THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia
ISSN: 0495-4548
theoria@ehu.es
Universidad del País Vasco/Euskal Herriko Unibertsitatea
España

BARRIO, Eduardo Alejandro

Modelos, autoaplicación y máxima generalidad

THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia, vol. 22, núm. 2, 2007, pp. 133-152
Universidad del País Vasco/Euskal Herriko Unibertsitatea
Donostia-San Sebastián, España

Modelos, autoaplicación y máxima generalidad *

(Models, self-application and absolute generality)

Eduardo Alejandro BARRIO

Manuscrito recibido: 27.01.2007
Versión final: 04.04.2007

BIBLID [0495-4548 (2007) 22: 59; pp. 133-152]

RESUMEN: En este artículo, me propongo exponer algunas dificultades relacionadas con la posibilidad de que la Teoría de Modelos pueda constituirse en una Teoría General de la Interpretación. Específicamente la idea que sostengo es que lo que nos muestra la Paradoja de Orayen es que las interpretaciones no pueden ser ni conjuntos ni objetos. Por eso, una elucidación del concepto intuitivo de interpretación, que apele a este tipo de entidades, está condenada al fracaso. De manera secundaria, muestro que no hay algún supuesto conjuntista que sea imprescindible para que surja la mencionada paradoja: sólo se necesita que las interpretaciones sean objetos. Voy a argumentar que si las interpretaciones son objetos, tal como lo supone la posibilidad de cuantificar sobre las mismas para poder dar una caracterización satisfactoria de consecuencia lógica, la autoaplicación (como un caso de aplicación de los recursos semánticos de la teoría de modelos para encontrar una interpretación con máxima generalidad) es imposible. Finalmente, discuto cada una de las dos soluciones que el propio Orayen imaginó frente a su paradoja, y muestro que cada una posee diferentes dificultades.

Descriptores: paradojas semánticas, Teoría de Modelos, autoaplicación, cuantificadores irrestrictos, dominio absolutamente inclusivo, universalidad absoluta.

ABSTRACT: *In this paper, I intend to present some problems to construe the Model Theory as a Theory of Interpretation. Specifically, I am going to defend that, according to the Paradox of Orayen, interpretations can not be neither set nor object. Thus, an explication of the intuitive concept of interpretation that appeals to these types of entities will be condemned to failure. Secondary, I will show that there is not any like-set assumption indispensable to get rise to that paradox: only all is needed is the assumption that interpretations are objects. I am going to argue that if interpretations are objects, as it is assumed by the possibility of quantifying over interpretations in order to offer an satisfactory characterization of logical consequence, self-application (as a case of application of semantics resources of the Model Theory for finding a interpretation with absolute generality) is not possible. Eventually, I will discuss both of the solutions provided by Orayen to his paradox and I will support that both have different difficulties..*

Keywords: *semantical paradox, Model Theory, self-application, unrestricted quantifiers, all-inclusive domain, absolute generality*

¿De qué hablan las teorías formales? Uno de los objetivos filosóficos centrales de la teoría de modelos (TM) es dar una caracterización de los diversos *modos de interpretar* las teorías formales. Al hacerlo, se intenta poner de manifiesto cuáles son los compromisos ontológicos necesarios para que la teoría, en esa interpretación, resulte ver-

* Algunas partes de este trabajo fueron presentadas originariamente en el Universidad de Barcelona en el 2003 y en el seminario de investigadores del Instituto de Filosofía de la UNAM en el 2004. Posteriormente, presenté algunas de las ideas en el *Coloquio Iberoamericano de Filosofía de la Lógica*, realizado en México DF en el Instituto de Filosofía de la UNAM en marzo de 2005 y en el *XII Congreso Nacional de Filosofía de la República Argentina* a fines del 2005. Finalmente, presenté parcialmente este trabajo en el encuentro entre integrantes de la SEFA y SADAF en la Universidad Complutense de Madrid. Quiero agradecer especialmente a los participantes de los mencionados eventos, en especial a David Calvo, Javier Castro Albano, Max Freund, Alberto Moretti, Alfonso García de la Sienra, Mario Gómez Torrente, Agustín Rayo y José Sagüillo por sus invalorables aportes.

dadera. Tal objetivo incluye la pretensión de autoaplicación: esto es, considerar el caso en el cual, la propia teoría de modelos, en tanto teoría formal, requiere una interpretación tal que sea posible que una de las colecciones de las que ella habla, sea el universo de discurso de esa interpretación. Raúl Orayen ha argumentado[1] que la posibilidad misma de hacer una teoría general de los modos de interpretar teorías (TI)[2] incurre en una inconsistencia. Específicamente, Orayen identificó un problema, conocido posteriormente como la *Paradoja de Orayen*, referido a la posibilidad de formalizar la semántica de los lenguajes formales a través de una teoría axiomática de conjuntos formalizada en primer orden. En este artículo, me propongo exponer algunas ideas relacionadas con la posibilidad de que TM cumpla aquel objetivo filosófico (el de poder constituirse en una teoría general de la interpretación de las teorías formales TI). Específicamente la idea que sostengo es que lo que nos muestra la paradoja es que una elucidación del concepto intuitivo de *interpretación* está condenada al fracaso, si toma a las interpretaciones como objetos. De manera secundaria, muestro que no hay algún supuesto conjuntista que sea imprescindible para que surja la paradoja. Mi estrategia general será mostrar que si las interpretaciones son objetos, tal como lo supone la posibilidad de cuantificar sobre las mismas para poder dar una caracterización satisfactoria de *consecuencia lógica* y *validez universal,* la autoaplicación (como un caso de aplicación de los recursos semánticos de la teoría de modelos para encontrar una interpretación con máxima generalidad) es imposible. Finalmente, discuto cada una de las dos soluciones que el propio Orayen imaginó frente a su paradoja, y muestro que cada una posee diferentes dificultades.

I

La capacidad expresiva de una teoría axiomática formal está delimitada por la colección de asignaciones de significados que satisfacen sus axiomas. Por ejemplo, en el caso de una teoría lógica de primer orden, cada una de esas asignaciones de significados a las expresiones no lógicas permite representar información acerca de individuos, sus características y sus relaciones mutuas, de la manera más general posible. La colección de cada una esas asignaciones de significados es determinante en la obtención de una caracterización apropiada de los conceptos intuitivos de *consecuencia lógica* y *validez universal.* Y la exhaustividad de los modos de interpretar sus fórmulas dentro de TM es esencial para la obtención de un mecanismo apropiado para ofrecer un análisis de los mencionados conceptos. Tal propósito se satisface recurriendo al concepto de *satisfacción por toda estructura conjuntista,* esto es, por todo modelo cuyo dominio es un conjunto. El desarrollo de TM permite demostrar que tal teoría es correcta y completa. El teorema de Kreisel muestra que en el caso de las teorías lógicas de primer orden, basta

---

[1] Raúl Orayen presentó sus argumentos en diversas ocasiones y trabajos. Aquí sigo Orayen, R. (2003).

[2] En Rayo (2005) se habla del proyecto de hacer *semántica generalizada.* En Moretti (inédito) se habla de "condiciones de adecuación filosófica" para una TM. Coincidente con el proyecto de tomar a TM como una TI en Rayo (2006) se discute, al mismo tiempo, tanto el carácter universal como la exhaustividad expresiva de TM. Cfr. Rayo A. (2005) y (2006) y Moretti, A (inédito).

con que las interpretaciones relevantes tengan como dominio un conjunto. La prueba del teorema depende de que la teoría sea completa y tal dependencia condiciona su aplicación a otros ordenes lógicos (Kreisel, 1967). TM permite establecer también, entre tantas cosas, diversos resultados acerca de la cardinalidad de las colecciones que se necesitan para obtener una formalización en primer orden de una teoría de conjuntos. Por supuesto, uno de ellos es el conocido teorema Löwenheim-Skolem. Otro es el teorema de Cantor[3]. En pocas palabras, en el caso de las teorías lógicas de primer orden, TM ofrece a la vez, tanto un mecanismo que delimita las colecciones (vistas estas como objetos dentro del alcance de los cuantificadores que usamos en TM) acerca de las que habla la teoría formal en cada una de las interpretaciones, como un medio para representar, a través de sus propias estructuras interpretativas, cada uno de los modos intuitivos en los que podemos interpretar esas teorías formales.

