algo sea el caso. Pero nuestro predicado no está diseñado para responder de una u otra manera en todos los casos, en particular en casos como el de Pierre.

Consideremos brevemente qué consecuencias tiene lo anterior respecto del Principio de Sustitución de Idénticos. Si efectivamente el predicado " __cree__ " es un predicado parcialmente definido, entonces habrá casos de atribuciones de creencia indefinidos, y PS nada podrá garantizar respecto de ellas. PS sólo garantiza que si sustituimos un nombre por otro correferencial en una oración verdadera, la oración resultante preservará el valor de verdad de la oración original, pero no dice nada respecto del caso de oraciones indefinidas. Más aún, si mi argumentación ha sido correcta, en una atribución de creencia la sustitución en la cláusula subordinada de un nombre por otro codesignativo puede no alterar el carácter indefinido de una atribución o convertir una atribución indefinida en una atribución falsa; también puede convertir una atribución falsa en indefinida. He argumentado que el problema no está en PS, sino en su aplicación a oraciones en las que figuran predicados parcialmente definidos.

Por último, quisiera destacar una consecuencia de la discusión anterior. Que un hablante llegue a una situación semejante (en los aspectos relevantes) a la que llega Pierre, no depende de que haya cometido algún error lógico o lingüístico, depende de que las cosas sean de la manera como son, depende, como dije antes, de que los hechos empíricos se muestren desfavorables para el hablante. En otras palabras, por más respetuoso que sea un hablante de las reglas de la lógica y por más cuidadoso que sea al hacer sus inferencias y al aprender la referencia de los

nombres que usa, nunca puede tener la garantía de no llegar a asentir a contradicciones, pues, como en el caso de Pierre, los hechos pueden resultarle adversos. Más aún, de acuerdo con la investigación que hemos llevado a cabo en este capítulo, aparentemente podemos concluir que los contenidos de creencia no se determinan internamente, sino externamente. Y, si es correcto lo que he argumentado, los hechos empíricos, la manera como es el mundo externo, puede hacer, incluso, que en situaciones sumamente desfavorables -como la de Pierre-, en un momento t_2 dejemos de tener alguna creencia que teníamos en t_1 .

Recapitemos la investigación realizada hasta aquí. Primeramente, admitimos que una teoría de la referencia directa de los nombres parece estar comprometida con la validez irrestricta de PS. Expusimos la estrategia de Kripke para defender PS y exploramos dos posibles respuestas al enigma planteado por él que parecen ser consistentes con la teoría de la referencia directa de los nombres propios y la validez de PS. Según la primera, en contra de lo sugerido por Kripke en su artículo, es posible atribuir a sujetos racionales creencias en proposiciones contradictorias; cuando una contradicción se expresa mediante una oración que no tiene una estructura superficial contradictoria (es decir, que no es una CLAP), un sujeto x plenamente racional, puede comprender tal oración y, con base en razones, asentir a ella en un contexto apropiado; en ese caso, x creería una proposición contradictoria que se expresa mediante una oración MI. Sin embargo, según la teoría de la referencia directa de los nombres toda oración MI expresa la misma proposición que una oración CLAP, de modo que si aceptamos que el objeto de la creencia de x se puede expresar mediante una MI, tenemos que aceptar que también se puede expresar mediante la CLAP correspondiente, pues, recordemos que toda MI y su CLAP correspondiente tienen exactamente el mismo valor semántico. Observamos, entonces, que este tipo de respuesta entra en tensión

con nuestras intuiciones acerca de la naturaleza de la creencia. Si creer es la actitud apropiada frente a lo tenido por verdadero, ¿cómo un sujeto puede creer una proposición expresada mediante una oración cuya falsedad necesaria puede descubrirse a priori? Esta dificultad nos llevó a explorar una segunda posible respuesta al enigma de Kripke cuando examinamos la posibilidad de que el predicado "__ cree __" sea parcialmente definido; según esta propuesta, las atribuciones que se hacen en casos como el de Pierre, con base en el asentimiento de Pierre a un par de oraciones incompatibles, o bien son falsas (cuando, como en (ii), en la atribución se usa una oración subordinada que es una contradicción lógica a priori) o carecen de valor de verdad (cuando, como en (i), la oración subordinada expresa una imposibilidad metafísica que quien hace la atribución sabe que es necesariamente falsa); en otras palabras, la investigación me condujo a concluir que no es posible hacer una atribución verdadera de creencia en una conjunción contradictoria a sujetos que ejemplifican situaciones como la de Pierre.

