



## REVISTA DISCUSIONES FILOSÓFICAS

ISSN 0124-6127

- Fundada en enero de 2000 -

Periodicidad semestral

Tiraje 300 ejemplares

Año 13, No 21, 274 p.

julio - diciembre, 2012

Manizales - Colombia

### Rector

Universidad de Caldas

Ricardo Gómez Giraldo

Vicerrectora Académica

Luz Amalia Ríos Vásquez

Vicerrector de Investigaciones y Postgrados

Carlos Emilio García Duque

Vicerrector Administrativo

Fabio Hernando Arias Orozco

Vicerrectora de Proyección

Fanny Osorio Giraldo

### DISCUSIONES FILOSÓFICAS

Es una publicación del Departamento de Filosofía de la Universidad de Caldas. Sus temas de interés son la Filosofía (en todas sus clasificaciones), y la Literatura (mirada desde un punto de vista filosófico). Nuestro objetivo es difundir resultados de investigación, propiciar el debate crítico sobre las tesis planteadas en los artículos, y contribuir a la literatura y la crítica filosófica mediante reseñas y traducciones sobre los temas ya enunciados. La revista está dirigida a estudiantes, investigadores, profesores y profesionales en filosofía y literatura, así como a otros lectores que tengan afinidad por nuestros temas de interés.

Los artículos que aparecen en esta revista están en la página Web:

**[discusionesfilosoficas.ucaldas.edu.co](http://discusionesfilosoficas.ucaldas.edu.co)**.

La revista Discusiones Filosóficas está indexada en Publindex categoría A2,

Philosopher Index, Scientific Electronic Library On-line SciELO,

Ulrich's Periodicals Directory, Latindex, MLA Index, EBSCO.

### COMITÉ EDITORIAL

Director

*Carlos Emilio García Duque* (Universidad de Caldas)

\*\*\*

*Julián Serna Arango* (Universidad Tecnológica de Pereira)

*Francisco Cortés Rodas* (Universidad de Antioquia)

*Juan Manuel Jaramillo* (Universidad del Valle)

*Alejandro Patiño Arango* (Universidad de Caldas)

### COMITÉ CIENTÍFICO INTERNACIONAL

*Carlos Ulises Moulines* (Institut für Philosophie Logik und Wissenschaftstheorie.

Universidad de München - Alemania)

*Garrett Thomson* (College of Wooster - USA)

*Andrés Rivadulla Rodríguez* (Universidad Complutense de Madrid - España)

*Anandi Hattiangadi* (Saint Hilda's College Oxford - Inglaterra)

*Décio Krause* (Universidad Federal de Santa Catarina - Brasil)

*Luis Salvático* (Universidad Nacional de Córdoba - Argentina)

**COMITÉ TÉCNICO DE APOYO A LA EDICIÓN**

*Juan David Giraldo Márquez* - Coordinación

*Juan David López González* - Diagramación

*Carlos Eduardo Tavera Pinzón* - Soporte Técnico

*Carlos Fernando Nieto Betancur* - Implementación Metodología SciELO

*Carlos Emilio García Duque* - Traducción de los resúmenes al inglés

*Raúl Andrés Jaramillo Echavarría* - Asistente editorial

**VENTAS, SUSCRIPCIONES Y CANJES**

Vicerrectoría de Investigaciones y Postgrados

Universidad de Caldas - Sede Central

Calle 65 No 26 - 10

Apartado Aéreo: 275

Teléfonos: (+6) 8781500 ext. 11222

E-mails: [vinves@ucaldas.edu.co](mailto:vinves@ucaldas.edu.co)

[revistascientificas@ucaldas.edu.co](mailto:revistascientificas@ucaldas.edu.co)

[discufilo@ucaldas.edu.co](mailto:discufilo@ucaldas.edu.co)

Manizales - Colombia

**EDICIÓN**

Universidad de Caldas

Vicerrectoría de Investigaciones y Postgrados

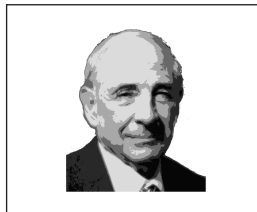


Imagen Portada:

Ilustración: John Searle

La responsabilidad de lo expresado en cada artículo es exclusiva del autor y no expresa ni compromete la posición de la revista.



# ph

## CONTENIDO

### CONTENTS

<b>PRESENTACIÓN EDITORIAL</b> <i>EDITORIAL</i>	<b>7</b>
<b>TRUTH-MAKING AND THE ALETHIC UNDECIDABILITY OF THE LIAR</b> <i>LOS HACEDORES DE VERDAD Y LA INDECIDIBILIDAD ALÉTICA DEL MENTIROSO</i> <i>Stephen Barker</i>	<b>13</b>
<b>SELLARS AND PRETENSE ON "TRUTH &amp; 'CORRESPONDENCE'"</b> <b>(WITH A DETOUR THROUGH MEANING ATTRIBUTION)</b> <i>SELLARS Y LA PRETENSÓN EN "VERDAD Y 'CORRESPONDENCIA'"</i> <i>(CON UN DESVÍO A TRAVÉS DE LA ATRIBUCIÓN DEL SIGNIFICADO)</i> <i>Bradley Armour-Garb</i> <i>James A. Woodbridge</i>	<b>33</b>
<b>BETTER TO BE THAN NOT TO BE?</b> <i>¿ES MEJOR SER QUE NO SER?</i> <i>Wlodek Rabinowicz</i> <i>Gustaf Arrhenius</i>	<b>65</b>
<b>NOMINALIST ANALYSES OF AN ENTITY BEING CHARACTERED</b> <i>ANÁLISIS NOMINALISTA DE UNA ENTIDAD QUE ESTÁ SIENDO CARACTERIZADA</i> <i>Michael Istvan</i>	<b>87</b>
<b>THE LIFE OF THE PHILOSOPHER: TESTIMONY OF PLUTARCH AND PORPHYRY</b> <i>LA VIDA DEL FILÓSOFO: TESTIMONIO DE PLUTARCO Y PORFIRIO</i> <i>Isha Gamlath</i>	<b>95</b>
<b>A CONCEPÇÃO DO VALOR: UMA ANÁLISE ACERCA DA OBJETIVIDADE DOS VALORES PROPOSTA POR P. GRICE</b> <i>THE CONCEPTION OF VALUE: AN ANALYSIS ABOUT THE OBJECTIVITY OF VALUES PROPOSED BY P. GRICE</i> <i>Lauren de Lacerda Nunes</i> <i>Gabriel Garmendia da Trindade</i>	<b>105</b>