De manera general, una *teoría exhaustiva* acerca de las asignaciones semánticas a una teoría formal tiene que poder representar, dada una teoría específica, cada uno de los integrantes de esa colección de asignaciones. Tiene que poder ofrecer, como parte de la misma, un *modo de interpretar* sus fórmulas para cada una de esas asignaciones semánticas. Esto es, cada *posible asignación de significado* tiene que poder corresponderse con una asignación que forme parte de nuestra teoría. Por eso, el deseo de lograr producir una *teoría general de la interpretación* (TI) se satisface sólo si, dada una teoría formal, hay tantos (y sólo tantos) modos de interpretar sus fórmulas dentro de nuestra teoría general, como asignaciones posibles de significados. Así, obtendríamos una caracterización que como mínimo sería extensionalmente equivalente a la colección de sus interpretaciones intuitivas. De esta manera, el principal interés filosófico que presenta TM es el de producir un análisis del concepto intuitivo de *interpretación:* uno en el cual, sin contradicción, podamos elucidar esos maneras intuitivas de interpretar las teorías formales.

Usualmente, con TM se adopta alguna suerte de *principio de optimismo semántico*: no hay teorías legítimas para las cuales no puede formularse explícitamente su semántica, esto es, para las cuales sea imposible dar una teoría exhaustiva acerca de todos los modos de interpretar tal teoría. Nótese, no obstante, que por las dificultades señaladas por Tarski (Tarski 1956), no puede exigírsele a TM, expresabilidad universal: surgiría una contradicción si pretendiéramos que, sin hacer distinciones de niveles de lenguaje, una TM pudiera expresar su propia semántica, en el sentido de ser capaz de dar una definición de *verdad* para una teoría suficientemente expresiva como para hablar de la aritmética, que incluya dentro de sus afirmaciones, el predicado veritativo de ese lenguaje. El predicado veritativo del metalenguaje utilizado para dar la semántica de su lenguaje objeto no puede tener un representante dentro de este lenguaje. Este resultado afecta de manera particular a la teoría de conjuntos (Rayo 2005): lo que se necesita para dar una semántica de una determinada TC es una TC´ más expresiva: una que

[3] El teorema Löwenheim-Skolem dice que si una colección de fórmulas de una teoría de primer orden tiene un modelo, entonces tiene un modelo contable. El teorema de Cantor dice que hay conjuntos que son incontables: conjuntos que son demasiado grandes como para ponerse en correspondencia uno a uno con los números naturales.

afirme que exista una totalidad cuya existencia no sea implicada por TC[4]. En este sentido, cada teoría de la jerarquía tendrá un poder expresivo mayor comparado con su teoría objeto. Tomemos una TC. Nada impide que haya una entidad dentro de otra TC (digamos TC´) que contenga todos los valores de las variables. TC´ tendrá un poder expresivo mayor. Lo que se impide es que esa entidad sea una de las que TC misma hable. Por supuesto, en los últimos años ha habido muchos intentos de llegar cada vez más cerca de ese límite (Kripke, 1975, Belnap & Gupta 1993, y Mc Gee 1991).

Ahora bien, se sabe que hace casi un siglo, Skolem presentó una colección de fórmulas pertenecientes a un lenguaje de primer orden, que con pequeñas variaciones, llegaron a ser aceptadas como una axiomatización estándar de la teoría de conjuntos (TC). Los axiomas de Skolem describen, en primer orden, los axiomas que informalmente habían sido presentados por Zermelo junto con un axioma agregado (el de reemplazo), el cual Fraenkel había mostrado como necesario. Estos axiomas son conocidos como ZF; como tales, ellos intentan hablar del universo conjuntista. Es decir, en tanto teoría formal axiomatizada, la teoría de conjuntos puede recibir distintas interpretaciones. Como sucede con toda teoría axiomatizada, la delimitación de la colección de interpretaciones que satisfacen los axiomas de la teoría es crucial para la elucidación de la colección de verdades e inferencias válidas dentro del alcance de la teoría. Pero, hay una asignación de significados a las fórmulas de esa teoría, una dentro de esa colección de posibles asignaciones, que pretende hablar acerca de todos o sólo aquellos individuos que consideramos conjuntos. Por ese motivo, la exhaustividad de TM incluye la necesidad de contar como parte de esos mecanismos representativos *uno* que permita describir *todos los conjuntos que hay*. La posibilidad de tener una interpretación con *máxima generalidad* representada en TM garantiza que esta pueda describir el universo conjuntista. Será esta, entonces, la segunda condición que debe cumplir TM en tanto TI: poseer una interpretación de *máxima generalidad* en donde quede representado el universo de la teoría. Por eso, debemos poder hacer como parte de TM afirmaciones cuantificadas cuyas condiciones de verdad dependan de todas esas colecciones.

El hecho de que TM, en tanto TI, sea una teoría *exhaustiva* acerca de los modos de interpretar las teorías formales incluyendo una con *máxima generalidad* es fundamental para tener una buena reconstrucción de la colección de inferencias y verdades propias de la teoría. No obstante, en tanto teoría acerca de *todos los modos de interpretar* teorías formales, TM no posee axiomas propios: los toma prestados usualmente de otra teoría formal con el propósito de precisar los modos en los que interpretamos, por medio de esas estructuras, las teorías formales. Como hemos visto, en el caso de las teorías lógicas de primer orden, TC ha sido empleada exitosamente para desarrollar una teoría exhaustiva acerca de los modos de interpretar las teorías de primer orden. Frente a tal resultado, es tentador pensar en la posibilidad de extender tal aplicación a otros tipos de teorías. Por supuesto, no contamos con una prueba del tipo de la de Kreisel. Pero

[4] Uso *TC* como una abreviatura de cualquier teoría de conjuntos que esté formalizada en primer orden. Por razones de simplicidad expositiva, supondré que es del estilo ZFC. Si en cambio TC fuera del estilo NGB, se requeriría afirmar la existencia de otras colecciones, ya que al haber clases propias el problema no surge con la totalidad que reúna todos los conjuntos de TC.

la postulación de un *principio de reflexión* [5] podría alentarnos en ese camino en el caso de las teorías de orden superior. Veremos más adelante que los resultados de Orayen limitan esta tentación.

Ahora bien, la utilización de una teoría de conjuntos como parte de TM alienta la necesidad de que como parte de los requerimientos expresivos que tiene que cumplir TM esté el de probar la existencia de su universo de discurso. Y a la vez que TM ha sido presentada como una teoría formal, ese universo de discurso sea lo suficientemente rico como para servir de marco para las diferentes teorías formales. Cuánto se requiere del dominio depende en parte de la riqueza del lenguaje. Como parte de estos requisitos expresivos, el universo de TM debe contener los medios suficientes como para poder hablar de todo aquello de lo que hablan las teorías formales, incluyendo al menos, todas las verdades de la aritmética[6]. Un modo usual de operar dentro de TM, como un medio para hacer inteligible la cuantificación sobre las entidades de esos "inmensos universos de discurso", es adoptar el principio denominado "todo en uno" [All-in-One Principle][7]. Según este principio, los objetos de los distintos dominios de interpretación de cada una de las estructuras que conforman TM, forman una entidad. Claro, al tratarse de un objeto colectivo, es natural pensarlo como un conjunto o alguna entidad que forma parte del universo conjuntista. Pero, este resultado carece de una prueba (Kreisel, 1967). Este principio está motivado en la idea de que la TM, y en particular, la caracterización modelo-teórica de la noción de *consecuencia*, requiere cuantificar sobre todos los dominios. Esta idea conduce a tomar a las interpretaciones mismas como *objetos* dentro del alcance de los cuantificadores de primer orden. Y no podemos cuantificar sobre algunas cosas sin que haya una cosa singular (una colección) de la cual ellas sean integrantes. Para poder cuantificar sobre ciertos objetos, se tiene que presuponer que esos objetos constituyen una colección completa, una única entidad de la cual ellos sean miembros[8].

En suma, TM, en tanto teoría de la interpretación, debe cumplir con un requisito de *exhaustividad expresiva*. Debe representar cada uno de los modos de interpretar las teorías formales, dentro de los cuales debe contener un modo en donde las afirmaciones de la teoría traten acerca de todas las entidades de las que habla en sentido intuiti-

---

5 Una formulación del *principio de reflexión* para segundo orden establece que cualquier fórmula verdadera φ de una teoría de segundo orden, es también verdadera en algún modelo conjuntista. Cfr. Shapiro, S. (1991) *Foundations without Foundationalismo: a Case for Second-Order Logic* (Oxford, Clarendon Press), y Rayo, A., & G. Uzquiano (1999), "Toward a Theory of Second-Order Consequence", *The Notre Dame J. of Phil. Logic* 40.

