

NIVELES DE LENGUAJE, AUTORREFERENCIA Y LAS PARADOJAS

Jesús Padilla Gálvez

Introducción

En este trabajo voy a reconstruir una *paradoja* que tiene que ver con los enunciados, que podemos poner en boca de un individuo ficticio, mediante los cuales se profiere o se aserta un contenido completamente contradictorio cuando aquél se *refiere a sí mismo*¹. Dicha paradoja se desarrolla en el marco de la teoría de la verdad. La paradoja de los individuos ficticios es, como veremos, arma poderosa que refuta, no sólo algunos de los argumentos que han sido propuestos, sino también algunas de las interpretaciones, con sus respectivos análisis, en las que se asienta la teoría de la verdad.²

El siguiente trabajo se estructurará del siguiente modo: primero plantearé los presupuestos tarskianos de "verdadera-en-un-lenguaje". Así pues, efectuaré una lectura de las propuestas de A. Tarski de modo que me permita reconstruir la metodología corriente que se lleva a cabo para

¹ Dicha paradoja tiene la *estructura autorreferencial*, también denominada en la bibliografía sobre el tema "SELF". Véase: Smullyan 1957, 37ss.; van Fraassen 1968, 136ss.; Kripke 1975, 690ss.; Martin y Woodruff 1975, 213ss.

² Este estudio es una ampliación del proyecto sobre el problema de la verdad. Véase: Padilla Gálvez 1988/89, 197ss.; Padilla Gálvez 1989, Cap. III; Padilla Gálvez 1991, 137ss. y Padilla Gálvez 1992.

introducir la convención (T). Seguidamente expondré, como contrapunto, el reto que me parece más significativo a la propuesta tarskiana³ y que consiste en la alternativa fundamentacional sugerida a partir de los análisis de S. C. Kleene y elaborada por S. A. Kripke. En palabras de A. Tarski, dicho proyecto es de carácter *relativo*. Su propuesta *absoluta* de la noción de verdad viene a ser analizada por aquélla como un caso especial.⁴ Las conjeturas que vamos a elaborar contra las propuestas relativas de la noción de verdad vienen de la mano del dato según el cual se hace imprescindible, para dilucidar adecuadamente la paradoja de los *individuos ficticios*, el poseer sistemas amplios que permitan esclarecer el lenguaje ficticio y los argumentos ficticios.

Abreviando, puedo sintetizar que las propuestas relativas reivindican para sí haber abordado objeciones de peso contra la propuesta tarskiana de verdad. Me interesa plantear la cuestión de si las objeciones expuestas pueden ser solucionadas por las propuestas kripkianas. Además, me inquieta la cuestión de si refutan los argumentos expuestos por la propuesta fundamentacional las consideraciones teóricas de Tarski. Sobre el resultado de dicha cuestión llegaremos a conclusiones determinadas.

1. 'Verdadera-en- \mathcal{L} '

La traducción de la tesis doctoral escrita en polaco del trabajo '*Der Wahrheitsbegriff in den formalisierten Sprachen*' supone una aportación considerable a los problemas concernientes a la verdad.⁵ La tesis presenta

³ El trabajo de A. Tarski sobre el concepto de verdad en los lenguajes formalizados contiene en su traducción alemana 144 páginas. Véase: Tarski 1935, 261-405. Hacer un comentario en el espacio de que se dispone en un artículo es una tarea árdua. Vale, sin embargo, la pena exponer planteamientos puntuales, que comparados con otros puntos de vista nos permitan dilucidar el problema de la verdad adecuadamente. Dos puntos han de ser subrayados a la hora de analizar dicho trabajo: uno de carácter material, podríamos decir. Otro de naturaleza formal. El aspecto material, plantearía el siguiente problema: ¿Qué interpretación semántica resulta más ajustada a nuestra noción de verdad? El aspecto formal se guiaría por cuestiones de índole sintáctica.

⁴ La distinción entre la *propuesta relativa* y la *propuesta absoluta* de la noción de verdad fue puesta por A. Tarski en 1935, 318 s.

⁵ Algunas consideraciones generales sobre la noción de verdad se encuentran en: Etchemendy

una formulación del concepto de verdad en términos semánticos y orientados a los lenguajes formalizados. Para ello, se formula la *condición tarskiana de verdad* (abreviadamente: T) que debe cumplir cualquier teoría de la verdad y que viene expuesta de la siguiente manera:⁶

Una definición formalmente correcta, en los términos del metalenguaje del símbolo '*Wr*' la denominaremos una *definición exacta de verdad* si trae consigo las siguientes consecuencias: (α) todo enunciado que se deduce de la expresión "*xεWr* si, y sólo si *p*", en la que se sustituye para el símbolo *x* un nombre descriptivo-estructural de un enunciado arbitrario del lenguaje considerado, y para el símbolo "*p*" la expresión que traduce dicho enunciado en el metalenguaje; (β) el enunciado "para un *x* arbitrario - si *xεWr*, entonces *xεAs*" (o en otras palabras "*Wr ⊂ As*").

Dicha definición viene a ser expuesta de manera escueta mediante la siguiente correlación:

(T) '*S*' es verdadera si, y sólo si *p*,

donde '*p*' puede sustituir cualquier enunciado del lenguaje para el cual se está definiendo la verdad y '*S*' tiene la *función nominal* del enunciado que reemplaza a '*p*'. Según A. Tarski, no es coherente que los *lenguajes sean cerrados*. Es decir, los lenguajes han de estar de tal manera contruidos que un lenguaje de un orden superior pueda definir un predicado de verdad para un predicado del lenguaje de orden inferior. Sea \mathcal{L}_o un lenguaje formal que venga definido mediante operaciones del cálculo de predicados de primer orden, a partir de un determinado grupo de predicados primitivos completamente definidos y en el cual se puede caracterizar su propia

1988, 51 ss.

⁶ La condición tarskiana de verdad viene a ser denominada por el propio A. Tarski como la condición *W* (*W de verdad, y que nosotros exponemos mediante T*) y dice en el original así: "Eine formal korrekte, in den Termini der Metasprache formulierte Definition des Symbols "*Wr*" werden wir eine *zutreffende Definition der Wahrheit* nennen, wenn sie folgende Folgerungen nach sich zieht: (α) alle Sätze, die man aus dem Ausdruck "*xεWr* dann und nur dann, wenn *p*" gewinnt, indem man für das Symbol *x* einen strukturell-deskriptiven Namen einer beliebigen Aussage der betrachteten Sprache und für das Symbol "*p*" den Ausdruck, welcher die Übersetzung dieser Aussage in die Metasprache bildet, einsetzt; (β) die Aussage "für ein beliebiges *x* - wenn *xεWr*, so *xεAs*" (oder mit anderen Worten "*Wr ⊂ As*")." Tarski 1935, 305s (la traducción del texto alemán es mía). Véase además: Padilla Gálvez 1989, Cap III y Padilla Gálvez 1991, 137ss.

sintaxis. Según A. Tarski, susodicho lenguaje no puede contener su propio predicado de verdad. En consecuencia, dicho predicado sólo puede ser definido en un lenguaje de orden superior. El proceso puede repetirse, desarrollándose una secuencia $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \dots$ de lenguajes cada uno de los cuales contendrá el predicado de verdad para el nivel anterior. Podemos comprobar que tienen vigencia las *leyes lógicas*. Por último, dichos enunciados no están reñidos con las *premisas empíricas*.

Si observamos el ejemplo propio que proporciona A. Tarski, entonces (T) viene a decir lo siguiente:⁷

- (1) "La nieve es blanca" es un enunciado verdadero si, y sólo si, la nieve es blanca.

La preferencia emitida en (1) es una instancia de (T) si se hace referencia al enunciado del lado derecho mediante su *nombre entrecomillado* del lado izquierdo. Así pues, la parte derecha del bicondicional está en relación con la oración que se encuentra a la izquierda en la que se dan las condiciones de verdad. El esquema expuesto por (T) es una *condición de adecuación material* por la cual queda determinada la extensión del término "verdad".⁸ De aquí resalta que podamos caracterizar de un modo absoluto el concepto de verdad, es decir, que todo enunciado adecuadamente bien formado debe ser, o bien verdadero, o bien falso, dependerá de si se acepta o se rechaza (T).