<b>O MODELO ONTOLÓGICO ESTRATIFICADO NO NATURALISMO BIOLÓGICO DE JOHN SEARLE: UNA CONTROVÉRSIA COM JAEGWON KIM</b>	<b>119</b>
<i>ONTOLOGICAL MODEL STRATIFIED IN THE NATURALISM BIOLOGICAL OF JOHN SEARLE: A DISPUTE WITH JAEGWON KIM</i>	
<i>Tárik de Athayde Prata</i>	
<b>LA REALIDAD BÁSICA Y LA REALIDAD HUMANA: CAPÍTULO INTRODUCTORIO PARA EL VOLUMEN DE MÜNSTER</b>	<b>139</b>
<i>THE BASIC REALITY AND THE HUMAN REALITY: INTRODUCTORY CHAPTER TO THE MÜNSTER'S VOLUME</i>	
<i>John R. Searle</i>	
<b>LA FILOSOFÍA DE LA HISTORIA DE LA CIENCIA DE THOMAS KUHN</b>	<b>167</b>
<i>THE PHILOSOPHY OF HISTORY OF SCIENCE OF THOMAS KUHN</i>	
<i>Alexander Bird</i>	
<b>LOS PROBLEMAS DEL SENTIDO-REFERENCIA EN LA SEMÁNTICA FILOSÓFICA CLÁSICA: DOS GRANDES CONCEPCIONES EN LAS TEORÍAS REFERENCIALISTAS DEL SIGNIFICADO</b>	<b>187</b>
<i>THE PROBLEMS OF SENSE-REFERENCE IN THE CLASSIC PHILOSOPHICAL SEMANTICS: TWO GREAT CONCEPTIONS IN REFERENTIALIST THEORIES OF MEANING</i>	
<i>Juan Manuel Jaramillo Uribe</i>	
<b>PROPUESTAS Y DIFERENCIAS PRAGMÁTICAS EN TORNO DEL LENGUAJE COMO INSTITUCIÓN: WITTGENSTEIN Y HABERMAS</b>	<b>207</b>
<i>PRAGMATIC PROPOSALS AND DIFFERENCES ON LANGUAGE AS INSTITUTION: WITTGENSTEIN AND HABERMAS</i>	
<i>Javier Alegre</i>	
<b>EL CAMPO DE LA ARGUMENTACIÓN ESENCIALMENTE POLÉMICA</b>	<b>225</b>
<i>THE FIELD OF ESSENTIALLY CONTESTED ARGUMENT</i>	
<i>Julder A. Gómez</i>	
<b>OBJETO, TIEMPO Y COLECTIVIDAD EN EL OTOÑO DEL PATRIARCA</b>	<b>245</b>
<i>OBJECT, TIME AND COLECTIVITY IN THE AUTUMN OF THE PATRIARCH</i>	
<i>Esteban Quesada</i>	
<b>COLABORADORES</b>	<b>264</b>
<i>COLLABORATORS</i>	

## PRESENTACIÓN EDITORIAL

EDITORIAL

Saludamos a nuestros lectores una vez más. El fascículo 21 de la revista *Discusiones Filosóficas* presenta, en esta ocasión, a la comunidad filosófica nacional y a nuestros lectores en el exterior 12 artículos de investigación y la traducción de un texto del conocido filósofo contemporáneo John Searle, quien amablemente ha autorizado su publicación en estas páginas. La presente entrega comienza con un artículo de Stephen Barker, titulado "Truth-Making and the Alethic Undecidability of the liar". En este trabajo, Barker defiende una novedosa solución a las paradojas semánticas, tema sobre el cual ya hemos publicado varias contribuciones en nuestros números anteriores. Barker intenta demostrar que a partir de una adecuada comprensión de los últimos hacedores de verdad o falsedad de los enunciados, es posible llegar a la conclusión de que oraciones como el mentiroso son alécticamente indecidibles. La indecidibilidad implica que no es posible, en principio, hacer ninguna afirmación sobre el valor de verdad de esas oraciones, ni sobre su estatus veritativo. El autor sugiere que por esta vía se pueden solucionar las paradojas semánticas sin tener que lidiar con los problemas de venganza, al tiempo que se conservan la lógica clásica y la validez del esquema-T.

A continuación se incluye el artículo "Sellars and Pretense on Truth & 'Correspondence' (with a detour through meaning attribution)" de Bradley Armour-Garb y James A. Woodbridge. Los autores parten de la aparente tensión interna de la explicación de la verdad que ofrece Sellars en "Truth & 'Correspondence'". Tal tensión podría resolverse, según los autores, tomando en cuenta la explicación de la atribución de significado que el mismo Sellars da en "Meaning as Functional Classification", la cual contiene los elementos requeridos para dicha tarea. Armour-Garb y Woodbridge muestran que tanto la tensión interna en la comprensión sellarsiana de la verdad, como la externa en su explicación de la atribución de significado se pueden resolver sin perder de vista el espíritu sellarsiano, apelando a las explicaciones *ficcionalistas* de habla sobre la verdad y habla proposicional (incluida la atribución de significado) desarrolladas por ellos mismos.

El fascículo continúa con un trabajo de Wlodek Rabinowicz y Gustaf Arrhenius titulado "Better to be than not to be?" En este trabajo Rabinowicz y Arrhenius defienden una respuesta afirmativa a esta clásica pregunta existencial. En particular, los autores se ocupan de examinar si es posible plantear la cuestión de si resulta mejor o peor existir para un individuo, sin que dicho planteamiento

nos haga desembocar en el absurdo. La solución de Rabinowicz y Arrhenius parte de la superioridad ontológica de la existencia sobre la no existencia, algo respecto a lo cual no nos permiten dudar nuestras intuiciones más familiares. Enseguida, hemos incluido el trabajo “Nominalist Analyses of an Entity Being Charactered” de Michael Istvan. El autor procura ofrecer un marco de referencia útil para comprender la disputa clásica entre el realismo y el nominalismo sobre el problema de los universales. Con la intención de esclarecer mejor la situación, Istvan proporciona un catálogo exhaustivo de los análisis básicos de una entidad que está siendo caracterizada y que los nominalistas pueden emplear tanto en una ontología no-constitutiva, como en una constitutiva. La serie de artículos en inglés de este fascículo culmina con una colaboración de Isha Gamlath, quien esta vez nos ofrece el trabajo “The Life of the Philosopher: Testimony of Plutarch and Porphyry”. Gamlath examina el rol y la función del filósofo cuyo modo de vida negocia la liberación del rango de los sentidos en los testimonios de Plutarco y Porfirio. Tal exploración se ocupa de investigar las particularidades de su vida ordinaria y el sentido de un sacrificio cuya culminación es la re-evaluación de su demarcación de las masas.

