

VINCENZO P. LO MONACO

UN EXAMEN DE LA ESTRUCTURA LÓGICA DE LAS ORACIONES DE CREENCIA DESDE LA INTERPRETACIÓN RADICAL DE DAVIDSON

Resumen. En este trabajo exploramos la dimensión lógica del concepto de creencia en la teoría de la interpretación radical de Davidson. En particular, tratamos de reconstruir el punto de vista davidsoniano sobre la forma lógica de los enunciados de creencia. Nuestro propósito es mostrar que la aproximación reconstruida no puede superar las dificultades formales que surgen en la estructuración lógica del concepto de creencia; en particular, aunque pueda tener éxito en evitar el surgimiento de la paradoja del “intérprete omnisciente”, no logra superar la antinomia tarskiana de la autorreferencia

Palabras clave: Creencia, teoría lógica, interpretación radical.

AN EXAMINATION OF THE LOGICAL STRUCTURE OF BELIEF STATEMENTS BASES ON DAVIDSON’S RADICAL INTERPRETATION

Abstract. In this paper, we explore the logical dimension of the belief concept in Davidson’s radical interpretation theory. Specifically, we propose reconstructing the davidsonian point of view regarding to the logical form of belief statements. We aim at showing that the reconstructed approach cannot solve formal difficulties connected with the logical structuration of the belief concept. In particular, we attempt to show how, although Davidson’s proposal has succeed in avoiding the “omniscient interpreter” paradox, it does not work for solving the tarskian self-reference antinomy.

Keywords: Belief, logical theory, radical interpretation.

La discusión en torno al valor y la naturaleza de nuestras creencias está estrechamente vinculada en la filosofía contemporánea a la discusión en torno al concepto de verdad. En efecto, el empleo del término “creencia”, tanto en filosofía como en el lenguaje común, presupone una clase de situaciones en las cuales comprender a los otros parece requerir no sólo que se les atribuya ideas y que se pueda reconstruir el estado mental interno que mantienen cuando conjugan y expresan tales ideas, sino también que disponemos de mecanismos para determinar si estas últimas son correctas o erróneas, verdaderas o falsas; en pocas palabras, podemos averiguar si nuestras creencias son racionales y justificadas. Si la creencia es una capacidad típicamente humana, deberíamos estar en condiciones de proporcionar una explicación unitaria de tal capacidad por medio de un análisis conceptual del término que permita no sólo verificar la existencia de ese estado mental y distinguirlo de otras formas de pensamiento, sino principalmente establecer cómo han de tratarse las oraciones que la expresan. Este último punto, en particular, plantea el problema de las condiciones semánticas necesarias para la atribución de creencias a un individuo por medio de la “reconstrucción lingüística” de las oraciones que emite.

Una de las propuestas más reconocidas en el panorama de la filosofía actual como intento de dar una explicación satisfactoria de los fenómenos lógico-lingüísticos envueltos en la atribución de creencias, se debe sin duda a Donald Davidson y es conocida bajo el nombre de “teoría de la interpretación radical”. Nuestro propósito en este trabajo es reconstruir, a partir de la teoría general de Davidson, su punto de vista sobre la estructura lógica de las oraciones de creencia y examinar si cumple o no con los *desiderata* de un sistema formal.

Según la teoría davidsoniana de la interpretación radical, el problema de la interpretación de enunciados proferidos por un hablante ha de replantearse como el problema de construir una teoría de los significados implícitos en las pro-

ferencias o emisiones lingüísticas y una teoría en torno a las creencias del hablante. Se manifiesta aquí con absoluta nitidez la interdependencia entre la atribución de creencias y la comprensión de enunciados: para comprender un enunciado es necesario formular hipótesis sobre los significados de las palabras empleadas por el hablante; mas no podremos jamás tener una idea de lo que sus palabras significan si no atribuimos previamente un *corpus* de creencias a nuestro hablante; a su vez, si nada sabemos acerca de los significados de las palabras proferidas, deberemos presumir que las creencias del hablante no son a la postre muy diferentes de las nuestras; deberemos, en otras palabras, *maximizar* el acuerdo entre sus creencias y las nuestras. De ahí el papel principalísimo del “principio de caridad”.¹ La caridad no es simplemente una opción; antes bien,

[...] nos es impuesta; nos guste o no, si queremos comprender a los demás, debemos darlos por acertados en la mayor parte de los asuntos. Si somos capaces de producir una teoría que reconcilie la caridad y las condiciones formales para una teoría, hemos hecho todo lo que pueda hacerse para asegurar la comunicación. No hay nada más que sea posible, ni hace falta nada más.²

¹ Sobre la trascendencia de este principio, Davidson ha afirmado: “El principio de caridad juega un papel crucial en el método de Quine y un papel aún más importante en mi propia variante. En ambos casos, el principio enseña al intérprete cómo traducir o interpretar de modo tal que algunos de sus propios criterios de verdad queden confirmados en la estructura de oraciones que el hablante reputa verdaderas. El propósito del principio es hacer inteligible al hablante, dado que las múltiples desviaciones respecto de la coherencia y de la corrección no permiten un terreno común desde el cual valuar el acuerdo o la diferencia. Formalmente, el principio de caridad ayuda a resolver el problema de la interacción del significado y la creencia restringiendo los grados de libertad concedidos a la creencia mientras se determina el modo de interpretar las palabras” (“A Coherent Theory of Truth and Knowledge”, en Lepore, E. (ed.), *Truth and Interpretation*, Oxford, Blackwell, 1986, p. 237. La traducción es nuestra).

² Davidson, D.: “On the Very Idea of a Conceptual Scheme”, en *Proceedings and Addresses of the A.P.A.*, 1973-74; recogido en Davidson, Donald: *Inquiries into Truth and Interpretation*, Oxford, Oxford University Press, 1984. Trad. cast. en *De la verdad y de la interpretación*, Barcelona, Gedisa, 1990, p. 202.

Una vez sentada la idea de compenetración entre lenguaje y estados cognitivos, la unidad mínima de significado para la interpretación es dada por un conjunto de creencias entrelazadas, es decir por la atención a aquellas oraciones que cada hablante considera verdaderas. Esto muestra cómo Davidson intenta posicionar su teoría de la interpretación dentro de una teoría general del comportamiento humano: podemos comprender y explicar lo que hace el hablante cuando sabemos qué cosas cree y cómo sus creencias se conectan con otras creencias. Por tanto, el principio de caridad obedece a una regla de benevolencia interpretativa y a una proyección de nuestra racionalidad, dos condiciones que tienden a maximizar el acuerdo y minimizar las diferencias. La condición para comprender al otro consiste en reconocerlo no sólo como sujeto dotado de palabra sino sobre todo como ser racional.³ Atiéndase, por ejemplo, a la interdependencia entre significados y creencias. Según Davidson, si no accedemos a éstas tampoco podremos determinar aquéllos y, viceversa, si no estamos en condiciones de atribuir creencias a un hablante resulta imposible establecer qué significan sus palabras. Cómo esto pueda en definitiva hacerse es difícil decirlo, pero Davidson intenta revelarlo en diversas partes de su obra, especialmente en los ensayos “Al decir *that*”,⁴

³ Ambas condiciones son comprensibles sólo en el trasfondo de un interés práctico, sin veleidades epistémicas. Esto es, reconocemos y comprendemos al otro no porque cognitivamente idéntico a nosotros, sino en tanto le atribuimos formas de coherencia, creencias y finalidades semejantes a las nuestras. No obstante, irónicamente respecto de sus intenciones, parece subsistir en la hipótesis de Davidson cierto forzamiento interpretativo; el “acto comunicativo” de Davidson, que aproxima entre sí a los dos interlocutores y conecta a cada uno con un mundo compartido, es sólo posible gracias al concepto objetivo de verdad que nos indica cuándo el lenguaje es usado correctamente y, por ende, cuándo la traducción ha resultado exitosa. Con esto, sin embargo, no se entiende cómo pueda ser posible el cambio lingüístico ni, mucho menos, cómo podamos comprender creencias para nosotros “falsas”. Sobre este último punto, puede verse Kirk, Ludvig, *Donald Davidson*, Nueva York, Cambridge University Press, 2003, pp. 107-109.

