

NOTAS BIBLIOGRÁFICAS

Manuel García-Carpintero, *Las palabras, las ideas y las cosas. Una presentación de la filosofía del lenguaje*, Ariel, Barcelona, 1996, xxix + 562pp.

Como su subtítulo anuncia, este libro es una exposición de algunos problemas y aportaciones importantes discutidos en la filosofía (analítica) del lenguaje. Los aspectos distintivos de la obra son seguramente la selección de los temas, la estructura de la presentación unificadora que se hace de ellos, y las propuestas originales que el autor hace sobre las diversas discusiones. En opinión del reseñador, se trata de una contribución ambiciosa e importante a la bibliografía original en español sobre filosofía del lenguaje, tanto en virtud de la interesante visión unificadora que subyace en la presentación de los temas como por la presencia de las propuestas originales del autor. En lo que sigue daré en una primera sección un resumen general del libro, mencionando los temas principales de que se ocupa e intentando describir algunas de las interconexiones que el autor establece entre ellos. En la segunda sección ofreceré algunas apreciaciones a modo de valoración general del libro. Finalmente, dedicaré una tercera sección a una serie de observaciones críticas sobre algunas de las propuestas sobre problemas específicos que el autor hace a lo largo de su obra.

I

El libro consta de un prólogo, una introducción y catorce capítulos. En el prólogo y la introducción el autor expone algunas ideas generales sobre la filosofía y la filosofía del lenguaje en particular. Según el autor, la filosofía analítica (no sólo la filosofía analítica del lenguaje) se distingue (de otros modos de hacer filosofía) por presuponer la tesis metodológica de que el estudio del lenguaje tiene prioridad filosófica sobre el estudio

del pensamiento (p. xvii). “La filosofía, en esta concepción, es una *actividad intelectual teórica*, coincidente con la lexicografía en particular y con la semántica de los lenguajes naturales en general en sus métodos y en su objetivo: la investigación del significado de las expresiones lingüísticas” (*ibid.*); específicamente, la filosofía se ocupa de investigar el significado de las expresiones de interés filosófico. Según el autor (simplificando algo), una “actividad intelectual teórica” es una actividad que busca ofrecer explicaciones a problemas planteados de forma independiente, explicaciones que usen conceptos teóricos nuevos (p. xix). Si al lector le suena esto como una descripción típica de la actividad científica, no es por casualidad; para el autor ciencia y filosofía son paradigmas de actividades teóricas, y los ejemplos que da de este tipo de actividades son ejemplos extraídos de la historia de la ciencia.

Según el autor, el problema principal de la filosofía del lenguaje es elucidar las relaciones entre el lenguaje, el pensamiento y la realidad extralingüística (p. xxix) (el título del libro responde a esta idea). La estructura de la obra la motiva en buena medida el deseo del autor de explicar una serie de ideas generales sobre estas relaciones, asociadas a los grandes filósofos de la tradición analítica; la presentación describe sucesivamente cada idea como una reacción (al menos en parte) a los defectos de las ideas expuestas previamente; al final del libro, el autor expone su propia concepción general de las relaciones entre lenguaje, pensamiento y mundo. En gran parte me refería a esta estructura “dialéctica” y a la propuesta de una “síntesis” final cuando hablaba antes de la visión unificadora de la obra. Es al hilo de la presentación de las grandes ideas sobre el problema principal de la filosofía del lenguaje que el texto expone también problemas más específicos y propuestas acerca de ellos, incluyendo las del propio autor.

El capítulo I lo dedica el autor a introducir algunos conceptos y distinciones característicos de la filosofía del lenguaje: la distinción entre tipos (*types*) y ejemplares (*tokens*), los conceptos de propiedad sistemática y productiva, de proposición, de oración gramatical, de sintaxis, semántica y pragmática, la distinción entre uso y mención de signos, entre lenguaje y metalenguaje, entre conocimiento tácito y explícito. También se introduce la noción, ideada por el autor, de “signo ostensivo” (explicada en la sección III de esta reseña). Esta noción se usa en el ca-

pítulo II, que es un repaso a las teorías más importantes del entrecomillado, junto con una propuesta original del autor acerca de este fenómeno lingüístico (en la que se usa la noción de signo ostensivo). Este capítulo no es un fin en sí mismo; el autor pretende ilustrar por medio de las teorías del entrecomillado su concepción de las teorías filosóficas como resultado de una actividad intelectual teórica en la que se investiga el significado de las expresiones lingüísticas. (Volveremos sobre el material del capítulo II en la sección III.)

El capítulo III presenta una serie de conceptos, distinciones e ideas epistemológicas y de filosofía de la mente. Se explica el concepto de intencionalidad (de ciertos estados mentales). Se introduce el nombre ‘realismo ingenuo’ para la idea de que la verdad de enunciados acerca de acaecimientos perceptibles es cognoscible por “contacto directo” con esos acaecimientos. Se introduce el nombre ‘realismo por representación’ para la tesis de que tal conocimiento es en todo caso “mediato”, presuponiendo una relación (también intencional) inmediata con ciertas entidades mentales, normalmente causadas por el mundo externo. A estas entidades de las que se tiene conocimiento directo, infalible e incorregible se las llama ‘vivencias’. El realismo por representación se presenta como superior al realismo ingenuo, que no explica satisfactoriamente, por ejemplo, la existencia de alucinaciones e ilusiones. Se explica la distinción entre internismo y externismo, de la siguiente forma: una teoría internista es una según la cual el objeto intencional de todos los estados mentales intencionales es “caracterizable” sin suponer la existencia de entidades del mundo externo; una teoría externista es una según la cual el objeto intencional de algunos estados mentales intencionales no es caracterizable de esa forma. Finalmente, en el capítulo se explican las distinciones entre verdades *a priori* y *a posteriori*, entre verdades analíticas y sintéticas, y entre verdades necesarias, contingentes y posibles.

El capítulo IV está dedicado a una exposición de la filosofía del lenguaje de Locke, a quien se presenta como un “realista por representación” (en el sentido del capítulo III), un defensor de la prioridad ontológica y epistemológica del pensamiento sobre el lenguaje y un “internista semántico”. Un internista semántico es alguien para quien las expresiones lingüísticas significan primariamente entidades mentales, y sólo secundariamente entidades

del mundo externo. Es en este sentido que se dice también de Locke que según su concepción “el lenguaje que cada uno de nosotros habla es *epistémicamente privado*: es imposible saber si, en su significación primaria, las palabras significan para nosotros lo mismo que significan para los demás” (p. 115).