6 Max Freund me llamó la atención sobre la necesidad de hacer explícitos los vínculos entre considerar a TM como una TI y su papel en la fundamentación de las matemáticas. Cfr. <http://www.accionfilo sofica.com/blog/mensaje.pl?id=90>.

7 Cartwright afirma "The general principle appears to be that to quantify over certain objects is to presuppose that these objects constitute a "collection" or a "completed collection" —some one thing of which those objects are members. I call this the All-in-One Principle".

8 Cartwright critica el principio all-in-one, ya que le parece inadecuada la idea de que "we cannot speak of the cookies in the jar unless they constitute a set..." R. Cartwright (1994), p. 7-8. Para una defensa, cfr. Parsons, C. (1984).

vo esa teoría. Al mismo tiempo, debe incluir dentro de los mismos, una interpretación con *máxima generalidad*: una que pueda hablar de *todas* las entidades de las que intuitivamente habla la teoría. En cuanto a la relación entre TM y TC, esta última usualmente es una herramienta fundamental para la construcción de estructuras interpretativas de las teorías formales. Lo es, al ser una teoría acerca de esas entidades colectivas imprescindibles para interpretar las teorías formales. Y tal requisito surge de la necesidad de cuantificar dentro de TM.

II

En "Una paradoja en la semántica de la teoría de conjuntos" (Orayen, 2003), Raúl Orayen presenta una grave dificultad al proyecto de dar una teoría general de la interpretación. El problema que descubre Orayen surge a partir de la pregunta por la posibilidad de formalizar la teoría informal de conjuntos[9] que se utiliza para interpretar los lenguajes de primer orden. Su análisis de la cuestión llega a un punto desconcertante: ninguna teoría formal de conjuntos es suficientemente rica como para contener una colección que sirva de dominio de su propia interpretación. De lo cual, tendríamos que concluir que no hay una formalización apropiada de la teoría informal que usamos en nuestra teoría de modelos, ya que sólo si pudiéramos obtener un dominio de interpretación apropiado para tal teoría, estaríamos en condiciones de formalizarla. Orayen argumenta que tal cosa es imposible, y con este argumento muestra que hay asignaciones posibles de significado que no corresponden a ningún modelo, ya que no hay ningún modelo cuyo dominio consista en *todos y sólo* las entidades de los que trata la teoría de conjuntos, lo que constituye una importante dificultad al proyecto de dar una teoría general de la interpretación de los lenguajes formales. De esta manera, al usar TM, como herramienta para interpretar TC, advertimos que no tenemos recursos expresivos suficientes dentro de la teoría de conjuntos como para poder hablar acerca del dominio de su propia interpretación pretendida, incumpliendo al mismo tiempo, tanto con el requisito de *máxima generalidad,* como con el de *exhaustividad expresiva*: no somos capaces de hablar acerca de todos los objetos del universo conjuntista, y por esto mismo somos incapaces de ser exhaustivos respectos de todos los modos de interpretar que deberíamos describir. Orayen nos muestra que no es posible encontrar, como parte de TM, un dominio capaz de servir de universo de discurso de la TC, uno que forme parte de una interpretación máximamente general, en el sentido de contener a *todas* las entidades de las que intuitivamente habla nuestra teoría de conjuntos. Tal cosa muestra que no contamos con una formalización apropiada.

Por supuesto, el problema que señala Orayen no implica que no podamos contar con un modelo, una estructura que determine las condiciones de verdad de las fórmu-

---

[9] Con "Teoría intuitiva de conjuntos" Orayen hace referencia a la colección de afirmaciones conjuntistas, preformales y consistentes, que usualmente se los lógicos utilizan al elaborar la semántica de los lenguajes formales. Por supuesto, no está haciendo referencia a las ideas conjuntistas previas a la paradoja de Russell a partir de la cual se elaboraron las diversas axiomatizaciones formales de las teorías de conjuntos.

las de la teoría. E incluso, tampoco querría decir que no contamos con una formalización que se *aproxima* a la interpretación intuitiva[10]. Lo que querría decir es que, dentro de TM, hay algo que "se nos escapa" con respecto a la capacidad expresiva de la teoría formal. Y, al mismo tiempo, al mostrar que hay *modos intuitivos de interpretar* las teorías formales a los que no corresponde ningún correlato en la teoría de modelos, el proyecto de dar una teoría exhaustiva de la interpretación, una que analice todos los diversos modos de asignar significados a las teorías formales a través de la teoría de modelos, correría serios riesgos[11].

La pretensión de Orayen es que tal dificultad tenga la fuerza de una *paradoja*, en el sentido de mostrar una incompatibilidad de supuestos comúnmente adoptados al hacer una reconstrucción racional de la práctica usual de los lógicos, sin dejarnos marcado un claro camino para su solución. Su estrategia consiste en prestar atención a las dos ideas siguientes, ambas, igualmente intuitivas. En primer lugar, cuando estamos interesados en interpretar una teoría de primer orden, lo que hacemos es asignar entidades apropiadas a cada una de las expresiones no lógicas de la teoría. Tal concepción parece comprometernos con la idea de que esas entidades pueden ser agrupadas en un conjunto (un dominio de interpretación) que es una construcción interna de la teoría intuitiva de conjuntos. En segundo lugar, la teoría intuitiva de conjuntos es presentada generalmente, en cualquiera de sus axiomatizaciones, como una teoría de primer orden, con un único predicado extra-lógico: el predicado binario de pertenencia. De esta manera, podríamos decir que todo lo que puede expresarse en una teoría intuitiva de conjuntos, podría formalizarse en una teoría de primer orden[12]. En palabras de Orayen[13]:

> (1) "La semántica de TQ se construye con la ayuda de T; en particular, se adopta la restricción de que solo conjuntos provistos por T pueden usarse como dominios de interpretación"
>
> (2) "Se dice que T puede ser formalizada dentro de TQ (i.e., puede ser expresada por una teoría de orden uno)" (Orayen 2003, p. 39)

---

[10] Tanto Áxel Barceló como Agustín Rayo exploran esta idea según la cual podemos ver a TM como una aproximación, como una reconstrucción parcial de los conceptos preteóricos (validez universal, consecuencia lógica, etc.).

[11] Por ejemplo, Etchemendy ha argumentado que, teniendo en cuenta las propiedades modales de la noción intuitiva de *consecuencia lógica*, no hay un argumento que garantice que hay tantos modos posibles de interpretar intuitivos como modos conjuntistas de interpretación, lo cual constituiría, según su perspectiva, un problema para la concepción tarskiana modelo teórica. Etchemendy, además, insiste en que al intentar brindar un argumento en esta dirección, Tarski cometió una falacia modal de alcance. *Cfr.* J. Etchemendy (1990), *The Concept of Logical Consequence*, Harvard: Harvard University Press.

[12] Boolos defiende la idea de que hay ciertas afirmaciones acerca de conjuntos que queremos hacer (quizás afirmar que hay una "totalidad" o "colección" que contiene todo y sólo los conjuntos que no se contienen a sí mismos sea una de estas afirmaciones), las cuales ciertamente no pueden hacerse usando fórmulas de primer orden, pero las cuales pueden ser expresadas por medio de fórmulas de segundo orden (entendidas como una cuantificación plural). *Cfr.* Boolos, G. (1984).

[13] Usaremos, siguiendo a Orayen, la expresión *TQ* como una abreviatura de una teoría axiomática cuantificacional clásica de primer orden; *T* como una abreviatura de la teoría intuitiva de conjuntos. Se entiende por "intuitiva" una teoría no formalizada.

Orayen hace dos afirmaciones con respecto a las tesis (1) y (2)

> (3) "Las definiciones contenidas en un desarrollo cuidadoso de (1) son incompatibles con la afirmación hecha en (2)"
>
> (4) "La razón es que si la semántica es construida como sostiene (1), T no puede ser formalizada, porque es imposible obtener un dominio apropiado para tal teoría". (Orayen 2003, p. 39)

(3) Afirma que las tesis (1) y (2) son incompatibles; (4) argumenta a favor de esa incompatibilidad.

El argumento con el que Orayen justifica (4) es el siguiente

> (a) (2) Establece que T puede ser formalizada adecuadamente dentro de TQ. Sea TC tal formalización.
>
> (b) Ya que una formalización de T por TC implica por sí sola una interpretación de TC a partir de T, una formalización de este tipo implica que hay una colección D que contiene las mismas entidades postuladas por T.