Ahora bien, el requisito que se establece mediante (T) concierne a:

LENGUAJE

- (i) Las expresiones del lenguaje \mathcal{L} ;
- (ii) Los nombres que vienen expuestos en \mathcal{L} ;
- (iii) Los términos semánticos de \mathcal{L} .

⁷ Véase: Tarski 1935, 269.

⁸ La propuesta de A. Tarski viene motivada por lo que se denomina *concepción aristotélica clásica de verdad*. Véase: Aristoteles 1957, 1011, b 25. Tal concepción ha de aceptar la condición de adecuación material propuesta en (T).

LEYES LÓGICAS

- (i) La aceptación de los axiomas y las reglas de cálculo de la lógica clásica.

PREMISAS EMPÍRICAS

- (i) Lo expresado en \mathcal{L} atiende a los datos empíricos.

Asentados dichos ámbitos de referencia, bien vale la pena indicar cuáles son los puntos más conflictivos de la propuesta tarskiana. Propongo para llevar a cabo dicho cometido, exponer su programa de la siguiente manera (abreviadamente: SPM):

SPM₁ El lenguaje usado contiene, además de sus expresiones, los medios de referirse a esas expresiones.

SPM₂ El lenguaje usado contiene, además de sus expresiones, predicados semánticos tales como "verdadero" y "falso".

SPM₃ Las leyes lógicas usuales tienen validez.

A. Tarski propone en SPM₁ que el lenguaje que se use para definir el concepto de verdad ha de ser lo bastante fuerte, que sugiera prioridades ontológicas, ya que si contuviese sólo y exclusivamente las expresiones usadas no reflejaría adecuadamente la contingencia del mundo. En conclusión, si este último prejuicio se infiltra en el ámbito de la presunta neutralidad de la definición de verdad, entonces la opresión lingüística se tornaría cognitiva, algo que ha de tenerse siempre en presente.

Ya que no se desea rechazar el supuesto expuesto en SPM₃, A. Tarski llega a la conclusión de que una definición formalmente correcta de la verdad ha de ser expresada en un lenguaje que no sea semanticamente cerrado. Este resultado se traduce en la definición de verdad-en- \mathcal{L}_o , donde \mathcal{L}_o es el *lenguaje objeto*, es decir el lenguaje para el cual se está definiendo la verdad. La definición de verdad de un enunciado de \mathcal{L}_o tendrá que presentarse en un *metalenguaje* \mathcal{L}_M , es decir el lenguaje en el que se defina 'verdad-en- \mathcal{L}_o '. Así (T) puede ser traducido del siguiente modo: 'p' reemplaza un enunciado en \mathcal{L}_o , es decir el lenguaje para el cual se está definiendo la verdad, y 'S' reemplaza una *descripción estructural* en la que aparece un predicado de verdad P(x) para \mathcal{L}_o .

Enumeramos los pasos de la estrategia seguida por A. Tarski para así dilucidar, de modo informal, el concepto semántico de verdad. Para ello definirá el predicado de verdad mediante el concepto de *satisfacción* que es un concepto no-semántico. La estrategia seguida se puede exponer *simpliciter* del siguiente modo:

- (a) Primeramente se especifica la estructura sintáctica de \mathcal{L}_o , sobre el que se va a enunciar el predicado "verdadero".
- (b) En un segundo paso se caracteriza la estructura sintáctica de \mathcal{L}_M en la cual se define "verdad-en- \mathcal{L}_o " y la cual debe contener:
 - (i) o las expresiones de \mathcal{L}_o , o traducciones de las expresiones de \mathcal{L}_o ;
 - (ii) el vocabulario sintáctico;
 - (iii) el aparato lógico usual.
- (c) Definir "satisface-en- \mathcal{L}_o "; y
- (d) Definir "verdadera-en- \mathcal{L}_o " en términos de "satisface-en- \mathcal{L}_o ".

La razón por la que A. Tarski define primero "satisface" es bien sencilla, no desea emplear en la definición de verdad un término primitivo semántico debido a la falta de nitidez que esto genera.

2. La paradoja de la ficción

La importancia que tienen las paradojas para la teoría de la verdad se puede caracterizar o bien de un modo negativo, o bien, positivo. La primera atañe a problemas de contradicción en un sistema. Desde un punto positivo nos pueden servir, por un lado, de indicador para aclarar si las *condiciones de adecuación formal* de las definiciones de verdad pueden evitarlas, y por otro lado indican la urgente necesidad de explicar los mecanismos de la *autorreferencia*.

El problema de los mecanismos autorreferenciales en la ficción puede ser introducido del siguiente modo: algunos enunciados que se expresan en nuestro lenguaje natural, pero que pertenecen al nivel ficticio, son, aparentemente obvios, y sin embargo conducen a contradicción mediante

un razonamiento válido.⁹ La paradoja que seguidamente desarrollaremos va a ser un instrumento sumamente importante para aclarar las propuestas tarskiana y repasar determinados argumentos relativos de los análisis de la verdad. La paradoja de la ficción no tiene sólo y exclusivamente una función sugestiva respecto de su capacidad de poner de relieve los problemas que envuelven los niveles de lenguaje, sino que además nos permite determinar de modo concreto problemas que tienen que ver con el contenido de los enunciados que se analizan.

El escenario de lo que voy a denominar la paradoja de la ficción se encuentra ejemplarmente expuesto en el complejo narrativo de Cervantes, en su libro *El Ingenioso Hidalgo don Quijote de la Mancha*, capítulo 74 de la segunda parte -es decir el capítulo final- titulado *De cómo don Quijote cayó malo, y del testamento que hizo y su muerte*.¹⁰ Si aceptamos la propuesta de Cervantes de que tanto se puede crear un personaje que describe como que vive, y por lo tanto llega un momento en que le interesa hacer como que muere pues describimos *in propria voce* que muere.¹¹ Sin embargo, al describir la muerte de dicho individuo, salta el problema siguiente: ¿Si don Quijote murió, cómo testó, como Alonso Quijano, o como don Quijote? Esta cuestión está formulada dentro del escenario de la acción narrativa, en donde el individuo ficticio "don Quijote" es una ficción dentro de la ficción de Alonso Quijano. De aquí salta pues la pregunta de si puede morir como don Quijote o máxime como Alonso Quijano. Es decir, la cuestión clave es saber si es posible que don Quijote muera aún

⁹ Un trabajo clásico sobre el problema de la verdad en aquellos casos ficcionales es el expuesto por D. Lewis. Véase: Lewis 1983, 261 ss.

¹⁰ Tomo como referencia la edición de John Jay Allen publicada en Cátedra.

¹¹ M. Cervantes tiene muchas razones para que muera su figura. Una de ellas viene expresada por el propio don Quijote que afirma: "Item, suplico a los dichos señores mis albaceas que si la buena suerte les trujere a conocer al autor que dicen que compuso una historia que anda por ahí con el título de *Segunda parte de las hazañas de don Quijote de la Mancha*, de mi parte le pidan, cuan encarecidamente se pueda, perdone la ocasión que sin yo pensarlo le di de haber escrito tantos y tan grandes disparates con ella; porque parto desta vida con escrúpulo de haberle dado motivo para escribirlos" (II, 72).

habiendo "existido" en la obra de Cervantes como ficción dentro de la ficción.¹²

Don Quijote es una forma de vida -la que adopta un tal Alonso Quijano- que se transforma en don Quijote, y que entre las muchas propiedades que podemos asignarle se encuentra la de adscribirse a sí mismo el haberse nombrado "don Quijote". La capacidad de autodenominarse la tiene, por lo tanto, Alonso Quijano y no don Quijote.¹³ Dicha capacidad se puede enunciar del siguiente modo:

(C) X cree que él mismo es Φ .

En donde X es sustituible por "Alonso Quijano" y Φ se entiende como un predicado que caracteriza una concepción estructurada de determinadas propiedades.

