

# VERDAD Y CIRCULARIDAD: EL PROBLEMA DE LA SUPERVENIENCIA SEMANTICA†

(*Truth and Circularity: The Problem of Semantic Supervenience*)

Eduardo Alejandro BARRIO\*

Manuscrito recibido: 2000.12.26.

Versión final: 2001.10.30.

\* Instituto de Filosofía, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Puán 470, 1406 Ciudad Autónoma de Buenos Aires, Argentina.

E-mail: ebarrio@filo.uba.ar

BIBLID [0495-4548 (2002) 17: 43; p. 63-79]

RESUMEN: Belnap y Gupta sostienen que la verdad es un concepto circular: su extensión no puede ser establecida si ella no es previamente hipotetizada. Yaqûb argumenta que esta circularidad es incompatible con la conjunción de la tesis de la superveniencia semántica y el principio de bivalencia. En este artículo, intento reforzar con nuevos elementos la posición de Yaqûb. Aunque se debilite la tesis de la superveniencia de una manera que Yaqûb no ha considerado, no parece fácil advertir cómo Belnap y Gupta pueden escapar a la mencionada incompatibilidad.

Descriptores: verdad, paradojas, concepto circular, superveniencia semántica, teoría de la revisión.

ABSTRACT: *Belnap and Gupta have recently maintained that truth is a circular concept: its extension cannot be established without being previously hypothesized. This has led Yaqûb to claim that the circular character in question cannot be made compatible with the thesis that semantic properties are supervenient ones. Belnap and Gupta have explicitly denied such a claim any plausibility. In this paper, I offer some new arguments in support of Yaqûb's position. Such arguments are based on an analysis of some aspects of Belnap and Gupta's theory that, as far as I know, had not been considered before.*

Keywords: *truth, paradoxes, circular concept, semantic supervenience, revision theory of truth.*

## SUMARIO

1. Los límites de la extensión de la verdad
2. El mentiroso pide la palabra
3. El honesto y el mentiroso: la tensión entre la superveniencia y la bivalencia

Bibliografía

La tesis según la cual las afirmaciones no semánticas constituyen el único sustento para el establecimiento de las afirmaciones semánticas parece fundarse en una simple intuición. Siempre que hayamos establecido los va-

lores de verdad de todas las oraciones que no contienen expresiones semánticas, habremos establecido, por esa simple razón, los valores de verdad de las oraciones que las contienen. El trasfondo redundantista que subyace a esta tesis puede advertirse con mayor claridad en la siguiente formulación en términos formales: según esta idea, para cualquier modelo básico de un lenguaje,<sup>1</sup> debería haber exactamente una sola interpretación del predicado veritativo. Este principio ha sido llamado por Michael Kremer *tesis de la superveniencia semántica* (M. Kremer 1988, p. 238) y contiene, tal como la presenta el autor, un costado ontológico. Dejando de lado las discusiones alrededor de la noción de *hecho*,<sup>2</sup> la idea muchas veces viene acompañada de la afirmación de que los hechos semánticos (aquellos hechos expresados mediante afirmaciones semánticas del tipo 'Es verdad que p', donde 'p' puede reemplazarse por cualquier oración con contenido asertórico) quedan determinados por los hechos no semánticos. Por supuesto, una vez que se ha dado este paso, la tentación a sacar consecuencias metafísicas adicionales de esta afirmación es inmediata: como Anil Gupta señala "no hay hechos semánticos" (Gupta 1984). O de otra manera, el mundo que nos rodea está constituido por hechos no semánticos. Esto puede tranquilizar a aquellos que tengan desconfianza en las entidades semánticas al verlas como entidades oscuras que deben ser eliminadas. Recientemente, sin embargo, Nuel Belnap y el propio Gupta han argumentado que la verdad es un concepto circular: que no puede determinarse su extensión sin conjeturarla previamente. Esto ha llevado a Aladim Yaqûb a sostener (Yaqûb 1993) que no es posible armonizar este carácter circular con la idea según la cual las propiedades semánticas son supervenientes. No obstante, Belnap y Gupta niegan explícitamente (Belnap, A. Gupta 1993) que no puedan armonizarse tales aspectos. En este artículo me propongo defender, al igual que Yaqûb, la tesis según la cual la circularidad y la superveniencia no pueden compatibilizarse. Tal defensa será realizada ofreciendo argumentos distintos que los que ofrece Yaqûb y contemplando aspectos de la posición de Belnap y Gupta que, según entiendo, no han sido tenidos en cuenta anteriormente.

### *1. Los límites de la extensión de la verdad*

Las ideas tarskianas alrededor de la verdad comparten la misma desconfianza acerca de las entidades semánticas (Tarski 1956). Esta concepción acerca de lo veritativo tiene un doble estatuto. Por un lado, en ella se intenta explicar nuestro uso cotidiano de las expresiones veritativas. En este nivel, se rescata un uso esencial: cualquier hablante normal de un lenguaje que

acepte una oración como 'la luna tiene cráteres', estará dispuesto a aceptar la verdad de esa oración. Lo mismo ocurre con la relación inversa: si un hablante normal acepta que es verdad que la luna tiene cráteres, él estará dispuesto a aceptar que la luna tiene cráteres. Frente a este dato, y con el propósito de dar cuenta de este aspecto, Tarski propone que la oración 'la luna tiene cráteres' es verdadera es equivalente a la luna tiene cráteres' debe ser una parte esencial para comprender el concepto de verdad. Usualmente se denomina *oraciones T* a las cláusulas de su definición con las que Tarski recoge el mencionado uso. Y con la incorporación de estas cláusulas se abre hacia el interior de nuestra explicación de la verdad la posibilidad de recoger la intuición superventista. El cumplimiento de la equivalencia entre ambos lados de la oración T sugiere fuertemente que para el establecimiento de nuestras afirmaciones semánticas como correctas no hace falta más que establecer lo dicho por una afirmación no semántica.