O en palabras de Orayen, si TC es una formalización adecuada de T, entonces TC dice *"las mismas cosas que T acerca de lo que T habla"* (Orayen 2003, p. 39). Claro que no es que cualquier interpretación que satisfaga los axiomas de TC incluirá un dominio D que contenga *todos y sólo* los objetos sobre los que habla T. TC, en tanto una TQ consistente, puede tener distintas interpretaciones que la satisfagan. No obstante, cuando intentamos formalizar T con TC, esa interpretación deberá tener un dominio que hable de lo mismo que T. Ahora bien, no obstante, por (1)

> (c) D Es uno de los conjuntos sobre los que habla T.

Pero, de (b) y (c), se sigue que

> (d) T debe afirmar que existe un conjunto D que contiene a todos los objetos sobre los que habla T.

Pero, como TC formaliza T

> (e) TC debe contener un teorema que afirme que existe un conjunto C tal que todos los conjuntos pertenecen a C.

Esto es así, ya que resulta conveniente que sobre el tema de lo que habla una teoría intuitiva (como lo es T), su formalización (esto es, TC) contenga las mismas inferencias y diga las mismas verdades que la teoría que se está formalizando.

> (f) Sin embargo, si TC es del tipo ZF, no existe tal conjunto.[14]

(e) y (f) muestran la contradicción: la que establece que hay y no hay un conjunto tal que *todos y sólo* los conjuntos que se necesitan para formalizar T con una TC son sus elementos.

En suma, podemos preguntarnos si existe un conjunto que pueda jugar el rol de brindar entidades apropiadas para interpretar una teoría intuitiva de conjuntos. Según Orayen, la respuesta a esta pregunta genera cierta perplejidad: este dominio debe ser o una construcción interna (un conjunto del que la teoría intuitiva de conjuntos hable) o una construcción externa (un conjunto que sirva de dominio de interpretación de la

---

[14] Algo semejante ocurre si TC es de tipo NGB o de tipo NF. Cfr. Hurt, W. (2003), "Carta de William Hart", en Moretti, A. & Hurtado, G. (2003).

semántica de primer orden). Pero, siendo una construcción interna, la teoría se vuelve inconsistente, siendo una construcción externa, no tenemos una construcción apropiada para interpretar la teoría de conjuntos[15].

III

Hemos visto que, según Orayen, su paradoja surge porque al interpretar *TC* como una formalización de T, necesitamos más conjuntos que aquellos permitidos por *TC*. (Orayen 2003 p. 47). Ese conjunto que necesitamos para formalizar T con una TC no puede ser ninguno de los que forman parte de TC. Las asignaciones semánticas de la TM, esto es, las diversas interpretaciones utilizadas en la construcción de la semántica de TQ, son asignaciones formales de estructuras apropiadas. La suposición de que esas estructuras, que como hemos visto, involucran colecciones, son conjuntos y que como tales deberían ser brindadas por alguna TC, conduce a contradicción. Todos los dominios de interpretación, todas las colecciones utilizadas en las estructuras que interpretan las TQ, incluyendo el suyo propio, deben existir dentro de una TC. Pero no puede haber ningún conjunto dentro de TC capaz de desempeñar ese papel. Por ese motivo, TM no logra tener una interpretación máximamente general para TC que pueda abarcar todas las entidades de las que habla T. Al mismo tiempo, TM no logra ser exhaustiva, esto es, no logra representar todos los modos de interpretar TC, uno que contenga todas las entidades de las que habla T.

El problema que Orayen encuentra al intentar brindar por medio de TM una interpretación conjuntista de TC pretende afectar a TM en general. No es únicamente un problema que encontramos al intentar obtener una interpretación conjuntista de TC. Sin duda, la pretensión de aplicación de TC para interpretar TC de manera tal que obtengamos una formalización apropiada de T, es decir, que autoapliquemos TC para obtener una interpretación de sí misma que formalice T, es parte central del problema. Pero, desde mi perspectiva, el problema planteado por Orayen no se circunscribe a la tarea de brindar una interpretación apropiada de TC que sea capaz de hablar de lo mismo que T. Quizás, por motivos semejantes, Adolfo García de la Sienra sostiene que la paradoja descubierta por Orayen es una "vejación a la teoría de modelos" (García de la Sienra 2003 p. 72). Sin embargo, en la línea opuesta a la supuesta vejación a la teoría de modelos, Mario Gómez Torrente ha tratado de mostrar que el argumento de Orayen no tiene la fuerza suficiente como para constituir una paradoja acerca de TM. (Gómez Torrente 2003). Desde su perspectiva, no es cierto que los modelos aceptables estén *siempre* constituidos por conjuntos.

En este punto, me interesa argumentar que, a diferencia de lo que Gómez Torrente sostiene, para que surja el problema que Orayen indica, hace falta un supuesto más

---

[15] Por supuesto, esto no quiere decir que no podamos encontrar modelos conjuntistas para una TQ. Para cualquier TQ consistente, siempre es posible construir un modelo, y por lo establecido en el teorema de completitud, un modelo conjuntista. Todo el punto es si hay una única teoría, en este caso, la de modelos, capaz de hablar acerca de todas las entidades necesarias para construir la semántica de todas las teorías formales.

débil: que las interpretaciones sean consideradas como objetos. Es decir, pienso que no hace falta suponer que los modelos sean siempre conjuntistas. Además, considero que, a diferencia de lo que ocurre con el supuesto que Gómez Torrente señala como inaceptable (acuerdo con él en que no contamos con un argumento general que muestre que las interpretaciones sean estructuras conjuntistas), en el caso del supuesto más débil, hay una manera plausible de defenderlo, si el propósito es usar a TM como una TI. Es decir, si queremos producir un análisis del concepto de *interpretación,* uno en el cual, por medio de una teoría formal podamos ser capaces de representar exhaustivamente cada una de las maneras de interpretar las teorías formales, incluyendo un modo absolutamente general (es decir, uno en el cual los cuantificadores sean interpretados como hablando de todo), entonces hay un modo de defender un supuesto más débil a partir del cual se generan las dificultades indicadas por Orayen.

Desde mi perspectiva, el supuesto más débil que se necesita para que surja la paradoja es que las interpretaciones sean representables en TM como objetos. Esto es, hemos visto que es usual al operar con TM la adopción del principio "todo en uno", según el cual, los objetos en los distintos dominios de interpretación de cada una de las estructuras forman algún conjunto o alguna entidad que forma parte del universo conjuntista. El principal argumento con el que se justifica esta idea es que la caracterización modelo-teórica de la nociones de *consecuencia* y de *validez universal* requiere cuantificar sobre todos los dominios. Esto permite tomar a las interpretaciones mismas como objetos dentro del alcance de los cuantificadores de primer orden. Y sólo así podemos decir que TM trata acerca de *todas* las colecciones necesarias para interpretar las teorías formales, obteniendo resultados satisfactorios (al menos extensionalmente) respecto de las mencionadas nociones lógicas.

Me parece que es este supuesto, y no el que Gómez Torrente le atribuye a la formulación de Orayen, el que hace posible plantar la paradoja: TM es una teoría formal, que como tal, tiene diversas interpretaciones, una de las cuales, la pretendida, tiene un dominio del que forman parte todas las colecciones necesarias para interpretarse. Pero, el dominio del modelo es una colección, y de acuerdo a las teorías de las colecciones, no existe una colección suficientemente grande como para incluir todas las entidades necesarias como para interpretarse a sí misma. Por lo tanto, ningún modelo puede capturar la interpretación deseada de TM. Es decir, de acuerdo a esta reconstrucción de Paradoja de Orayen, lo crucial es que:

> si las interpretaciones caen bajo el alcance de las afirmaciones cuantificadas de TM, es posible mostrar que hay más modos intuitivos de interpretar teorías formales que estructuras que caigan bajo el alcance de las afirmaciones cuantificadas de la propia TM.