⁴ *Synthese*, 18(1968-69), pp. 130-146, recogido en Davidson, *De la verdad y... cit.*, pp. 108-122.

“La creencia y el fundamento del significado”,⁵ “Interpretación radical”⁶ y “Meaning, Truth and Evidence”.⁷ La idea clave queda resumida en el método de la interpretación radical:

[...] Primero buscamos la mejor manera de ajustar nuestra lógica, hasta el punto requerido para obtener una teoría que satisfaga la Convención T, en el nuevo lenguaje; esto puede significar leer la estructura lógica de la teoría de la cuantificación de primer orden (más la identidad) dentro del lenguaje, no tomando las constantes lógicas una por una, sino tratando a esta gran porción de la lógica como una gran red que se instalará sobre el lenguaje en una única operación relámpago. La evidencia consiste aquí en clases de oraciones que siempre se consideran verdaderas y siempre se consideran falsas por casi todas las personas casi todo el tiempo (verdades potenciales) y en patrones de inferencia. El primer paso identifica predicados, términos singulares, cuantificadores, conectivos e identidad; en teoría, establece cuestiones de forma lógica. El segundo paso se concentra en oraciones con deícticos; aquellas oraciones que a veces [son] consideradas verdaderas y a veces falsas de acuerdo con cambios descubribles [*sic*] en el mundo. Este paso en conjunción con el primero limita las posibilidades de interpretar predicados individuales. El último paso se ocupa de las oraciones restantes, aquellas para las cuales no hay un acuerdo uniforme, o cuyo valor de verdad considerado no depende sistemáticamente de cambios en el medio ambiente.⁸

Como puede apreciarse, Davidson integra su teoría de la comprensión lingüística a una teoría más amplia de la interpretación de la acción no lingüística. En esta teoría, las emisiones lingüísticas son acciones, las acciones son eventos y los eventos pueden ser descritos de diversas maneras. Así, el proceso de interpretar las emisiones lingüísticas de los hablantes de otras lenguas es parte del proceso de interpre-

⁵ *Synthese*, 27(1974), pp. 309-323, recogido en Davidson, *De la verdad y...*, cit., pp. 141-163.

⁶ *Ibidem*.

⁷ En Barrett, R. y Gibson, R. (eds.): *Perspectives on Quine*, Oxford, Blackwell, 1993.

⁸ Davidson, D., “Radical Interpretation”, *Dialectica*, 27(1973), pp. 313-328; reimpresso en Davidson, *De la verdad y...*, cit., p. 147.

tar sus acciones, de modo tal que interpretar una emisión del hablante equivale a re-describirla *qua* acción en términos del metalenguaje, esto es, insertándola dentro del esquema “dice que *p*”, donde ‘*p*’ es una oración del metalenguaje. Por esta misma razón, para comprender las emisiones del hablante, el intérprete radical de Davidson construye una oración-T en forma tal que “*s* es verdadera sii *p*”, donde *p* establece las condiciones de verdad de *s*. “A diferencia de la teoría de Quine,” –escribe Susan Haack– “la de Davidson *no* pretende poner en correlación las frases expresadas en la lengua del hablante con las frases expresadas en la lengua del intérprete, ni combinar relaciones de asentimiento con condiciones de asentimiento; pretende combinar condiciones de asentimiento con condiciones de *verdad* hallando equivalencias al estilo de Tarski en las que, si la frase expresada de la parte de la izquierda es considerada verdadera por el hablante, la frase empleada de la derecha *es verdadera*”.⁹ Por otra parte, dada la estrecha conexión existente entre considerar verdadera a una oración y creer en ésta, se establece una íntima relación entre la atribución de creencias al hablante nativo y la asignación de significados a sus emisiones lingüísticas.

Hemos hablado de la teoría; hablemos ahora de la estrategia de Davidson. La estrategia de la interpretación radical consiste básicamente en aceptar la perspectiva de la tercera persona, la perspectiva del proceso por el que cada uno de nosotros trata de entender a los demás atribuyéndoles creencias que hagan inteligible su conducta. Pero para ello es menester asumir dos supuestos sin los cuales la interpretación no sería posible, a saber: que las creencias sobre el mundo que asignamos a los hablantes son en general objetivamente verdaderas, en el sentido de que al atribuir significados a sus emisiones presuponemos que sus estados mentales se corresponden en principio con hechos y situaciones

⁹ Haack, S.: *Evidence and Inquiry. Towards Reconstruction in Epistemology*, Oxford, Blackwell, 1993; trad. cast., *Evidencia e investigación. Hacia la reconstrucción en epistemología*, Madrid, Tecnos, 1997, p.96

del mundo mismo (En pocas palabras, el hablante cree que las oraciones expresadas por medio de sus emisiones son verdaderas); segundo, que la atribución de creencias a los hablantes está regida por los principios de la racionalidad y la coherencia. Si las emisiones son acciones y replican estados mentales, entonces habremos de suponer que la conducta del agente es racional en la medida en que su acción puede explicarse haciendo referencia a sus creencias: “un hablante sostiene que una oración es verdadera debido a lo que la oración (en su lenguaje) significa, y debido a lo que él cree”.¹⁰ Sin embargo, esta última conclusión carece de sentido sin la premisa adicional de que la justificación de una creencia depende de su coherencia con un cuerpo sustancial de otras creencias: “... lo que distingue a una teoría de la coherencia” –escribe Davidson– “es simplemente su pretensión de que nada puede contar como una razón de una creencia excepto que otra creencia”.¹¹ Resulta entonces innegable pensar que la noción de creencia no sólo constituye la pieza principal del bagaje de conceptos semánticos del intérprete radical, sino que se convierte en la noción central de la teoría del lenguaje de Davidson.

Esto último nos lleva razonablemente a preguntarnos qué son realmente las creencias. La posición de Davidson sobre este punto no es del todo clara; incluso en el transcurso de un mismo párrafo llega a proporcionar dos versiones bastantes distintas de lo que son las creencias: primero afirma que son “oraciones que son verdaderas para alguien que las entiende”;¹² y luego, que son “estados de las personas que son causados por –y que causan a su vez– eventos internos y externos al cuerpo de sus poseedores”.¹³ Davidson habla entonces de dos cosas en principio distintas: habla primero de *contenidos* de creencias que podemos evaluar semánticamente en el contexto de sus relaciones con (los

¹⁰ Davidson, “Interpretación radical” ..., cit., p. 145.

¹¹ Davidson, “Afterthoughts. 1987”, en Malachowski, A.R. (ed.): *Reading Rorty*, Oxford, Blackwell, 1990, p.135.

¹² Davidson, “A Coherence Theory of Truth and Knowledge”, cit., p. 121.

¹³ *Ibidem*.

contenidos de) otras creencias y que son fijados por la situación o evento externo que las causa. Pero también habla de *estados* de creencias como eventos mentales regidos por normas objetivas de coherencia que atribuimos a un sujeto para justificar y hacer inteligible su conducta. Y en esta situación de incertidumbre el uso de la expresión “actitud proposicional” como referida a una creencia no parece servirnos de mucho, pues en el discurso de Davidson la expresión parece referirse indistintamente sea a contenidos que a estados de creencia.

Sin embargo, a pesar de la dificultad no parece haber nada contradictorio en esto. Si reparamos en el *desideratum* davidsoniano de una “teoría unificada del significado y de la acción”, podemos perfectamente aceptar que la interpretación consiste en un proceso global en el que asignamos condiciones de verdad a los contenidos de creencias y atribuimos paralelamente estados de creencias al agente en un único acto recíprocamente condicionado. ¿Son entonces las creencias pensamientos verbalizados hechos públicos como contenidos de los estados mentales de algún sujeto?