Por otro lado, se intenta que la definición, además de explicar nuestro uso de las expresiones semánticas, sea formalmente correcta: sus cláusulas no deben conducir a contradicción. De acuerdo a las ideas de Tarski, para que una definición del predicado veritativo de un lenguaje determinado sea juzgada adecuada, se requiere que tenga, al menos, como consecuencia todas las oraciones (T). Este requisito garantiza que se cumpla el mencionado objetivo según el cual el predicado veritativo definido debe tener una extensión correcta, es decir, que se aplique al menos a todas las oraciones verdaderas del lenguaje. Tarski llama a las oraciones (T) *definiciones parciales de verdad*,<sup>3</sup> y frecuentemente describe su proyecto como brindando una definición equivalente a la conjunción de todas las instancias de (T) para el lenguaje correspondiente:

(...) toda equivalencia de la forma (T) (...) puede considerarse una definición parcial de verdad que explica en qué consiste la verdad de una frase individual. La definición general debe ser, en cierto sentido, una conjunción lógica de todas estas definiciones parciales (Tarski 1944; la traducción es mía).

La corrección formal de la definición se relaciona con el deseo de Tarski de asegurar no sólo que su definición de verdad implique todas las instancias de (T), sino que las implique evitando todo tipo consecuencias no deseadas, en particular, evitando inconsistencias.

Aunque parezca lo contrario, la armonía entre estos dos aspectos no es fácil de lograr. Sin reflexionar demasiado, es fácil ver que este uso esencial, que parece presentar al predicado veritativo como una expresión vacua, conduce a contradicción. Una de las cosas que parece surgir de las pa-

radojas semánticas es que nos muestran que no podemos aceptar consistentemente *todas* las oraciones T.

En efecto, parece perfectamente posible, frente a una persona muy confiable, decir:

- (1) Todo lo que Déborah dice es verdadero

Supongamos que las afirmaciones relevantes de Déborah sean

- (2) La luna tiene cráteres

y

- (3) La luna es el único satélite terrestre

Si lo anterior es lo único que Déborah ha dicho, de acuerdo con nuestra tesis de la superveniencia semántica, para determinar las condiciones de verdad de nuestra atribución veritativa (1) es suficiente determinar las condiciones de verdad de (2) y (3). De acuerdo a la concepción tarskiana de la verdad antes expuesta, la explicación de las condiciones de verdad de (2) y (3) puede efectuarse a través de las dos oraciones (T) correspondientes:

- (2) es verdadera ssi la luna tiene cráteres

y

- (3) es verdadera ssi la luna es el único satélite terrestre

Si se cumplen que la luna tiene cráteres y que es el único satélite terrestre, habremos establecido la verdad de nuestra afirmación semántica (1). Ahora bien, supongamos que Déborah dice algo más:

- (4) La oración (1) es falsa

Nuevamente, de acuerdo a la concepción tarskiana, las condiciones de verdad de (4) quedan establecidas afirmando la oración (T): 'la oración (1) es falsa' es verdadera ssi la oración (1) es falsa.

Pero es importante destacar dos cuestiones:

En primer lugar, y en relación con la desconfianza en los elementos semánticos, la afirmación de cualquiera de las dos oraciones hace una referencia a la verdad. A diferencia de nuestro caso anterior, pareciera que a primera vista, si todos los hechos no semánticos aludidos en las afirmacio-

nes de Déborah son establecidos, no podemos establecer la verdad de (4), porque encontramos que para evaluar su valor de verdad hay que recurrir a nuestra evaluación del valor de verdad de (1). Pero, para establecer tal valor, es imprescindible establecer el valor de verdad de (4).

Pero, en segundo lugar, la aplicación de T a (1) y a (4), en ciertas circunstancias, conduce a contradicción. Supongamos que lo único relevante que Déborah dice es (4).<sup>4</sup> Esta es una circunstancia empírica perfectamente admisible. Supongamos, en esta circunstancia que (1) es verdadera. Entonces (4) (que es lo único que Déborah dice) será verdadera. Pero, por T, si (4) es verdadera, (4) es el caso (es decir, la oración (1) es falsa). Por lo cual, (1) es verdadera y falsa (lo que constituye una contradicción). Abandonemos el supuesto de que (1) es verdadera. Supongamos que (1) es falsa. Entonces, (4) es falsa. Pero, por T, (4) es falsa si lo que dice no es el caso. Por lo cual, ya que lo que dice (4) es que la oración (1) es falsa, se sigue que la oración (1) es verdadera. Esta asignación contradice nuevamente nuestra suposición original de valor de verdad. Por lo tanto, en esta circunstancia, (1) y (4) resultan paradójicas.<sup>5</sup>