La presente entrega continúa con un artículo de Lauren de Lacerda Nunes y Gabriel Garmendia da Trindade titulado “A concepção do valor: Uma análise acerca da objetividade dos valores proposta por P. Grice”. En este trabajo los autores analizan ciertos temas del capítulo inicial del libro de Grice *Conception of value*. A través de este análisis, ellos intentan adelantar una discusión metaética sobre la posibilidad de una perspectiva objetiva o una escéptica acerca de los valores morales. Los temas más importantes que se discuten en este artículo son: la lectura que hace Grice sobre la teoría del error de J. Mackie, la contribución de R. M. Hare a la discusión de la teoría del error de J. Mackie; las críticas de Grice al enfoque de R. M. Hare y el análisis de Grice al anti-objetivismo radical de J. Mackie en tanto opuesto a su propia visión objetivista. Los autores presentan, a manera de conclusión, sus desacuerdos con el tratamiento que hace Grice de la teoría del error de J. Mackie, que califica los valores morales como falsos cuando el propio Mackie los habría tratado metodológicamente como “sinsentidos”. De este modo, el escepticismo respecto a los valores que propone Mackie resulta descalificado cuando enfrenta la posibilidad de un objetivismo de valores tal como lo propone Grice.

A continuación hemos incluido el artículo “O modelo ontológico estratificado no naturalismo biológico de John Searle: Uma controvérsia com Jaegwon Kim” de Tárík de Athayde Prata. En este trabajo se examina la controversia entre Jaegwon Kim y John Searle sobre la manera como Searle usa el modelo ontológico estratificado en su filosofía de la mente. Luego del contexto necesario, que incluye una exposición general del modelo ontológico estratificado, se discuten las respuestas de Searle a algunas de las críticas que Kim le hace al naturalismo biológico. Según el autor, Searle termina por sucumbir ante un dualismo de propiedades, pese a usar un modelo estratificado e intentar la construcción de un fisicalismo no reduccionista.



Entre esta sección y la siguiente hemos intercalado la traducción del capítulo introductorio al volumen de Muenster, de John Searle, según la cuidadosa versión castellana de Laura Giraldo, docente del departamento de Filosofía de la Universidad Minuto de Dios, a quien agradecemos desde estas páginas por su contribución y por gestionar el permiso de Searle para publicar su texto en nuestra revista. Este capítulo constituye una verdadera primicia para nuestra comunidad filosófica, pues el libro no ha sido vertido a nuestra lengua aún. La serie de artículos se reinicia con el trabajo “La filosofía de la historia de la ciencia de Thomas Kuhn” del reconocido especialista Alexander Bird, gracias a la traducción (autorizada por el propio Bird) del filósofo colombiano Juan Carlos Aguirre. En este artículo Bird sostiene que Kuhn fue un historicista bifronte: por un lado, fue un conservador en el sentido de Mannheim –con su exaltación de la tradición para comprender el cambio científico y la creencia de que el contexto histórico es vital para la evaluación de una idea científica–; y por otro lado fue un defensor del determinismo –acepta la existencia de un modelo de cambio científico, similar a las leyes del desarrollo científico–. Bird considera que esta suscripción del determinismo le impone a Kuhn un internalismo sobre las causas del cambio científico; y que dicho internalismo contrasta con el externalismo típico de la mayor parte de los estudios de la ciencia post-kuhnianos. El artículo se cierra con una explicación sobre el vínculo entre el historicismo de Kuhn con los propósitos filosóficos de la historia de la ciencia kuhniana.

A continuación, hemos incluido un artículo del filósofo colombiano Juan Manuel Jaramillo, quien se ocupa esta vez de “Los problemas del sentido-referencia en la semántica filosófica clásica: dos grandes concepciones en las teorías referencialistas del significado”. Jaramillo ofrece un importante contexto que arranca desde la introducción del neologismo “*ontoepistemosemántica*” por C. U. Moulines para dar cuenta de los estrechos vínculos que hay entre las reflexiones ontológicas, epistemológicas y semánticas. Se aclara que tal relación ya había sido formulada y reconocida por Frege, cuyo interés por la semántica definió la naturaleza de sus principales contribuciones a la filosofía. Jaramillo muestra y hace el contraste entre dos grandes teorizaciones de la reflexión ontoepistemosemántica inicial sobre el significado-referencia de los términos singulares: la teoría de la *referencia indirecta* de Frege y la teoría de la *referencia directa* de Russell, al menos para lo que este último denomina “*nombres lógicamente propios*”. El artículo demuestra la singular importancia de estas reflexiones pioneras para cualquier investigación sobre temas óntico-ontológicos, epistemológicos y semánticos.

Tras esta incursión en los fundamentos de la filosofía estructuralista, hemos incluido un artículo de Javier Alegre titulado “Propuestas y diferencias pragmáticas en torno del lenguaje como institución: Wittgenstein y Habermas”. Alegre analiza las concepciones *institucionalistas* del lenguaje de cuño pragmático que se encuentran en la filosofía contemporánea, dedicando especial atención a las contribuciones de Wittgenstein y a la reapropiación y las propuestas de desarrollo que hizo Habermas. El autor precisa los núcleos centrales del *giro*

*pragmático*, para estudiar, con cierto detalle, los ejes conceptuales e implicancias de la teoría de los *juegos de lenguaje* del segundo Wittgenstein y su recepción y modificación en el marco de la *teoría de la acción comunicativa* habermasiana. Mediante esta estrategia busca analizar los lineamientos y consecuencias que se desprenden de ambas teorías y, en particular, trata de establecer si la reapropiación de Habermas permite retomar la potencia heurística y teórica propia de la concepción pragmática institucionalista del lenguaje que se halla en Wittgenstein.

A continuación hemos incluido el artículo de Julder A. Gómez “El campo de la argumentación esencialmente polémica”. El profesor Gómez nos presenta los resultados de un proyecto de investigación, uno de cuyos objetivos es determinar una clase de desacuerdos comprensible como un campo de argumentación que no se produce porque haya malentendidos, sino por la naturaleza de algunos fenómenos sociales. El autor se aproxima a dicho campo a través de la noción de concepto esencialmente polémico de Gallie y plantea que este fenómeno se puede comprender satisfactoriamente, en la perspectiva de la retórica clásica, como un campo constituido por una clase de cuestiones de definición. Sus planteamientos están organizados en dos partes, la primera se ocupa de la noción de concepto esencialmente polémico de Gallie e incluye algunas notas sobre la recepción de su propuesta; la segunda expone la noción de campo de argumentación, el debate en torno al modo en que los campos de argumentación pueden ser comprendidos y la cuestión de cómo entender el campo de argumentación en relación a los conceptos esencialmente polémicos.