En el capítulo V se introducen varios conceptos y distinciones metafísicas y epistemológicas. Se distingue entre relaciones causales y de participación y entre propiedades primarias y secundarias (en términos de la distinción entre propiedades disposicionales y categóricas). Se distingue entre eliminativismo, realismo fingido y proyectivismo acerca de las propiedades mencionadas en un ámbito de discurso (como el discurso acerca de las relaciones causales): el eliminativista dice que esas propiedades no existen; el realista fingido dice que podemos actuar como si existieran, aunque no sepamos ni podamos saber si existen; el proyectivista dice que la ejemplificación de esas propiedades “depende de reacciones producidas dadas ciertas condiciones en un ser racional, o en un grupo de tales seres” (p. 167). Al final del capítulo se ofrece un análisis mackiano de la causalidad.

El capítulo VI explica la distinción de Frege entre sentido y referencia. El autor presenta a Frege como un realista por representación lockeano, si bien insiste en que “según Frege, el origen último del problema de Locke está en tomar a las palabras como aquello que primariamente tiene significado, cuando lo que primariamente tiene significado son más bien las oraciones” (p. 179). El autor expone los principios fregeanos del contexto y de composicionalidad, pasando luego a explicar el argumento de Frege en su clásico artículo “Sobre sentido y referencia”, mediante el cual introdujo la distinción correspondiente. Los sentidos son las entidades intencionales inmediatas cuya aparición en la teoría de Frege convierte a éste en un realista por representación.

En el capítulo VII se presenta además a Frege como un inter-nista semántico, pero no en virtud del argumento en su artículo clásico, sino en virtud de la observación de que hay términos sin referencia que aparecen en oraciones a las que no obstante atribuimos intuitivamente significado; según el autor, Frege concluye de esto que incluso el objeto intencional de los términos con referencia (o de las oraciones que los contienen) es caracterizable sin mencionar esa referencia. También según el autor,

la famosa polémica entre Frege y Russell acerca del Mont-Blanc y sus nieves es una polémica entre internismo y externismo. El capítulo continúa con una exposición de las conocidas críticas de Kripke a la concepción fregeana de los sentidos de los términos singulares. El autor propone entonces una concepción neofregeana de esos sentidos que argumenta está a salvo de esas críticas. Finalmente se presentan ciertas ideas de Kaplan sobre las atribuciones *de re* de actitudes proposicionales, ideas que según el autor permiten reconstruir esas atribuciones como atribuciones *de dicto* en la línea internista; estas ideas también se aplican en una explicación propuesta por el autor del llamado “enigma de Kripke”. (Volveremos sobre las propuestas originales de este capítulo en la sección III.)

En el capítulo VIII se expone la teoría de las descripciones de Russell, cuyo desarrollo llevó a éste, según el autor, a cambiar radicalmente del externismo manifestado en su polémica con Frege a una forma extrema de internismo, el fenomenalismo. (Debe observarse, sin embargo, que el autor no usa este término para la tesis epistemológica o semántica de que el significado de los enunciados acerca de objetos físicos es equivalente al de enunciados acerca de sensaciones, sino para la tesis metafísica más fuerte “de que sólo existen las vivencias y sus constituyentes” (p. 153).)

El capítulo IX contiene, en primer lugar, una presentación de la teoría “figurativa” o “isomorfista” del significado del *Tractatus* de Wittgenstein. El autor se apoya luego en la teoría wittgensteiniana de la necesidad lógica derivada de la concepción figurativa para ofrecer sus propios puntos de vista acerca de la necesidad lógica. El capítulo X está dedicado a defender una interpretación de la teoría del significado del *Tractatus* según la cual es una teoría internista y además fenomenalista (en el sentido del autor). Seguramente la parte más interesante y potencialmente más polémica de este capítulo es precisamente aquella en que el autor interpreta los (escasos) textos wittgensteinianos que según él contienen el argumento que lleva a Wittgenstein al rechazo del realismo por representación, rechazo que, conjugado con el internismo, le lleva al fenomenalismo. (Comentaremos brevemente este argumento más adelante.)

En el capítulo XI el autor ofrece una reconstrucción de las consideraciones de Wittgenstein en las *Investigaciones filosófi-*

cas que le llevan a negar la existencia de lenguajes privados (en el sentido explicado en el capítulo IV). El núcleo de estas consideraciones es la tesis de que los significados son normativos. Por esto el autor parece entender que han de ser entidades cuyo conocimiento permite a un hablante (competente) determinar si las expresiones correspondientes se aplican correctamente en cada caso. (Habitualmente por ello parece entenderse algo aparentemente más débil, a saber, que el hecho de que una palabra signifique algo determina (en algún sentido no necesariamente dependiente de las habilidades de los hablantes) su correcta aplicación en todos los casos.) El argumento contra los lenguajes privados se presenta como una aplicación de estas consideraciones: los significados no pueden ser entidades mentales, pues el contacto con una entidad tal “asociada” a un término en un momento dado no permite por sí solo a un hablante determinar si el término se está aplicando correctamente (con el “mismo significado”) en otros casos (el hablante no puede confiar del todo en su memoria, por ejemplo). La tesis wittgensteiniana de que el significado es el uso se ve entonces como una propuesta que sí recoge el elemento normativo del significado; simplificando, de acuerdo con la reconstrucción de la propuesta wittgensteiniana hecha por el autor los significados son ciertas disposiciones generales comunitarias a la conducta lingüística; esta concepción excluye la posibilidad de que hablantes propiamente dispuestos apliquen incorrectamente los términos de su lenguaje. (Sería imposible discutir en esta reseña las virtudes o problemas exegéticos y filosóficos de la propuesta del autor sobre este espinoso tema.) Pero la propuesta wittgensteiniana es criticada por sus aparentes consecuencias antirrealistas: si el significado es el uso, entonces parece que no es posible que una palabra signifique algo a menos que los usuarios competentes sean capaces de reconocer que es así. Sin embargo, el autor no es explícito a la hora de dar argumentos que muestren que sí es posible; ciertos conocidos argumentos externistas (que muestran, por ejemplo, que usuarios competentes de ‘agua’ podrían aplicar incorrectamente el término —a “XYZ”) podrían usarse aquí, pero posiblemente no sean los que tiene en mente el autor.