Y como he argumentado, que las interpretaciones caigan bajo el alcance de los cuantificadores que hacemos al hacer afirmaciones en TM, es una necesidad que surge al intentar hacer generalizaciones que nos permitan dar, entre otras cosas, una definición apropiada de *consecuencia* y de *validez universal.*

Si se me ha seguido hasta aquí, se debe advertir que mi punto contra la posición de Mario Gómez Torrente es que lo que se necesita para que surjan los problemas señalados por Orayen es proponerse que TM sea capaz de brindar la ontología y el poder expresivo necesario como para interpretarse a si misma. Y lo que es más importante, no hace falta suponer, para que surjan tales dificultades, que *todo modo de interpretar* usado por TM tenga que ser *un modo de interpretar conjuntista*, esto es, uno que tenga como dominio un conjunto. Sólo hace falta que necesitemos cuantificar sobre objetos colectivos, cuantificación que requiere que "juntemos" esas entidades en un nuevo objeto. Para apreciar la fuerza de este punto, conviene introducir la siguiente prueba de Williamson (Williamson 2004):

Supongamos que L es un lenguaje de primer orden cuyos cuantificadores son interpretados de manera irrestricta. Sea P un predicado monádico de L. Sea F cualquier predicado significativo del metalenguaje. Según Williamson, debe ser posible interpretar P como significando F. Esto es,

(1) $\forall x$ (i_F es una interpretación bajo la cual P se aplica a x ssi Fx)

Por ejemplo, si el metalenguaje es el español, un caso de este tipo consiste en `$\forall x$ ($i_F$ es una interpretación bajo la cual $P$ se aplica a $x$ ssi $x$ es un hombre)´, donde `$\forall x$´ tiene un rango irrestricto.

Dada la suposición de que las interpretaciones son objetos[16], parece razonable permitir definir un nuevo predicado R tal que

(2) $\forall x$ (Rx ssi x no es una interpretación bajo la cual P se aplica a x)

Si en (1), R reemplaza a F y al mismo tiempo, aplicamos la definición de R en (2) obtenemos:

(3) $\forall x$ (i_R es una interpretación bajo la cual P se aplica a x ssi x no es una interpretación bajo la cual P se aplica a x)

Ya que `$\forall$x´ en (3) es absolutamente irrestricto, podemos instanciarlo y obtener:

(4) i_R es una interpretación bajo la cual P se aplica i_R ssi i_R no es una interpretación bajo la cual P se aplica i_R.

(4) es una contradicción que no depende de que las interpretaciones sean conjuntos. Sólo depende de que sean objetos. Y la cuantificación sobre dominios parece suponer esto último.

Por supuesto, no estoy sosteniendo la tesis según la cual no hay maneras de interpretar las teorías formales que no supongan que las interpretaciones sean objetos. Incluso más adelante en este mismo trabajo consideraré la posibilidad del recurso al lenguaje natural como un modo alternativo de interpretar teorías formales. Lo que simplemente estoy diciendo es que la necesidad de cuantificar sobre interpretaciones parece dejarnos sin salida: al considerar a estas como objetos, no hay suficientes objetos como para representar todos los modos de interpretar una TI.

---

[16] En la formulación de Linnebo: Sem 1: una *i* es un objeto. Cfr. Linnebo, O. (2006), en Rayo, A. & Uzquiano (2006).

Una manera de responder a la objeción acerca de lo que supone la paradoja[17] consiste en señalar que Gómez Torrente también objetaría que este supuesto más débil sea una creencia compartida por los lógicos. Tampoco la paradoja llegaría a constituirse como tal, ya que no sería cierto que las únicas estructuras aceptables usadas al interpretar una teoría sean objetos. Frente a esta posible objeción me gustaría señalar con mayor claridad cuál es el interés de la paradoja.

Es claro que la paradoja de Orayen no afecta a diversos propósitos que podemos tener al usar TM: no afecta en absoluto, por ejemplo, nuestro uso de TM para dar una caracterización correcta del concepto de *consecuencia* para las teorías de primer orden, ni el resultado de Kreisel respecto de los modelos conjuntistas, ni muchos otros propósitos que podemos tener al utilizar TM[18]. Lo que afecta es el propósito filosófico de utilizar TM como una TI, esto es, considerar si a través de las estructuras que utilizamos en TM estamos en condiciones de brindar una caracterización del concepto de *interpretación.* Al ser un objetivo filosófico, considero que no se trata de discutir simplemente acerca de lo que los lógicos creen, sino más bien, al intentar ofrecer una reconstrucción racional de la práctica de interpretar lenguajes formales, cuáles son los supuestos necesarios para ofrecer un análisis satisfactorio. Insisto, subyace al proyecto de dar una teoría general de la interpretación (TI), el intentar caracterizar de un modo preciso la noción intuitiva de *modo de interpretar* una teoría formal. Como en otros casos en los que se busca una elucidación, la noción en consideración posee cierto grado de vaguedad e imprecisión. En un sentido relativamente no técnico, es posible interpretar las expresiones no lógicas de una teoría de diversas maneras, haciendo asignaciones intuitivas de objetos, o equiparándolas con las de otra teoría, e incluso traduciéndolas al lenguaje natural. Este procedimiento, nos permite tener una idea intuitiva acerca de las verdades de una teoría formal. Ahora bien, elucidar ese concepto es el objetivo central de una TI: vincular el concepto intuitivo de *modo de interpretar* una teoría formal con un concepto más claro y preciso, que permita extraer correctamente conclusiones interesantes acerca del primero. En este punto, aquel que se haya propuesto esta tarea, recurre a TM. Por supuesto, como en toda caracterización de este tipo, se requiere cierto grado de idealización: aquel que vincula en un análisis dos conceptos, no pretende, por lo general, que todo aquel que crea que el concepto analizado tiene ciertas características, también tendrá que creer que el concepto que lo analiza tenga esas mismas características. Esto es, usualmente no es posible pedir la equivalencia analítica entre las nociones involucradas. Y es claro que, en nuestro caso, en el intento de ofrecer una caracterización precisa del concepto intuitivo de interpretación, encontramos diversos elementos que tienen los modos de interpretar intuitivos que no tienen las estructuras que usamos en TM.

[17] Un réferi anónimo de *Theoria* me ha señalado esta objeción. Agradezco su comentario.

[18] Hay distintos objetivos que pueden seguirse al construir una (TM) para un lenguaje formal: simplemente, puede existir un propósito lógico: definir una noción adecuada de *consecuencia lógica,* u epistémico tal como explicar la relación entre el lenguaje y la experiencia, o incluso cognitivo, intentando modelar aspectos matemáticos de la cognición.

Hemos visto, entonces, que para que una TI pueda realizarse correctamente, hace falta que dentro de TM podamos tener una representación exhaustiva de los modos de interpretar intuitivos las teorías formales. Tal requisito nos aseguraría la coextensionalidad entre el concepto intuitivo y el concepto que lo representa en TM, dando de esta manera, condiciones necesarias y suficientes de aplicación[19]. Claro que uno de los casos de aplicación, cuando usamos TM como una TI, es aquel en el que fuera posible delimitar una colección (un objeto) como parte de TM que sea el D apropiado para brindar las entidades necesarias para interpretar a la misma TM. La paradoja muestra que no pueden existir ni suficientes conjuntos ni suficientes objetos como para representar todos los modos intuitivos de interpretar teorías. Y de esta manera, la paradoja marca límites a la caracterización usual dentro de TM de lo que es una interpretación: una interpretación no puede ser un conjunto, ya que hay más modos de interpretar TC que conjuntos capaces de ser su D. Tampoco una interpretación puede ser un objeto, ya que hay menos objetos que modos de interpretar predicados.

IV

He insistido, en varias oportunidades, que es importante advertir que la paradoja de Orayen sale al proponernos brindar una caracterización del concepto intuitivo de *interpretación* a partir de una única teoría general que intente representar *todos los modos posibles de interpretar* los lenguajes formales, incluyendo la necesidad de analizar cómo esta misma teoría puede interpretarse a sí misma. He llamado a este proyecto TI y he intentado mostrar que si se recurre a TM con este propósito, surge la dificultad señalada por Orayen. He sostenido, en el punto anterior, que el problema que descubre Orayen no parece estar en si hacemos o no una suposición conjuntista al elaborar una TI. En este apartado, en cambio, quiero enfatizar que el problema que presenta Orayen se debe, en parte, a que este proyecto es autoplicativo: TI debe poder contar dentro de sus propios compromisos ontológicos, con una colección (sea o no ésta un conjunto) capaz de reunir a todas aquellas entidades de las que habla ella misma.