La propuesta de Tarski para recortar el número de oraciones T aceptadas como parte de la definición es bien conocida: el predicado veritativo no debe formar parte del mismo lenguaje que las oraciones de las que se predica, y por ello, no habrá casos de oraciones T que se apliquen a oraciones que contengan a su vez al predicado veritativo de ese lenguaje. La desconfianza en las afirmaciones semánticas subyacente en tal prohibición, apoyada por el riesgo que conlleva cualquier afirmación de este tipo, sustenta su apego radical al fisicalismo (doctrina que supone, entre otras cosas, que no hay hechos semánticos). La estructura de la definición de Tarski (formulada desde un metalenguaje) indica en forma sistemática cómo eliminar las apariciones de la expresión 'es verdadera' por otras expresiones que se suponen claras. Y esto significa que cada vez que tenemos en nuestro lenguaje una afirmación que supuestamente remita al plano semántico, podemos reemplazar esta afirmación por otra que no remita a ese plano.

La concepción tarskiana ha sido criticada insistentemente en los últimos años (Cfr., entre otros, Kripke 1975, Gupta 1984, Belnap, Gupta 1993 y Mc Gee 1991). En muchos de estos intentos se sostiene que su propuesta es demasiado extrema. La estrategia es mostrar que hay ciertos tipos de lenguajes que pueden contener una definición de su propio predicado veritativo. La supuesta ventaja que estos enfoques poseerían tiene que ver con la posibilidad de acercarnos aún más a nuestro uso cotidiano de nuestras expresiones veritativas. En efecto, si se pretende dar cuenta de nuestras atribucio-

nes de verdad en el lenguaje natural, cualquier rechazo de algún rasgo aparente de su uso tiene que estar argumentativamente fundado. Y es claro, como Kripke ha señalado (Kripke 1975), que en el lenguaje natural tenemos un sólo predicado veritativo y no, tal como Tarski propone, una infinidad de predicados con subíndices. ¿En qué sentido, entonces, la definición de Tarski analiza nuestro predicado veritativo? Si hay aspectos de ese predicado que no son representados en el sustituto tarskiano, ¿cómo podremos sostener que el segundo es un análisis del primero? Los planteos posteriores a Tarski han intentado entonces salvar este hiato entre los dos predicados ampliando el conjunto de las oraciones T aceptadas en la definición y con ello ampliaron la extensión del predicado veritativo intentando que esta se acerque todo lo posible a la extensión del predicado veritativo del lenguaje natural. El punto que quiero discutir es si en este movimiento de ampliación se salva la superveniencia semántica. ¿En qué sentido si hay que recurrir para el establecimiento de nuestras afirmaciones semánticas a otras afirmaciones semánticas es posible salvar la superveniencia?

## 2. *El mentiroso pide la palabra*

Lo que está en juego en los planteos postarskianos es la posibilidad de salvar en toda su dimensión la idea según la cual afirmar la verdad de una oración es equivalente a afirmar esa oración y permitir, al mismo tiempo, atribuciones veritativas dentro del propio lenguaje sin que esto signifique transformar en contradictorias a nuestras expresiones veritativas. Una propuesta provocativa en esta línea puede encontrarse en los trabajos de Nuel Belnap y Anil Gupta.<sup>6</sup> La clave para ellos la encuentran en la reivindicación de las definiciones circulares.

La pregunta que debemos hacernos para comprender mejor este enfoque es la siguiente: ¿Por qué se sostiene que la verdad es un concepto inevitablemente circular? El punto de partida para llegar a esta conclusión es afirmar que el conjunto total de las oraciones T del lenguaje, aún aquellas que atribuyen condiciones veritativas a oraciones que contienen expresiones veritativas como el mentiroso o el honesto, *definen* la verdad para ese lenguaje (Cfr. Belnap, Gupta 1993, p. 4). De aquí, se da un paso más: tomarse en serio la idea de definición a través de la conjunción de cada una de las oraciones T implica aceptar, en todas sus consecuencias, la idea de que dichas oraciones *establecen el significado* del concepto de *verdad*. El supuesto que autoriza este paso es la idea de que la función principal de las defini-

ciones es el establecimiento del significado del concepto definido. Debe aclararse en este punto que los autores relacionan la noción de *significado* con la de *intensión*, intentando aislar aspectos del significado relacionados con las modalidades aléthicás. De esta manera, ellos se comprometen con la idea tradicional según la cual la intención de una expresión se determina por su extensión a través de todos los mundos posibles. El modo en que ellos relacionan las dos nociones (la de *significado* y la de *intensión*) queda expresado en lo que ellos denominan *Tesis de la Intensión*:

Las oraciones T fijan la intención de la verdad.<sup>7</sup>

De la tesis según la cual las oraciones T establecen esta manera particular de entender el significado del concepto de *verdad*, y de la idea según la cual las oraciones son o bien verdaderas o bien falsas, Belnap y Gupta concluyen que la verdad es un concepto circular (Cfr. Belnap, Gupta 1993, cap. 4, #2). Para ellos esto es así, ya que la determinación de la extensión del predicado veritativo, es decir, la determinación del conjunto de las oraciones verdaderas del lenguaje, *no puede hacerse sin tener ya fijada la mencionada extensión*. La circularidad es inevitable, ya que para establecer el valor veritativo de las oraciones que contienen el predicado veritativo hace falta suponer con anterioridad el valor veritativo de esas oraciones. La naturaleza de los conceptos circulares exige, entonces, una explicación de su funcionamiento semántico distinta de la habitual. Cuando estamos en presencia de un concepto circular,<sup>8</sup> nunca podremos fijar categóricamente, tal como se hace en las explicaciones semánticas habituales, la extensión del concepto. Podremos determinar su significado, pero nunca, salvo de manera hipotética su extensión.