Sostengo, pues, en este trabajo, que la paradoja de la ficción se lleva a cabo en el momento en que don Quijote se adscribe a sí mismo

¹² W. Künne sale al paso de dichas cuestiones alegando que toda preferencia aseverativa-ficcional ha de entenderse como acto mimético. Esta viene a ser caracterizada como una "cuasi-aseveración". De este modo si se reconoce una expresión como ficcional, entonces se encuentra a nivel superior de entendimiento. La cuestión que deja sin despejar dicho autor es: ¿Qué ocurre con una potenciación de los niveles ficcionales -como ocurre en nuestro ejemplo- se reitera con ello los niveles de entendimiento?, ¿Se convierten los enunciados ficcionales aseverativos que enuncia don Quijote, expresados por la boca de Alonso Quijano, en una "cuasi-cuasi-aseveración"? Así afirma dicho autor: "Hier ist der singuläre Term in einer fiktionaler Rede so verwendet, daß gilt: es gibt einen Gegenstand auf den der Erzähler mit dem Term Bezug nimmt" (Véase: Künne, 1983, 292). Esta explicación caracteriza a individuos ficticios como "Alonso Quijano", "Gregor Samsa", etc. pero no a los individuos ficticios como "don Quijote" ya que la estructura de estos individuos es más compleja al haberse autorreferido a sí mismos y haberse impuesto una caracterización determinada de modo que cambia paradójicamente su sistema referencial. La propuesta de W. Künne ha de rechazarse porque no permite explicar adecuadamente la estructura referencial que procede en dicho metabolismo referencial.

¹³ Hay una confusión sobre los nombres de don Quijote ya que al comienzo se habla del sobrenombre de Quijada, Quesada, o Quejana (I, 1). Más adelante, un labrador lo denominará "Señor Quijana" (I, 5) y don Quijote mismo se referirá a su descendencia por línea directa de un tal "Gutierre Quijada" (I, 49). Sin embargo, al final de la segunda parte no tiene reparos en identificarse como "Alonso Quijano" (II, 74).

determinadas propiedades como pudiera ser la que afirma en un *memento mori*, cuando dice:¹⁴

(2) me muero.

En (2) un individuo ficticio ha desarrollado la capacidad (C) de tal modo que afirma de sí mismo mediante una autoadscripción (abreviamos: A):

(A) Y se adscribe a sí mismo la propiedad Γ .

En donde 'Y' indica en este caso a "don Quijote", es decir, 'X' después de haber desarrollado su capacidad (C), y aquel se adscribe a sí mismo una propiedad Γ . Supongamos que dicha propiedad viene descrita por la preferencia enunciada en (2), entonces nos encontramos ante un problema, a saber, el de dilucidar si se da el caso de que (2) sea o no sea verdadero. Es decir, si muere o no muere don Quijote. Tal cuestión se traduce en la pregunta de si es verdadero el enunciado que asevera don Quijote cuando se adscribe (2).¹⁵ Si afirmamos que se cumple el predicado, y realmente la propiedad se cumple, entonces podemos afirmar que Y se muere, pero con ello no tenemos por qué afirmar irremediamente que X se muere, ya que debido a que las propiedades Φ y Γ son diferentes y no intercambiables *salva veritate*, aunque corresponda al mismo individuo con diferente nombre, bien pudieran los nombres de dichos individuos o bien, (i) no ser traducibles, o bien, (ii) incompatible o bien, (iii) insustituibles.

La cuestión es, por lo tanto, saber cuándo se puede aseverar el enunciado (2). Una posible respuesta sería la siguiente: el enunciado (2) sólo lo puede aseverar don Quijote "habiendo dejado hacer de don Quijote". Por lo tanto, (2) sólo puede ser proferido por Alonso Quijano, con respecto a don Quijote. Consecuentemente, tendrfa validez la siguiente preferencia:

¹⁴ Téngase presente que Cervantes se refiere a don Quijote y éste, ya cuerdo, enuncia (2) de la siguiente manera: "Yo, señores, siento que me voy muriendo a toda priesa" (II, 74).

¹⁵ El lector entendido se habrá dado cuenta que lo único que he hecho es haber planteado la paradoja ficcional en términos de la autorreferencia. Nuestro ejemplo arroja alguna luz sobre los mecanismos que llevan inmersos los enunciados autorreferenciales. La descripción y reconstrucción de dicha paradoja nos permite introducirnos en dichos problemas (Véase: Tarski 1969, 67 s.).

(3) murió don Quijote.

Una explicación complementaria es que, aparentemente, se nos presenta obvio ante la insistencia de conocer la vida y andanzas de don Quijote, que el escenario final conduzca a la afirmación, aparentemente obvia, expuesta en (2). Una explicación sería la siguiente: Cervantes crea un personaje de tal relevancia que absuelve al individuo que soporta; es capaz de hacer olvidar, en tanto que llega el momento en que nadie advierte que murió Alonso Quijano, y enunciamos (3) como verdadero. Desde este punto de vista, la cuestión que legitima el error estaría en la *singularidad*. La falta de espacio no sólo me impide describir este problema, sino también me impide mencionar las dificultades formales que convierten el escenario final de don Quijote en algo sumamente no trivial.

Vamos a resaltar dos problemas diferentes y que tienen que ver con las propuestas que se han llevado a cabo para dilucidar adecuadamente la noción de verdad que estamos barajando. El primer problema parte del hecho que la cuestión descrita tiene un paralelismo con la concepción jerárquica de A. Tarski y su propuesta de los lenguajes abiertos, ya que puede ser extendida a niveles transfinitos. El enunciado (2) se encontraría expresado para un individuo ficticio en una situación ficticia y por lo tanto no tendría la misma referencia que al nivel real. Esto así dicho se podría enunciar del siguiente modo: Alonso Quijano tiene derecho a aseverar:

(4) El enunciado "me muero" es verdadero,

y llega a unas conclusiones como es la de pedir un confesor y un notario.

Ahora bien (4) está referida por Cervantes sobre don Quijote y esta viene extendida de un supuesto *memento mori* de un hecho real en el que un individuo note su real pérdida de conciencia. De aquí resulta que (4) ha transgredido: primero el *memento mori* contingente y ha sido elevado al nivel ficticio y aplicado a la figura de don Quijote. Sin embargo, si bien Cervantes se refiere a Alonso Quijano, es decir, un individuo de supuesto estado *de jure* con respecto a don Quijote sigue nombrándolo "don Quijote". Por si esto fuera poco, ha transgredido un meta-nivel en tanto que don Quijote ha extendido para si el enunciado (2) y lo entiende como (4). La cuestión central es si (4) alcanza un valor de verdad en dicho proceso por el hecho de provenir de (2).

El segundo problema tiene que ver con la gramática de una adscripción autorreferencial. Cervantes nos presenta aquí un problema profundo cuando intenta describir el fin de la vida de un individuo ficticio y este se refiere a sí mismo y se refiere a su propio *memento mori*. La cuestión que de aquí resulta sería saber si es correcto decir que el enunciado "sé que estoy muerto", que se le podía adscribir a don Quijote, describe una aserción epistémica de algún acontecimiento privado. Dicha cuestión ha quedado sin plantear en la literatura reciente.

Creo que A. Tarski tiene razón cuando afirma que el predicado "verdadero" debe ser definido, es decir, que todo enunciado adecuadamente bien formado debe ser o verdadero o falso. El problema es saber cómo se asigna dicho predicado a enunciados como (2)-(4). Siendo su carácter, sin duda, extraordinario, ya que se presentan insertos en un contexto ficticio, configurando de algún modo la verdad o falsedad de dichos enunciados.

3. El problema esquematizado

Hemos enunciado una serie de problemas en el apartado anterior, de modo que es lícito saber cómo se disuelven. Me parece que el hecho de que Cervantes cometa diferentes errores en lo que respecta a la referencia induce a pensar que él mismo se encontraba ante un nudo de niveles que llevan a presentar la susodicha paradoja de la ficción.¹⁶

La convención de Tarski (abreviadamente: (T)) se expresa en un lenguaje \mathcal{L}_0 que no envuelve "verdad" ya que es definido en un lenguaje de orden superior (verdad-en- \mathcal{L}_0). Consecuentemente, para todo Φ en \mathcal{L}_0 (en donde Φ es una propiedad) vale:

(T) $\lceil \lceil \text{"}\Phi \text{"} \rceil \text{verdadero} \leftrightarrow \Phi \rceil$ es verdadero.