Para finalizar, publicamos un artículo sobre literatura desde una mirada filosófica. En esta ocasión hemos escogido el artículo “Objeto, tiempo y colectividad en *El Otoño del Patriarca*” de Esteban Quesada. El autor describe la construcción de esta novela de nuestro premio nobel de literatura a través de la aplicación de una “arqueología objetual” que retoma elementos del análisis noético-noemático de la fenomenología eidética y que se aplica a la lectura de la novela. Según el autor, este tipo de análisis revela que en esta novela del boom de la literatura latinoamericana la presencia de un complejo de relaciones entre los conceptos de eternidad, historia y muerte, que son resultado de la manera como la colectividad se enfrenta a la reconstrucción de los espacios y de los objetos que permanecieron a través de la dictadura, particularmente, del palacio presidencial y del cuerpo muerto del patriarca. Quesada muestra como esa “reconstrucción de la realidad” plantea un problema fenomenológico y una posibilidad política para los pueblos de América Latina.

Como siempre, saludamos la excelente acogida que ha tenido nuestra revista y reiteramos la invitación a los potenciales autores para que nos envíen sus manuscritos y a nuestros amables lectores a compartir con nosotros sus inquietudes a través del correo de la revista.

*Carlos Emilio García Duque*  
*Universidad de Caldas*

ph

DISCUSIONES FILOSÓFICAS



# TRUTH-MAKING AND THE ALETHIC UNDECIDABILITY OF THE LIAR

*LOS HACEDORES DE VERDAD Y LA INDECIDIBILIDAD ALÉTICA DEL MENTIROSO*

STEPHEN BARKER

University of Nottingham, United Kingdom. [stephen.barker@nottingham.ac.uk](mailto:stephen.barker@nottingham.ac.uk)

RECIBIDO EL 28 DE SEPTIEMBRE DE 2012 Y APROBADO EL 23 DE NOVIEMBRE DE 2012

## RESUMEN    ABSTRACT

En el presente artículo, sostengo que es posible una nueva solución a las paradojas semánticas, basada en los hacedores de verdad. Demuestro que, con base en una comprensión adecuada de cuáles son los últimos hacedores de verdad o falsedad de las oraciones, se puede demostrar que oraciones como el mentiroso son aléticamente indecidibles. Esto significa que no se puede decir en principio si tales enunciados son verdaderos, no verdaderos, falsos, no falsos, ni verdaderos ni falsos, verdaderos y falsos, y así sucesivamente. Sostengo que esto conduce a una solución de las paradojas semánticas que parece estar libre de los problemas de venganza, nos permite conservar la lógica clásica y la validez del esquema-T.

I argue that a new solution to the semantic paradoxes is possible based on truth-making. I show that with an appropriate understanding of what the ultimate truth and falsity makers of sentences are, it can be demonstrated that sentences like the liar are alethically undecidable. That means it cannot be said in principle whether such sentences are true, not true, false, not-false, neither true nor false, both true and false, and so on. I argue that this leads to a solution to the semantic paradoxes that appears to be free of revenge problems, allows us to maintain classical logic and the validity of the T-schema.

## PALABRAS CLAVE    KEY WORDS

indecidibilidad alética, indeterminación, el mentiroso, venganza, verdad, condiciones de verdad, verdad de las decisiones, esquema-T.

alethic undecidability, indeterminacy, Liar, revenge, truth, truth-conditions, truth-making, T-schema.

This paper sketches a new solution to the Liar and other semantic paradoxes based on the phenomenon of truth-making. According to the solution sentences like the Liar are undecidable with respect to all their alethic properties. That means it's undecidable—it cannot be said in principle—whether such sentences are true, false, either true or false, neither true nor false, both true and false, and so on. This *alethic undecidability* is not due to lack of information or some verification transcendent fact concerning such sentences. Rather, as I shall argue, it's an objective fact about these sentences that arises from their meanings, their referential properties, and the nature of truth itself. The aspect of the nature of truth that concerns us is that the facts of truth (and falsity) are *dependent facts*, that is, that a sentence *S* is true, or that *S* is false, etc, is a fact that is made the case by reality, and so is never a basic fact. In other words, truths are made. I will argue that truth's dependency in this sense, given facts about the referential properties of paradoxical sentences, entails that such sentences are alethically undecidable, and that this is the basis of a solution to the semantic paradoxes. Call this the *TM-solution* (truth-making solution). The TM-solution allows us to preserve classical logic, or at least involves no revision of logic as such and places no limitation on the expressive power of a language. What it entails is that we rethink the theoretical role and nature of truth-conditions, in relation to linguistic-meaning and reference, inference, and validity.

The paper proceeds by articulating the basic truth-making principle, *TM*, and the concept of alethic undecidability §(I). In §II, we look at the class of *ground-unspecifiable* sentences, which are sentences that exhibit certain looping or infinitely descending referential chains. This class includes, but is not exhausted by, paradoxical sentences. I show that by appealing to ground-unspecifiable sentences, an apparent refutation of *TM* becomes available. In §III, I show that we can rebut the refutation of *TM* by recognition of a specific idea of what truth-conditions are, given *TM*. I show in detail how this works in §IV. Finally in §5, we assemble the results accumulated in the course of paper, which deliver a solution to the semantic paradoxes.

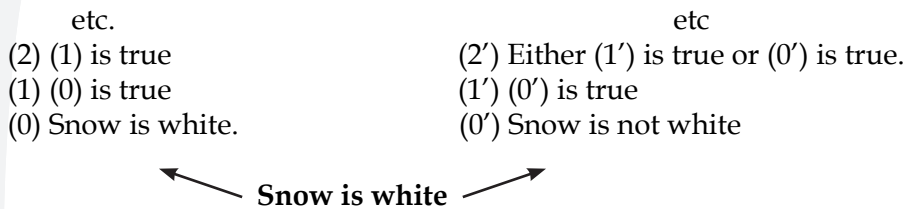
## I

### Truth-making and alethic undecidability

The first two concepts that we need to be acquainted with, and which play a crucial role in what follows, are truth-making and alethic undecidability. Here is the core contention regarding truth-making.

Truths (and falsities) are made true (made false) by how things are. Truth and falsity is dependent. Facts of truth are never basic facts. So, a sentence *Snow is white* is true in virtue of the fact that snow is white. Where *S* is *snow is white*, the fact that *S is true* is true, ultimately, in virtue of the fact that snow is white. Likewise, the fact that *Snow is pink* is false holds in virtue of the fact that snow is white. The fact that snow is white grounds a vast hierarchy of sentences.