El objetivo de Belnap y Gupta consiste en dar una explicación satisfactoria del funcionamiento semántico de las definiciones circulares. La clave es contraponer los conceptos de *regla de aplicación* y de *regla de revisión*. Para advertir el punto, hace falta prestar atención a las oraciones del lenguaje que contienen al predicado veritativo. Ejemplos de esas oraciones son el honesto ('Esta oración es verdadera') y el mentiroso ('Esta oración es falsa'). También lo es una oración como (4) ('Todo lo que Déborah dijo es verdadero) y como muchos de nuestros principios lógicos (por ejemplo 'Toda oración es verdadera o falsa'). Lo que proponen frente a estas oraciones es que si queremos capturar el significado que una definición circular atribuye a la expresión que se está definiendo es imposible pensar al significado como una regla que determina completamente su extensión. No podemos

pensar al significado como una regla de aplicación (tal como se realiza en la semántica tarskiana tradicional).<sup>9</sup> Una definición circular, en cambio, fija el significado de la expresión que se define de *manera hipotética*. No tenemos una *regla de aplicación* sino una *regla de revisión*: una regla que aplicada a una extensión hipotética, nos ofrece el mejor candidato para determinar la extensión del concepto a definir.

Formalmente, la *semántica de revisión* se presenta de la siguiente manera. Sea M un modelo ordinario de primer orden que asigna una extensión al predicado veritativo de nuestro lenguaje. Las reglas de revisión nos brindan un procedimiento para obtener un modelo perfeccionado M\*. Nominalmente, para cualquier oración que tenga el nombre S, las reglas asignan a S a la extensión del predicado veritativo del lenguaje en M\*, si la oración se evalúa como verdadera en M. Así, si (2) formó parte de la extensión del predicado veritativo en M, (2) formará parte de la extensión de ese predicado en M\*. Por supuesto, la idea es que nuestro predicado veritativo sea parte de nuestro lenguaje. Así podremos afirmar cosas como '(2) es verdadera' y preguntarnos si forma o no parte de la extensión de nuestro predicado veritativo. Esta es una afirmación de tipo semántico cuya verdad se explica fácilmente a partir de la verdad de una afirmación no semántica. La reiteración sucesiva de este proceso en el que nos preguntamos por la extensión del predicado veritativo genera una serie de modelos M, M\*, M\*\*, M\*\*\*, etc., que son construidos usando las reglas de revisión que evalúan a las oraciones a partir de cada uno de los elementos de la serie.

Este proceso se genera a partir de un modelo M de base. Debe admitirse, también que pueden formularse distintos modelos de base, de acuerdo a las distintas oraciones que resulten parte de la extensión del predicado veritativo. Esto generará, a su vez, una serie de secuencias de modelos M, M\*, M\*\*, etc. que reflejará, al menos, la contingencia de algunas de nuestras verdades según lleguen o no a formar parte de la extensión de nuestro predicado semántico. Algunas de nuestras verdades resultan ser especiales: ellas llegan a formar parte de la extensión en cualquiera de nuestras secuencias. Ellas son, por supuesto, las verdades lógicas.

Por supuesto, en los casos ordinarios, es decir, en aquellos en los cuales las oraciones no contienen al predicado veritativo del lenguaje (casos como (1) y (2)), las reglas de revisión no toman en cuenta las hipótesis que se consideran: es irrelevante cuáles son las extensiones que se le van asignando al predicado veritativo para establecer si esas oraciones forman parte de su extensión. En ese sentido, para este sector del lenguaje, la semántica gene-



rada es equivalente a la semántica tarskiana: dada cualquier oración de este sector del lenguaje (oraciones que no contienen términos semánticos), éstas podrán pertenecer (o no) a la extensión del predicado veritativo. Por ejemplo, seguramente, (1) 'la luna tiene cráteres' pertenece a la extensión de nuestro concepto de verdad. Y por supuesto, si (1) pertenece a la extensión del predicado veritativo en una fase cualquiera del proceso de revisión, seguirá perteneciendo a la extensión de ese predicado, conjetúrese lo que se conjeture, en las diversas fases del sucesivo proceso de revisión. Seguramente, no pertenecerá

(5) la luna es de queso

aunque su comportamiento es similar al de (1). Si no pertenece en una fase, no pertenecerá nunca.