Para un enunciado de la convención (T) se usa la noción de verdad. Supongamos por ejemplo que para un lenguaje en el que hemos definido

¹⁶ El primer error viene de la mano de querer hacer morir a don Quijote cuando la única criatura que podría "morir", según el argumento, es Alonso Quijano. El segundo, cuando escribe que "replicó" o respondió don Quijote cinco veces, cuando ya había dejado de hacer de don Quijote y afirmaba a gritos que no era tal, sino Alonso Quijano. Tercero, se equivoca Cervantes al referirse a la "sobrina y ama de don Quijote".

sus caracterizaciones sintácticas y su universo puede axiomatizarse la teoría T_0 en \mathcal{L}_0 extendida a "verdad" para todo Φ en \mathcal{L}_0 de modo que:

(T₀) \vdash ("Φ")verdadero \leftrightarrow Φ),

en donde " \vdash " es la convención introducida por Quine.¹⁷

Nuestro ejemplo es el siguiente.¹⁸ Para el enunciado (2) descienden determinadas secuencias de lenguajes del siguiente modo:

nivel ₁	Cervantes enuncia (2) --- en --- \mathcal{L}_0
	--- adscribe que ...
nivel ₂	Alonso Quijano enuncia (2) --- en --- \mathcal{L}_f
	--- se adscribe ser ...
nivel ₃	don Quijote enuncia (2) --- en --- \mathcal{L}_{f+1}

Esquema 1

El esquema 1 nos permite exponer de un modo figurado los diferentes niveles a los que puede atender un mismo enunciado. La secuencia de lenguajes $(\mathcal{L}_i)_{i \in \omega}$ expuesta a diferentes niveles se define del siguiente modo:

$\mathcal{L}_f := \mathcal{L}_0$

$\mathcal{L}_{f+1} :=$ al resultado de sumar una característica especial de \mathcal{L}_f .

Según la convención (T), el nivel₁ sería verdadero, si se diera el caso de que al enunciar Cervantes (2) es el caso que se encuentra en un *memento mori*. En el nivel₂, si Alonso Quijano enuncia "me muero", dicho enunciado es verdadero de acuerdo con el nivel₁, es decir, si Cervantes escribe que Alonso Quijano se muere, sabiendo que dicha aserción se encuentra en un lenguaje ficticio (\mathcal{L}_f). En el nivel₃, si don Quijote enuncia (2), es verdadero, si Alonso Quijano realmente se muere porque Cervantes escribe que dicho individuo ficticio se muere. El nivel₃ se lleva a cabo, aparentemente, en un lenguaje ficticio como el de Alonso Quijano, con la salvedad de una propiedad, el que Alonso Quijano se autodenomina don

¹⁷ Véase: Quine 1983, 88.

¹⁸ En los siguientes esquemas altero algunos rasgos y los presento con caracterizaciones propias como es presentado por: Visser 1989, 692.

Quijote por lo que el nivel₃ tiene, lo que hemos denominado \mathcal{L}_{f+1} , es decir, consta de las mismas propiedades que \mathcal{L}_f más, al menos, una propiedad determinada, a saber, la de autoadscripción de un determinado nombre. La peculiaridad del esquema 1 es que es descendiente y finito. Este planteamiento está por lo tanto de acuerdo con la propuesta de S. A. Kripke que afirma:

Aun cuando estén en cuestión las definiciones irrestrictas de verdad, los teoremas estándar nos permiten fácilmente construir una cadena *descendente* de lenguajes de primer orden L_0, L_1, L_2, \dots , tal que L_i contiene un predicado de verdad para L_{i+1} .¹⁹

Si esto es aceptado entonces podemos describir de una manera ejemplar los niveles tarskianos del siguiente modo:

- nivel₁: Cervantes enuncia (2) en \mathcal{L} .
- nivel₂: Alonso Quijano enuncia (2) en $\mathcal{L}_{f \text{ Descr.}}$. Cervantes enuncia (2) en boca de Alonso Quijano en \mathcal{L}_f .
- nivel₃: Don Quijote enuncia (2) en $\mathcal{L}_{f+1 \text{ Descr.}}$. Cervantes enuncia (2) en boca de don Quijote refiriéndose a Alonso Quijano.

Vamos a completar los datos expuestos en el Esquema 1 atendiendo a un marco que admita niveles transfinitos del siguiente modo:

$$\begin{aligned} \delta_0 &:= && \text{"(2)"} \\ \delta_f &:= && \text{"}\delta_0\text{" es verdadero} \\ \delta_{f+1} &:= && \text{"}\delta_f\text{" es verdadero} \\ & \cdot && \\ & \cdot && \\ & \cdot && \end{aligned}$$

$$\delta_\omega := \text{"para todo } f \delta_{f+n} \text{ es verdadero"}$$

de modo que lo podríamos exponer mediante un segundo esquema del siguiente modo:

¹⁹ Véase: Kripke 1975, 697. Traducción al castellano en: Kripke 1984, 17.

"(2)"	δ_0	δ_f	δ_{f+1}	...	δ_w
nivel ₁	f	-	-	...	-
nivel ₂	-	v	-	...	-
nivel ₃	-	-	v	...	-
.
.
nivel _w	-	-	-	...	v

Esquema 2

Así pues, es altamente improbable que Cervantes emita (2) en el nivel₁ debido a razones de carácter empírico. Si Alonso Quijano hiciese lo mismo, tendría un resultado diferente ya que la preferencia la escribe Cervantes a un nivel más bajo. Lo mismo ocurre con respecto al nivel₃. Si Cervantes emite (3), este enunciado no sería ni verdadero ni falso. Su consistencia viene de la mano de que tenga un criterio adicional, es decir, el que se tenga presente que el predicado de verdad no puede estar contenido en su mismo nivel. Para nuestro caso se cumplen los requisitos de A. Tarski de que la "verdad" de un enunciado para un nivel depende del predicado del siguiente nivel. De este modo el único enunciado arriesgado sería el del nivel₁ en tanto que si está bien formado, ha de contrastarse con la realidad. El problema ahora es saber cuáles son las condiciones a las que atienden los índices de riesgo del nivel₁.

4. La propuesta fundamentacional

Veamos si aplicando la *propuesta fundamentacional* a la paradoja de la ficción se pueden resolver los problemas arriba expuestos. La propuesta kripkiana parte del rechazo de una definición del predicado de verdad como arriba fue definido en tanto que transforma el aparato referencial de tal

modo que SPM_1 es caracterizado de un modo genuino;²⁰ segundo, se rechaza SPM_2 al modo como es entendido por A. Tarski,²¹ y por último se restringe el *tertium non datur*, delimitando así SPM_3 en base a una lógica trivalente.²²

De este modo parece ser que la propuesta fundamental es el opositor más importante a la propuesta tarskiana por lo que valdría un análisis detenido.²³ Esta posición parte de la situación nada anormal de tener que explicar a un individuo el significado de "verdadero". Para ello un modo coherente de dilucidar dicho significado puede venir de la mano de que podamos aseverar de un enunciado que tiene derecho a ser aseverado por lo que en consecuencia un enunciado sería verdadero siempre y cuando tenga derecho a ser aseverado y falso cuando no se tiene derecho a ser aseverado.²⁴ Consecuentemente, nociones como "aseveración", "justificación" y "fundamentación" pertenecen a una misma familia de conceptos que permiten explicar el significado de "verdadero". Dado que un interlocutor tiene derecho a aseverar:

(5) La nieve es blanca.

Según la propuesta fundamental se puede aseverar de (5) con derecho:

(6) "La nieve es blanca" es verdadera.

²⁰ Vease: Padilla Gálvez 1989, Cap. III.

²¹ Vease: Padilla Gálvez 1991, 137ss.

²² En concreto, se analizan aquellos lenguajes que permiten vacíos de verdad. Vease: Padilla Gálvez 1992.

²³ No ha de olvidarse al respecto las críticas llevadas a cabo por H. Field, en la que se achaca a Tarski el que su propuesta no tiene sentido por el hecho de que no se aplica al lenguaje natural y su definición sólo abarca el lenguaje formal (Field, 1972, 347ss). S. Soames ha argumentado contra las críticas de Field contra la propuesta tarskiana, que su crítica no atiende a las propiedades semánticas de los enunciados sino de las expresiones (Soames 1984, 411ss.). Así pues, A. Tarski nunca planteó problemas que tuviesen que ver con las expresiones lingüísticas.