Pienso que la de arriba es una buena definición de “creencia” que encaja satisfactoriamente en el discurso de Davidson, aunque tengo serias dudas en torno a si él hubiese podido compartirla, y más dudas aún sobre la necesidad de admitirla para conferir coherencia a su razonamiento. En efecto, en “A Coherence Theory of Truth and Knowledge” hay un par de pasajes muy reveladores al respecto:

[...] cualquiera que tenga pensamientos, y en particular cualquiera que se pregunte si tiene alguna razón para suponer que está generalmente en lo cierto acerca de la naturaleza de su entorno, ha de saber qué es una creencia y cómo han de detectarse e interpretarse las creencias en general. Puesto que éstas son hechos perfectamente generales que no podemos dejar de usar cuando nos comunicamos con otros, o cuando tratamos de hacerlo, o incluso cuando simplemente pensamos que lo estamos haciendo [...].¹⁴

¹⁴ *Ibid.*, p. 127 (La traducción es nuestra).

La argumentación parece aquí tener una orientación negativa con respecto a nuestra definición de arriba, pues es obvio que se considera a la creencia como una noción primitiva, una noción intuitivamente o trascendentalmente clara en sentido kantiano, no susceptible de ser definida en términos de otras nociones más claras o elementales. Sin embargo, es igualmente claro que cuando en el texto se emplea el término creencia Davidson no se está refiriendo a estados de creencias en cuanto eventos mentales, sino que lo hace en su acepción de contenidos de creencias en cuanto pensamientos verbalizados, en cuanto emisiones de oraciones; ¿qué otra cosa sino ésta podría entenderse por “hechos generales que no podemos dejar de usar cuando nos comunicamos”?

La dificultad en el uso de la noción puede mejor apreciarse en este otro pasaje unas pocas líneas más abajo:

Puesto que estas relaciones [entre oraciones verdaderas] se traducen directamente en creencias, es fácil ver cómo el significado depende de la creencia. La creencia, sin embargo, depende directamente del significado, pues el único acceso a la fina estructura e individuación de las creencias lo constituyen las oraciones que los hablantes y sus intérpretes usan para expresar y describir creencias.¹⁵

Nótese que en el texto se emplea cinco veces la palabra creencia, aunque es claro que no se usan en la misma acepción. En la frase “se traducen directamente en creencias”, el término parece indicar contenidos de creencias, mientras que en “el significado depende de la creencia” y en “La creencia [...] depende igualmente del significado” es razonable pensar que se refiere a estados de creencias, y no a contenidos de creencias, en consonancia con la descripción de lo mental que Davidson mantiene cuando afirma que “... las creencias y las demás actitudes pueden “estar en la cabeza” aun cuando en parte sean identificadas como tales

¹⁵ *Ibid.*, p. 128 (La traducción es nuestra).

en términos de lo que no está en la cabeza”.¹⁶ Pienso que el uso de la expresión en la frase “la fina estructura e individuación de las creencias” tampoco es el mismo que en “expresar y describir creencias”; mientras que en esta última, “creencia” apunta a “estados de creencia”, en la primera se emplea, de nuevo, el significado de “contenidos de creencia”. Así las cosas, ¿por qué Davidson insiste en sostener que la noción de creencia es intuitiva o naturalmente clara? Creo que en el fondo la respuesta más obvia es la siguiente: aunque la creencia como noción parece poder descomponerse en *contenidos* y *estados*, las relaciones entre estos elementos son científica y epistemológicamente tan intrincadas y confusas que resulta una mejor elección filosófica adoptar la noción como primitiva o inanalizable.

Quiero en este punto hacer mía una suerte de prescripción metodológica que Davidson adopta de Burge y parafrasea en los siguientes términos: “... cada vez que una persona esté equivocada, confundida o parcialmente mal informada sobre el significado de una palabra, está asimismo equivocada, confundida o parcialmente mal informada acerca de cualquier creencia suya que se exprese [...] mediante el uso de esa palabra”.¹⁷ Supongo que, de concederse algún crédito a esta idea, y creo que la idea es básicamente correcta, es hora de dejar de debatir sobre el carácter intuitivo de la noción de creencia, sobre contenidos y estados de creencias, y comenzar a buscar un punto de apoyo más firme. A decir verdad, el mismo Davidson reconoce algo de nuestra conclusión cuando sentencia que “... tener actitudes proposicionales implica aceptar el requisito de la evidencia global”, el cual, dada la preeminencia que aquí concede a la primera persona, sólo puede satisfacerse por el rasero de la imitación social:

¹⁶ Davidson, “Knowing One’s Own Mind”, *Proceeding and Addresses of the American Philosophical Association*, 1987, p. 452.

¹⁷ *Ibid.*, p. 445.

[...] averiguo lo que yo creo de forma muy similar a como lo averigua otra persona: reparando en lo que digo y hago.¹⁸

El hecho de que en el razonamiento de Davidson pueda encontrarse un reconocimiento explícito de la prioridad del lenguaje para la determinación de creencias –a pesar de que generalmente defiende el monismo anómalo como una doctrina epistemológica y ontológica más que lingüística–, indica algo del rol decisivo que se otorga a la interpretación de oraciones. Esta última reflexión hace posible, a mi juicio, reconducir todo el asunto al ámbito adecuado para su análisis: los contextos de actitud proposicional. Si, como quiere Davidson, significado y creencia son conceptualmente interdependientes, en el sentido de que “...interpretar las oraciones tenidas por verdaderas [...] no debe distinguirse de atribuir creencias”,¹⁹ entonces no sólo se hace necesario disponer de una teoría de las creencias, sino levantar asimismo una lógica que la sostenga.

Aunque parezca irónico, los lineamientos de semejante lógica los encontramos en el propio Davidson, pero en trabajos anteriores a los que hemos estado escudriñando, concretamente en “Pensamiento y habla”²⁰ y “Al decir *that*”.²¹ En el primero de los trabajos la creencia se caracteriza específicamente por las maneras generales en que los pensamientos de un hablante son determinados por la asignación de valores de verdad a sus emisiones bajo circunstancias dadas y por su posición en el patrón de las oraciones que tanto intérprete como hablante consideran verdaderas. En el segundo, Davidson se adentra directamente en el análisis lógico de los contextos de actitud proposicional, pasando revista a las teorías del discurso indirecto de Carnap, Church, Scheffler y Quine. Dado el papel central que, como

¹⁸ *Ibid.*, p. 441.

¹⁹ Davidson, “A Coherence Theory...”, cit. p. 123.

²⁰ En Guttenplan, S. (ed.) *Mind and Language*, Oxford, Oxford University Press, 1975; recogido en Davidson, *De la verdad y...* cit., pp. 164-178.

²¹ *Synthese*, 19 (1968-69), pp. 130-146; recogido en Davidson, *De la verdad y...* cit., pp. 108-122.

se ha visto, tienen los contextos de creencia en su teoría de la interpretación, Davidson explora detenidamente la posibilidad de construir una lógica de la creencia de su propia cosecha, concebida como una forma de integrar en la estructura semántica de las oraciones–contenido de los contextos de actitud proposicional la asignación de verdad a esas oraciones según el método de la convención–T de Tarski.

El requisito citativo o inscripcional aplicado a la forma lógica de los contextos de creencia es crucial para lo que Davidson describe como “la posibilidad de caracterizar recursivamente un predicado de verdad que satisfaga los requerimientos de Tarski”.²² El punto de partida de su teoría es, en efecto, la aproximación citativa o inscripcional de la *oratio obliqua* dada por Quine en *Word and Object*. Como se sabe, la idea de Quine consiste en analizar los contextos de actitud proposicional como relaciones monádicas entre un hablante y una oración–contenido, esta última descrita lingüísticamente como inscripción material que hace uso de nombres de letras y signos de puntuación, unificándose el todo por medio del signo de concatenación. Así, por ejemplo, la estructura lógica de una oración del tipo “Davidson cree *that* la creencia es un estado” es identificada en los términos de una relación entre un creyente y el nombre de una oración representada mediante una inscripción notacional en la forma siguiente:

Davidson cree que “L”∩“a”∩“”∩“c”∩“r”∩“e”∩“e”∩
 “n”∩“c”∩“i”∩“a”∩“”∩“e”∩“s”∩“”∩“u”∩“n”∩“”∩
 “e”∩“s”∩“t”∩“a”∩“d”∩“o”.