CONCLUSIONES

Me propuse al inicio de este estudio investigar la manera como los nombres refieren, esto es, el modo como el nombre se relaciona con su referente. Luego de examinar las dificultades que presenta la teoría que sostiene que los nombres refieren indirectamente, es decir, que refieren sólo en la medida en que contienen descripciones, conceptos o representaciones que se ajustan únicamente al individuo que designan, pasé a considerar la concepción contraria, según la cual los nombres refieren directamente y no a través de "sentidos", ni por el intermediario de algún otro tipo de contenido descriptivo o representacional. Consideré específicamente la versión que defiende Saul Kripke de la teoría de la referencia directa de los nombres propios ordinarios y examiné tres cuestiones problemáticas que parecen seguirse de dicha concepción: la primera en el terreno de la epistemología, la segunda en el terreno de la metafísica y la tercera en el de la filosofía del lenguaje y de la mente. Presenté esas tres cuestiones como dificultades genuinas que tienen que superarse si hemos de adoptar (razonablemente) la teoría de la referencia directa de los nombres propios ordinarios. Así, dediqué los capítulos III, IV y V a examinar cada una de las dificultades aludidas.

En el capítulo III discutí la cuestión de si efectivamente, como sugiere Kripke, la verdad de las tesis de la referencia directa y rígida de los nombres

propios implica la posibilidad de conocer a priori verdades contingentes. La cuestión es importante porque la tesis supuestamente implicada parece chocar con nuestras intuiciones respecto de nuestras capacidades epistémicas; en efecto, no parece verosímil que podamos conocer un hecho contingente sin necesidad de observar cómo es el mundo, sin necesidad de percibir cómo son las cosas. Kripke considera que tal posibilidad existe y que se deriva de otras dos tesis suyas: la de que es posible introducir un nombre mediante una descripción accidental que fije su referencia y la tesis de que la relación entre el nombre y la descripción que fija su referencia no es una relación de sinonimia. Concluí que cuando un hablante h introduce un nombre "N" mediante una descripción accidental que fija su referencia -digamos, mediante la descripción "el Φ "- y, luego, afirma saber a priori que N es el Φ , h no puede tener en realidad tal conocimiento a priori, pues o bien sabe empíricamente que algo satisface la descripción "el Φ ", y en ese caso su conocimiento de que N es el Φ es empírico, o bien no sabe si algo o alguien satisface la descripción "el Φ ", en cuyo caso h carece de todo conocimiento formulable mediante oraciones en las que se use esencialmente el nombre "N".

En el capítulo IV consideré la cuestión de si las tesis acerca de la designación rígida y la necesidad metafísica de la relación de identidad implican efectivamente la falsedad de la teoría materialista de lo mental. La cuestión es también importante porque el dualismo mente/cuerpo, que es la alternativa frente al materialismo, no parece ser una propuesta cuya verdad pueda establecerse sólo a partir de consideraciones semántico-filosóficas y, peor aún, parece ser una propuesta, si no ya superada, sí en vías de serlo. Por una parte, los conocimientos aportados por las neurociencias, la informática y las ciencias cognitivas y, por otra, el punto de vista cada vez más extendido y más convincente acerca de la posibilidad de contar con una teoría física general omniabarcante que se aplique