²⁴ La forma reflexiva "aseverarse" indica un momento inicial definido en la lógica como aquello que se denomina "instanciación universal".

La idea intuitiva de la propuesta fundamentacional es que un enunciado está fundamentado para siempre en el caso de que a partir de un determinado momento alcance un valor de verdad. Dicha idea intuitiva viene a ser caracterizada mediante la exigencia de la *monotonidad*.²⁵ Dicho argumento descansa en una tesis realista según la cual sólo hay una descripción verdadera de la manera como es el mundo, y esta manera es fija. Cuando alcanzamos dicho nivel nos encontramos ante el modo genuino y verdadero de enunciar algo sobre el mundo. En el marco de los sistemas autorreferenciales nos encontramos con tipos muy diferentes. Los sistemas autorreferenciales pueden ser del tipo de los *relatores de verdad*.²⁶

(7) (7) es verdadero.

En los relatores de verdad, el enunciado se refiere a sí mismo, enunciando su propia verdad. En (9) tenemos un relator de verdad en la que 'a' denota 'Ta'. De este modo, 'Ta' afirma sobre sí mismo, que es verdadero. Si \mathcal{L} contiene una sintaxis adecuada, no parece problemática la construcción de oraciones que dicen de sí mismas que son verdaderas. Formalmente, los relatores de verdad pueden ser formalizados del siguiente modo. Sea Pa un predicado que satisface únicamente su propia oración, entonces vale:

(8) $\forall x(Px \rightarrow T\langle AFIR \rangle x)$,

es decir, todos los enunciados que satisfacen o se afirman a sí mismos son verdaderos.²⁷

²⁵ Kripke, 1975, 700, nota 18.

²⁶ Traduzco la expresión inglesa "Truth Teller" por el de "relator de verdad" al castellano. Los relatores de verdad dicen que es verdad lo que se dice. Por lo tanto dice la verdad reduplicativamente. Es por esta razón por la que voy a formalizar el tipo de relatores de verdad mediante la expresión siguiente: $T\langle AFIR \rangle x$. En el caso de la paradoja descrita en el Quijote, los *relatores de verdad* (RV) tendrían la forma de enunciado del tipo:
(RV) Estoy vivo.

Dichos enunciados son por lo general, en una situación descontextualizada, de carácter trivial.

²⁷ El enunciado (8) envuelve una deducción interesante que ilustra algunos elementos de trabajo claves, puesto que: $\forall x(Px \rightarrow T\langle AFIR \rangle x) \vdash \forall x T\langle AFIR \rangle x$, esto puede deducirse de la siguiente manera:

- | | |
|---|---------|
| (1) $\forall x(Px \rightarrow T\langle AFIR \rangle x)$ | Premisa |
| (2) Pa | Premisa |

Sin embargo, no todos los enunciados alcanzan un valor de verdad. Las paradojas como:

(9) (9) es falso,

son un tipo de aquellos enunciados que no alcanzarán ni a ser verdaderos ni falsos. La paradoja del mentiroso puede formularse de la siguiente manera:

(10) $\forall x(Px \rightarrow T(\text{NEG})x)$,

en este caso, Px es un predicado que satisface únicamente la propia oración [(10)].²⁸

Según Kripke, (9) nunca será verdadera ni se podrá estipular el que (9) no sea verdadera. Por lo tanto (9) no recibirá ningún valor de verdad. Esto supone pues que ha de recurrirse a la lógica trivalente. Con (9) le ocurriría algo así como con (2), a saber, que el enunciado pertenezca al ámbito del lenguaje ficcional, es decir, a un nivel más alto, por lo que, como hemos

(3) $Pa \rightarrow T(\text{AFIR})a$	E.G., 1
(4) $T(\text{AFIR})a$	M.P., 2 y 3
(5) $\forall xT(\text{AFIR})x$	I.G., 4

²⁸ La prueba del enunciado (10) nos permite conocer algunos rasgos del tipo de operaciones que proponemos, a saber:

(1) $\forall x(Px \rightarrow T(\text{NEG})x)$	Premisa
(2) Pa	Premisa
(3) x	Premisa
(4) $Px \rightarrow T(\text{NEG})x$	Por instanciación
(5) $T(\text{NEG})x$	M.P., 4 y 2
(6) $\neg x$	Aplicación del esquema-T
(7) $\neg x$	Mediante el lema, según el cual $\{\neg x\} \vdash \neg x$
(8) $\neg \forall x(Px \rightarrow T(\text{NEG})x)$	Premisa
(9) $\exists x(Px \wedge \neg T(\text{NEG})x)$	Leyes de la interdefinición de los cuantificadores
(10) $Pa \wedge \neg T(\text{NEG})x$	E.P., 9
(11) $T(\text{NEG})x$	Mediante el esquema-T
(12) $a \neq x$	Principio de identidad
(13) $\exists x(Px \wedge x \neq x)$	I.P.

La formulación de la paradoja del mentiroso no se debe a que termina en una contradicción. La conclusión es totalmente consistente. La dificultad radica en que la misma variable puede ser nombrada de modo diferente y tener diferentes propiedades, de modo que conduce a (13). Esa es el resultado al que nos ha conducido también la paradoja de don Quijote. Téngase en cuenta que la conclusión a la que nos orientamos introduce diferentes tipos de negación.

podido comprobar, se pone en un aprieto a don Quijote al forzar el lenguaje y aseverar que si que se muere.

Según S. A. Kripke, tiene sentido decir que poseemos determinadas intuiciones sobre la referencia y que estas intuiciones son básicas para toda teoría semántica. Sin embargo dicha propuesta altera nuestro sentido de plantear las cuestiones acerca de la referencia. Nos exige exponer los problemas arriba expuestos mediante la siguiente pregunta: (α) ¿Cuál es la mejor forma de expresar la verdad o falsedad de las creencias habituales de don Quijote -aduciendo que no trataban de nada en absoluto- que son verdades sobre un hecho ficticio? Dicha cuestión sólo se puede contestar si partimos de la respuesta a la siguiente cuestión: (β) ¿Hay una entidad del mundo que esté conectada con el uso de "don Quijote" por la relación de referencia?

5. Breve excursio sobre la teoría de modelos

Según el fundamentacionismo, responder a (α) es un hecho que está determinado por la respuesta que demos a (β). Supongamos pues que para dar una respuesta coherente a (β) tenemos que introducir algunos cambios cualitativos en nuestra concepción estándar. Estos cambios pueden ser indicados de un modo más plausible si introducimos la *teoría de modelos*.²⁹ Para responder adecuadamente a (β) no podemos definir:

(T_{MT}) [Verdad-en- \mathcal{L} relativo a un modelo],

que es, dicho de paso, el modo usual con el cual hemos intentado exponer adecuadamente lo que ocurre con la paradoja de la ficción.³⁰ S.A. Kripke introduce una variación con respecto a la teoría de modelos estándar e introduce:

(T_{MK}) [Verdad-en- \mathcal{L} (*simpliciter*) para una determinada concepción del lenguaje].

²⁹ Véase al respecto el trabajo de Vaught 1986, 869.

³⁰ A. Tarski define la noción de verdad en términos de satisfacción, y con la ayuda del concepto de satisfacción, el concepto de modelo (Tarski 1936, 9).

Veamos dicho cambio detenidamente. Examinemos el lenguaje estándar de primer orden, \mathcal{L} , como un tripló $\langle C_{\mathcal{L}}, D_{\mathcal{L}}, F_{\mathcal{L}} \rangle$, donde $C_{\mathcal{L}}$ es una función de conjuntos que representan diferentes categorías de \mathcal{L} -expresiones bien formadas; $D_{\mathcal{L}}$ es el dominio de objetos; y $F_{\mathcal{L}}$ es una función que hace corresponder los nombres de \mathcal{L} a los objetos de $D_{\mathcal{L}}$, a los predicados monádicos de \mathcal{L} hace corresponder subconjuntos de $D_{\mathcal{L}}$, etc.³¹ Si \mathcal{L} es una clase de susodichos lenguajes, entonces, según la propuesta fundamentacional, se demuestra que se pueden definir en conceptos no-semánticos para variables de \mathcal{L} , en \mathcal{L} , según los modos expuestos por A. Tarski. La única salvedad respecto a éste es que los conceptos de *referencia primitiva* no son definidos mediante niveles específicos, como los elaborados en los esquemas 1 y 2, sino para variables de \mathcal{L} mediante las funciones interpretadas en determinados lenguajes. En concreto: un nombre n refiere a un objeto o en un lenguaje \mathcal{L} si, y sólo si se cumple:

$$F_{\mathcal{L}}(n) = o.^{32}$$

Según dicha definición, los lenguajes aparecen caracterizados como objetos abstractos, de los cuales se puede suponer que portan las *propiedades semánticas* de un modo esencial. No existe la posibilidad de que las expresiones de dichos lenguajes refieran algo diferente de lo que realmente denotan. Cada variación en los niveles ficcionales de las propiedades semánticas supone una variación del lenguaje. El resultado es que las propiedades semánticas no se encuentran en dependencia alguna.