Pero, para aquellas oraciones que contienen al predicado veritativo la situación es completamente distinta. Cuando queremos evaluar oraciones como '(2) es verdadera' las hipótesis que se toman en cuenta son determinantes del comportamiento semántico de esas oraciones. Este hecho es el que determina el origen de la conducta (ordinaria o patológica) del predicado veritativo. En este sentido, decir que 'Esta oración es falsa' es verdadera' es hacer una afirmación tan patológica como afirmar la oración del mentiroso (Esta oración es falsa) (Cfr. Belnap, Gupta 1993, p. 140 y 255). Y esto se explica observando que el valor veritativo que se aplica a esta oración a partir de las reglas de revisión no se estabiliza nunca: de la hipótesis de que el mentiroso es verdadera en un modelo  $M^n$  podemos concluir que ella es falsa en la fase  $n + 1$ , y así sucesivamente. La conducta patológica del mentiroso consiste en la continua oscilación de su valor veritativo en las distintas fases del proceso de revisión. En la misma línea puede explicarse el comportamiento patológico del honesto: su valor veritativo se estabiliza a partir de la primera hipótesis inicial. Ella será verdadera, si está incluida en la extensión inicial y será falsa, si no lo está. Pero su patología está en que no tenemos fundamento semántico alguno para avalar ninguna hipótesis inicial.

Vemos, entonces, que las reglas de revisión, que explican qué quieren decir nuestras oraciones T, nos brindan un método para obtener aproximaciones cada vez mejores de la extensión del predicado veritativo. Volvamos para enfatizar a nuestro ejemplo sobre las afirmaciones de Déborah. Tratemos de fijar la extensión del predicado veritativo de ese lenguaje que contenga sólo a (1), (2) y (3) como componentes. En ese caso, la con-

junción de las tres nos brinda la extensión de nuestro predicado semántico. Por supuesto, aunque generemos la secuencia de modelos  $M$ , las hipótesis que tomemos en consideración en la fase anterior, serán, como hemos dicho antes, absolutamente irrelevantes para la fase que estemos considerando. Pero supongamos ahora que permitimos que Déborah diga cosas tales como '(2) es verdadera'. En este caso, si resultó que (2) forma parte de la extensión del predicado veritativo en  $M$ , entonces '(2) es verdadera' formará parte de la extensión de nuestro predicado semántico en  $M^*$ . Pero, qué sucede con la extensión de ese predicado si agregamos referencias cruzadas entre las afirmaciones de Déborah o simplemente permitimos que afirme el mentiroso. Nunca lograremos estabilizar a la extensión de nuestro predicado en todo el proceso de revisión. Dada cualquier extensión inicial que asignemos a nuestro predicado, se irán generando a partir de ella nuevas extensiones con la siguiente característica: si por ejemplo, el mentiroso es conjeturada como falsa en  $M$ , resultará verdadera en la nueva extensión resultante en  $M^*$ . Y este proceso no se estabilizará nunca. Por esta razón, las oraciones (T) nunca nos brindan *la* extensión del predicado veritativo, sino más bien nos ofrecen una respuesta temporal para cualquier extensión inicial de la que partamos. Ellas deben ser consideradas como afirmaciones hipotéticas: sólo establecen las condiciones de verdad bajo la suposición de que hemos conjeturado cuál es la extensión de la verdad.

Es importante admitir que aceptar que el predicado veritativo pueda tener una conducta patológica no es lo mismo que admitir que su aplicación genera contradicciones en el lenguaje. Como ya he destacado, para los autores las oraciones (T) establecen el significado del predicado veritativo. Por esta razón, ellas deben ser interpretadas como equivalencias definicionales y *no* como meras equivalencias materiales (Cfr. Belnap, Gupta 1993, p. 138). De esta manera, se captura la mencionada intuición en forma plena, sin que su reconstrucción teórica implique que existan instancias de ellas que nos lleven a contradicción. En efecto, si las oraciones (T) son interpretadas como estableciendo una equivalencia material, ellas tienen instancias contradictorias tales como el bicondicional material ' $p$  es verdadera ssi  $p$  no es verdadera' donde  $p$  es idéntica al mentiroso. Por el contrario, el propósito central de la concepción de Belnap y Gupta es reconocer y explicar la conducta patológica del predicado veritativo manteniendo su consistencia. Por ello, las oraciones (T) deben interpretarse de otra manera. Ellas remiten al proceso de revisión, por lo que los autores proponen leerlas usando índices. Por ejemplo la lectura de

(2) es verdadero =<sub>def</sub> la luna tiene cráteres<sup>10</sup>

implica que de la hipótesis de que '(2) es verdadero' vale en la fase de revisión 1, podemos concluir que la luna tiene cráteres vale en la fase 0 y conversamente. Las características patológicas de oraciones como el honesto o el mentiroso son el resultado de que las oraciones (T) establecen el significado del concepto de *verdad*: no producen contradicciones, sino un proceso que no se estabiliza nunca. Con ello, se restablece la armonía perdida en la semántica tarskiana entre la totalidad de las oraciones (T) como dadoras de significado al concepto de *verdad* y la necesidad de evitar las contradicciones.

En suma, las paradojas veritativas surgen como un fenómeno natural que subyace a la conducta del predicado veritativo a partir de su naturaleza circular. La extensión de los conceptos circulares sólo puede ser establecida de manera hipotética y bajo ciertas circunstancias esas hipótesis pueden generar un comportamiento patológico. El punto tiene una gran importancia filosófica: el significado de un predicado circular, es decir, su intención, es capturado por una regla de revisión. Y si bien no puede establecerse de manera categórica si ciertas oraciones que poseen este comportamiento patológico forman parte de la extensión del predicado veritativo del lenguaje, puede establecerse categóricamente que hay oraciones cuyo valor veritativo es fluctuante en las distintas fases del proceso de revisión (las oraciones paradójicas como el mentiroso) y hay oraciones cuyo valor veritativo depende exclusivamente del valor veritativo que le hemos asignado hipotéticamente en la etapa inicial (como el honesto).