La situación “dialéctica” en este punto es la siguiente (véase la p. 421): el mentalismo representacionista de Locke y Frege ha quedado mal parado por el argumento del *Tractatus* que lleva

al fenomenalismo; a su vez, la teoría del *Tractatus* ha quedado refutada por el argumento contra los lenguajes privados de las *Investigaciones filosóficas*; la teoría del significado de las *Investigaciones* sigue siendo, sin embargo, una teoría internista (comunitaria, no individualista como la de Locke o el *Tractatus*), combinada con un proyectivismo general en metafísica. El capítulo XII explica las ideas más conocidas de la filosofía de Quine, a quien se presenta como un internista y proyectivista con argumentos emparentados con los del segundo Wittgenstein. Se exponen primero las críticas de Quine a la distinción entre verdades analíticas y sintéticas y su propuesta de la naturalización de la epistemología. Después, el núcleo del capítulo lo constituye la exposición de las ideas de Quine sobre el significado y la traducción. Las conocidas tesis de este filósofo sobre la indeterminación de la traducción y la inescrutabilidad de la referencia se discuten en detalle, llamando el autor la atención sobre algunas aparentes dificultades en su enunciación. De esas tesis se sigue que, en cierto sentido, el significado (entendida esta noción intuitivamente) de las expresiones usadas por los demás está indeterminado observacionalmente para un hablante en general. Quine cree además que cualquier noción aceptable de significado es observacional, lo cual tiene, como en el segundo Wittgenstein, indeseables consecuencias proyectivistas y antirrealistas: habrá muchas cosas que intuitivamente podrán ser significados de nuestras expresiones (ciertos géneros naturales, por ejemplo, en los argumentos externistas más conocidos; de nuevo, quizá no sean estos los ejemplos que el autor tiene en mente), pero la mera observación de la conducta lingüística no permitirá determinar que lo son.

En el capítulo XIII el autor explica algunas nociones e ideas de la teoría de la acción y de pragmática, que se usarán luego en el siguiente capítulo (el último) en el esbozo de una teoría del significado favorecida por el autor. Se tratan aquí la teoría de las acciones de Davidson, la teoría de los actos de habla de Austin y sus seguidores, y el análisis griceano de las implicaturas conversacionales y los significados no literales en general. El capítulo XIV ofrece una reelaboración personal del programa de Grice, en la que se esboza un análisis de la noción de significado convencional en términos de la noción de significado ocasional del hablante y la noción de convención lingüística tal como se

explica en la teoría de Lewis. Éste es un análisis mentalista (tanto la noción de significado ocasional del hablante como la de convención lingüística lewisiana se explican por referencia a complejos estados mentales de los hablantes). Sin embargo, el autor sostiene que no se le aplica la objeción wittgensteiniana sobre la normatividad de los significados. Al parecer, la idea del autor es que la normatividad del significado queda recogida al derivarse del análisis la consecuencia de que los significados son funciones teleológicas de los enunciados que los poseen: las funciones de satisfacer en toda ocasión ciertas intenciones comunicativas uniformes. La derivación de estas consecuencias no es enteramente perspicua, y tampoco es claro por qué el autor cree que su propuesta no tiene consecuencias antirrealistas (parece que debería decirse algo al respecto, dada la importancia que se ha dado a esta objeción al discutir las teorías del segundo Wittgenstein y Quine, pero no es claro que el autor lo haga); la sección final de este capítulo (junto con las del capítulo XI) es seguramente de las más difíciles del libro, y probablemente también una de las que merecen un estudio más profundo; tal estudio está fuera de los límites de esta reseña.

II

Como dije al principio, mi valoración global del libro es muy positiva. Su reconstrucción unificadora de las principales aportaciones de la filosofía analítica del lenguaje, las relaciones que establece entre estas aportaciones y varios grandes temas de la epistemología y de la metafísica tradicionales, y las propuestas específicas sobre cuestiones (comparativamente) de detalle revelan siempre un conocimiento profundo de los problemas y la bibliografía, y una reflexión personal subyacente no menos profunda. La obra ha de ser, como se suele decir (aunque pocas veces tan justificadamente como en este caso), lectura obligada para los interesados en la filosofía del lenguaje en el ámbito hispanohablante, y merecería ser difundida en otros ámbitos. El filósofo que esté trabajando sobre un tema relacionado con alguno tratado en el libro hará sin duda muy bien en leer los capítulos correspondientes como parte de su preparación. Y el especialista en esos temas hará también muy bien en leer el texto

y ofrecer sus reacciones a las propuestas originales de éste, todas ellas de notable interés.

Quizá haya algunos reproches de tipo general (aparte de las críticas estrictamente filosóficas) que se le puedan hacer al libro, aunque seguramente éstos tendrán que ver con cuestiones de gustos. Por ejemplo, posiblemente muchos lectores objetarán a la elección de temas hecha por el autor. Recuérdese que según el autor la filosofía es una actividad emparentada con la lexicografía y con la semántica de los lenguajes naturales en cuanto que investiga el significado de las expresiones lingüísticas (de interés filosófico). Sin embargo, el espacio dedicado precisamente a aspectos concretos de semántica de expresiones de interés filosófico (la semántica de las construcciones modales, de las construcciones epistémicas, del predicado de verdad, las paradojas a que algunas de estas construcciones dan lugar bajo supuestos razonables, etc.) es muy escaso y en ocasiones inexistente; por otro lado, se dedica un capítulo entero a la semántica de las comillas, cuyo interés filosófico es escaso —como reconoce el autor. (Dicho sea de paso: parece dudoso que todas las empresas intelectuales que se presentan en el libro se ajusten bien a la caracterización (aproximada) de la filosofía como semántica, o aun como semántica normativa de términos en existencia; muchas parecen más bien consistir en la formulación de problemas, o en la de nuevos conceptos, o en el descubrimiento de verdades no analíticas en las que aparecen viejos conceptos.) Y hay muchos otros temas que algunos pueden echar de menos en el libro: por ejemplo, una discusión de las ideas de Chomsky sobre la naturaleza del lenguaje o de las de Davidson sobre teorías del significado para lenguajes naturales habrían sido bienvenidas por algunos lectores.