³¹ Téngase en cuenta que la noción de interpretación de un lenguaje formal y sus frases, en un universo concreto de objetos, que aquí estamos barajando, se adapta al concepto de verdad matemático. Si presentamos una interpretación exacta de los signos de \mathcal{L} , propios de la teoría que manejamos, y una interpretación estándar de los signos lógicos $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$ y los cuantificadores $\{\exists, \forall\}$; las variables $\{x, y, z, \dots\}$ se referirán a objetos de un determinado dominio de la interpretación y los signos propios de la teoría se traducen en relaciones entre objetos del dominio. Así pues, las expresiones formales que desarrollemos se convierten en expresiones que especifican los objetos de determinados dominios y, en ellos, si determinamos las propiedades de que gozan los objetos de dichos dominios, serán verdaderas o falsas.

³² Téngase en cuenta que $F_{\mathcal{L}}$ es concebido como un objeto matemático puro, es decir un conjunto de pares ordenados. De este modo se refiere a un concepto determinado definido semánticamente.

6. La propuesta fundamentacional formalizada

Sea un lenguaje de primer orden, como el arriba descrito, cuyo dominio D_x contiene los números naturales y puede ser extendido a un lenguaje \mathcal{L} mediante un predicado monádico adicional Px . Un subconjunto parcial de D_x , $S = (S_1, S_2)$ es un par de subconjuntos que no están relacionados entre sí en D_x . $S_1(S_2)$ contienen los elementos de S . S se encuentra sin definir. $S \leq S'$ significa $S_1 = S'_1$ y $S_2 = S'_2$. Si se interpreta Px mediante un subconjunto parcial de \mathcal{L} entonces, según la propuesta fundamentacional, se puede interpretar como verdadero, falso o indeterminado.³³ Puder ser determinado de la siguiente manera:

Dado un dominio no vacío D , un predicado monádico $P(x)$ se interpreta mediante un par (S_1, S_2) de conjuntos disyuntos de D . S_1 es la extensión de $P(x)$ y S_2 es su antiextensión. $P(x)$ ha de ser verdadero de los objetos en S_1 , falso de aquéllos en S_2 , de otra manera será indefinido. La generalización de esto para predicados n -ádicos es obvia.³⁴

S. A. Kripke reconoce explícitamente la lógica trivalente expuesta por S. C. Kleene.³⁵ Si $P(x)$ es interpretado mediante $S=(S_1, S_2)$ y por lo tanto vale: $\phi(S)=S'=(S'_1, S'_2)$, en donde $S'_1(S'_2)$ es el conjunto de la numeración de Gödel de enunciados verdaderos (falsos) de \mathcal{L} en la interpretación propuesta, entonces se cumple $S_1 \leq S_2$, $\phi(S_1) \leq \phi(S_2)$.³⁶ Expresado de otra manera esto supone afirmar:

Sea S'_1 el conjunto de (códigos de) [es decir: los números de Gödel de] las oraciones verdaderas de $\mathcal{L}(S_1, S_2)$ y sea S'_2 el conjunto de todos los elementos de D que o no son (códigos de) [es decir: los números de Gödel de] oraciones

³³ Véase: Padilla Gálvez 1992.

³⁴ Véase: Kripke 1975, 700. En la traducción al castellano: Kripke 1984, 20. La interpretación que llevamos arriba es decididamente más escueta pero igual de apropiada.

³⁵ Véase al respecto: Kripke 1975, 711, y Kleene 1971, 332 ss. trad. esp. 1974, 301 ss. Parte dicha tesis de la base de que $P(x)$ es vaga e intuitiva por lo que no puede ser demostrada. Sin embargo permite garantizar determinadas funciones particulares para lo cual es sumamente efectiva, y por otro lado permite reconocer que nuestro conocimiento acerca de muchas otras funciones es insuficiente.

³⁶ Como se sabrá hay una numeración de Gödel de los objetos formales tal que los predicados definidos de un modo determinado son expresables numeralmente en el sistema formal. A esta caracterización se está haciendo referencia en estas páginas.

de $\mathcal{L}(S_1, S_2)$ o son (códigos de) [es decir: los números de Gödel de] oraciones falsas de $\mathcal{L}(S_1, S_2)$. La elección de (S_1, S_2) determina la manera única a S'_1 y S'_2 . Si $T(x)$ ha de interpretarse como la verdad para el lenguaje mismo \mathcal{L} que contiene al propio $T(x)$, obviamente debemos tener $S_1=S'_1$ y $S_2=S'_2$... Un par (S_1, S_2) que satisface esta condición se llama un *punto fijo*.³⁷

De lo antes expuesto resulta un teorema sencillo, a saber: si tiene ϕ un *punto fijo* entonces, definiendo $S_0 = \{\text{subconjunto completamente sin definir}\}$, entonces se da el caso $S_{\alpha+1} = \phi(S_\alpha)$ siempre y cuando β sea un límite ordinal para el que vale $\phi(S_\beta) = S_\alpha$ ($\alpha < \beta$) y, entonces, ha de considerarse α (α_0) como un mínimo, de modo que $S_\alpha = S_{\alpha+1}$ es el mínimo punto fijo de ϕ .³⁸ Cada uno de los puntos fijos son en un *método de cadena* (a saber, el que se expresa mediante ' \leq ') un *punto fijo maximal*.

...si $(S_1, S_2) \leq (S_1^+, S_2^+)$ entonces cualquier oración que sea verdadera (o falsa) en $\mathcal{L}(S_1, S_2)$ retiene su valor de verdad en $\mathcal{L}(S_1^+, S_2^+)$. Lo que esto significa es que la interpretación de $T(x)$ se amplía dándole un valor de verdad definido a algunos casos previamente indefinidos, ningún valor de verdad previamente establecido cambiará ni se hará indefinido; cuando mucho, algunos valores de verdad previamente indefinidos se vuelven definidos.³⁹

Según la propuesta fundamentacional no se puede presentar el mayor punto fijo, pero sí existe el *mayor punto fijo intrínseco*. La estructura de lo que se denomina punto fijo intrínseco, entonces es una red completa bajo el método de cadena en el que se aplica ' \leq '.

De este modo hemos completado todos los requisitos para llegar a una primera conclusión que afecta a los niveles tarskianos, a saber: si $P(x)$ es interpretado mediante un punto fijo de ϕ , entonces \mathcal{L} es un lenguaje con su propio predicado de verdad.

³⁷ Véase: Kripke 1975, 702 y la traducción al castellano: Kripke 1984, 24s. En dicha traducción he introducido los corchetes con el fin de entender mejor que la noción de "códigos de" hace referencia a los números de Gödel que representan la sintaxis en \mathcal{L} mediante su artificio codificador.

³⁸ Este procedimiento ha sido denominado por M. Fitting el "*fixed point approach*" y sirve, entre otras cosas, como solución de la asignación del valor de verdad de un enunciado en un lenguaje que contiene autorreferencias (Véase: Fitting 1989, 225 ss.).

³⁹ Véase: Kripke 1975, 703. Traducido en: Kripke 1985, 26.