Hemos visto que un aspecto crucial para que surja la paradoja al elaborar TI es la pretensión de usar TM como una teoría acerca de *todos los modos de interpretar las teorías formales*, incluyendo uno que posea máxima generalidad. Sólo así las estructuras interpretativas que TM brinda podrían ser consideradas una caracterización correcta del concepto intuitivo de *interpretación*. Esto es, que las afirmaciones cuantificadas dentro de TM puedan ser interpretadas sin ninguna restricción contextual (es decir, que "toda interpretación" se interprete como toda interpretación). El punto a destacar es que al tener esta pretensión, debemos aceptar que la autoaplicación es un caso de aplicación

[19] Incluso, se podría sostener que exigir coextensionalidad quizás sea demasiado, ya que los conceptos intuitivos poseen siempre cierto grado vaguedad y como sucede con todo concepto vago, hay efectivamente casos en donde se aplica, casos en los que no se aplica y casos intermedios (en la frontera) en los que no está claro si este se aplica o no. Quizás entonces, lo que tiene sentido exigir es que TM brinde un concepto preciso de *interpretación* en cuya extensión caigan sólo casos claros de modos de interpretar teorías, en un número suficientemente amplio de ellos y en cuya extensión no caigan casos que no sean casos claros de aplicación.

de TI. Por eso, cuando esperamos que TM brinde un análisis general de la noción de *(modo de) interpretación*, deseamos que haya *uno* que incluya como caso aquel que consiste en autoaplicarse. Lo que la paradoja muestra es que no hay una manera apropiada de analizar el concepto de *modo de interpretación* a partir del concepto de *estructura,* ya que no hay una colección que forme parte de alguna teoría de las colecciones que sea capaz constituir una estructura apropiada que se autointerprete. Es decir, sea cual sea el modo en el cual se realice esa autoaplicación, la paradoja surge al intentar reunir en una *única entidad* que sirva como dominio de una interpretación (sea esta un conjunto, clase propia, una hiperclase o pluralidad de cualquier tipo) todo aquello que sea necesario para hablar de *todas* las entidades de las que presuntamente habla una teoría (incluyéndose a sí misma como una de esas entidades). La sola utilización dentro de TM de cuantificadores cuyo rango pueda incluir las distintas colecciones necesarias para caracterizar apropiadamente todos los modos de interpretar TM, incluyendo un modo de interpretar TM que hable de todo, requiere que *uno de los objetos de los que hable TM* sea su propio universo de discurso. Es justamente esta necesidad la fuente de la perplejidad que encuentra Orayen.

Considero, entonces, que al igual que lo que sucede con el concepto de *pertenencia*, el concepto de *interpretación* involucra casos de autoaplicación. Por eso, el propósito de brindar una teoría exhaustiva acerca de los modos de interpretar teorías formales y que posea la capacidad expresiva suficiente como para poder hablar de todo lo que habla la aritmética, conduce a una paradoja.

Anticipando algunas posibles objeciones, me gustaría aclarar que con lo anterior no estoy sosteniendo la tesis general de que todo caso de autoaplicación conduce a una paradoja[20]. La autoaplicación no es una fuente suficiente (como muestran los resultados de Gödel) y probablemente tampoco necesaria (como sugiere la paradoja de Yablo) de origen de paradojas. Lo que estoy sosteniendo en que bajo ciertos propósitos expresivos, la autoaplicación produce problemas. Tampoco estoy sosteniendo que el hecho de que TM tenga que tener como parte de sus entidades a una entidad impredicativa sea suficiente para producir problemas. El D de una TM que se proponga como una TI sería una entidad impredicativa, ya que ella misma tendría que formar parte del universo de discurso que ella establece. Pero, el sólo hecho de que una teoría de las colecciones se comprometa con una entidad universal no conduce a contradicción: el sistema de Quine *New Foundations* es de este tipo (Quine 1963). No obstante, tal como argumenta Hart en contra de Alchourrón (Hart, 2003), el recurso a una teoría que, a diferencia de NGB o ZF, se comprometa con colecciones universales tampoco ayuda en la resolución de la paradoja de Orayen, ya que la necesidad de bloquear la instancia del axioma de comprensión que produce la paradoja de Russell, obliga a limitar la relación de *pertenencia* en dicho sistema, con el propósito de evitar que sea autoaplicante.

En suma, he argumentado que si nos proponemos analizar los modos en los que interpretamos las teorías formales, necesitamos una TI que sea autorreferencialmente coherente: necesitamos una teoría que incluya una colección vista como un objeto,

[20] Max Freund me ha hecho notar este punto.

que contenga todas las entidades del universo de la teoría, incluyendo esa colección que es el universo de la teoría en su interpretación pretendida. Pero, si lo que estoy sosteniendo es correcto, lo que muestra la paradoja, sin la necesidad de suponer en ningún punto que esa totalidad es un conjunto, sino sólo la suposición según la cual esa totalidad debe ser una construcción interna de la TM propuesta como una TI, es que no existe una colección capaz de realizar esa tarea. Claro, el punto principal para que mi argumento se siga, es aceptar que dentro de los propósitos fundamentales de TI debemos incluir el requisito de *exhaustividad expresiva*. De otra manera, no hay dificultad alguna. La Paradoja de Orayen muestra, entonces, que nunca tendremos una TM exhaustiva, una que sea capaz de caracterizar completamente el concepto de *interpretación* y al hacerlo constituirse en una adecuada TI.

V

Como en toda paradoja, en la de Orayen se obtiene una contradicción a partir de la reconstrucción racional de algunos supuestos establecidos en la práctica de la disciplina. Claro que una paradoja no implica imposibilidad de búsqueda de solución, generalmente emparentada con el abandono de alguno de esos presupuestos. Sin embargo, es deseable que ese abandono esté fundado en alguna razón adicional a la que pudiera darse mostrando que ese es simplemente un camino para evitar la mencionada dificultad. Reformar la reconstrucción racional de una práctica arraigada exige razones. No dar ninguna razón más que la de ofrecer una solución a la dificultad provoca suspicacias relacionadas con las consideraciones que generan los enfoques ad hoc.

Ahora bien, el planteo mismo de la Paradoja de Orayen supone que tendría que haber *una* teoría que brinde la ontología necesaria para la construcción de la interpretación de las teorías formales. Un primer camino posible de solución es abandonar el supuesto de que hay una única teoría que puede dar la ontología necesaria para la construcción de la semántica de TQ. Una opción dentro de esta estrategia es continuar aceptando el enfoque conjuntista de la semántica, abandonando tal supuesto unificador: estamos en presencia de la adopción del enfoque ortodoxo de las jerarquías (Orayen 2003. p. 47 y siguientes). En tal caso, podemos construir una interpretación de *TC* usando los conjuntos de otra teoría de conjuntos (llamemos a esta teoría *TC´*) y luego podemos usar una teoría *TC´´* para construir una interpretación de *TC´*, y así sucesivamente, cada vez que necesitemos más conjuntos que los provistos por la teoría misma que estamos usando. En palabras de Orayen "deshacerse de la idea de que una sola teoría de conjuntos puede proveer todos los conjuntos que son requeridos en la semántica" (Orayen 2003 p. 49). Esta respuesta implica que por más que ascendamos en esta jerarquía, nunca nos detendremos. El análisis jerárquico nos dice que en cada nivel hay algo *fuera* de la extensión del predicado "es un conjunto" que siempre se nos escapa. Por ese motivo, cada una de las teorías de la jerarquía debe ser ontológicamente más rica respecto de los conjuntos con los que nos vamos comprometiendo a medida que ascendemos. Y por ello, cada una de las teorías de conjuntos de la jerarquía sufre una profunda *limitación expresiva*: ninguna puede autoaplicarse. Ni es capaz de hablar, directa o indirectamente, acerca del dominio de interpretación a partir del cual sus fórmulas recibirán asignaciones semánticas. Nos quedamos sin la posibilidad de tener una teoría de conjuntos máximamente comprensiva. Una que nos brinde la mayor y más rica diversidad posible de estructuras que nos permitan interpretar nuestras fórmulas de *TQ*. A lo anterior, podría agregarse la siguiente objeción usual en las discusiones de este enfoque: ¿desde dónde se habla al elaborar la jerarquía? Toda noción de *conjunto* debe pertenecer a un nivel de la jerarquía de teorías, de manera tal que podamos interpretar esa noción a través de un conjunto del cual hable una teoría de nivel superior. Pero, al enunciar esta tesis debemos usar una noción de *conjunto* que no pertenezca a ningún nivel de tal jerarquía. El propio Orayen señala que la cuestión parece remitir al punto de vista del *Ojo de Dios*, hay que ubicarse *fuera* de toda la jerarquía para formular la afirmación de que la jerarquía existe. Debemos estar fuera de la jerarquía conjuntista para formular la tesis de que hay una jerarquía entre los distintas *TC* que permite interpretar a través de la inclusión entre sus objetos, los distintos conjuntos que servirán de dominios de interpretación. Pero, si tomamos seriamente la afirmación, debemos admitir que hay más conjuntos que los admitidos por cualquiera de las *TC*. Y si esto es así, las *TC* de la jerarquía no hablarían acerca de *todos los conjuntos*, incumpliendo el requisito de *máxima generalidad*, y nuevamente no tendríamos garantías acerca de que este conjunto del cual habla la tesis sea un conjunto admisible.