Las ventajas de la *teoría del deletreo* con respecto a otras formas de describir las oraciones de actitud son múltiples y obvias. En primer lugar, se coordina perfectamente con una teoría recursiva de la verdad en el sentido de Tarski; segundo, permite sin restricciones importantes el uso de los co-

²² Davidson, “Al decir *that*”, cit., p. 114.

nectivos proposicionales y el aparato de la cuantificación; tercero, nos aleja de la tentación de considerar a las oraciones–contenido como frases que refieren o expresan semánticamente *pensamientos* o *proposiciones* en tanto *intensiones* o *entidades abstractas*; finalmente, es adecuada para dar cuenta del fenómeno de la competencia lingüística desde el punto de vista, por ejemplo, de un Chomsky o de un Davidson, esto es, como una explicación del modo como un hablante *idealizado* representa o proyecta su conducta lingüística.

La teoría *in commento* presenta, no obstante, una clara desventaja, pero tan significativa que basta por sí sola para inducirnos a desconfiar de ella, a saber: no pasa el *test* de la traducción; si el éxito de la aproximación citativa descansa en la teoría del deletreo y ésta supone a su vez una determinada materialización inscripcional de la oración–contenido, habrá de presuponerse la relativización de todas la oraciones de actitud –con independencia del lenguaje en que de hecho están formuladas– a un lenguaje particular, cosa que exigiría de inmediato una especial y complicada versión de la sinonimia interlingüística para satisfacer el requisito de la sustituibilidad *salva veritate*. Aún así, dadas las restricciones que derivan del punto de partida, tal versión se vería en notables dificultades a la hora de evitar que una misma oración–contenido pueda tener diferentes significados en diferentes lenguajes, resultado claramente indeseable.²³

El reconocimiento de la posibilidad de admitir inferencias no deseadas es, a juicio de Quine, una consecuencia directa de la indeterminación de la traducción y es básicamente la razón por la que en el curso mismo de *Word and Object* abandona esta versión de la aproximación citativa, refugiándose en una versión menos poderosa pero más sólida, que aplica importantes restricciones a la consecuencia lógica. En esta nueva versión, que preserva la teoría del deletreo, se analiza los contextos de actitud como relaciones entre una persona y un predicado monádico en tanto expre-

²³ Cf. *ibid.*, p. 112.

sión singular inanalizable, para seguir con nuestro ejemplo: “x-cree-*that*-la-creencia-es-un-estado”. Se restablece por esta vía la aplicación universal del modelo a expensas de la consecuencia lógica, pues sería ahora imposible inferir, por ejemplo, “Davidson cree algo” de nuestro predicado monádico de arriba. Y es ahí donde Davidson levanta su voz de protesta, porque es claro que la versión final del análisis quineano no deja espacio a la teoría de la interpretación, en la medida en que elimina casi toda la estructura requerida por la teoría de la verdad: “Puesto que [...] las citas son términos singulares sin estructura semántica significativa, y dado que debe haber un número infinito de citas diferentes, ningún lenguaje que las contenga puede tener un predicado de verdad definido recursivamente”.²⁴

¿Qué es exactamente lo que Davidson rescata de la idea de Quine y cómo lo reelabora? No es difícil colegir lo que rescata: retoma de ésta, como se ha visto, la aproximación citativa al discurso indirecto y el método del delecto; adopta también la hipótesis del *hablante finito*, una criatura con poderes lingüísticos limitados y capaz no obstante de exhibir competencia en un escenario signado por un número indefinidamente grande de oraciones; asume, en fin, una concepción puramente extensional del lenguaje. Menos fácil es sin embargo explicar cómo construye su propia versión citativa. La estrategia davidsoniana en este punto es asumir la primera versión quineana de la aproximación citativa, pero concibiendo los contextos de actitud como una relación triádica entre un hablante, una oración y el sujeto que realiza la emisión de atribución, de modo tal que nuestro ejemplo de arriba pueda ahora parafrasearse como sigue:

Si x = “La creencia es un estado”, entonces
 “ $(\exists x)$ (la emisión x de Davidson y mi emisión y
 nos hace igual-dicentes)”.²⁵

²⁴ *Ibid.*, p. 111.

²⁵ Al lanzar la hipótesis de la aproximación citativa a los contextos de actitud proposicional en general, Davidson trata como emblemático el caso de las

Ahora la nueva paráfrasis sólo requiere algunos añadidos puntuales. Requiere, en primer lugar, interpretar todo el contexto como un acto de habla, como un realizativo en el que se tiene éxito en igualar inscripciones deletreadas de emisiones concretas. Por otra parte, se garantiza el *desideratum* de la determinación de la estructura semántica de la oración, respetando la consecuencia lógica y explicando el fallo aparente en la sustituibilidad *salva veritate*. “Puesto que una emisión de [“Davidson cree *that*”] y cualquier emisión que le siga son lógicamente independientes, no hay razón para predecir, basados sólo en su forma lógica, ningún efecto *particular* sobre la verdad de la primera a partir de un cambio en la segunda”.²⁶ El último añadido concierne a la relación de sinonimia entre las emisiones *x* e *y* de los hablantes, la cual se habrá de entender no en el sentido de la justificación lógica de la noción de identidad de significado interlingüístico, sino “...como una parte inanalizada del contenido de la forma idiomática familiar del discurso indirecto”.²⁷

aserciones, del “decir *that*”, sin abordar específicamente los contextos de creencia, aquellos expresados en el “creer *that*”. Como hubo de advertirlo Quine a propósito de la propuesta de Scheffler, puede que el análisis inscripcional se tope con nuevas dificultades cuando es aplicado a la creencia, dadas las características peculiares que ésta presenta y que la distingue de otras actitudes proposicionales. En particular, la reformulación de la estructura del ejemplo de arriba requiere de un operador de identidad (“Lo que Davidson dice y lo que yo digo son lo mismo”), cosa que Davidson no hace explícita. Desde luego, eso no significa que no pueda producirse una explicación adecuada de las inferencias que involucran cuantificación más identidad en tales contextos, sino simplemente que la misma no se da automáticamente. Para ello, es menester introducir un análisis de la identidad como el que lleva a cabo, por ejemplo, C. Williams en *What is Identity?*, Oxford, Clarendon Press, 1989. Davidson es plenamente consciente de ello, aunque se exime de abordar específicamente esta última cuestión para allanar el camino y facilitar la exposición de su punto de vista. Por idéntica razón, aunque se ha adaptado el análisis de Davidson a los contextos de creencia, se prefiere aquí soslayar este otro aspecto polémico de la cuestión.

²⁶ Davidson, “Al decir *that*”, cit., p. 121. La oración entre corchetes es producto de la sustitución de “Galileo dijo *that*” por nuestro ejemplo de arriba, para efectos de comprensión del discurso.

²⁷ *Ibid.*, p. 118.

Debo sin embargo apresurarme a advertir que ni la interpretación relacional de Quine que establece el carácter descriptivo del operador de creencia, ni la versión pragmatista de Davidson que lo concibe como un *modificador intralingüístico* que hace que el significado de la subordinada sea dado a través de las conexiones inferenciales con otros significados, logran superar las clásicas dificultades que generan los operadores de creencia en la teoría lógica.