La propuesta fundamentacional considera que la interpretación de \mathcal{L} use para los puntos fijos de ϕ un modelo aproximado para el concepto intuitivo que expresa 'verdad-en-el-lenguaje-natural'. El punto fijo más bajo es probablemente el más natural: sin embargo, los demás niveles son útiles para llevar a cabo diferenciaciones intuitivas o efectuarlas de un modo más preciso. De este modo (2) expresada por boca de don Quijote no se puede asignar, según Kripke, en ningún punto fijo un valor de verdad. En base a estos datos se lleva a cabo una variación de la propuesta tarskiana de las jerarquías. Así si Sancho Panza afirma:

(11) Todo lo que afirma don Quijote es verdadero,

entonces ha de elegir un nivel con el cual pueda medir el predicado "verdad". Si algunas de las expresiones de don Quijote se encuentran en un nivel demasiado alto, entonces (11) puede cubrir todo lo que afirma don Quijote. Según dicha propuesta no puede conceder al predicado "verdadero" ningún nivel. Si (11) tiene un valor de verdad en un punto fijo minimal como es el caso normal, entonces tiene en un sentido determinado, un sentido descrito como el más pequeño β de modo que se cumple: $(11) \in S_\beta$. Sin embargo, según Kripke el nivel depende de lo que afirme don Quijote y no tanto del predicado de verdad de (11).

Si \mathcal{L}_0 está sujeto a la aritmética, entonces podemos construir una jerarquía tarskiana para los predicados de verdad y los metalenguajes para \mathcal{L}_0 en el marco de \mathcal{L} . El planteamiento es el siguiente: Si $A(x)$ es verdadero para la fórmula de \mathcal{L}_0 , entonces $P_0(x) = P(x) \wedge A_0(x)$, en la que \mathcal{L}_1 es un sublenguaje que se construye mediante la agregación de $P_0(x)$ a \mathcal{L}_0 . Dicha construcción puede proseguirse mediante números de cifras superiores a α_0 . Enunciados como el expuesto en (11) no aparecen en ningún nivel de la jerarquía de Tarski. De aquí resulta que al establecerse un punto fijo sobre una extensión de $P(x)$ su valor de verdad ni cambiará ni se hará indefinido. De este modo la convención tarskiana puede ser caracterizada en el modo propuesto por Kripke (abreviadamente: T_K) del siguiente modo:

$(T_K): T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner). \supset . (A \equiv T(\ulcorner A \urcorner) \wedge (\neg A \equiv T(\ulcorner \neg A \urcorner)))$.

Mediante (T_K) se consigue una teoría axiomatizada de "verdad".⁴⁰ En

⁴⁰ Según G. Priest, se puede extender el mentiroso a enunciados de la forma $T\ulcorner A \urcorner \wedge *T\ulcorner A \urcorner$,

(T_K) se afirma simplemente que se puede desarrollar un algoritmo para obtener el valor de verdad de $T(\ulcorner A \urcorner)$ o $T(\ulcorner \neg A \urcorner)$ a partir de $A \equiv T(\ulcorner A \urcorner)$ y $(\neg A \equiv T(\ulcorner \neg A \urcorner))$. Un algoritmo para $T(\ulcorner A \urcorner)$ o $T(\ulcorner \neg A \urcorner)$, *ab initio* tendrá acceso a información acerca de $T(\ulcorner A \urcorner)$ y $T(\ulcorner \neg A \urcorner)$ solo por el mero uso de algoritmos para dichas caracterizaciones, y que están incorporadas a él.⁴¹ Una decisión alcanzada por aplicación del algoritmo para un enunciado, por caso, en el que su valor sea verdadero, debe haberse basado en informaciones acerca de uno u otro de los componentes, que haya sido producido en algunos estadios finitos de la aplicación de los algoritmos correspondientes a $T(\ulcorner A \urcorner)$ y $T(\ulcorner \neg A \urcorner)$. En cualquier estadio del algoritmo correspondiente a $T(\ulcorner A \urcorner)$, o bien habremos hallado que $T(\ulcorner A \urcorner)$ es "verdadero", o bien que $T(\ulcorner A \urcorner)$ es "falso" o simplemente no habremos averiguado el valor de verdad de $T(\ulcorner A \urcorner)$. Si $T(\ulcorner A \urcorner)$ es realmente no definible, sólo podrá averiguarse por un razonamiento metateórico acerca del algoritmo.⁴² Es aquí donde la propuesta fundamentacional ve legítimo introducir un *tercer valor de verdad*.⁴³

Si $D_{\mathcal{L}}$ contiene una sucesión finita de objetos del dominio, entonces se puede elaborar una construcción análoga de la extensión de \mathcal{L} , extendiéndose a un lenguaje que tiene un predicado de satisfacción.

7. Crítica y propuestas

Hemos comenzado exponiendo una paradoja en la ficción como un caso curioso en el que, sin ser óptima, su visión resulta fascinante. La paradoja es aparente: ciertamente las cuestiones que genera son tan sólo brillantes, como mucho puede aportar algunas notas de interés para los niveles de lenguaje. Lo que Cervantes expone, aunque con técnica literaria adecuada,

pero siempre que valga $\{T(\ulcorner A \urcorner) \wedge *T(\ulcorner A \urcorner)\} \models B$. En donde, '*' es un functor, según el cual $\ulcorner *A \urcorner$ es verdadero precisamente si $\ulcorner A \urcorner$ no es verdadero, o falso (Priest 1984, 161).

⁴¹ Una consecuencia de la formalización que aquí expongo es que difiero de la propuesta de Priest 1984, 161.

⁴² Estoy pues de acuerdo con los razonamiento de W. Rabinowicz según los cuales es preferible, en base a las conclusiones a las que llegamos, el modelo de Beth al de Kripke, véase: Rabinowicz 1985, 209 ss.

⁴³ Véase: Padilla Gálvez 1992.

hurga con explicitud en la teoría tarskiana acerca de la noción de verdad, que sigue abierta, como se puede deducir de las discusiones recientes, planteando nuevas incógnitas. El entramado argumental de la paradoja de la ficción nos plantea aquellos problemas que son capitales para toda propuesta semántica a saber, de qué tipo es la relación entre el lenguaje-objeto y metalenguaje.

He propuesto en otro lugar el argumento según el cual la propuesta fundamentacional es insatisfactoria en tanto que es un tanto desilusionador encontrar que la innovación que propone rechaza elementos importantes de la teoría tarskiana. Dichos argumentos atenderían a lo que podríamos denominar un acuerdo de intenciones. La falta de la propuesta fundamentacional se encuentra en que la solución dada a la paradoja de don Quijote en el lenguaje objeto aparece de nuevo en la formulación de los resultados básicos en el metalenguaje de manera diferente. Así pues, en vez de interpretar (2) mediante la variante del mentiroso desarrollada en (9), se puede generar un mentiroso débil que refuta todos los argumentos en contra suya y que viene expuesto por la negación de (11), afirmando:

(12) (12) no es verdadero.

Tanto el mentiroso que viene a ser expuesto en (2), así como el mentiroso fuerte que asevera (12) aparecen representados mediante la reconstrucción kripkiana en el tercer valor de verdad. Debido al aspecto parcial del predicado de verdad en el sistema de Kripke, (12) es equivalente a afirmar:

(13) (13) no posee valor de verdad.

De este modo (13) expresa de manera intuitiva el resultado de la reconstrucción kripkeana de la paradoja del mentiroso, lo que supone reconocer que el mentiroso es infundado. Lo que S. A. Kripke no asume en su propuesta es que el predicado "verdadero" se entiende de un modo ambiguo. Según su propuesta se presupone un mentiroso que reflexiona y un mentiroso que no reflexiona. En la lógica autorreferencial valen pues dos actitudes: (i) si no reflexionamos, entonces el valor de verdad de un enunciado está infundado, siendo su prototipo de enunciados el expuesto en el caso (10). En el momento que esto viene a ser aceptado, comienza a oscilar el valor de verdad. El mentiroso que reflexiona -como el caso tratado en Alonso Quijano/don Quijote- no es verdadero de modo definitivo

en ningún nivel finito (a este resultado hemos llegado en el esquema 2). Por todo esto, el mentiroso que reflexiona emite un enunciado que se convierte en verdadero en el nivel- ω bajo la condición de que δ_ω permita algún tipo de reflexión. Dicha reflexión está de acuerdo con la intuición que se expone en la jerarquía tarskiana, según la cual ha de usarse un nuevo predicado de verdad en cada nivel. Dicha alteración del predicado de verdad tiene la función de ser un ajuste y sus razones han sido esbozadas en el esquema 1.