### 3. *El honesto y el mentiroso: la tensión entre la superveniencia y la bivalencia*

Las anteriores consideraciones podrían llevarnos a pensar que la circularidad es un obstáculo obvio para el cumplimiento de la superveniencia. En este sentido, Aladdin Yaqûb argumenta (Yaqûb 1993, p. 39) que si no se rechaza la bivalencia (es decir, la idea de que dada cualquier oración de un lenguaje, ella es o bien verdadera o bien falsa), hay que abandonar la tesis de que los valores semánticos de todas las oraciones que hacen afirmaciones acerca de las condiciones de verdad de otras oraciones dependen de los valores de las oraciones que no hacen afirmaciones semánticas.<sup>11</sup> Lo anterior se cumpliría porque el conjunto de hechos que determinan si una oración pertenece a la extensión del concepto de *verdad* debería *estar incluido* en la extensión inicial del predicado que representa el concepto veritativo

en el proceso secuencial de generación de modelos  $M$ , cuando en el lenguaje contamos con expresiones como (4). El valor de verdad del honesto y del mentiroso no depende de los valores de verdad de ningún conjunto de afirmaciones no semánticas. Sostiene Yaqûb que si se toma en serio el carácter circular, "la imagen metafísica que emerge en un lenguaje bivalente es la de una propiedad (veritativa) no superveniente" (Yaqûb 1993, p. 42).

Si se acepta la reconstrucción que hace Yaqûb de la posición de Belnap y Gupta, su crítica me parece perfectamente efectiva. Sin embargo, su línea argumentativa depende de una interpretación fuerte de la superveniencia semántica: lo que se busca es que el estatuto de todas las oraciones (incluyendo al honesto y al mentiroso) esté completamente determinado por los hechos no semánticos. Sin embargo, Yaqûb no parece advertir que dentro de la semántica de la revisión, es claro que Belnap y Gupta pretenden que haya un sentido en el cual la verdad sea reductible, aunque no eliminable (dado su aspecto circular). Hay un sentido, según ellos, en el cual cada vez que damos una interpretación para todas las constantes no semánticas del lenguaje, cualquier interpretación de las constantes semánticas (es decir, cualquier interpretación adicional del predicado veritativo) es redundante (Gupta 1982, p. 201). Dicho brevemente, hay un sentido en el cual la verdad es superveniente. ¿Qué sucede, entonces, con la objeción de Yaqûb según la cual para que el estatuto veritativo del honesto y del mentiroso esté completamente determinado hace falta comprometerse con la existencia de hechos semánticos que no están determinados por la base no semántica? Considero que si queremos continuar aceptando la objeción de Yaqûb, hay que mostrar que no hay un sentido plausible más débil de superveniencia compatible con la conjunción del principio de bivalencia y del carácter circular de la verdad. En suma, comparto con Yaqûb que surge algún tipo de incompatibilidad entre las mencionadas ideas, pero intentaré tomar en cuenta algunos matices de la semántica de revisión para volver la crítica más efectiva.

En efecto, lo que parece subyacer a las ideas de Belnap y Gupta con la que se bloquearía esta argumentación es una reinterpretación más débil de la noción de *superveniencia semántica*. Lo que parecen sostener es que *el estatuto* de una oración como el honesto o el mentiroso está determinado por los hechos no semánticos, y no *el valor veritativo*. En el caso del mentiroso, por ejemplo, no sería ya su valor de verdad sino su continua fluctuación en las distintas etapas de revisión lo que queda determinado por la base no semántica del modelo. La idea con la que se pretendería salvar la superveniencia es que no necesitamos ninguna información adicional que la brin-

dada por la base no semántica para fijar este estatuto fluctuante y por ello patológico. Por supuesto, una explicación similar se podría elaborar para el honesto. En este caso, su estatuto es su tendencia a estabilizarse en el valor que tome en la hipótesis inicial, cualquiera sea esta. Es este punto el que Yaqûb deja de advertir y con ello vuelve vulnerable su objeción que comparto.

Bajo esta reinterpretación de la superveniencia semántica en términos del comportamiento y no de un valor de verdad determinado, podría pensarse que el honesto y el mentiroso comparten una importante característica con las oraciones lógicamente verdaderas: su estatuto es el mismo cualquiera sean los hechos del mundo. Como argumenta Carolina Sartorio (Sartorio 1998), en defensa de la idea de que el planteo de Belnap y Gupta puede salvar la superveniencia, así como las verdades lógicas no violan la superveniencia, las oraciones patológicas tampoco lo hacen: en ambos casos su estatuto queda completamente determinado por la base no semántica.