A partir de *An Inquiry into Meaning and Truth* de Bertrand Russell, hay al menos dos teorías o maneras de aproximarse al problema de la expresión lógica de las actitudes de creencia. La primera, genéricamente conocida por el nombre de “gramática ortodoxa” o simplemente “análisis relacional”, asume que la preferencia de oraciones que contienen un verbo de actitud proposicional expresa una relación de primer orden entre dos o más objetos, y postula que la proposición expresada mediante la preferencia estaría compuesta de una relación entre el sujeto al que adscribimos la actitud y el objeto al que denota el contenido de la adscripción encabezado por la cláusula “que”. Dependiendo de la forma en que analicemos la constitución del objeto representado por la oración subordinada, la relación podrá ser simple (*monádica*) o múltiple (*diádica, triádica, etc.*).²⁸ El segundo tipo de enfoque, considerado como no relacional o adverbial, analiza la preferencia de oraciones que contienen un verbo de actitud proposicional como la expresión de una oración simple (la subordinada) modificada adverbialmente por el verbo de la actitud en conjunto con el individuo al que la adscribimos,²⁹ o bien como una oración subordinada que modifica adverbialmente la propiedad expresada por el verbo de actitud proposicional.³⁰

²⁸ Ejemplos de análisis relacional son casi todos los tratamientos clásicos de las actitudes proposicionales como los de Carnap y Church.

²⁹ Véase, por ejemplo, Prior A., y Geach P., *Objects of Thought*, Oxford, Clarendon Press, 1971; también Recanati F., *Oratio Obliqua, Oratio Recta. An Essay on Metarepresentation*, Cambridge, MIT Press, 2000.

³⁰ Cf. Kiteley M., “The Grammar of ‘Believe’”, *The Journal of Philosophy*, 61(1964), pp. 244-259.

Adviértase que nuestra cuestión no es ya la constitución de una teoría para interpretar las actitudes proposicionales, sino más bien el problema de la elección de la lógica adecuada para expresar lógicamente los contextos de actitud, cuestión ampliamente independiente de la primera. Pero es ahí justamente donde la distinción entre teorías relacionales y no-relacionales adquiere relevancia para nuestro asunto. En efecto, los enfoques relacionales parecen indefectiblemente ligados a las semánticas filosóficas del lenguaje natural de tipo fregeano, y en particular a aquellas teorías semánticas asociadas con la teoría de modelos.

Nacida con los trabajos de Tarski sobre la semántica extensional para el cálculo de predicados de primer orden, la teoría de modelos alcanza su nivel actual de desarrollo con la teoría de Montague basada en una generalización de la teoría de los mundos posibles que Kripke implementara para las lógicas modales. Una lógica de la creencia anclada en los mundos posibles es un lenguaje proposicional clásico extendido con un operador modal de creencia B que asume como argumentos fórmulas del lenguaje: si α es una fórmula, entonces $B\alpha$ es también una fórmula. La teoría de los mundos posibles comporta que los objetos de la creencia sean proposiciones, por lo que $B\alpha$ pasa a leerse como “es creído que α ”. En esta lógica valen tanto el *axioma de distribución* como la *regla de necesidad*, el primero es generalmente introducido como esquema de axioma para la creencia y se representa simbólicamente

$$\text{AD. } B\alpha \wedge B(\alpha \rightarrow \beta) \rightarrow B\beta$$

para toda fórmula α y β del lenguaje. La regla de necesidad, en símbolos

$$\text{Nec. de } \alpha \text{ sigue } B\alpha,$$

es un equivalente de la regla correspondiente en los sistemas de lógica modal alética.

Simplificando, la idea central es asignar a cada sujeto epistémico un conjunto de mundos que representan –en la formulación de Hintikka,³¹ por ejemplo– los posibles estados de cosas compatibles con lo que es creído por el sujeto, de modo tal que un enunciado α es creído por un sujeto x si, y sólo si, α es verdadero en todos los mundos posibles compatibles con las creencias del sujeto mismo.

La dificultad más seria que las actitudes proposicionales plantean a una solución de tipo modelo–teórica obedece a la así llamada “asunción de omnisciencia lógica”.³² Si los mundos posibles son estructuras semánticas completas y consistentes, es decir estructuras que asignan a cada enunciado un solo valor de verdad, entonces es necesario que todo sujeto epistémico sea lógicamente omnisciente, esto es, que crea todas las fórmulas lógicamente válidas y todas las consecuencias lógicas de sus creencias. Así, por ejemplo, si el enunciado β se sigue lógicamente del enunciado α , entonces “Davidson cree que β ” se seguirá lógicamente de “Davidson cree que α ”. Esto se explica fácilmente si se considera que la regla de necesitación obliga a que el sujeto crea todas las fórmulas válidas y el axioma de distribución se encarga de que las creencias sean fórmulas cerradas respecto de la consecuencia lógica.

En verdad, este último resultado constituye una idealización difícilmente aceptable a los fines de las teorías semánticas de raigambre cognitiva, un pesado fardo incluso para cualquier teoría de la creencia basada en la interpretación del lenguaje natural. En efecto, no sólo en un enfoque modelo–teórico sino en cualquier enfoque que conciba en general a la semántica como teoría de las relaciones entre expresiones lingüísticas y hechos o estados del mundo, difícilmente pueda admitirse que la asunción de la omnisciencia descienda de la teoría semántica misma, que sea una

³¹ Cf. Hintikka, J., “Semantics for Propositional Attitudes”, en Davis, J.W. *et al.* (eds.), *Philosophical Logic*, Dordrecht, Reidel, 1969.

³² Sobre este punto, puede verse Grim, P., “On Omniscience and the ‘Set of all Truths’”, *Analysis*, 5(1990), pp. 271-276.

cuestión de lógica y no una simple cuestión factual. El mismo Davidson hubo de entretenerse un tiempo con la idea del *intérprete omnisciente* para apuntalar el principio de caridad, afirmando que “no hay nada de absurdo en la idea de un intérprete omnisciente”.³³ Pero, además de que tal hipótesis –según lo muestra Susan Haack–³⁴ no es ni necesaria ni suficiente para el objetivo prefijado por Davidson, queda el hecho de que el proyecto davidsoniano de una teoría general del lenguaje natural se revela abiertamente incompatible con una idea de la semántica basada en una teoría del significado del lenguaje –en palabras de Barbara Partee– “...como si fuera hablado por Dios”.³⁵ Si a esto se agrega que Davidson defiende una concepción de la semántica como modelo de la competencia lingüística de los hablantes, entonces podrá tenerse una clara idea del porqué no puede compartir la idea implícita de que “...todos los sujetos de creencia sean sujetos dotados de capacidades lógicas ‘sobrenaturales’”.³⁶

En honor a la verdad, es menester decir que desde la semántica modelo–teórica misma se ha hecho numerosos esfuerzos por evitar la asunción de omnisciencia lógica, pero ninguna de las propuestas –casi todas encaminadas a debilitar la regla de necesidad– ha logrado hasta ahora proporcionar una solución plenamente satisfactoria y definitiva al problema³⁷. Entre las más prometedoras, cabe sin

³³ Davidson, “The Method of Truth in Metaphysics”, en French, P.A., Uehling, T.E. y Wettstein, H.K. (eds.), *Midwest Studies in Philosophy*, II, Minneapolis, The University of Minnesota Press, 1977; reproducido en Davidson, *De la verdad y...*, cit. p. 205.

³⁴ Cf. Haack, *Evidence and Inquiry...*, cit., pp. 92-94.

³⁵ Partee, B., “Belief sentences and the limits of semantics”, en Peters, S. y Saarinen, E. (eds.), *Processes, Beliefs and Questions*, Dordrecht, Reidel, 1982.

³⁶ *Ibidem*.