De este modo se genera una nueva disquisición entre lo fundado y lo infundado. Consecuentemente, no se puede ajustar adecuadamente el resultado principal de la investigación llevada a cabo por Kripke en el sistema kripkeano. El mentiroso vuelve a campar a sus anchas por su sistema, lo que equivale a decir que Cervantes es víctima de él mismo, en su confusión.

Hemos reconstruido la propuesta anterior partiendo de la base de un dato importante que resulta de la respuesta a la siguiente cuestión, ¿cómo podemos decir de un enunciado que se tiene derecho a aseverarse tal como verdadero o falso? La intuición fundamentacional expuesta sobre la oración típicamente fundamentada se puede reducir a la relación $F_\varphi(n) = o$. Es para mí una sorpresa que la propuesta fundamentacional no garantice de manera obvia la fundamentación ni la exigencia de la monotonicidad, en el caso del mentiroso reflexivo, como arriba hemos reconstruido. Consecuentemente, no es del todo evidente que en todas las instancias donde se presenta la referencia de una categoría lingüística determinada sea rígida aún cuando se caracterice como tal. Una de las razones para dudar de la efectividad de tal caracterización parte del argumento siguiente: no es posible asertar el derecho de un enunciado, si en una de las partes se sustituye uno de los referentes. Así podemos afirmar que las propiedades que caracterizan el predicado adscrito a X y a Y son diferentes a pesar de que sean un mismo individuo ficticio. El método puede exponerse del siguiente modo (teniendo en cuenta (C) y (A) para lo que vale el siguiente esquema):

- (14) El individuo X se caracteriza mediante una relación \mathfrak{R} en tanto que asume Φ con Y.
- (15) El individuo ficticio X = el individuo ficticio Y.

- (16) La propiedad Γ que se adscribe Y a sí mismo no tiene por qué ser adscrita por X.

De este modo la identidad de las propiedades expresadas en un predicado es notoriamente problemática. Si para la presente cuestión de la verdad se asume una versión por la cual la propiedad Γ que se adscribe un individuo no puede adscribir a ese mismo individuo con diferente nombre.

Se sobrentiende que si rechazamos la referencia objetual tendremos que presentar un principio análogo por el cual determinemos las estructuras semánticas. Las razones para un tal rechazo se pueden exponer del siguiente modo: Un aspecto relevante de la teoría de la verdad tarskiana, es que caracteriza la verdad de los enunciados mediante una relación -es decir, la de "satisfacción"- entre lenguajes -enunciados abiertos- y el mundo -secuencias de objetos. Sobre el concepto de correspondencia, que constituye una pieza clave de la teoría de la verdad, expresó satisfactoriamente todo lo que en su sistema podía decir Tarski para los lenguajes formales. En resumidas cuentas, contra la crítica -así como la propuesta- llevada a cabo por el programa fundamentacional que intenta echar por tierra el modelo tarskiano sólo hay una posible salida: demostrar que dicho programa es defectuoso.

Universidad de León

Bibliografía

- ARISTOTELES, 1957: *Metafísica*. (Rec. W. Jaeger). Oxford: Oxford University Press.
- BLAU, U., 1983: "Vom Henker, vom Lügner und Ihrem Ende", *Erkenntnis*, 19, 27-44.
- BLAU, U., 1985: "Die Logik der Unbestimmtheiten und Paradoxien", *Erkenntnis*, 22, 369-459.
- BLAU, U., 1986: "Die Paradoxien des Selbst", *Erkenntnis*, 25, 177-196.

- BLAU, U., 1987: "Zur natürlichen Logik der Unbestimmtheiten und Paradoxien". Manuscrito.
- BLAU, U., 1990: "Zur natürlichen Logik der Unbestimmtheiten und Paradoxien", *Synthesis Philosophica*, 10, 593-615.
- ETCHEMENDY, J., 1988: "Tarski on truth and logical consequence", *Jour. Symb. Logic*, 53, 51-79.
- FIELD, H. 1972: "Tarski's Theory of Truth", *The Journal of Philosophy*, 69, 347-375.
- FITTING, M., 1989: "Bilattices and the theory of truth", *Jour. Phil. Logic*, 18, 225-256.
- FRAASSEN, B. van, 1968: "Presuppositions, implication, and self-reference", *The Journal of Philosophy*, 65, 136-152.
- KLEENE, S. C. 1971: *Introduction to Metamathematics*. Amsterdam-New York: North-Holland. (Trad. cast. *Introducción a la metamatemática*, Madrid: Ed. Tecnos, 1974).
- KRIPKE, S. A. 1975: "Outline of a Theory of Truth", *The Journal of Philosophy*, 72, 690-716. (Trad. cast. *Esbozo de una teoría de la verdad*. Cuadernos de *Crítica*. México: UNAM, 1984).
- KRIPKE, S. A., 1976: "Is there a Problem about Substitutional Quantification?", en: G. Evans/J. McDowell (Eds.) *Truth and Meaning. Essays in Semantics*. Oxford: Oxford University Press, 325-419.
- KÜNNE, W. 1983: *Abstrakte Gegenstände. Semantik und Ontologie*. Frankfurt a.M.: Suhrkamp.
- LEWIS, D., 1983: "Truth in Fiction", en: *Philosophical Papers*. Vol. I. New York: Oxford Univ. Press, 261-275.
- MARTIN, R. L., (Ed.), 1984: *Recent Essays on truth and the liar paradox*. Oxford: Oxford University Press.
- MARTIN, R. L. y Woodruff, P. W., 1975: "On representing 'True-in-L' in L", *Philosophia*, 5, 213-217. Publicado en: Martin, 1984, 47-51.
- PADILLA GALVEZ, J. 1988-89: "Referencia y verdad", *THEORIA*, 10, 197-213.
- PADILLA GALVEZ, J., 1989: *Referenz und Theorie der möglichen Welten*. Frankfurt a. M.: Peter Lang.
- PADILLA GALVEZ, J., 1991: "Wahrheit und Selbstrückbezüglichkeit", *Jour. for Gene. Phil. of Science*, 22, 137-154.

- PADILLA GALVEZ, J., 1992: "La concepción semántica de la verdad, la paradoja del mentiroso y los sistemas autorreferenciales", en: *Lenguajes Naturales y lenguajes formales*, VII. 1. Barcelona. (En prensa).
- PRIEST, G., 1979: "The Logic of Paradox", *Jour. Phil. Logic*, 8, 219-241.
- PRIEST, G., 1984: "Logic of Paradox revisited", *Jour. Phil. Logic*, 13, 153-179.
- QUINE, W.v.O., 1983: *Mathematical Logic*. Edición revisada. Cambridge Mass: Harvard University Press.
- RABINOWICZ, W., 1985: "Intuitionistic truth", *Jour. Phil. Logic*, 14, 191-228.
- SMULLYAN, R., 1957: "Language in which self-reference is possible", *Jour. Symb. Logic*, 19, 37-40.
- SOAMES, S. 1984: "What is a Theory of Truth", *The Journal of Philosophy*, 81, 411-429.
- STACHNIAK, Z., 1988: "Two theorems on many-valued logics", *Jour. Phil. Logic*, 17, 171-179.
- STACHNIAK, Z., 1989: "Many-Valued Computational logic", *Jour. Phil. Logic*, 18, 257-274.
- TARSKI, A., 1935: "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philosophica*, 1, 261-405.
- TARSKI, A., 1936: "Über den Begriff der logischen Folgerung". *Actes du Congrès International de Philosophie Scientifique*, Paris 1935, Vol. VII, ASI 394, 1-11.
- TARSKI, A., 1969: "Truth and Proof", *Scientific American*, 220, Nr. 6, 63-77.
- TARSKI, A., 1986: *Alfred Tarski, Collected Papers*. 4 Vols. Basilea: Birkhäuser.
- VAUGHT, R., 1986: "Alfred Tarski's work in model theory", *Jour. Symb. Logic*, 51, 869-882.
- VISSER, A., 1984: "Four valued semantics and the liar", *Jour. Phil. Logic*, 13, 181-212.
- VISSER, A., 1989: "Semantics and the liar paradox", en: D. Gabbay y F. Guenther (Eds.), *Handbook of Philosophical Logic*, Vol. IV. Dordrecht: Reidel, 617-706.