Frente a estos problemas del enfoque jerárquico, una opción diferente es continuar adoptando el enfoque unificador y abandonar la semántica conjuntista, adoptando un enfoque alternativo. Las opciones aquí son variadas. La menos radical consiste en adoptar un enfoque substitucional. En líneas generales, en una semántica substitucional[21], las expresiones componentes de una fórmula cualquiera integrante de un lenguaje formal que no son términos lógicas son interpretadas a través de reemplazos por otras expresiones preservando la forma lógica de la expresión original. Así, las letras de predicado son interpretadas vinculándolas con predicados de un lenguaje previamente interpretado, y de esta manera, se intenta evitar que estos tengan extensiones. Y, en el caso de las constantes, se les asigna un objeto al que se aplica el predicado asociado con las variables. Las condiciones veritativas de las fórmulas en las que aparezcan cuantificadores, no presuponen que exista una *totalidad* de la cual extraer los valores de las asignaciones a las variables, sino simplemente casos de sustitución de las variables por expresiones del lenguaje interpretador. De esta manera, un *modo de interpretar* una fórmula es un reemplazo preservador de la forma lógica de expresiones componentes por otras expresiones apropiadas de otro lenguaje ya interpretado. En nuestro caso, por la necesidad de tener máxima capacidad expresiva, esos objetos que pueden ser valores de las variables serán los que caigan bajo el alcance de un predica-

---

[21] No obstante, Tarski y otros han dado buenas razones para pensar que este enfoque es insuficiente para reconstruir la totalidad de las interpretaciones necesarias para dar cuenta de las nociones de *validez universal* y de *consecuencia lógica.* El recurso a los resultados de Kreisel no parece ayudar demasiado, ya que su aplicación se circunscribe al primer orden. Cfr. Tarski, A. (1936), "On the Concept of Logical Consequence", p. 417, en Tarski, A. (1956). Para una discusión general del enfoque substitucional, cfr. H. Leblanc "Alternatives to standard first-order semantics", en D. Gabbay & F. Guenthner, *Handbook of Philosophical Logic*, Vol. 2.

do del lenguaje natural. El análisis de los modos de interpretar fórmulas de un lenguaje formal no se realiza por medio de TC sino del lenguaje natural. Con matices y distintos grados de adhesión, simpatizan con este enfoque, Alberto Moretti (Moretti inédito) y Max Freund (Freund inédito). Adolfo García de la Sienra y Mario Gómez Torrente han mostrado algunas dificultades al mismo (García de la Sienra, 2003 y Gómez Torrente 2003). Dar una respuesta al argumento de Williamson presentado unos párrafos más arriba, también constituye en obstáculo para este enfoque.

Un punto importante a destacar es que se supone la predicación, y de manera general, la interpretación, como algo *primitivo*, que funciona en TM como condición para la construcción misma de TM. No obstante, este recurso a un *modo de interpretación* carente de explicación, que recurre al uso de expresiones de nuestro lenguaje natural, más que ser una solución del problema, parece implicar una renuncia a dar una respuesta al mismo. Esto es así, ya que si queremos que TM sea exhaustiva, no debería quedar ninguna asignación sin una representación dentro de la misma. Pero, el recurso al lenguaje natural muestra que queda un modo de interpretación, aquel que en última instancia, asigna máxima generalidad a las expresiones de nuestra teoría, que está más allá de la teoría de la interpretación que estamos construyendo.

Ahora bien, más allá de esta limitación filosófica del recurso al lenguaje natural como herramienta dentro de una TM que se proponga como una TI, el propio Orayen señalo dos dificultades adicionales a esta opción (Orayen, 45-48). En primer lugar, advirtió que es difícil llegar a aceptar que podamos encontrar siempre predicados interpretadores del lenguaje natural que puedan dar los valores de verdad cuando son aplicados a objetos que están entre los valores de las variables. Moretti señala que "esta no parece ser una idealización más grave que las que son frecuentes en las aplicaciones de la matemática para entender fenómenos físicos (...) En particular, que la propia aplicación de TC para entender las propiedades referenciales de los lenguajes" (Moretti inédito, p. 7). Sin embargo, si dentro de las pretensiones que tenemos con TM incluimos la de ofrecer una teoría exhaustiva acerca de todos los modos de interpretar las teorías formales, cuando incluimos como parte de estas, teorías cuyo poder expresivo esté más allá del primer orden, el recurso al lenguaje formal no parece ser sólo una idealización. Parecen haber casos en los cuales directamente no es posible, para obtener una interpretación apropiada, la apelación al lenguaje natural. Una manera de ejemplificar esto último es prestar atención a los lenguajes plurales. Boolos (Boolos 1984) ha argumentado que la cuantificación plural permite la siguiente lectura plural del axioma de comprensión

> *Hay algunos conjuntos* tales que un conjunto es uno de ellos justo en el caso en el que no es miembro de sí mismo.

cuyo contenido no puede ser representado a través de una lectura en primer orden sin caer en contradicción. En tal representación es crucial el recurso al predicado plural del lenguaje natural "es uno de los". Dejemos a un lado la discusión acerca de si la cuantificación plural supone el compromiso con entidades plurales[22]. Mi punto tiene que ver con el recurso al lenguaje natural y si siempre hay correlatos en este último para interpretar lenguajes formales. Cuando tenemos un lenguaje de plurales que incluya predicados que sean verdaderos de pluralidades cuyos miembros son demasiado grandes como para ser un conjunto, la única manera de entender esos predicados parece ser el recurso a superpluralidades (pluralidades de pluralidades). Pero, a diferencia de lo que sucede con las pluralidades, no se puede construir una TI enriqueciendo TM con el lenguaje natural, ya que "no hay términos o cuantificadores superplurales en el inglés" (Rayo 2006, p. 227) (ni en ningún otro lenguaje natural)[23]. En la misma dirección, Linnebo (Linnebo 2006) analizando la posibilidad de encontrar interpretaciones para cuantificadores plurales de nivel superior, enfatiza las dificultades de encontrar un modo de interpretar los mismos (tal como Boolos hace con la cuantificación plural ordinaria) dentro del lenguaje natural.

Otro ejemplo, es el siguiente (Simmons 1993): consideremos el universo de los conjuntos de la interpretación con máxima generalidad en toda la jerarquía ZF. Para cada conjunto, hay un concepto distinto, digamos: *ser un miembro de este conjunto particular*. Pero dadas ciertas limitaciones acerca de los límites superiores acerca del tamaño de las expresiones y la longitud de las oraciones del lenguaje natural, estos conceptos superan la capacidad expresiva de cualquier lenguaje natural. Presento, entonces, tanto el recurso a los superplurales como a los conceptos de la jerarquía conjuntista del universo de ZF como casos donde el lenguaje natural podría tener medios insuficientes como para completar TM, y así posibilitar que esta se pudiera convertir en una TI aceptable. Claro, únicamente si se justificara la tesis según la cual el lenguaje natural es universal en el sentido según el cual todo tiene que poder ser expresado en el lenguaje natural, mi punto dejaría de tener sustento. Pero aunque no poseo un argumento ge-

---

[22] En Rayo, A. (2005), se conecta la propuesta de Boolos con el propósito de tomar a TM como TI, y así intentar una solución a la paradoja de Orayen: la tesis es que se logra, utilizando la cuantificación plural dentro de la TM, mayor capacidad expresiva, sin agregar compromisos ontológicos con más colecciones que las que ya teníamos al tener una TC. El recurso a la *cuantificación plural* es un elemento central en este enfoque. El aparato de Boolos, aplicado a la TM vista como una TI, nos da mayor riqueza ideológica sin tener que adquirir mayor riqueza ontológica. Sin embargo, el propio Rayo (2006) señala limitaciones a esta propuesta. También se pueden consultar C. Parsons, "The Problem of Absolute Universality", en A. Rayo & G. Uzquiano (2006), y O. Linnebo (2003), "Plural Quantification Exposed", *Noûs* 37.