Here is a fragment of that hierarchy of sentences, all grounded in that fact about snow:



There is no limit on the logical complexity of the sentences in the hierarchy. The truth and falsity of all the sentences in this infinite hierarchy are grounded, in the sense that they have truth-makers or false-makers.<sup>1</sup>

We now look at an important distinction regarding *grounding-facts*. In the hierarchy above the grounding-fact is a *non-alethic fact*, viz., a fact that has nothing to do with truth and falsity. The fact that snow is white is not a fact about a proposition's being true or false. We can allow that non-alethic reality includes facts of meaning and reference, and use, facts about speech acts, and so on, that is, linguistic and semantic facts excluding facts of truth and falsity. Now consider all non-alethic facts, in this sense, in one vast super fact. Then, a vast hierarchy of sentence truth and falsity will be determined by that fact.

There are questions, with respect to the metaphysics of truth-making and this great hierarchy. For example should non-alethic fact include negative facts, like snow's not being black, or totality facts, such as that all the objects in a certain position are white, and so on? Or indeed, should we think of the entities that make sentences true/false as facts, or indeed, as

<sup>1</sup> I assume in what follows that both sentences (given that they express thoughts) and thoughts themselves, or propositions, can be truth-bearers.

entities at all? I am sweeping all these issues aside here, as more or less, orthogonal to our concerns in this paper. (See Barker and Jago (2012) for ideas about what can be said regarding some of these issues.) In what follows, I assume for convenience, no logical restrictions on the kinds the facts there are, though nothing hangs on this for my purposes below.

By *grounding-fact* we mean any fact that can be amongst the ultimate truth/falsity-makers for some given sentence. We have seen that grounding-facts must include non-alethic facts, that is, facts that have nothing to do with truth and falsity of sentences or propositions, like the fact that snow is white. The question now is: do grounding facts include some alethic facts? The answer is *yes*. Some alethic facts are grounding facts, though not all; the vast majority are not – they are non-basic. Consider sentences about truth-making, like:

$S^T$ : The sentence *Snow is white* is true in virtue of the fact that snow is white.

What makes  $S^T$  true is not just the non-alethic fact that snow is white, though this plays a role. What  $S^T$  describes is a relation between a fact of truth and the fact that snow is white. That relation holds because of the nature of truth itself. It's a necessary fact about truth that the truth/falsity of sentences/propositions are dependent. This is the fact that truth (and falsity) is *inherently dependent*. The complete truth-maker for  $S^T$  is then:

- (a) Snow is white.
- (b) Truth is inherently dependent.

$S^T$ 's truth-maker is a mix of non-alethic fact, and the metaphysically necessary fact of truth's dependency. The fact of truth's dependency can itself function, all by itself, as a truth-maker. For example, the truth that all truths are dependent is itself made true by this fact about the nature of truth. Search in vain and you will not find the truth-maker of this sentence amongst the non-alethic facts.

Given truth is inherently dependent, what makes sentence true/false will be non-alethic fact or the fact of truth's dependency. The alethic facts that are not in the class of ultimate grounding facts are the facts that simply correspond to the facts of truth and falsity. These are facts like: *The fact that it's true that snow is white*, or *The fact that it's true that it's true that it's true that snow is white*, and so on. All these sorts of alethic fact are non-basic. They are not grounding-facts, but rather, grounded facts.



We can now sum up the view about truth (and falsity) being proposed:

*TM*: If a sentence *S* is true/false, then it's true/false, ultimately, in virtue of non-alethic fact or the fact that truth/falsity is inherently dependent.

People may take *TM* to be a heavy duty metaphysical thesis. I don't think it is. It's just that metaphysicians make heavy weather of it. However, as already indicated, I am not concerned with these storms here. In particular, I leave aside issues as to whether there really are negative facts or totality facts, or facts at all. All we need is the bare idea that truth (falsity) is dependent.

Nevertheless, there is another issue about how to understand *TM*. How are we to think about the truth-making relation? I take truth/falsity-making to involve a *grounding* or *in-virtue-of* relation between a proposition's truth or falsity and facts. Grounding is an asymmetric determinative relation – Rodriguez-Pereyra (2005). The form of a truth-making fact, then is this, where *X* is some fact and *S is true* is the condition or state of affairs that *S* is true:

$$X \implies S \text{ is true.}$$

This construal of truth-making assimilates truth-making to a more general phenomenon of grounding. For example, falsity is in as much need of grounding as is truth. So facts of falsity are grounded, which means, there are facts like this:

$$X \implies S \text{ is false.}$$

This way of thinking about truth/falsity-making makes sense. Facts of truth and falsity are grounded in the same way that facts of personal identity, or facts mental life, and so on, are grounded.<sup>2</sup> Grounding is a relation between facts. Again, what follows does not depend precisely on this way of understanding truth-making, but it's a convenient way of thinking of it for our purposes.

<sup>2</sup>See Barker (2012) for development of these ideas. This way of understanding things is at odds with the standard idea that truth-making is a primitive relation, *truthmaking*, which links a fact and proposition (see Armstrong 2005.) This approach ignores false-making. We could develop the argument below using this understanding, but it's more complicated.

Finally, a quick note on facts: I take facts to be the grounders of truth and falsity. Again, this will disturb some metaphysicians. I take facts to be *obtaining conditions*, where a *condition* is a state (*possibillum*) that may or may not obtain. A certain realism about facts and states of affairs makes the formulation of arguments much simpler.

### Alethic undecidability

We now look at the thesis of *alethic undecidability*. Consider the Liar, *L*, which says of itself that it's not true. According to the TM-solution, *L* is alethically undecidable. That means we cannot say in principle that *L* is true. We cannot say, in principle, that *L* is false. Nor can we affirm that *L* is bivalent, viz., that it is either true or false. However, our inability to affirm bivalence is not the result of our affirming a truth-value gap. On the contrary, according to the TM-solution, we cannot affirm that *L* lacks truth-value, that is, that it is neither true nor false. So, the indeterminacy we are proposing is not the familiar truth-value gap indeterminacy, proposed by Kripke (1975), Goldstein (1999), Field (2007), Maudlin (2007). Since we cannot affirm that *L* is true and cannot affirm it's false, we cannot affirm that it is both true and false, that is, we cannot affirm that *L* suffers a truth-value *glut*. So, Priest's (1987) dialethic position cannot be affirmed.

Let's try and fix the ideas in play here a bit more precisely. Let an *alethic predicate* be any one-place open-sentence of the form (...*x* ...*true*...) or (...*x* ...*false*...), where the other constituents of the open-sentence are standard logical operators: '&', '∃*x*' and '¬', etc. So, alethic predicates include, *x is true*, *x is false*, ( $\neg\neg x$  is true), and so on. Alethic predicates do not include certain intentional verbs like *believe* or *say*. So, (*believes of x that x is true*), (*x is said to be true*), or (*x cannot be said to be true*) are not alethic predicates. Call an *alethic property* any property corresponding to an alethic predicate. We shall be relaxed about the metaphysics of properties in what follows and allow that each meaningful predicate corresponds to a property. Again, nothing hangs on this except convenience of exposition.