³⁷ Los diversos intentos emprendidos en esa dirección, desde las lógicas *minimales* de Chellas hasta la lógica del *razonamiento local* de Fagin y Halpern, pasando por la lógica de la *relevancia* de Levesque y la de la *inconsistencia* de Rescher y Brandon, tropiezan invariablemente bien con problemas de inconsistencia lógica, bien con alguna variante de las paradojas epistémicas (Para una descripción de las diversas estrategias empleadas y de las dificultades que se generan, puede verse Thomason R., “Paradoxes and semantic representa-

duda señalar la propuesta de lógica modal intensional que adopta Montagüe.³⁸ En oposición a las semánticas atomistas de tipo informacional, Montague defiende una lógica intensional compatible con un enfoque modelo-teórico. En esta lógica se interpreta a los constructos epistémicos no como predicados, sino como operadores proposicionales, introduciendo un conjunto de variables de tipo sintáctico que permiten cuantificar directamente sobre proposiciones, sin que sea necesario asumir que los términos del lenguaje denoten objetos de las actitudes proposicionales.

Aunque la solución de Montague ha sido en general bien acogida entre los defensores del significado de inspiración modelo-teórica y reúne los *desiderata* metódicos auspiciados por Davidson para una teoría de la interpretación, es no obstante vulnerable en un punto: emplea una lógica de orden superior, con las consabidas dificultades que de ello se desprenden, en particular, la imposibilidad de una axiomatización completa y finita. Y ésta es una consecuencia grave para teorías semánticas como la de Davidson, orgullosamente confesas de su ascendencia tarskiana. En efecto, abandonar la lógica de primer orden equivale a renunciar a la posibilidad de hacer uso de la teoría de la demostración en el tratamiento modelo-teórico de los significados; en otras palabras, implica renunciar a la posibilidad de generar recursivamente las consecuencias lógicas de una teoría. Soslayando las implicaciones ontológicas, lo cierto es que los enfoques de este tipo, aunque parezcan constituir una manera viable de rodear la asunción de omnisciencia lógica, no "...están sujetos" –según nos refiere Davidson– "a la teoría en el sentido de una definición de la verdad que responda a los requisitos de Tarski. Lo que se interpone en el camino es que [...] hay un número infinito de expresiones primitivas; esto entorpece directamente la posibilidad de caracterizar

tion", *Proceedings TARK-86*, 1986, pp. 225-239; también Turner, R.: *Truth and Modality for Knowledge Representation*, Londres, Pitman, 1990).

³⁸ Cf. Montague R., *Formal Philosophy*, New Haven, Yale University Press, 1974 (en especial, pp. 286-302).

recursivamente un predicado de verdad que satisfaga los requerimientos de Tarski”.³⁹

Lo que he planteado hasta aquí es precisamente que la lógica de la creencia de Davidson no parece encajar cómodamente en un enfoque de tipo relacional, ni en el sentido de la arquitectura extensional modelo-teórica clásica ni el sentido de la variante intensional excogitada por Montague, debido respectivamente a la aparición de la paradoja del creyente omnisciente y a un exceso descriptivo que entraba la relación de consecuencia lógica. La cuestión que quiero mostrar ahora es que si la interpretación citativa de las expresiones de creencia que Davidson defiende se identifica con las teorías no-relacionales o adverbiales acerca del análisis de la estructura lógica de las actitudes proposicionales, parece ser igualmente vulnerable en materia de consistencia lógica y semántica.

Tal como se dijo antes, hay dos tipos de teorías adverbiales que no conciben los contextos de creencia como expresión de una relación de primer orden entre dos o más objetos. Las teorías del primer tipo, introducidas originariamente por Kiteley y Matthews, asumen que la expresión “creer *that*” en los contextos de actitud constituye un predicado monádico que expresa la propiedad que resulta adverbialmente modificada por la oración subordinada que le sigue, la cual deja de denotar la relación entre uno o más objetos para convertirse en un simple índice gramatical o marcador sintáctico de un determinado estado mental del sujeto. De modo que nuestro ejemplo de arriba pasaría ahora a parafrasearse como:

[que la creencia es un estado] CREER (Davidson),

donde la oración entre corchetes modifica al predicado CREER, cuyo hueco de argumento es saturado por el individuo Davidson.

³⁹ Davidson, “Al decir *that*”, cit., pp. 113-114.

Escrúpulos mentalistas al margen, no parece haber buenas razones para asimilar el análisis de Davidson a estos nuevos parámetros, en especial si se considera que, a diferencia de su propuesta, estos últimos no requieren forzosamente de una aproximación sintáctica o inscripcional. A decir verdad, aunque presentan fuertes afinidades con las propuestas de tipo citativo en virtud de la inserción de elementos sintácticos en la semántica, estos modelos adverbiales no tratan en realidad a los objetos de creencia como oraciones o expresiones notacionales de un lenguaje público, sino que los conciben como expresiones de un sistema de representación interno.⁴⁰

Cuando Davidson escribe “De la idea misma de un esquema conceptual”, estas veleidades mentalistas no tenían ya la influencia que habían tenido a principios del siglo XX, por lo que no las atacó directamente en su ensayo. Pero las critica abundantemente en su obra posterior, en especial a través de dos de las dogmas fundamentales que las habrían alimentado y alimentarían asimismo su renacimiento *in modern dress*—por ejemplo, en Fodor o en Dreske—, a saber: la creencia en que los conceptos son contenidos mentales

⁴⁰ Ejemplo palmario es la lógica de Konolige, que interpreta los objetos de creencia como entidades de tipo sintáctico empleando no obstante un lenguaje modal de primer orden con predicados para las actitudes proposicionales. En un lenguaje así no es necesario establecer que los términos a los que se aplica los predicados “creer” o “decir” denoten enunciados o fórmulas del lenguaje. En el sistema de Konolige se presupone que las creencias son manifestaciones de estados mentales de un agente y que puede describirselas asignándole un subsistema de creencias básicas (*belief subsystem*) a partir de las cuales se definen mecanismos lógicos de derivación. Se asume por tanto que todo agente tiene un lenguaje interior, cuyos mecanismos inferenciales son formulables en términos de reglas lógicas. Tales mecanismos, que Konolige llama “deduction model”, tienden a conformar un modelo abstracto de las capacidades concretas de agentes epistemológicos reales, bien sea naturales o artificiales. A tal efecto, se asigna a cada sujeto de creencia una estructura de deducción (*deduction structure*) i , compuesta de un par ordenado $\langle \mathcal{A}(i) = LB(i) \ p(i) \rangle$, donde $B(i)$ es un conjunto de fórmulas de un lenguaje lógico L cualquiera. La finalidad principal del modelo es mostrar que las personas pueden extraer consecuencias de las cosas que creen, aunque no puedan derivar *todas* las consecuencias lógicamente posibles (Cf. Konolige, K., *A Deduction Model of Belief*, Los Altos (CA), Morgan Kaufmann, 1986).

que expresan propiedades a éstos directamente correlacionadas en el mundo, y la idea de que las actitudes proposicionales se identifican con estados mentales cuyos términos relacionan un organismo y una representación mental, entendida esta última como *oración en el lenguaje del pensamiento* provista de propiedades semánticas y sintácticas.

No es éste, como se ha adelantado, el camino que transita Davidson. Su aversión al atomismo y a la concepción representativa de la mente le impide expresamente esta posibilidad, a la que considera como resabio de la tradición mentalista y subjetiva que nace de la filosofía de Descartes y se prolonga en la actualidad en importantes corrientes de la ciencia cognitiva. En clave davidsoniana, la posibilidad de la interpretación es incompatible con el supuesto de que haya objetos de las representaciones mentales –y menos aún podría ser compatible con la peregrina idea de que de tales representaciones, privadas y conceptualmente neutras, pudieran brotar por acción relacional nuestras creencias. A mi modo de ver, si la principal razón por la que Davidson rechaza el análisis relacional y opta en consecuencia por el enfoque sintáctico o inscripcional es la imposibilidad del primero de dar cuenta de las actitudes proposicionales como contextos lingüísticos de carácter público o social, el enfoque adverbial deberá, pues, ser capaz de integrar el análisis lógico de la creencia en una concepción del lenguaje donde el contenido de las creencias más básicas de los sujetos esté formado no por representaciones mentales privadas, sino por situaciones y eventos comunes e intersubjetivos que se tornan manifiestos gracias a la producción de actos de habla que convergen en un espacio público compartido. En la concepción adverbial que se ha estado discutiendo parece que una y otra exigencia apuntan en direcciones opuestas e inconciliables, por lo que deberemos concluir que este ámbito no se presta para dar respuesta a las dificultades que hemos detectado en relación con la teoría de la creencia de Davidson.