Lo que quiero sostener es que esta reinterpretación tampoco es aceptable. Para ello, me interesa argumentar en contra de esta posibilidad que hay una diferencia importante que cuestiona la analogía entre las oraciones patológicas y las verdades lógicas. El punto clave para bloquear la analogía es que estas últimas tienen un valor de verdad determinado. Ellas *son* verdaderas. Como toda oración con un comportamiento "normal", las verdades lógicas se estabilizan en el proceso de generación de secuencias de modelos  $M$ .<sup>12</sup> Lo mismo no ocurre ni con el honesto ni con el mentiroso. Sólo puede determinarse su valor, si se supone una extensión inicial. Y es claro que en este punto surge una tensión importante en el planteo de la revisión: o se salva la idea de que todas las oraciones tienen un valor de verdad determinado y con ello la bivalencia, o se salva la idea de que la base no semántica determina por completo el valor de verdad de todas las oraciones (incluso las patológicas). Ambas ideas parecen chocarse. Y no veo como puede evitarse tal tensión. La armonía entre la tesis de la superveniencia y la bivalencia parece requerir una formulación de la primera, no en términos del estatuto de las oraciones patológicas, sino en términos de valores de verdad determinados por la base no semántica del lenguaje. Reinterpretar la superveniencia parece entonces suponer algún tipo de abandono implícito de la bivalencia.

Quizás Belnap y Gupta podrían replicar que una cosa es *hipotetizar que una oración que contenga al predicado veritativo forma parte de la extensión de este predicado* y otra cosa es *afirmar categóricamente* que ello sucede. Nuestra hipótesis no representa hecho semántico alguno. Pero, nuevamente,

tal salida parece plausible pero nos conduce a la idea de que algunas oraciones, las patológicas, carecen de un valor de verdad determinado a partir del valor de verdad de las afirmaciones no semánticas de nuestro lenguaje. Aceptar lo anterior es aceptar algún tipo de limitación para la bivalencia, limitación que los autores parecen negar en toda su discusión con Kripke (Cfr. Belnap, Gupta 1993, cap. 3).

Un punto que no quiero dejar de notar es que mi objeción, si se acepta, muestra que la circularidad de la definición de verdad obliga (para lenguajes como el natural que contengan al mentiroso y al honesto) a rechazar la posibilidad de tomar a las equivalencias T como estableciendo la extensión del predicado veritativo. Salvar la superveniencia semántica en términos del estatuto de las oraciones patológicas (en el caso del mentiroso su continua oscilación en las distintas etapas de revisión) supone que no es posible determinar de manera categórica la extensión del mencionado concepto semántico. Como hemos visto anteriormente, podríamos pensar que aquel rechazo conduce a la idea según la cual tales equivalencias establecen el significado del predicado veritativo. Y ésto, visto a la luz de la mencionada tesis de la intensión, significaría que no tendríamos *la* extensión de la verdad sino su intensión. Pero en qué sentido tal establecimiento puede ser explicado en términos compatibles con la superveniencia: ¿Cuáles son los hechos físicos sobre los que supervienen esos hechos semánticos? O de otra manera, ¿a qué conjunto de afirmaciones no semánticas se reduce este hecho acerca de la intensión de nuestro predicado semántico? La imposibilidad de apelar a la extensión de la verdad parece abrir la puerta a misteriosas entidades intensionales. Entidades que no determinan *la* extensión de la verdad y que parecen poco aceptables desde el punto de vista de aquel que se ha esforzado por salvar la superveniencia semántica por sus reparos ontológicos. Nótese, además, que esa apelación a ese tipo especial de entidad intensional es justamente el resultado de haber admitido el carácter circular del concepto de *verdad*.

En suma, he intentado mostrar que la presunta circularidad del concepto de *verdad* a favor de la que Belnap y Gupta argumentan resulta incompatible con el cumplimiento de la superveniencia semántica, siempre que no se quiera abandonar la bivalencia. Para tal fin, he presentado una reinterpretación de la noción de *superveniencia* que resulta favorable a la posición de Belnap y Gupta. En esa reinterpretación, tal característica se define en términos del comportamiento o estatuto de una oración en las distintas etapas del proceso de revisión. Si se fija la base no semántica del lenguaje, queda determinado el comportamiento de todas las oraciones que hacen afirma-

ciones semánticas. Sin embargo, esta reinterpretación no resulta compatible con la idea de que todas las oraciones (incluyendo las oraciones que hacen afirmaciones semánticas) tienen que tener un valor de verdad determinado. Por otra parte, el abandono de la pretensión de fijar categóricamente la extensión del concepto de *verdad* parece hacer surgir un compromiso con entidades intensionales: el conjunto de los bicondicionales T establecen el significado del mencionado concepto y no su extensión (de manera categórica). Pero ahora, ¿en qué sentido podemos cumplir con la tesis de la superveniencia de las afirmaciones semánticas a las afirmaciones no semánticas si debemos comprometernos con significados? Si el significado del concepto de *verdad* no se agota en la extensión, ¿cómo puede cumplirse con la superveniencia mencionada?<sup>13</sup>

### Notas

† El término *superveniencia semántica* alude a la idea de que no puede haber diferencias en lo semántico sin diferencias en lo no semántico. Esta expresión es usual en las discusiones sobre la visión estratificada de la realidad y en el presente artículo hará referencia a la supuesta relación entre el ámbito semántico y el no semántico.