That's the concept of alethic property. Before we proceed, we need to know more about the sense of *cannot-say* being used here. Let assertion be a speech-act whose only norm is to represent how things are. That's not an implausible idea about assertion. Perhaps there are other norms, like knowledge, or certainty, etc, but I am not concerned with those

here, nor are they required for my argument. Let's suppose we have a body of information *I* that is perfectly informative about non-alethic reality and inherent dependency of truth, and about the reference and linguistic meaning of a sentence *S*. There is no epistemic constraint on *I* whatsoever. Let *P* be any alethic property, such as being true, being false, being either true or false, being neither true or false, and so on. What I am proposing is the following conception of alethic undecidability:

*Alethic undecidability:* An ideally functioning assessor, furnished with perfect information *I*, cannot assert that *S* has *P*, for whatever alethic property *P*.

This conception of alethic undecidability does not involve accepting some kind of anti-realism about alethic reality. There are no epistemic constraints on *I*. It does, however, involve accepting that alethic properties of a sentence *S* are dependent on *S*'s meaning and reference, and facts about the world. That's uncontroversial. For surely, a given sentence's truth does depend on what it means, and the facts, out there.

Another, related, way of understanding alethic undecidability is this. A perfect representation of reality will have a gap in it with respect to the alethic status of *L*. It will lack the representations: *L is true*, *L is false*, *L is either true or false*, *L is neither true nor false*, *L is both true and false*, *It's not the case that L is not true*, and so on. This is not because our language cannot express something about its truth-predicate. We assume it's as powerful as a language can be. We also assume that the perfect presentation reflects perfectly non-alethic reality. There is no epistemic restriction or verification transcendent fact. In short, the ideal representation has a *representation gap*, but there is no truth-value gap. The representation gap is the fact that in the ideal representation there is a gap with respect to the alethic status of *L*. In contrast, for the truth-value gap approach, there is no representation gap, since the ideal representation will represent *L* in these terms: *L is neither true nor false*, which is the representation of an alethic property. That's not present in the ideal representation, according to the TM-solution.

The TM-approach is not the proposal of an intuitionist logic, in which, say, we might affirm that  $\neg\neg(L \text{ is true})$  but not be able to infer *S is true*. To affirm  $\neg\neg(L \text{ is true})$  is to fill the representation gap about *L*. It's to represent *L* lacking the property of lacking truth. But given alethic undecidability, the ideal representation will lack this representation as well.

Below, phrases like *One cannot say that S is true*, and so forth, will be used to express this fact about an ideal assertor, given information *I*. In saying that it's undecidable what *S*'s alethic properties are we mean just this: we cannot say what its alethic properties are. Given that we cannot affirm (in principle) a gap in the case of the Liar, *L*, the familiar revenge problem for the truth-value gap approach cannot arise for the TM-approach. We cannot say that *L* is neither true nor false, and so, we cannot say that *L* is not true. Whereas, on the truth-value gap approach, we must say that *L* is not true (as a commentary claim about *L*.) Therefore, we are committed to saying, by the T-schema, that *L* is true, after all. However, we cannot say this on the TM-approach.

You might object that there will be other forms of revenge that can threaten this position. And indeed there are forms of apparent revenge that we must consider below. Take the predicate: *cannot be said to be true*. This is not an alethic predicate so doesn't correspond to an alethic property. However, there is a seeming revenge using it: a sentence that says of itself that it cannot be said to be true. We examine this case below in §V.

How then does *TM* generate this alethic undecidability? To answer that we need to move to our next topic: the class of sentences that exhibit what I call *ground-unspecifiability*.

## II

### Ground-unspecifiability and the refutation of TM

My case for the alethic undecidability of the liar is based on rejection of an attempted refutation of *TM*. I outline this argument below. The refutation involves appeal to the properties of a specific class of sentences, which I call *ground-unspecifiable sentences* (for reasons to that will become clear below.) Semantically paradoxical sentences are members of this class of sentences. I now explain the property of being ground-unspecifiable, and then display the refutation of *TM* that's based on reasoning about this class of sentences.

Consider the sentences below. *L* is a semantically paradoxical sentence: a Liar. The other sentences, *Infinite Descent* and *T*, are *truth-tellers*. These sentences are not considered paradoxical. Nevertheless, they are in the same class that I am calling *ground-unspecifiable sentences*:

*Infinite Descent:* (0) (1) is true  
 ↓  
 (1) (2) is true  
 ↓  
 (2) (3) is true  
 ↓  
 (3) (4) is true  
 ↓  
 (4) (5) is true...  
 ↓  
 etc

Looping: ↓  
 T : T is true.  
 ↓  
 L : L is not true.  
 ↓

These sentences are perfectly meaningful in the sense that they are grammatical, have meaningful predicates, and suffer no apparent reference failure. What each exhibits is a specific loopy/infinately descending referential chain. We might put it this way: if you follow the reference of these sentences around, or down, to determine what they are about, then, you always encounter a specification of a condition about the truth, lack of truth, falsity, or lack of falsity of a sentence, and so on.

Let's refine this idea. Take *Infinite Descent*. The sentence (0) says (1) is true. (1) says (2) is true. (2) says (3) is true, and so on. Each sentence in the list simply says that the sentence below it is true. At no point do we find a sentence that says something about non-alethic reality – like *snow is white*, or *grass is purple* – or about the nature of truth, such as it's dependent. In short, no sentence in the list specifies a *grounding-condition* of the kind required by *TM* for truth or falsity, even though there is no reference failure or meaningless predicate. Similar comments apply to *T* and *L*. In sum: The linguistic and referential facts about these sentences don't determine a specification of a grounding-condition, whose obtaining or non-obtaining could be said to make them true or false. That's the property of being *ground-unspecifiable*.

Being *ground-unspecifiable* is about not determining any *specification* of a possible grounding-fact. Being *ungrounded* is the distinct property of actually lacking a grounding-fact: there is no grounding fact that makes the sentence true or false. The main premise in the argument threatening *TM* links these two properties thus:

*Kripke's Thesis:* If a sentence is ground-unspecifiable, it's ungrounded.

One can discern *Kripke's Thesis* in Kripke's (1975) theory, hence the name of the principle. I won't justify that interpretation here.

The refutation of *TM* that I mentioned above as threatening *TM*, has this form – here *S* is a ground-unspecifiable sentence:

### *TM-Refute*

1. Suppose *TM*.
2. *S* is ungrounded. [By *Kripke's Thesis*]
3. *S* is not true. [By *TM* from 2.]
4. Sentence *step 3* above is true. [By the *T*-schema, from 3.]
5. *Step 3* is ground-unspecifiable and so ungrounded. [By *Kripke's Thesis*.]
6. *Step 3* is not true. [By *TM* and 5.]
7. *Step 3* is true and not true. [From 4 and 6.]
8. *TM* is false. [Reductio 1-7].