a todo lo real, hacen sumamente poco probable que existan estados psicológicos de las personas que no sean idénticos a ciertos estados de su cerebro o de su sistema nervioso central. Examiné la argumentación que conduce a Kripke a sostener que los estados mentales de una persona no son idénticos a sus estados cerebrales (o neuronales o corporales) y encontré un salto en su argumentación. Kripke, en efecto, sostiene que todo término usado para referir a un estado físico considerado idéntico a un estado mental de una persona, es un designador rígido, y supone que la referencia de ese designador se fija mediante una descripción esencial del referente. Argumenté que lo anterior no parece ser el caso, pues, por ejemplo, la descripción que usamos para fijar la referencia del término "estimulación de las fibras C" (el cual designa el estado físico que suponemos es idéntico al estado mental de dolor) es algo así como lo siguiente: "el estado cerebral correlacionado con cierto tipo de observaciones cuando se llevan a cabo ciertas pruebas de laboratorio", y parece muy poco verosímil sostener que esa descripción tenga que ser satisfecha en todo mundo posible por el referente de "estimulación de las fibras C" (es decir, por el estado de dolor). Las observaciones que constituyen nuestra vía de acceso epistemológico a ese estado cerebral no son una característica esencial de tal estado, ya que es enteramente posible que exista ese estado sin que haya tales observaciones: si no tuviéramos los órganos sensoriales que tenemos o si no existieran los aparatos con los que contamos, la estimulación de las fibras C (= el dolor) no sería detectada de la manera como se le detecta en el laboratorio. Esto me permitió rechazar la tesis de Kripke, según la cual la concepción de los nombres propios como designadores rígidos y la idea de que toda identidad es necesaria implican la verdad del dualismo cartesiano.

Por último, en el capítulo V abordé la que considero ser la cuestión más urgente y desafiante para quienes concebimos la referencia de los nombres como

una relación no mediada: el problema de la aparente falla de sustitución de los nombres codesignativos en contextos de actitudes proposicionales. Admití que la verdad de la teoría de la referencia directa de los nombres implica la verdad del Principio de Sustitución de Idénticos (PS) e intenté explicar la aparente falla de PS de dos maneras diferentes. Primero, consideré la posibilidad de que nuestro sistema de creencias pudiera admitir la posibilidad de creer contradicciones lógicas. Observé, sin embargo, que intentar salvar PS de esta manera choca con nuestras intuiciones acerca de lo que es la actitud proposicional de creencia. En efecto, si la creencia es la actitud frente a lo tenido por verdadero, parece imposible adoptar tal tipo de actitud frente a una contradicción lógica cuya falsedad puede detectarse a priori. Más aún, argumenté que la actitud proposicional adecuada frente a una contradicción lógica es, por el contrario, la actitud de no creencia. Este tipo de dificultad me llevó a explorar otra posibilidad de explicar lo que sucede cuando PS parece fallar, a saber, la posibilidad de que el predicado de creencia, C, sea un predicado parcialmente definido y, por lo tanto, la posibilidad de que haya casos en los que ni C ni no-C se apliquen y la pretendida atribución de creencia resulte ser una oración que no es ni verdadera ni falsa, sino indefinida. Argumenté que los casos en los que PS parece fallar pueden explicarse, precisamente, como casos en los que se intenta aplicar PS a contextos indefinidos y en donde el resultado de tal aplicación es una oración falsa. Destaqué la congruencia de la solución propuesta con el esbozo de una teoría de la verdad que ofreció Kripke en su célebre artículo de 1979.

Si mi argumentación ha sido correcta, creo poder concluir que la teoría de la referencia directa de los nombres, lejos de constituir una visión ingenua o simplificadora de la referencia de los nombres propios ordinarios como se le ha solido presentar, se nos revela como una teoría poderosa y viable; es decir, como

una teoría que cuenta con los recursos necesarios para responder a las objeciones que se le suelen plantear y que contesta de manera satisfactoria muchas de las preguntas que cualquier teoría de la referencia aceptable debe contestar. Con base en este estudio creo tener razones para poder afirmar que las dificultades que comúnmente se le suelen achacar a la concepción de los nombres como términos singulares que refieren directamente a sus portadores son más aparentes que reales.