Una aproximación más cercana al análisis davidsoniano consiste en analizar los contextos de actitud proposicional como estructuras lingüísticas cuyo significado es dado por el papel gramatical que juegan dentro del lenguaje al que pertenecen. Es éste el análisis de las actitudes propuesto por las teorías adverbiales del segundo tipo. Aunque los lineamientos generales de esta solución se deben a Urmson,⁴¹ fue Prior quien en verdad desarrolló el análisis con precisión en *Objects of thought*. No es fácil explicar directamente el análisis originario de Prior, por lo que me tomaré la licencia de introducir algunos cambios simplificadores en un sentido relevante para la cuestión de la estructura lógica de las adscripciones de creencia.

Hay algo más en la explicación de Prior que la simple búsqueda de la estructura lógica. Creo que el objetivo de su caracterización –que se asemeja bastante, pero no en todo, a las teorías no-relacionales del primer tipo– consiste en proporcionar una explicación de las actitudes proposicionales enmarcada en una teoría del significado al estilo de Grice para las transacciones interlingüísticas de los hablantes. En esta caracterización la adscripción de creencias es concebida como un procedimiento gramatical que refleja la manera como utilizamos nuestras palabras y nuestros conceptos. Así, en los contextos de creencia, como en general en todos los contextos de actitud proposicional, lo que hacemos es emplear los verbos de actitud como dispositivos gramaticales que contextualizan las preferencias de adscripción dentro del entramado lingüístico en el que funcionan, incitando al hablante a asignar a tales verbos la función de operadores intralingüísticos, a conferirles el significado que tienen dentro del conjunto de expresiones con las que se conectan. No son, pues, los verbos que funcionan gramaticalmente como operadores los que proporcionan significación a las adscripciones, sino que más bien son las propias relaciones lingüísticas entre las expresiones envueltas en la adscripción las que conceden significado a los operadores y justifican su

⁴¹ Cf. Urmson, J., “Parenthetical verbs”, *Mind*, 61, 1952, pp. 480-496.

rasgo característico de marcadores de actitud. Esto no añade mucho al viejo análisis de Grice del significado no natural, pero marca el tránsito de una concepción descriptiva de las aptitudes proposicionales a una aproximación sintáctica que pone en primer plano la naturaleza gramatical del análisis de la creencia y subraya la relevancia de las conexiones gramaticales en la explicación del funcionamiento de los verbos epistémicos. Y este aspecto de la cuestión es importante para nuestros propósitos, pues parece ponernos sobre la pista de un análisis lógico de la creencia que congenia con el punto de vista de Davidson según el cual las adscripciones de creencia son realizativos que "...poseen, sí, una estructura, [...] pero se trata de una estructura familiar y no ocasiona problema alguno para la teoría de la verdad...".⁴²

Veamos ahora los detalles. Debería quedar claro que si la aproximación filosófica de Prior es correcta, las adscripciones de creencia son estructuras lingüísticas complejas compuestas de la unión de dos oraciones regidas por el verbo de actitud, el cual funciona a su vez como un adverbio que modifica o determina la circunstancia de la proposición expresada por la preferencia de la oración que sigue a la cláusula "que". Puesto que el verbo de actitud, junto con el sujeto de la adscripción, se analiza como un único operador lógico, un *predicado intralingüístico* cuyo significado es dado por su comportamiento inferencial, es fácil ver cómo los contextos de actitud pasan ahora a interpretarse pragmáticamente como preferencias de oraciones que toman *nombres* como argumentos, sin el menor vestigio de referencia a relaciones entre objetos o conexiones entre mentes y mundo que causaran, en las *Investigaciones Filosóficas*, la ira de Wittgenstein en contra de las concepciones *descriptivistas* como modos de dar cuenta del funcionamiento lógico-semántico de los verbos de actitud proposicional. En particular, la forma lógica general de las oraciones de creencia sería la de una función monádica provista de dos huecos de argumento, uno para el sujeto de la adscripción y otro para

⁴² Cf. Davidson, "Al decir *that*", cit., p. 121.

la oración subordinada expresada mediante la preferencia. Así, nuestro ejemplo de arriba podría parafrasearse como:

[(Davidson) cree que] ES UN ESTADO (la creencia).

Obsérvese que aquí la expresión “[Davidson] cree que” se interpreta como un operador intralingüístico de circunstancia que marca sintácticamente la oración que cae bajo su alcance, “ES UN ESTADO (la creencia)”, cuya emisión es concebida ahora como *nombre de la oración*.

Mayores detalles técnicos no vienen ahora al caso; lo cierto es que tenemos ahora entre manos una manera de aproximarnos al problema de la forma lógica de la adscripción de creencia que parece compatible con el modelo inscripcional o citativo de Davidson. En efecto, hemos logrado separar un poco el elemento subjetivo o mental, dejando a salvo el carácter público o social de las actitudes proposicionales; no precisamos de la arquitectura extensional modeloteórica de la lógica modal, evadiendo así la asunción de omnisciencia lógica, ni necesitamos para definir la verdad nada que no sea los recursos iterativos básicos, los conectivos proposicionales y el aparato de cuantificación con identidad; en fin, tampoco se requiere asumir que las expresiones que caen bajo el alcance del operador de creencia denoten objetos de tipo intensional como las proposiciones, sino sólo que se refieren a *nombres de enunciados o fórmulas*.

A estas alturas, digamos de una buena vez que no hay en la literatura generada por la teoría lógica demasiados candidatos para dar con el cometido prefijado en las condiciones arriba estipuladas. A mi juicio, existe una sola alternativa: adoptar un modelo sintáctico de las actitudes proposicionales basado en el uso de un lenguaje predicativo de primer orden en lugar de un lenguaje modal. Para fortuna nuestra, un modelo de este tipo ha sido desarrollado por McCarthy en el contexto de la discusión sobre el lenguaje

ideal para la inteligencia artificial.⁴³ El propósito inicial de McCarthy es identificar combinaciones de representaciones finitas de los significados lingüísticos que permitan construir modelos de las representaciones y los procedimientos que emplea un agente finito –y por tanto, no omnisciente– para explicar o proyectar su conducta lingüística. En la versión más refinada, se emplea un lenguaje de predicados de primer orden que hace posible expresar la creencia por medio de un predicado monádico –por ejemplo, *bel(x)*– que tiene como argumentos nombres de enunciados o fórmulas del lenguaje, de modo tal que si “ α ” es el nombre de la fórmula α , “creer α ” se expresa mediante la fórmula *bel* (“ α ”). Además, McCarthy menciona otros dos criterios básicos que considera indispensables para una formulación estrictamente sintáctica de la lógica de la creencia. Dado que en la formulación predicativa de primer orden no hay en sentido estricto objetos de creencia sino sólo términos singulares como nombres de fórmulas –en Davidson, todos son subsumidos bajo un solo término, *that*–, es posible cuantificar sobre éstos. Así, por ejemplo, podemos emplear la cuantificación en fórmulas como:

$$\exists x(\alpha(x) \wedge \textit{bel}(x)),$$

o bien

$$\forall x(\alpha(x) \rightarrow \textit{bel}(x)),$$

donde $\alpha(x)$ es una fórmula cualquiera del lenguaje, con x como variable libre. Por otra parte, podemos definir una cierta noción sintáctica de la consecuencia lógica, garantizando la recursividad en la generación de las inferencias pero con las restricciones que se derivan de asumir las ora-