Note that the sentence *step 3* is a ground-unspecifiable sentence because it's a truth-claim about a ground-unspecifiable sentence *S*.

To conclude: *TM* looks like it's refuted, given *Kripke's Thesis*. Some might say: so much the worse for *TM*! But I think *TM* is a correct principle about truth. I want to demonstrate how we can save it. My way of doing so is by questioning *Kripke's Thesis*. Observe, however, that this might, at first, look like a hopeless strategy.

If we deny *Kripke's Thesis*, we must find a counterexample to it: a ground-unspecifiable sentence *S* (one like *Infinite Descent*, *T* or *L* above) that, nevertheless, has a grounding-fact, that is, a truth-maker of the kind required by *TM*. But how could we find any such thing? Take (0) of *Infinite Descent*. None of the sentences beneath it, upon whose truth it would depend if it were true, describes any grounding-condition, like *snow is white*. So, how can we link (0), in the end, to a grounding-fact as its truth/false-maker? We cannot. The moral seems to be general: No ground-unspecifiable sentence *S* can be associated with a grounding-fact as its truth/falsity-maker. So it seems we cannot deny *Kripke's Thesis*.

That might suggest my strategy for saving *TM*—questioning *Kripke's Thesis*—is hopeless. But not so. It's here that the twist of this paper comes in. My position isn't that we should deny *Kripke's Thesis*. Rather, it's that *Kripke's Thesis* is alethically undecidable: it's essentially undecidable whether it's true, false, either true or false, and so on. If *Kripke's Thesis* is alethically undecidable, we cannot use it in *TM-Refute*, since we cannot affirm it.

The challenge for this solution is to find an independent reason for thinking that *Kripke's Thesis* is undecidable. The rest of this paper provides that independent reason.

## III

## Truth-conditions and inference

The key is how to think of sentence truth-conditions given *TM*. If you accept *TM* you ought to hold that truth-conditions are linked to *possible truth-makers*. This means that if a sentence has truth-conditions, then, given reference and meaning, there is a *condition* that would, if it obtained, make it true, and another condition that, if it obtained, would make it false. Call this principle *TC*. *TC* is not a standard idea about truth-conditions, but, I will argue, it goes with *TM*. Given *TC*, we can show that the sentence *S* is *ungrounded*, that is, *step 2* in *TM-Refute*, cannot be said to have truth-conditions or not. Its truth-conditional status is essentially undecidable. If so, the truth-conditional status of *Kripke's Thesis* is undecidable. The argument against *TM* is then blocked, but not by denying *Kripke's Thesis*, but by affirming its undecidability.

In short, *Kripke's Thesis* goes beyond what can be said, not because our language has some expressive deficiency, but because of the nature of truth-conditions themselves, given *TC*. Once that's established, we get the result that we can keep *TM*, but ground-unspecifiable sentences, like *Infinite Descent*, *T* and *L* are alethically undecidable. Consequently, the Liar, *L*, and its kin, are alethically undecidable. This leads to a new kind of solution to the Liar and other semantic paradoxes: alethic undecidability.

Let's examine *TC* more closely. By *grounding-condition* we mean the kind of condition whose obtaining (as a fact) is, by *TM*, an ultimate truth/false-maker. The phrases  $x \Rightarrow S$  is true and  $x \Rightarrow S$  is false, as we noted above, signify, respectively, that  $x$ 's obtaining makes it that *S* is true and  $x$ 's obtaining makes it that *S* is false. *TC* then is:

*TC*: *S* has truth-conditions iff there is a grounding-condition  $x$  such that if  $x$  obtains,  $x \Rightarrow S$  is true and a grounding-condition  $y$  such that if  $y$  obtains,  $y \Rightarrow S$  is false.

Since the conditions  $x$  and  $y$  are not themselves *facts* but states that may or may not obtain, *TC* does not imply that a given sentence *S* is bivalent. Both  $x$  and  $y$  might fail to obtain. I note again that there is nothing particularly metaphysically loaded about any of these ideas. The metaphysical posits in this paper are merely convenient ontology for expressing the core ideas directed towards semantic matters.



My purpose is to show that *TC* is very plausible given *TM*. Of course, *TC* is not a standard idea about truth-conditions. Truth-conditions, in standard semantics, are associated with the biconditionals, like *S is true iff P*, where *P* expresses the content of *S*. This is a *correspondence conception* of truth-conditions. It's this that we cannot accept if we accept *TM*. I want to indicate how this is so.

Let's think of a truth-condition for a sentence *S* is a condition *C*, that may or may not obtain, under which *S* is true. It's the condition described by *S*. The condition *C* that is the truth-condition of a sentence *S* is either alethic or non-alethic. If *C* is non-alethic, then, by *TM*, if *C* obtains, it must be the truth-maker for *S*. If *C* is alethic, then, by *TM*, if it obtains, there must be a non-alethic condition *x*, in virtue of whose obtaining *C* obtains, and in virtue of which *S* is true. This holds likewise for the false-condition for *S*. We have just shown that if *S* has truth-conditions, then the right-hand side of *TC* holds. The converse also holds (obviously).

One might object that a ground-unspecifiable sentence like (0) in *Infinite Descent*, has truth-conditions, namely, that (1) is true. That is, we express these conditions thus:

$T^{(0)} : (0) \text{ is true iff } (1) \text{ is true.}$

In other words, one might contend, that (0) is associated, by virtue of its meaning and reference, with a *condition*, a state of affairs, that if it obtains, will necessitate (0)'s truth, and such that, if (0) is true, must obtain. Surely then (0) has truth-conditions. But note we cannot associate any non-alethic condition with (0). And so, this is not in accord with *TC*.

This objection should not affect the proponent of *TM*. The objector's idea of what it is to have truth-conditions involves a biconditional,  $T^{(0)}$ . To affirm that (0) has as its truth-conditions, we have to affirm this biconditional, but that requires we can make sense of supposing that (1) is true. But, if (1) is true, then (2) must be true, and so on ad infinitum. In short, if (1) is true, then an infinite chain of facts of truth must underlie its truth, but we cannot say, at any stage, that there is any non-alethic condition in the chain of facts descending downward. But affirming this is clearly at odds with affirming *TM*. If so, you cannot accept *TM* and accept the bi-conditional conception of truth-conditions, which is the standard view. So what view ought you to accept? I say *TC*.