Es inevitable reseñar este libro sin hacer alguna observación sobre cuestiones de estilo. No creo equivocarme si digo que una de las mayores dificultades que puede encontrar un lector de esta obra es la a menudo enrevesada y siempre algo dificultosa forma de escribir del autor, así como su manera de disponer el material, sin apenas pausas para reformular lo recién dicho o resumir lo que se avecina o lo que ha quedado atrás. Así, por ejemplo, los “sumarios” que aparecen al final de los capítulos son raramente resúmenes, y se convierten con frecuencia en pretextos para añadir consideraciones nuevas. El autor es consciente de que su

estilo es difícil, pero no se excusa por ello y de hecho bosqueja una teoría según la cual el tipo de profundidad requerido por la filosofía queda mejor servido por un estilo sin concesiones al lector (p. xii).

Dice el autor que una de las funciones de su libro “habría de ser la de servir de manual introductorio a personas que desean o precisan iniciarse en la filosofía contemporánea del lenguaje” (p. xi). La dificultad del estilo y la presentación del material hacen concebir serias dudas de que el libro pueda ser usado con provecho en cursos introductorios de filosofía del lenguaje. Incluso el lector con un interés serio por la disciplina pero sin preparación previa encontrará, me parece, grandes dificultades para un estudio personal del libro. Pero de lo que no cabe duda es de que será una lectura provechosísima para estudiantes aventajados y especialistas.

III

Una valoración filosófica detallada del conjunto de esta obra, y especialmente de su visión unificadora de los grandes temas y las contribuciones más importantes de los autores de la tradición analítica, sería una empresa tan ambiciosa como la obra misma, y está fuera de los límites de esta reseña. Cabe esperar y desear que, después de la lectura y el estudio detenidos a que el libro es acreedor, los filósofos hispanohablantes lleven a cabo una discusión cuidadosa de sus propuestas de mayor alcance exegético y filosófico, que permita apreciar en la justa medida el interés que tienen y las objeciones que fundadamente se les pueda hacer. Aquí sólo es posible discutir críticamente algunas de las contribuciones del autor sobre ciertas cuestiones que, dentro del conjunto de la obra, aparecen inevitablemente como cuestiones de detalle. Por razones de espacio, las cuestiones que es posible discutir serán sólo tres.

La primera de estas cuestiones es la semántica de los entrecomillados, o, para ser más precisos, los entrecomillados con la función de referir a la expresión entrecomillada. El autor suscribe los argumentos de Davidson en contra de la teoría del entrecomillado de Tarski y Quine (véanse la sección 3 del capítulo I y la 1 del capítulo II), y luego ofrece una versión personal de la teoría de Davidson (con la que quiere subsanar algunos

defectos que percibe en la teoría original). En opinión del reseñador, los argumentos de Davidson contra la teoría de Tarski y Quine sólo se aplican a una teoría “de paja” que seguramente nadie ha propuesto (otros autores, como Mark Richard y Scott Soames, han llegado de forma independiente a la misma conclusión); además, las teorías de Davidson y del autor tienen que afrontar serios problemas.

La teoría de Tarski dice que la parte de un entrecomillado que refiere a la expresión entrecomillada es el entrecomillado mismo. (Un entrecomillado lo forma una expresión entrecomillada junto con las comillas que están a sus lados.) Tarski dice en su clásica monografía sobre el concepto de verdad que un entrecomillado puede verse como una palabra sintácticamente simple, y que los entrecomillados tienen el mismo carácter que los nombres propios. Davidson interpreta este texto de Tarski como queriendo decir que los entrecomillados funcionan como los nombres propios en todos los aspectos semánticamente significativos. Basándose en esta lectura, Davidson (y el autor) presentan dos argumentos contra la teoría de Tarski. Es importante explicarlos y explicar por qué no afectan a la teoría de Tarski, pues son seguramente la justificación principal de quienes han propuesto teorías alternativas. El primero de los argumentos consiste en observar que si los entrecomillados fueran nombres propios, entonces nada se perdería si sustituyéramos cada entrecomillado, de forma biunívoca, por un nombre propio de la expresión entrecomillada (un nombre sin relación con esa expresión); pero claramente algo se perdería: los usuarios de este nuevo lenguaje tendrían que aprender el significado de cada nombre mediante actos de aprendizaje individuales, mientras que el dominio del mecanismo semántico del entrecomillado no requiere nada semejante. El segundo argumento es ligeramente más complejo, y se apoya en la observación de que el número de entrecomillados (como el número de expresiones entrecomillables) es potencialmente infinito (el primer argumento no depende de este hecho): si los entrecomillados son en todos los respectos como nombres propios, entonces la única forma de dar sus denotaciones (o referencias) en una teoría recursiva es por medio de una cláusula particular para cada entrecomillado. Entonces una teoría de la verdad para un lenguaje que contuviera entrecomillados habría de tener infinitas cláusulas; no sería una teoría enunciable por

seres finitos. Pero es plausible suponer, como hace Davidson, que un lenguaje cuya teoría de la verdad no fuese enunciable por seres finitos tampoco sería aprendible por ellos.

Estos argumentos son convincentes, pero está claro que la teoría que atacan no es la teoría de Tarski. Una lectura cuidadosa del texto de Tarski muestra que, cuando éste dice que los entrecomillados tienen el mismo carácter de los nombres propios, se refiere a que comparten con ellos la característica que acaba de mencionar, a saber, la característica de ser sintácticamente simples. A su vez, esto quiere decir, en el sentido técnico de Tarski, que su denotación no es una función de la denotación de ninguna de sus partes propias con denotación (esto es por razones sabidas: ‘Marco Tulio’ no denota la misma expresión que ‘Cicerón’). Es claramente inaceptable interpretar el texto de Tarski como diciendo que los entrecomillados son como nombres propios en todos los aspectos semánticamente significativos. Es como si al explicar algunas características de las ballenas, y decir que son mamíferos, dijéramos que son del mismo carácter que los ratones: lo son, en cuanto que ambos son mamíferos, pero nadie inferirá de esto que creemos que tienen el mismo carácter en todos los aspectos.