BIBLIOGRAFÍA

- Almog, J., Perry, J., y Wettstein, H., comps. (1989), Themes from Kaplan, OUP, Oxford.
- Bradley, R., y Swartz, N., (1979), Possible Worlds: An Introduction to Logic and its Philosophy, Basil Blackwell, Oxford.
- Carnap, R., (1958) Introduction to Symbolic Logic, Dover Publications, Nueva York.
- Castañeda, H.N., (1972), "Thinking and the Structure of the World", Crítica. Revista Hispanoamericana de Filosofía, vol. VI, nº 18.
- Chapell, V.C., comp., (1962), The Philosophy of Mind, Prentice Hall, Nueva Jersey.
- Chisholm, R., (1977), Theory of Knowledge, Prentice Hall, Nueva Jersey.
- Church, A., (1956), Introduction to Mathematical Logic. I, Princeton University Press, Princeton, Nueva Jersey.
- Davidson, D., (1970) "Mental Events", en Foster, L., y Swanson, J.W., comps. (1970). Versión castellana de Lourdes Valdivia, Sucesos mentales, Cuadernos de Crítica, Nº 11, UNAM, México.
- Donnellan, K., (1979 a) "Speakers Reference, Descriptions and Anaphora", en French, P. et alt., comps., (1979).
- (1979 b) "The Contingent A Priori and Rigid Designation", en French,

- P., et alt., comps., (1979).
- Dummett, M., (1981) Frege. Philosophy of Language, segunda edición Duckworth, Londres.
- Evans, G., (1978) "Reference and Contingency", The Monist, 1978.
- (1982) Varieties of Reference, Clarendon, Oxford University Press, Oxford y Nueva York.
- Ezcurdia, M., (1994), Sense, Indexicals, and Action, Tesis Doctoral, Kings College, Universidad de Londres. Inédito.
- Foster, L., y Swanson, J. W., comps. (1970), Experience and Theory, The University of Massachusetts Press, Amherst.
- Frege, G., (1879), Begriffsschrift. eine der arithmetischen nachgebildete Formalsprache des reinen Denkens, Halle. Citado Begriffsschrift. Traducción al castellano de Hugo Padilla, en Frege, G., (1972)
- , (1884), Die Grundlagen der Arithmetik. Versión castellana de Hugo Padilla, Los Fundamentos de la Aritmética, en Frege, G., (1972).
- , (1892) "Über Sinn und Bedeutung", Zeitschrift für Philosophie und philosophische Kritik, n.f. 100. Versión castellana de Eugenio Bulygin y Eduardo Rabossi, en Simpson, T.M., comp. (1973).
- , (1892-1895) "[Comments on Sense and Meaning]" en Frege, G., (1979). Versión castellana de C. U. Moulines, "Consideraciones sobre Sentido y Referencia", en Frege, G., (1971).
- , (1918-1919) "Der Gedanke", en Beiträge zur Philosophie des deutschen Idealismus, 1, Versión castellana de Carlos Pereda, "El Pensamiento: una investigación lógica", en Valdés, M. M., (1996) (Los números de

página en las citas refieren a esta traducción.)