- 1 Una interpretación que haga verdaderas sólo a oraciones que no contengan predicados semánticos.
- 2 Quizás el argumento más relevante en contra de la noción de *hecho* sea el que Barwise y Perry han denominado "el hondazo" (slingshot). Tal argumento puede rastrearse en los trabajos de Church y Davidson.
- 3 Hay que aclarar que aun cuando cada oración T puede verse como una definición parcial de la verdad para una oración, no obstante, las oraciones T no necesitan ser cláusulas de la definición. Ellas sólo deben ser consecuencias de la definición.
- 4 Eliminemos (2) y (3) de lo que Déborah dice.
- 5 Es usual denominar *el mentiroso* a la oración 'Esta oración es falsa' y *el honesto* a la oración 'Esta oración es verdadera'. La primera genera instancias contradictorias de oraciones (T) tales como 'El mentiroso es verdadera ssi esta oración es falsa' donde el nombre 'el mentiroso' hace referencia a 'Esta oración es falsa'. La segunda genera instancias frente a las cuales podemos atribuir consistentemente cualquier valor de verdad tales como 'El honesto es verdadera ssi esta oración es verdadera' en donde 'El honesto' hace referencia a 'Esta oración es verdadera'.
- 6 La exposición más completa de sus ideas se encuentra en el ya citado Belnap, Gupta (1993) También en Belnap (1982) y Gupta (1982).
- 7 La tesis de la intensión debe diferenciarse de la tesis de la implicación. Según esta última tesis, una definición de verdad debe implicar una oración T para cada una de las oraciones del lenguaje. Belnap y Gupta sostienen que ambas tesis son diferentes y que

- la primera de ellas es más adecuada que la segunda al ser más plausible y poseer mayor neutralidad filosófica. Ellos consideran que la tesis de la significación es neutral con respecto a la polémica entre correspondentistas y deflacionistas. Cfr. Belnap, Gupta (1993, p. 25).
- 8 Otros conceptos circulares serían el de referencia, el de satisfacción, pertenencia, ejemplificación, necesidad, creencia y conocimiento.
  - 9 Subyace a esta propuesta la idea de que hay dos tipos de conceptos: los ordinarios, tales como rojo y estrella, cuyas condiciones de aplicación nos permiten separar a aquellos objetos a los cuales el concepto se aplica de aquellos a los que no se aplica y los conceptos circulares, para los cuales necesitamos una conjetura inicial acerca de la extensión del concepto para decidir acerca de cuáles son los objetos a los que se aplica el concepto y cuáles no.
  - 10 Belnap y Gupta usan '=def' en lugar de '≡' para indicar que las oraciones T, en la teoría de la revisión, exhiben el contenido del concepto veritativo. El '=def' es interpretado como la equivalencia definicional. Como en el caso de otros conceptos circulares, la definición de la verdad muestra una conducta patológica asociada con su significado.
  - 11 Yaqûb sostiene que la verdad es superveniente si y sólo si el estatuto veritativo de cada oración está plenamente determinado por los hechos no semánticos. Cfr. Yaqûb (1993, p. 39, nota 15).
  - 12 En realidad, toda verdad lógica se estabilizará como verdadera en toda secuencia de generación de modelos M, sean cuales sean las asignaciones a las constantes no lógicas en el modelo M de partida. Esta es una ventaja de esta propuesta frente a otros enfoques postarskianos acerca de la verdad, ya que garantiza la validez de los razonamientos clásicos.
  - 13 Quiero agradecer los valiosos aportes de Juan Comesaña, Eleonora Orlando, Federico Penelas y Carolina Sartorio realizados durante la elaboración de este artículo. También quiero agradecer los comentarios efectuados por los asesores académicos anónimos de la revista *Theoria*, y a la *Fundación Antorchas* por su apoyo económico.

## BIBLIOGRAFIA

- Belnap, N.: 1982, 'Gupta's rule of revision theory of truth', *J. of Phil. Logic* 11.
- Belnap, N., Gupta, A.: 1993, *The revision theory of truth*, Cambridge, The MIT Press, 1993.
- Gupta, A.: 1982, 'Truth and paradox', *J. of Phil. Logic* 11.
- Gupta, A.: 1984, 'Truth and Paradox', in R. Martin: *Recent Essays on Truth and the Liar Paradox*, Oxford, Oxford UP.
- Kremer, M.: 1988, 'Kripke and the Logic of Truth', *J. of Phil. Logic* 17.
- Kripke, S.: 1975, 'Outline of the Theory of Truth', *J. of Phil.* 72, n. 19.
- Mc Gee: 1991, *Truth, vagueness, and paradox*, Indianapolis, Hackett.
- Sartorio, C.: 'Circularidad, elucidación y lenguajes naturales', presentado en el *IV coloquio Bariloche de filosofía*, 24-28 de junio de 1998, Bariloche.
- Tarski, A.: 1956, *Logic, Semantics and Metamathematics*, Oxford, Oxford University Press. Segunda edición, 1990.



- Tarski, A.: 1944, 'The Semantic Conception of Truth and the Foundations of Semantics', *Phil. and Phenomenological Research*, 4, p. 344.
- Yaqúb, A.: 1993, *The Liar Speaks the Truth*, Oxford, Oxford UP.

*Eduardo Alejandro Barrio* es Doctor en Filosofía (UBA) y Profesor de Lógica (UBA). Director de proyectos de investigación en semántica y filosofía de la lógica. Director de la Colección Enciclopedia Lógica (Eudeba). Autor de *La verdad desestructurada* (Buenos Aires, Eudeba, 1998) y de numerosos artículos en revistas hispanoamericanas sobre la verdad, los supuestos filosóficos de la lógica modal y la lógica de los condicionales contrafácticos.