La teoría de los entrecomillados de Tarski dice meramente que la parte referencial de un entrecomillado es el entrecomillado mismo, y que lo referido es la expresión entrecomillada. Es claro que la primera objeción de Davidson no se aplica a esta modesta teoría, pues no se sigue de ella que “nada se perdería” si los entrecomillados fueran sustituidos por nombres propios sin relación gráfica con la expresión entrecomillada. Muy al contrario, la teoría de Tarski consiste en poco más que la enunciación de la regla que permite (generalmente de modo implícito; los usuarios de las comillas no tienen por qué habérsela formulado conscientemente) descubrir la referencia de cualquier posible entrecomillado.

La segunda objeción davidsoniana tampoco se aplica a la teoría de Tarski. Es perfectamente posible dar teorías recursivas de la verdad para lenguajes que contengan un número infinito de expresiones denotativas sintácticamente simples en el sentido de Tarski. Considérese por ejemplo un lenguaje típico de primer orden en el que las variables se forman por sufijación de un número positivo de pequeños palotes a la letra ‘x’. Estas variables

tienen el mismo “carácter” que los nombres propios, son sintácticamente simples en el sentido de Tarski: su denotación (con respecto a una secuencia) no es una función de la denotación de ninguna de sus partes propias (la letra ‘x’, los palotes y las variables con un número inferior de palotes). Y naturalmente, son infinitas en número. Sin embargo, una cláusula que da la denotación de toda variable se da de la forma conocida. (Hay dificultades de un tipo enteramente distinto para dar una teoría de la verdad para lenguajes que contienen entrecomillados, que tienen que ver con ciertos fenómenos de ambigüedad estructural generados por el uso habitual de las comillas; pero no es posible discutir estas dificultades aquí.) Es inevitable concluir que los argumentos de Davidson no tienen fuerza alguna contra la teoría de Tarski.

La teoría del entrecomillado propuesta por el autor es introducida, al igual que la de Davidson, como una superación de la supuestamente defectuosa teoría tarskiana, y coincide con la de Davidson en la tesis de que la expresión referencial en un entrecomillado son las comillas. Éstas abrevian un término singular, una descripción que contiene un demostrativo, descripción que a su vez es parte de lo que el autor llama “signo ostensivo” (un signo compuesto de entidades lingüísticas, que incluyen un demostrativo, y un objeto u objetos señalados por el demostrativo, y relacionados de una forma “natural” con el significado del signo). Esta descripción varía con el contexto; ejemplos típicos son: ‘el tipo Baskerville del que esto es un ejemplar’ (donde ‘esto’ demuestra una expresión), ‘el sonido del que esto es un ejemplar escrito’, etc. (La principal diferencia con la teoría de Davidson es que en ésta la descripción abreviada no varía con el contexto.)

En opinión del reseñador esta teoría ha de vérselas con un buen número de dificultades. La más fácil de explicar es la siguiente. Una de las características más familiares que relacionan la noción de referencia con los entrecomillados es la naturaleza “desentrecomilladora” de la referencia. Por ejemplo, los usuarios competentes del entrecomillado y de la noción de referencia asentirán a la oración

(1) ‘‘Sócrates’’ refiere a ‘Sócrates’.

Pero según las teorías de Davidson y del autor la oración (1) es falsa, pues abrevia una oración no esencialmente diferente de 'La expresión de la que esto es un ejemplar refiere a la expresión de la que esto es un ejemplar' (donde el primer 'esto' demuestra a un ejemplar de "Sócrates" y el segundo 'esto' a uno de 'Sócrates'); y esta última oración es claramente falsa según la teoría del autor (y según la de Davidson), pues de acuerdo con ella un entrecomillado como 'Sócrates' no refiere a nada; sólo las comillas son términos referenciales.

La segunda cuestión sobre la que haremos unas breves observaciones críticas es la propuesta del autor sobre la semántica de los nombres propios y su explicación del llamado "enigma de Kripke". El autor propone que el sentido de un nombre-tipo viene dado por una regla que asigna un sentido descriptivo (dado por una descripción) a cada ejemplar suyo; esta descripción contiene una referencia a ese ejemplar: los nombres son, en este sentido, expresiones "ejemplar-reflexivas" (*token-reflexives*), como en la propuesta de Reichenbach acerca del significado de los deícticos. Para tomar un ejemplo, la regla asociada al nombre-tipo 'Aristóteles' sería: *dada una preferencia π en que aparece un ejemplar de 'Aristóteles', el referente de ese ejemplar es el objeto cuyo "etiquetado" mediante algún ejemplar de la expresión-tipo 'Aristóteles' es relevante en el contexto en que se ha emitido π* (véase la p. 252, donde el autor da una regla idéntica cambiando 'Aristóteles' por otro nombre). El sentido asociado a un ejemplar de 'Aristóteles' lo daría la descripción *el objeto cuyo etiquetado con 'Aristóteles' es relevante en el contexto en que se ha emitido ese ejemplar de 'Aristóteles'* (véase la p. 257, donde el autor da una descripción idéntica como sentido de un ejemplar del mismo nombre usado antes por él —con los cambios oportunos); y parece que no sólo la descripción, sino también una indicación de la referencia de la expresión 'ese ejemplar de 'Aristóteles'' que aparece en ella, pues la descripción por sí misma no basta para determinar un referente. El autor abandona la tesis de que son los nombres-tipo los que tienen referencia, y sólo concede referencia a ejemplares. También propone en esbozo una teoría similar de los sentidos para otras expresiones, como los términos de género natural. (La relación de "etiquetado" es compleja, y cubre una larga serie de prácticas mediante las cuales un objeto adquiere un nombre y lo conserva,

o conserva “sucesores” de ese nombre (véase la n. 7 en la p. 238). La noción de “relevancia” es compleja también, e incluye la idea de que hay alguna conexión causal entre el ejemplar cuyo sentido está en cuestión, el objeto referido y el ejemplar del nombre-tipo —o un ejemplar de otro nombre-tipo— usado en el “etiquetado” original.)