- , (1971) Estudios sobre semántica, versión castellana de Ulises Moulines, Editorial Ariel, Barcelona.
- , (1972), Conceptografía. Los fundamentos de la Aritmética. Otros estudios filosóficos. Versión castellana de Hugo Padilla, UNAM, 1972.
- , (1979) Posthumous Writings, edición de Hermes, H., Kambartel, F., y Kaulbach, F., Basil Blackwell, Oxford.
- , (1980) Philosophical and Mathematical Correspondence, editado por Gabriel, G., Hermes, H., Kambartel, F., Thiel, C., y Veraart, A., University of Chicago Press.
- French, P., et al., comps., (1979), Contemporary Perspectives in the Philosophy of Language, University of Minnesota Press, Minneapolis.
- , et al., comps., (1981), Midwest Studies in Philosophy, vol. vi.
- García-Carpintero, M., (1996), Las palabras, las ideas y las cosas, Editorial Ariel, Barcelona.
- Geach, P., (1962), Reference and Generality, Cornell University Press, Ithaca y Londres.
- Goldschmidt, V., (1982) Essai sur le "Cratyle", Contribution à l'histoire de la pensée de Platon, Librairie Philosophique J. Vrin, Paris.
- Gomperz, T., (1950), Pensatori Greci, Storia de la Filosofia Antica, vol. III, 3a. edición.
- Gosling, J.C.B., (1973 a) "Falsity in the Cratylus", en Gosling, J.C.B. (1973 b).
- , (1973 b) Plato, The Arguments of the Philosophers, Routledge and Kegan Paul, Londres.
- Grice, H.P., (1957) "Meaning", Philosophical Review, vol. LXVI.
- , (1968) "Utterer's Meaning, Sentence-Meaning and Word-Meaning",

Foundations of Language, 4.

- Hansberg, O., (1978) "Monismo anómalo", Dianoia, 1978.
- Harman, G., y Davidson, D., (1972), Semantics of Natural Language, Reidel, Dordrecht y Boston.
- Holtzman, S., y Leich, C., comps. (1981), Wittgenstein: To Follow a Rule, Routledge and Kegan Paul, Londres y Boston.
- Kant, I., (1787), Kritik der reinen Vernunft, 2a edición, Riga. Versión inglesa de Norman Kemp Smith, Critique of Pure Reason, Mc Millan, Londres, 1963.
- Kaplan, D., (1968) "Quantifying In", Synthese, vol, 19, nº1/2. Versión castellana de T. M. Simpson, "Cuantificación, modalidad y creencia (reflexiones en torno a Quine)", en Simpson, T. M., comp. (1973).
- , (1979) "On the Logic of Demonstratives", Journal of Philosophical Logic, 8.
- Ketchum, R.J., (1979), "Names, Forms, and Conventionalism: Cratylus 383-395", Phronesis, vol.24.
- Kretzman, N., (1971), "Plato on the correctness of names", American Philosophical Quarterly, vol. 8, nº 2.
- Kripke, S., (1971) "Identity and Necessity", en Munitz, M., (1971). Versión castellana de Margarita M. Valdés, "Identidad y necesidad", Cuadernos de Crítica, nº7, UNAM, 1978.
- , (1980) Naming and Necessity, edición corregida y aumentada, Basil Blackwell, Oxford. Versión castellana de Margarita M. Valdés, El nombrar y la necesidad, UNAM, 1985.

- , (1975), "Outline of a Theory of Truth", The Journal of Philosophy, vol.72, nº 19. Versión castellana de Margarita M. Valdés, "Esbozo de una teoría de la verdad", Cuadernos de Crítica, nº 36, UNAM, 1984.
- , (1979 a) "A Puzzle about Belief", en Margalit, A., comp., (1979). Versión castellana de Margarita M. Valdés, "Un enigma sobre la creencia", en Valdés, M. M., comp., (1996).
- , (1979 b), "Speakers Reference and Semantic Reference", en French, P., et al., comps., (1979)
- , (1982), Wittgenstein on Rules and Private Language, Harvard University Press, Cambridge MA. Versión castellana de Alejandro Tomasini, Wittgenstein: Reglas y lenguaje privado, UNAM, 1989.
- Lewis, D., (1969), Convention. A philosophical study, Harvard University Press, Cambridge MA.
- , "An Argument for the Identity Theory", The Journal of Philosophy, vol. 63, nº1. Versión castellana de Enrique Villanueva, "Un argumento en favor de la teoría de la identidad", Cuadernos de Crítica, nº 30, UNAM, 1984.
- Marcus, Ruth Barcan, (1981), "A Proposed Solution to A Puzzle about Belief", en French, P., et al., Midwest Studies in Philosophy, vol vi.
- ,(1983), "Rationality and Believing the Impossible", The Journal of Philosophy, vol. lxxx, n.6.
- Margalit, A.,(1979), comp., Meaning and Use, Reidel, Dordrecht.
- Martí, G., (1995) "The Essence of Genuine Reference", Journal of Philosophical Logic, 24.
- Mates, B., (1972), Elementary Logic, Segunda edición, OUP, Oxford.