[23] Es cierto que tal como objeta uno de los réferis de *Theoria,* la inexistencia de expresiones plurales de orden superior en el lenguaje natural podría arrojar dudas sobre la idea misma de que sea genuina la supuesta interpretación plural dentro de TM. Quizás como el mismo réferi señala puede recurrirse a una cuantificación sobre clases de algún tipo (hiperclases, etc.). En este caso, yo replico que, en primer lugar, no está claro que este recurso, tal como lo muestra el ejemplo que presento a continuación, permita recurrir al lenguaje natural. Y en segundo lugar, el punto de Boolos es que hay sentidos que pueden ser expresados utilizando cuantificación plural que no pueden ser expresados utilizando cuantificación sobre clases. Una manera de apreciar el punto es que dado que los únicos valores que se permiten para esas variables son de tamaño conjuntista, si los cuantificadores variaran sobre todo lo que hay, ese método está limitado por modelos cuyos dominios son de tamaño conjuntista. Y si el universo de las pluralidades es más grande que el de las entidades conjuntistas, no todo hablar acerca de pluralidades podría ser expresado en un hablar conjuntista.

neral en contra de la misma, considero que la adopción de tal posición es al menos discutible.

En síntesis, tanto el recurso a los plurales como el recurso a los conceptos determinados por el universo conjuntista de ZF, parecen servir para mostrar que no siempre el lenguaje natural puede tener tanto poder expresivo como para representar los modos de interpretar las teorías formales. Y en estos casos, tal como sostiene Orayen, no siempre podremos encontrar expresiones del lenguaje natural capaces de interpretar las de la teoría formal. Pero, tal cosa es necesaria, si el lenguaje natural se toma como fuente última de máxima expresabilidad[24].

En segundo lugar, Orayen señala que al afirmar que los valores de las variables son los objetos a los cuales se les aplica cierto predicado, puede estar presuponiéndose, de manera inapropiada, que existe *una totalidad dada de objetos* de la cual extraer algunos mediante ese predicado. Dice Orayen "Si este fuera el caso, quizás estaríamos suponiendo nuevamente algo así como el conjunto universal" (Orayen 2003, p. 49.). En este punto, Moretti dudando acerca del carácter insuperable de esta dificultad, se pregunta *por qué* no creer que existe una extensión universal (Moretti A. 2003 p. 8.). Y agrega que TC, por la paradoja de Russell, no puede aspirar a ser una teoría de todas las cosas, y por lo mismo, no puede prohibir tal existencia. Tal cosa es cierta, pero me parece que esta opción es demasiado arriesgada. Al comprometernos con tal extensión universal, con una pluralidad universal que no sea un conjunto, necesitamos que el lenguaje natural sea capaz de hablar acerca de ella. De otra manera, tal pluralidad sería innombrable. La clave, entonces, está en que el lenguaje natural tenga predicados máximamente expresivos capaces de hacer referencia a esas totalidades. Esto es, queda abierta la pregunta por la capacidad expresiva misma del lenguaje natural: ¿posee el lenguaje natural los recursos necesarios para hablar acerca de todas las colecciones que necesitamos en el desarrollo de los distintos sistemas lógicos y matemáticos? El punto es al menos controversial. Si los predicados y cuantificadores del lenguaje natural estuvieran siempre restringidos contextualmente, no tendríamos manera de hacer plausible tal exigencia. Glanzberg y Parsons, entre otros, defienden la tesis de que aparentes apariciones de cuantificadores irrestrictos en el lenguaje natural son siempre casos, de hecho, restrictos. Williamson, en cambio, argumenta que para entender la cuantificación restricta hay que presuponer el cuantificador irrestricto (Glanzberg 2004, Parsons 2004 y Williamson 2004). Y más allá de esa discusión sobre el carácter relativo de los cuantificadores del lenguaje natural está siempre la sospecha acerca de la necesidad de comprometernos con una extensión que esté más allá de lo que podemos hablar en nuestras mejores teorías de las colecciones. La opción es, al menos, osada. Y en todo caso, aún cuando se considere que es discutible tanto que no haya siempre correlatos expresivos en el lenguaje natural capaces de interpretar cualquier expresión de un lenguaje formal, como que no sea posible presuponer que el lenguaje natural tenga recur-

[24] Nótese que aún cuando no se acepte esta tesis, se puede continuar sosteniendo que el lenguaje natural es universal, el sentido según el cual en él hay recursos expresivos como para expresar su propia semántica. Por supuesto, como Tarski ha argumentado ese aspecto genera problemas.

sos como para hablar de una extensión universal, resta mi principal objeción a esta salida: el recurso al lenguaje natural como un medio primitivo de ofrecer interpretaciones *anula* la posibilidad de dar una explicación que no esté más allá de la teoría que estamos construyendo. En suma, una teoría que recurra a una interpretación que no se interprete.

REFERENCIAS

Belnap, N., & A. Gupta (1993), *The Revision Theory of Truth*. Cambridge: MIT Press.

Boolos, G. (1984). "To Be us to be a Value of a Variable (or to be Some Values of Some Variables)", *Journal of Philosophy* 81, 430-449. Reimpreso en G. Boolos (1998), *Logic, logic and Logic*, Cambridge: Cambridge U.P.

Cartwright, R. (1994). "Speaking of everything", *Noûs* 28, 1-20.

Freund, M. (inédito). "Irresolubilidad de la paradoja de Orayen", en García de la Sienra, A. (inédito).

García de la Sienra, A. (2003). "La paradoja de Orayen", en A. Moretti & G. Hurtado (2003), 71-81.

——— (inédito). *Filosofía de la lógica: reflexiones sobre la Paradoja de Orayen*. México: UNAM.

Gómez Torrente, M. (2003). "Notas sobre la Paradoja de Orayen", en A. Moretti & G. Hurtado (2003), 83-94.

——— (inédito). "Interpretaciones y conjuntos", en García de la Sienra, A. (inédito).

Hart, W. (2003). "Carta de William Hart", en A. Moretti & G. Hurtado (2003), 99-100.

Kripke, S. (1975). "Outline of the Theory of Truth", *J. of Phil.* 72, 690-710.

Kreisel, G. (1967). "Informal Rigour and Completeness Proofs", en I. Lakatos (1967), *Problems in the Philosophy of Mathematics*. Amsterdam: North Holland, 138-171.

Linnebo, O. (2006). "Sets, Properties, and Unrestricted Quantification", en A. Rayo & G. Uzquiano (2006), 152-178.

McGee, V. (1991). *Truth, vagueness, and paradox*. Indianápolis: Hackett.

Moretti, A. (inédito). "La interpretación de los lenguajes de primer orden", A. en García de la Sienra (inédito).

——— & G. Hurtado (2003). *La paradoja de Orayen*. Buenos Aires: EUDEBA.

Orayen, R. (2003). "Una paradoja en la semántica de la teoría de conjuntos", en A. Moretti & G. Hurtado (2003), *La paradoja de Orayen*. Buenos Aires: EUDEBA, 35-59.

Parsons, C. (1974). "Sets and classes", *Noûs* 8, 1-12.

——— (2006). "The Problem of Absolute Universality", en A. Rayo & G. Uzquiano (2006), 203-219.

Quine, W. (1963). *Set Theory and Its Logic*. Harvard: Harvard U.P.

Rayo, A. (2005). "Nota crítica sobre la paradoja de Orayen", *Crítica* 37, 99-115.

Rayo, A. (2006). "Beyond Plurals", en A. Rayo & G. Uzquiano (2006), 220-254.

——— & G. Uzquiano (2006). *Absolute Generality*. Oxford: Oxford U.P.

Simmons, K (2006). *Universality and The Liar*. Cambridge: Cambridge U.P.

Tarski, A. (1956). *Logic, Semantics and Metamathematics*. Oxford: Oxford University Press. 2.º ed. 1990.

Williamson, T. (2004). "Everything", *Philosophical Perspectives* 17, 415-465.

Eduardo Alejandro BARRIO es Doctor en Filosofía (UBA), Profesor de Lógica (UBA) e Investigador Adjunto del CONICET. Codirector del GAF (www.accionfilosofica.com) y de proyectos de investigación en 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, el concepto de consecuencia lógica, la teoría de modelos y los supuestos filosóficos de la lógica modal y los condicionales contrafácticos.

DIRECCIÓN: Instituto de Filosofía, Universidad de Buenos Aires. Puán 430, 4to piso. Ciudad Autónoma de Buenos Aires, Argentina. C.P. 1406. E-mail: ebarrio@fibertel.com.ar.