⁴³ Cf. McCarthy, J., “First-order theories of individual concepts and propositions”, en Hayes, J.E., Michie, D. y Mikulich, L.L. (eds.): *Machine Intelligence*, 9, Nueva York, Halsted Press, 1979, pp. 120-147.

ciones como términos singulares, cosa que lejos de ser una desventaja prohíbe por el contrario la producción de inferencias indeseadas.⁴⁴

Esto es todo lo que da de sí nuestra reconstrucción lógica del concepto de creencia adelantada en concordancia con el análisis davidsoniano de la noción de creencia y su motivación. ¿Qué es entonces lo que podría objetarse a una solución lógica como la de McCarthy sobre la cual lo hemos apuntalado? Dijimos más arriba que para proporcionar una semántica de las expresiones del lenguaje natural compatible con las exigencias de la formalización de las adscripciones de creencia, era necesario disponer de un poder expresivo tal que permitiera la cuantificación sobre los objetos de las actitudes proposicionales, además de garantizar un buen arsenal de operadores sintácticos para asegurar la posibilidad de realizar las manipulaciones formales indispensables en la estructura de las fórmulas del lenguaje. Es indudable que la formulación de McCarthy cumple con las dos exigencias sin incurrir en la hipótesis del creyente omnisciente que se interpone en las soluciones modelo-teóricas. Desafortunadamente, tropieza con una dificultad de distinta índole, aunque no de menor monta: es lógicamente inconsistente. Su mayor mérito, el gran poder expresivo, trócase irónicamente en su gran desventaja. En efecto, como lo ha demostrado Montague,⁴⁵ todos los tratamientos sintácticos de las actitudes proposicionales basados en un lenguaje predicativo de primer orden y que incluyen axiomas equivalentes a los usuales en los sistemas modales para las actitudes epistémicas, además de los axiomas de Robinson para la representación de las funciones recursivas de la teoría, son lógicamente inconsistentes. La fuente de la inconsistencia reside en el hecho de que la capacidad de representar la aritmética en la teoría facilita la formación de fórmulas au-

⁴⁴ En Konolige, por ejemplo, las restricciones quedan reflejadas en la estructura misma de los conjuntos de creencias, al estipularse que en cada $p(v)$ valga la regla de $\vdash_{p(v)} \alpha \vee \beta$ sigue $\vdash_{p(v)} \beta \vee \alpha$.

⁴⁵ Cf. Montague, *Formal Philosophy...* cit., pp. 286-302.

toferenciales, por lo que se hace posible derivar en la teoría una paradoja análoga a la que empleara Tarski para demostrar la imposibilidad de introducir consistentemente un predicado de verdad en un lenguaje afecto de autorreferencia. Observa Montague que la demostración del teorema de Tarski descansa en el hecho de que en un sistema formal de primer orden que incluya a su propio predicado de verdad y en condiciones de referirse a sus propias expresiones, puede demostrarse una fórmula análoga a la antinomia del mentiroso, esto es, un enunciado que afirma su propia falsedad.⁴⁶

Aunque la demostración de Montague es dirigida originalmente a la lógica modal alética, sus resultados son fácilmente extrapolables a la lógica de la creencia. Thomason, en efecto, ha extraído conclusiones análogas a las de Montague en el tratamiento de teorías de la creencia expresadas en un lenguaje de primer orden.⁴⁷ En síntesis, el procedimiento de Thomason consiste en mostrar que, para toda teoría de primer orden P_{ws5} con símbolos para la aritmética de Robinson y un predicado monádico bel , es demostrable para toda fórmula α y β del lenguaje lo que sigue:

(i) $bel("Q")$,

donde Q es la conjunción de los axiomas para la aritmética de Robinson;

(ii) $bel("α \rightarrow β") \rightarrow bel("α") \rightarrow bel("β")$;

(iii) $bel("α") \rightarrow bel("bel("α)")$;

(iv) $\neg bel("α") \rightarrow bel("bel("α)")$;

⁴⁶ En los sistemas de lógica epistémica este resultado es representado mediante la fórmula $\alpha \leftrightarrow se\ que(\neg\alpha)$, que ha dado lugar al célebre *knower paradox*. Para saber más sobre el asunto, puede verse Grim, P., "Operators in the Paradox of the Knower", *Synthese* 94, 1993, pp. 409-428; también Égré P., "The Knower Paradox in the Light of Provability Interpretations of Modal Logic", *Journal of Logic, Language and Information* 14, 2005, pp. 13-48.

⁴⁷ Cf. Thomason R., "A note on syntactical treatments of modality", *Synthese* 44, 1980, pp. 391-395.

y valga además la regla:

PNec: de α se sigue $bel("a")$,

el equivalente de la regla de necesitación correspondiente en los sistemas de lógica modal alética. El resultado es simple: Pws5 no es en sí un lenguaje inconsistente, pero es internamente inconsistente, por decirlo así, *dentro* de los contextos de creencia, toda vez que permite derivar, para cada fórmula α o β , respectivamente, $bel("a")$ o $bel("b")$; desde luego, parece innecesario agregar que Pws5 se torna abiertamente inconsistente tan pronto se introduzca un operador de sustitución para los nombres de las fórmulas, necesario para correlacionar un término " α " –el nombre de la fórmula α – con el término " $\alpha \rightarrow \beta$ " –el nombre de la fórmula $\alpha \rightarrow \beta$. En tal caso Pws5 resulta obviamente inconsistente al incluir entre sus axiomas la fórmula $bel("a") \rightarrow \neg bel("\neg a")$.

En conclusión, la exigencia de una representabilidad finita de los significados lingüísticos, necesaria para dar cuenta de la competencia semántica del hablante, encuentra dificultades que, al menos desde el punto de vista de Davidson, no eran fácilmente previsibles en aquel temprano ensayo de 1968. Tal vez se deba a la incómoda presencia de tales dificultades el hecho de que en su obra posterior haya ignorado obstinadamente la intensa controversia desatada en la literatura contemporánea alrededor de la estructuración formal de la adscripción de creencias, y se contentara más bien con tratar la noción como "cuestión filosófica". Quienes, como Montague y Thomason, sí se han ocupado tenazmente de la cuestión formal, interpretan los problemas envueltos en el surgimiento de las paradojas epistémicas y antinomias semánticas como un claro síntoma de las dificultades de conciliar los objetivos de la semántica filosófica y de la teoría de la interpretación con el destino del tratamiento sintáctico de las actitudes proposicionales. Si en verdad, como quiere Davidson, "la sintaxis es muchísimo más social

que la semántica”,⁴⁸ la justificación del postulado habrá que buscarla no ciertamente en “el patrón de inferencia y estructura creado por las constantes lógicas”.⁴⁹ En efecto, en la vertiente más formalista del tratamiento de las creencias, los lógicos intentan ocultar a duras penas el desconcierto que produce saber que ese mismo “patrón de inferencia y estructura” sea responsable de que nuestras capacidades aritméticas acaben generando autorreferencia. No deja de producir cierta desazón admitir que se siga de la teoría semántica que el hablante ideal deba saber menos aritmética que la necesaria para poder derivar la antinomia.

Al fin y al cabo, la mejor manera de dejar de ser omniscientes es paradójicamente intentando ser conscientes de las implicaciones que las dificultades que hemos estado discutiendo tienen para el proyecto davidsoniano de “una teoría del significado y de la acción”. De otro modo, nuestra conducta equivaldría a disponernos a resolver la *paradoja del gato* prescindiendo por entero de las leyes de la teoría cuántica.

Instituto de Filosofía
Universidad Central de Venezuela
italianos@tutopia.com

⁴⁸ Davidson, “Communication and Convention”, en Dascal, M.(ed.): *Dialogue. An Interdisciplinary Approach*, Amsterdam, 1982; recogido en Davidson: *De la Verdad y...* cit., p. 275.

⁴⁹ *Ibidem*.