- McDowell, J., (1977) "On the sense and reference of a proper name", Mind, vol. LXXXVI, n° 342. Versión castellana de Lourdes Valdivia, "Sobre el sentido y la denotación de un nombre propio", Cuadernos de Crítica, UNAM, 1983.
- McGinn, C., (1977), "Anomalous Monism and Kripke's Cartesian Intuitions", Analysis, 37.
- Millikan, R.G., (1991) "Perceptual Content and the Fregean Myth", Mind, vol. C, n°4.
- , (1993 a), "White Queen Psychology or the Last Myth of the Given" en Millikan, R.G., (1993 b)
- , (1993 b), White Queen Psychology, MIT Press.
- Munitz, M., (1971), comp., Identity and Individuation, New York University Press.
- Nagel, T., (1965), "Physicalism", Philosophical Review, vol. lxxiv, n° 3. Versión castellana de Julieta Lascuráin y Enrique Villanueva, "Fisicalismo", Cuadernos de Crítica, n°2, 1977.
- Nussbaum, M., y Schofield, M., (1981), comps., Logos and Language, Cambridge University Press.
- Peacocke, C., (1981), "Rule Following: The Nature of Wittgenstein's Argument" en Holtzman, S., y Leich, C., comps., (1981).
- Platón, Cratilo, texto establecido y versión francesa de Louis Méridier, Cratyle, Collection des Universités de France, Association Guillaume Budé, París, 1969.
- Platts, M., (1979) Ways of Meaning. An Introduction to a Philosophy of Language, Roudledge and Kegan Paul. Segunda edición, aumentada, (1997), MIT Press. Versión castellana de Antonio Zirión, Sendas del

- significado, UNAM-FCE, México (1992).
- Place, U.T., (1962) "Is Consciousness a Brain Process?", en Chapell, V., C., comp., (1962).
- Quine, W.V.O., (1943), "Notes on Existence and Necessity", The Journal of Philosophy, vol. XL.
- , (1947), "The Problem of Interpreting Modal Logic", Journal of Symbolic Logic, vol. XII, nº 2.
- , (1953a), "Reference and Modality", en Quine, W.V.O., (1953b).
- , (1953b), From a Logical Point of View, Harvard University Press, Cambridge, MA.
- , (1960), Word and Object, MIT Press, Cambridge, MA.
- Récanati, F., (1979), La transparence et l'énonciation, Editions du Seuil, Paris.
- Robinson, R., (1956) "A Criticism of Plato's Cratylus", The Philosophical Review, vol. LXV.
- , (1969), Essays on Greek Philosophy, OUP, Oxford.
- Russell, B., (1905), "On Denoting", Mind, XIV. Reimpreso en Russell, B., (1956).
- , (1910-1911), "Knowledge by Acquaintance and Knowledge by Description", Proceedings of the Aristotelian Society, vol. 11. Reimpreso en Russell, B. (1917).
- , (1917), Mysticism and Logic, George Allen and Unwin Ltd., Londres.
- , (1956), Logic and Knowledge, Londres, George Allen and Unwin Ltd., Londres. Versión castellana de Javier Muguerza, Lógica y conocimiento, Taurus, Madrid, 1966.
- Ryle, G., "Letters and Syllables in Plato", The Philosophical Review, vol. 69.