El propósito principal del autor al hacer esta propuesta es que sea inmune a la objeción kripkeana según la cual las descripciones hipotéticamente “asociadas” como sentidos a nombres propios como ‘Aristóteles’ (*v. gr.*, ‘el maestro de Alejandro Magno’) no determinan un único objeto como referencia. El autor incorpora a las descripciones generadas por su propuesta condiciones que precisamente fueron sugeridas por Kripke como explicación de los mecanismos que normalmente permiten que los nombres tengan referencias determinadas (la noción de “relevancia” usada por el autor resume las ideas de la llamada “teoría causal de la referencia”); si, como se suele hacer, se aceptan esas ideas, se aceptará que las descripciones propuestas como sentidos por el autor determinan efectivamente referencias únicas para los ejemplares de los nombres propios usados en preferencias apropiadas (el autor no observa que hay ejemplares, por ejemplo entre comillas, que no tendrán referencia a pesar de aparecer en preferencias típicas).

Ahora bien, la propuesta del autor es claramente vulnerable a otras objeciones kripkeanas a las que apenas se presta atención en su libro (que yo vea, sólo en un párrafo en la p. 240). El famoso argumento fregeano en “Sobre sentido y referencia” se basa en la premisa de que ciertas diferencias en nuestro uso de pares de expresiones correferenciales en la atribución de actitudes proposicionales han de explicarse como diferencias en algún componente del significado de esas expresiones; a ese componente del significado se lo llama sentido, y es el responsable de la identidad o la diferencia de significado de pares de expresiones correferenciales. Si la descripción de más arriba (junto con una indicación de la referencia de ‘ese ejemplar de ‘Aristóteles’’) da el sentido de un ejemplar de ‘Aristóteles’, entonces (algún ejemplar de esa descripción) ha de tener el mismo significado que ese ejemplar (pues tienen la misma referencia, por hipótesis). Sin embargo, no son intercambiables en todos los contextos, en contra de lo que cabría esperar si fueran sinónimos. En particular,

Kripke observó que los nombres y las descripciones que dan su sentido fregeano “tradicional” no son intercambiables en ciertos contextos modales; adaptando su ejemplo: aunque ‘Es posible que Aristóteles no fuera idéntico a Aristóteles’ es falsa, la larga oración ‘Es posible que Aristóteles no fuera idéntico al objeto cuyo etiquetado con ‘Aristóteles’ es relevante en el contexto en que se ha emitido ese ejemplar de ‘Aristóteles’ (donde ‘ese ejemplar de ‘Aristóteles’ refiere al segundo ejemplar de ‘Aristóteles’ en la primera oración) es verdadera. Parece inevitable concluir que la teoría del autor es deficiente como teoría del sentido de los nombres propios. (Nótese que esta objeción es paralela a la objeción usual a la teoría de Reichenbach para los deícticos. Por ejemplo, un ejemplar de ‘yo’ no puede significar ‘la persona que emite esto’ (donde ‘esto’ refiere a la proferencia en que aparece), pues ‘Es posible que yo no fuera yo’ es falso, pero ‘Es posible que yo no fuera la persona que emite esto’ es verdadero.)

Algunos neofregeanos niegan que la relación de “dar el sentido” que se da entre una descripción “asociada” a un nombre y ese nombre sea la relación de sinonimia. Incluso si el autor recurriera a esta estrategia, otra grave objeción la proporcionaría el hecho de que su teoría tiene la consecuencia antiintuitiva de que no hay ningún par de nombres o términos de género natural diferentes (ya sean expresiones-tipo o expresiones-ejemplar) que tengan el mismo significado. Por ejemplo, ‘Aristóteles’ y ‘Aristotle’ no tienen el mismo significado, pues a pesar de ser correferenciales, la regla que da su sentido es diferente; lo mismo ocurre con cualquier par de ejemplares suyos, pues tendrán asociadas descripciones diferentes. La situación es idéntica con ‘arvejo’ y ‘guisante’, o con ‘asno’ y ‘burro’. (Si la propuesta se extiende al sentido de otros términos, como por ejemplo ‘albéitar’ y ‘veterinario’, tendrá también presumiblemente la consecuencia de que no son sinónimos.) Más aún, incluso ejemplares distintos de una misma expresión-tipo tendrán distinto sentido, aunque intuitivamente muchos pares de tales ejemplares tienen el mismo significado.

La propuesta del autor sobre los sentidos de los nombres propios hay que relacionarla con su propuesta acerca de cómo entender el llamado “enigma de Kripke”, que cabe llamar muy propiamente (Kripke lo hace ocasionalmente) ‘paradoja’: una paradoja derivable a partir de ciertos principios incontrovertibles

subyacentes en nuestra práctica de atribución de creencias, en nuestro uso de la noción de “creencia”. Entre estos principios están los *principios de desentrecomillado* (uno para cada lengua); el del español dice que todas las oraciones que resultan de poner una oración declarativa española en lugar de ‘*p*’ en el esquema siguiente son verdaderas:

si un hablante normal del español asiente sinceramente a ‘*p*’, entonces cree que *p*.

Además de los principios de desentrecomillado, en la derivación de la paradoja se usa el *principio de traducción*: si una oración de un lenguaje expresa una oración verdadera en ese lenguaje, entonces una traducción de esa oración a otro lenguaje expresa una oración verdadera en ese otro lenguaje. Estos principios permiten derivar pares de atribuciones de creencias inconsistentes a un mismo sujeto, tales como ‘Pedro cree que Londres tiene parajes lindos’ y ‘Pedro cree que Londres no tiene parajes lindos’ bajo supuestos concebibles acerca de Pedro, a pesar de que intuitivamente no hay razones para censurarlo. Pedro, en una situación como la descrita por Kripke, puede ser un hablante monolingüe del español que asiente a ‘Londres tiene parajes lindos’, de lo cual podemos concluir, usando el principio de desentrecomillado para el español, que Pedro cree que Londres tiene parajes lindos; por otro lado, sin dejar de asentir a aquella oración, puede viajar a una parte desagradable de Londres, aprender inglés sin diccionario y asentir a la oración inglesa ‘London has no nice places’, de lo cual podemos concluir, usando primero el principio de desentrecomillado para el inglés y luego el principio de traducción, que Pedro cree que Londres no tiene parajes lindos. Pedro cree que Londres tiene parajes lindos y Pedro cree que Londres no tiene parajes lindos.