- Salmon, N., (1982), Reference and Essence, Basil Blackwell, Oxford.
- , (1986), Frege's Puzzle, MIT Press. (Reimpreso por Ridgeview Publishing co., Atascadero, CA., 1991).
- , (1989a), "Illogical Belief", en Tomberlin, J., comp. (1989).
- , (1989b), "How to become a Millian Heir", Noûs, 23.
- y Soames, S., comps. (1988), Propositions and Attitudes, Oxford Readings in Philosophy, OUP, Oxford.
- Schiffer, S., (1978), "The Basis of Reference", Erkenntnis, 13.
- , (1979), "Naming and Knowing", en French, P., et al., (1979).
- , (1987), Remnants of Meaning, MIT Press, Cambridge MA. y Londres.
- , (1992), "Belief Adscription", The Journal of Philosophy, vol. 89, n° 10. Versión castellana de Margarita M. Valdés, en Valdés, M., M., comp. (1996).
- Schofield, M., "The Dénouement of the Cratylus", en Nussbaum M. y Schofield, M., comps., (1981).
- Searle, J., "Proper Names", Mind, vol. 67, 1958. Recogido en Strawson, P., (1967).
- Simpson, T. M., comp. (1973), Semántica Filosófica. Problemas y discusiones, Siglo XXI, Argentina Editores, Bs. Aires.
- , (1975), Formas Lógicas. Realidad y Significado, Segunda Edición, corregida y aumentada, EUDEBA, Bs. Aires.
- Sluga, H., (1980), Gottlob Frege, The Arguments of the Philosophers, Routledge and Kegan Paul, Londres y Boston.
- Smart, J.J.C., (1959), "Sensations and Brain Processes", Philosophical Review, vol. LVIII.
- Soames, S., "Direct Reference and Propositional Attitudes", en Almog, J., Perry,

- J. y Wettstein, H., comps. (1989). Versión castellana de José Antonio Robles en Valdés, M.M., comp., (1996).
- , (1995), Truth and Meaning, Cátedra Gaos 1995, manuscrito, Biblioteca Eduardo García Maynez, Instituto de Investigaciones Filosóficas, UNAM.
- Sosa, E., (1992), Conocimiento y Virtud Intelectual, UNAM-FCE, México.
- Strawson, P., (1959), Individuals, Methuen, Londres.
- , (1964), "Intention and Convention in Speech Acts", The Philosophical Review, vol. LXXIII, n° 4.
- , comp. (1967), Philosophical Logic, Oxford Readings in Philosophy, OUP, Oxford.
- , (1974), Subject and Predicate in Logic and Grammar, Methuen, Londres.
- Tomberlin, J., comp. (1979), Philosophical Perspectives, 3, Philosophy of Mind and Action, Ridgeview Publishing Company, Atascadero, CA.
- Valdés, M. M., (1987), "Dos teorías de la referencia en el Cratilo", Análisis Filosófico, vol VII, n° 2.
- , (1993) "Fallas de sustitución de términos correferenciales", Revista Latina de Pensamiento y Lenguaje, vol. 1, n° 2.
- , comp. (1996), Pensamiento y Lenguaje. Problemas en la atribución de actitudes proposicionales, UNAM, México.
- Valdivia, L., (1989), Introducción a la semántica y la ontología de Frege, UNAM, México.
- Villoro, L., (1982), Creer, saber, conocer. Siglo XXI, México.
- Wiggins, D., (1975), "Identity, Designation, Essentialism, and Physicalism", Philosophia, vol. V, n° 1-2.

- Williams, B., (1981), "Cratylus' Theory of Names and its Refutation", en Nussbaum, M. y Schofield, M., comps. (1981).
- Williamson, T., (1994), Vagueness, Routledge, Londres y Nueva York.
- Wittgenstein, L., (1853), Philosophical Investigations, Blackwell, Oxford. Versión castellana de Alfonso García Suárez y C.U. Moulines, Investigaciones Filosóficas, Crítica-UNAM, Barcelona, 1990.