La propuesta del autor para aliviar la complejidad en que nos sume la paradoja es que esas atribuciones son atribuciones *de re*, y que en general todas las atribuciones *de re* (donde el designador que refiere a la *res* tiene una aparición puramente extensional y cuantificable existencialmente), tienen la forma lógica o semántica profunda de ciertas cuantificaciones existenciales. (Es discutible que las dos atribuciones a Pedro recién mencionadas hayan de construirse en todos los contextos imaginables como atribuciones *de re*; lo que no parece discutible, como veremos,

es que la paradoja puede formularse para atribuciones *de dicto*.) Respectivamente, ‘Pedro cree que Londres tiene parajes lindos’ y ‘Pedro cree que Londres no tiene parajes lindos’ tienen las formas:

- (2) Hay un α tal que (Pedro cree que $\# \alpha$ tiene parajes lindos $\#$ y $\Delta(\alpha, \text{Londres})$);
- (3) Hay un α tal que (Pedro cree que $\# \alpha$ no tiene parajes lindos $\#$ y $\Delta(\alpha, \text{Londres})$).

‘ $\#$ ’ es una “comilla de sentidos”, lo cual quiere decir que las expresiones que aparecen entre dos signos ‘ $\#$ ’ denotan (o denotan con respecto a una asignación a las variables) sentidos (en el caso de las expresiones constantes, el sentido asociado a ellas). ‘ Δ ’ denota la relación que se da entre un sentido y un objeto cuando aquél “presenta” a éste. La idea es que la paradoja no se reproduce al nivel de la “forma semántica” de las atribuciones, pues (2) y (3) no atribuyen creencias inconsistentes a Pedro, sino meramente una relación con sentidos distintos que “presentan” a Londres de maneras diferentes. En general, todas las atribuciones *de re* tienen esta forma real; “si el hablante utiliza este recurso [la atribución *de re*] es porque no está en disposición de utilizar un término que refiera directamente al constituyente proposicional que define con precisión el pensamiento atribuido al sujeto, quizás por desconocimiento —o porque no quiere, por las razones que sean” (p. 266).

No está claro cuál es la naturaleza de la propuesta del autor, pues no está claro cuál sea la naturaleza de la relación de “creencia” que aparece en las “formas semánticas” que utiliza. Como Kripke subrayó, su paradoja sólo puede resolverse modificando los principios que aceptamos acerca de la noción de creencia (y que permiten derivar las atribuciones de creencias inconsistentes), o bien desechando la noción por completo. Parece que el autor opta por la primera alternativa, usando una noción modificada de creencia; pero si es así no deja claro cuáles son los principios pertinentes que rigen su uso, y por tanto no está claro si una paradoja similar a la de Kripke puede derivarse acerca de esa noción. (Un hablante normal del castellano no sabría cómo usar la notación “creer que” + “comillas de sentidos” del autor. No sabría, por ejemplo, si del asentimiento de Pedro a ‘London

has no nice places' podría concluir 'Pedro cree que #Londres no tiene parajes lindos#' (además de poder concluir (3)). Si puede hacerlo, una aparente paradoja para la nueva noción no está lejos, pues seguramente de su asentimiento a 'Londres tiene parajes lindos' podremos concluir 'Pedro cree que #Londres tiene parajes lindos#'.)

En opinión del reseñador, hay problemas más específicos en la propuesta del autor. Kripke señaló que no se podía buscar alivio a su paradoja en la idea de que Pedro estaba relacionado con dos sentidos distintos (que presentan a Londres), entre otras cosas porque 'Londres' y 'London' son intuitivamente sinónimos además de correferenciales. No sólo eso: Kripke describe una situación en la que Pedro asocia con ellos las mismas descripciones, si bien una "en español" y la otra "en inglés" (sin que Pedro sepa que las dos descripciones son sinónimas); es decir, que (según la teoría atribuida a Frege) son sinónimos en el idiolecto de Pedro. Pero ya vimos que según el autor no hay dos ejemplares de un nombre-tipo con el mismo sentido, ni por tanto con el mismo significado. Como dijimos, ésta es una tesis que contradice abiertamente nuestras intuiciones acerca de la noción de significado, si bien en este caso parece tener al menos la virtud de ofrecer al fregeano (*lato sensu*) una salida no imaginada por Kripke: la persona que hace las atribuciones alude a un sentido con el ejemplar de 'Londres' en su primera atribución y a otro con el ejemplar de la segunda atribución; incluso si funde sus dos atribuciones en una, 'Pedro cree que Londres tiene parajes lindos y Londres no tiene parajes lindos', el primer ejemplar de 'Londres' alude a un sentido y el segundo a otro.

Sin embargo, creo que ni siquiera es ésta una salida que el autor tenga siempre disponible. Usando los principios del razonamiento de la paradoja (bajo los supuestos típicos acerca de Pedro) más nuestro conocimiento de la lógica y de las transformaciones sintácticas permisibles en español podemos derivar la atribución 'Pedro cree que Londres tiene y no tiene parajes lindos'. (Los ejemplares de) esta atribución sólo contiene(n) un ejemplar de 'Londres'. Las instrucciones del autor para hallar la forma semántica profunda de las atribuciones *de re* (añadir comillas de sentidos y cuantificar existencialmente sobre los sentidos asociados a los nombres que aparezcan en la cláusula de atribución) generan la cuantificación

- (4) Hay un α tal que (Pedro cree que $\# \alpha$ tiene y no tiene parajes lindos $\#$ y $\Delta(\alpha, \text{Londres})$),

que difícilmente puede leerse de otra forma que atribuyendo nuevamente una creencia contradictoria a Pedro. Debe observarse además que incluso si el autor, al conocer esta objeción, ofreciera instrucciones adicionales para computar una forma distinta para la atribución, el ejemplo parece refutar su concepción de las atribuciones *de re*. El hablante que dice ‘Pedro cree que Londres tiene y no tiene parajes lindos’ no lo hace por imposibilidad de referirse a un supuesto sentido que hace verdadera a (4); simplemente no hay un sentido tal: no hay un sentido para Pedro mediante el que éste acepte una contradicción.

En cualquier caso, es claro que la propuesta del autor no sirve para explicar todas las versiones posibles de la paradoja de Kripke, pues hay algunas en las que es claro que las atribuciones de creencias inconsistentes no son *de re*. Es concebible, por ejemplo, que Pedro asienta a ‘Papá Noel lleva un cinturón negro’ y también a ‘Santa Claus does not wear a black belt’, en cuyo caso podremos llegar al par de atribuciones ‘Pedro cree que Papá Noel lleva un cinturón negro’ y ‘Pedro cree que Papá Noel no lleva un cinturón negro’. Pero la solución a la paradoja aquí no puede ser que ambas atribuciones son correctas pues tienen las formas

- (5) Hay un α tal que (Pedro cree que $\# \alpha$ lleva un cinturón negro $\#$ y $\Delta(\alpha, \text{Papá Noel})$);
(6) Hay un α tal que (Pedro cree que $\# \alpha$ no lleva un cinturón negro $\#$ y $\Delta(\alpha, \text{Papá Noel})$).

(5) y (6) son falsas, pues no hay ningún objeto presentado al mismo tiempo por el sentido de Pedro para ‘Papá Noel’ y el sentido de Pedro para ‘Santa Claus’. Para empezar, no hay ningún objeto presentado por ‘Papá Noel’ o ‘Santa Claus’.

La última cuestión sobre la que haremos algunos comentarios críticos es la reconstrucción por el autor del supuesto argumento de Wittgenstein que lleva a éste al rechazo del realismo por representación y con ello al fenomenalismo. El autor apenas cita algún texto de Wittgenstein que sólo muy indirectamente parece tener algo que ver con dicha reconstrucción. Pero no nos ocuparemos de cuestiones exegéticas aquí, sino de evaluar la reconstrucción

misma. (Si estamos en lo cierto al creer que el argumento reconstruido es deficiente, entonces naturalmente esto será una razón indirecta para no atribuirlo a Wittgenstein, siquiera sea por buena voluntad hacia él.) El argumento reconstruido es el siguiente. Consideremos los casos particulares del esquema

(V) S es verdadero en L si y sólo si p,

“donde en lugar de ‘S’ colocamos el nombre de un enunciado significativo cualquiera de L (por ejemplo, su cita [entrecomillado]), y en lugar de ‘p’ ponemos ese mismo enunciado” (p. 369). Estos casos particulares son, según el autor, verdades analíticas (ésta es la primera premisa del argumento). Sin embargo, no serían analíticas “si aquello secundariamente significado por lo que sustituye a ‘p’, de cuyo darse o no depende la verdad del enunciado nombrado por lo que sustituye a ‘S’, fuese algo accidentalmente relacionado con lo que entendemos al comprender el enunciado” (*ibid.*) (este condicional es la segunda premisa del argumento). Pero lo supuestamente significado secundariamente por lo que sustituye a ‘p’ estaría “accidentalmente relacionado” con lo entendido al comprender el enunciado (tercera premisa). Un fácil *modus tollens* nos lleva a una contradicción y a la conclusión deseada, según el autor, por Wittgenstein: las supuestas significaciones secundarias de los enunciados no existen.

La tercera premisa parece muy discutible. Naturalmente, cómo se la entienda depende muy especialmente de cómo se entiendan las vagas expresiones ‘accidentalmente relacionado’ y ‘lo entendido al comprender el enunciado’. Dada una comprensión intuitiva de estas expresiones, la tercera premisa parece falsa. El hecho de que la mesa tenga cuatro patas (si es que existe) no parece accidentalmente relacionado con lo que yo entiendo al leer ‘la mesa tiene cuatro patas’; pues es precisamente el hecho de que la mesa tenga cuatro patas (sean lo que sean los “hechos”) lo que convierte en verdadero a eso que yo entiendo, y lo que lo haría verdadero en cualquier circunstancia posible en que la mesa tuviera cuatro patas; la relación entre el hecho y lo entendido parece pues una relación modal esencial. Por tanto, si hemos de aceptar esa premisa como tal, un argumento independiente parece necesario.

Obsérvese sin embargo que si en lugar de ‘lo entendido al comprender el enunciado’ ponemos ‘el enunciado’ lo que resulta

sí es verdadero: el enunciado está sólo accidentalmente relacionado con sus condiciones de verdad, pues podría haber tenido otras. Pues bien, el autor parece usar de forma intercambiable en este contexto ‘lo entendido al comprender el enunciado’ y ‘el enunciado’, ya que, inmediatamente después de formular el argumento, reformula su segunda premisa con estas palabras: “Las condiciones de verdad de los enunciados no pueden estar accidentalmente relacionadas con los enunciados [...] cuando ‘verdad’ es un concepto casi redundante” (p. 369). Como ‘lo entendido al comprender el enunciado’ y ‘el enunciado’ se usan de forma intercambiable, no es sorprendente que se acepte la tercera premisa, en virtud de una falacia de equívocación.

Que el autor parece haber sido víctima de una falacia de este tipo se ve aún más claramente si se observa que la primera premisa es obviamente falsa (aunque parece verdadero que los casos particulares de

(V') Lo entendido al comprender que p es verdadero si y sólo si p

son verdades analíticas). Para ver que es obviamente falsa, nótese que los casos particulares de (V) ni siquiera son verdades necesarias. Parece claramente absurdo pensar que mi idiolecto tiene como propiedad esencial que ‘la mesa no tiene cuatro patas’ significa que la mesa no tiene cuatro patas. Por tanto, hay circunstancias posibles en que ‘la mesa no tiene cuatro patas’ es verdadero en mi idiolecto (porque en esa circunstancia posible el enunciado significa en mi idiolecto que dos más dos son cuatro) pero donde la mesa sigue teniendo cuatro patas. (Otra posible fuente de confusión para el autor quizá sea el hecho de que la tesis general de que los casos particulares de (V) son verdaderos sí parece una verdad analítica.)*

MARIO GÓMEZ-TORRENTE

* Agradezco a Maite Ezcurdia, Max Kölbel, Alejandro Tomasini y especialmente a Ignacio Vicario sus comentarios a una versión anterior de este texto.