## Inference and truth-conditions

Below I show that given *TC*, we can conclude that it's undecidable whether sentences like *step 2* in *TM-Refute* have truth-conditions or not. It follows from this, I suggest, that we cannot use *step 2*, and similar sentences, in inferences. In drawing that particular conclusion, I am invoking a regulative principle governing inference:

*Inference-Rule*: If it's undecidable whether a sentence *S* has truth-conditions or not, then we ought not to use *S* in logical inference.

*Inference-Rule* looks totally plausible. Indeed, if you cannot say a sentence has truth-conditions or not, you ought not to use the sentence in inference. Of course, you might use such a sentence *S* because you are ignorant of this alethic undecidability. Or you might explicitly assume you can affirm *S*'s having truth-conditions. That's consistent with *Inference-Rule*. But once you uncover a contradiction based on that assumption, then in withdrawing the assumption, you should cease using *S* in logical arguments.

## IV

### Saving TM and Semantic Entanglement

We have assembled all the required elements to block the argument *TM-Refute*, which threatens *TM*. Let's see how this works. *TM-Refute*, recall, concerns ground-unspecifiable sentences, like *Infinite Descent* (0) and *Looping T*, above. The premise in the argument that concerns us is *Kripke's Thesis* and *step 2*. *Kripke's Thesis* amounts to accepting the inference:

**K**: *S* is ground-unspecifiable. Therefore, *S* is ungrounded.

It's this inference we are questioning. My contention isn't that *K* is invalid. It's that *K*'s validity is *undecidable*. That's why *Kripke's Thesis* is alethically undecidable. If so, one cannot use *TM-Refute* against *TM*, since it's main premise, *Kripke's Thesis*, is undecidable.

The argument for undecidability of *K*'s validity focuses on the sentence *step 2*, which is the conclusion of *K*: *S is ungrounded*. My argument is that given *TC* — our principle about truth-conditions — it's undecidable whether *step 2* has truth-conditions or not. If the premise of *K* is true,

but we cannot say whether its conclusion, *step 2*, is true or false, then we cannot say whether *K* is valid or not.

I now show that *step 2*'s truth-conditional status is undecidable given *TC*. *Step 2* is the claim that no grounding-fact makes *S* true or false. We can represent *step 2* more perspicuously as below:

*Step 2*:  $\neg\exists x(x \text{ is a grounding-fact} \ \& \ x \Rightarrow S \text{ is true})$ .<sup>3</sup>

Note now three things about *step 2*:

- (i) The sentence *S is true* is a constituent of *step 2*.
- (ii) *S is true* is a ground-unspecifiable sentence.
- (iii) If it's undecidable that *S is true* has truth-conditions, it's undecidable that *step 2* has truth-conditions.

Facts (i) and (ii) are obvious. Let me argue for (iii). Say it's undecidable that *S is true* has truth-conditions. Then, it's undecidable, evidently, that the open-sentence,

$x \Rightarrow S \text{ is true,}$

has *satisfaction-conditions*. If so, it's undecidable, obviously, that the open-conjunction,

$(x \text{ is a grounding-fact} \ \& \ x \Rightarrow S \text{ is true}),$

has satisfaction-conditions. Then it's undecidable, clearly, that the existential quantification,

$\exists x(x \text{ is a grounding-fact} \ \& \ x \Rightarrow S \text{ is true}),$

has truth-conditions, and therefore, it's undecidable that its negation, *step 2*, does.

What I show now is that, given *TC*, it's undecidable that *S is true* has truth-conditions. In which case, it's undecidable that *step 2* has truth-conditions, and consequently, it's undecidable whether *K* is valid, and thus whether *Kripke's Thesis* is true or false.

---

<sup>3</sup> *Step 2* is actually the more complicated:  $\neg\exists x(x \text{ is a grounding-fact} \ \& \ (x \Rightarrow S \text{ is true} \vee x \Rightarrow S \text{ is false}))$ . The argument below applies equally to this more complicated version, but I use the simpler version in the main text.

Here's the argument that it's undecidable that *S is true* has truth-conditions. There is no specification of a grounding-condition for *S is true*, since the latter is a ground-unspecifiable sentence. But that means that a perfect assertor, searching for what's assertable given perfect information *I* about *S*, cannot uncover a specification of a grounding-condition for *S is true*. There just isn't one. So, the perfect assertor cannot affirm *T*:

*T*: There is an *x* such that *x* is a grounding-condition and if *x* obtains,  $x \Rightarrow 'S \text{ is true}'$  is true.

So the perfect assertor cannot affirm that *S is true* has truth-conditions. So, *S is true* cannot be said to have truth-conditions.

A very similar argument can be given that it cannot be said that *S is true* lacks truth-conditions. For, by *TC*, that requires affirming: *There is no y, such that y is a grounding-condition, and if y obtains,  $y \Rightarrow 'S \text{ is true}'$  is true.* For this to be assertable, we have to affirm '*S is true*' is true has truth-conditions, and so, there must be, by *TC*, a condition, *x* such that  $x \Rightarrow 'S \text{ is true}'$  is true is true. If there is such a condition, there is a condition, *x* that is grounding, such that, if *x* obtains,  $x \Rightarrow S$  is true. But *S* is a ground-unspecified sentence. Therefore, that condition cannot be specified. In which case, the perfect assertor, given information *I*, cannot assert '*S is true*' lacks truth-conditions.

To conclude: it's undecidable that *S is true* has truth-conditions. Therefore, it's undecidable that *step 2* has truth-conditions. If it's undecidable that *step 2* has truth-conditions, then, by *Inference-Rule*, we cannot use it in any logical conclusion. If so, we cannot say that *K* is valid, nor that *Kripke's Thesis* is true or false. If so, the proof, *TM-Refute*, against *TM* cannot proceed. We have shown its soundness is undecidable.<sup>4</sup>

We have strong, intuitive reasons, for thinking *TM* is correct. There is no threat from *TM-Refute*. So *TM* stands.

---

<sup>4</sup> Note that if we assume that *Kripke's Thesis* can be said to have truth-conditions, we can affirm that *Kripke's Thesis* entails *TM* is false, and so we can infer *TM* entails that *Kripke's Thesis* is false. Fortunately, the assumption in question, that *Kripke's Thesis* can be said to have truth-conditions, cannot be discharged. So we cannot draw the conclusion that *Kripke's Thesis* is false via this route.

## V

**Paradox and alethic undecidability**

Given *TM*, we cannot infer that ground-unspecifiable sentences lack truth-makers, or false-makers. Therefore, contra Kripke (1975), we cannot infer that they suffer truth-value gaps. But we cannot infer that they are bivalent either, given *TM*. Indeed, given *TC*, it's undecidable that ground-unspecifiable sentences have truth-conditions. So they are alethically undecidable. What then can we conclude about semantic paradoxes?

Paradoxical sentences are ground-unspecifiable sentences. (That's easily verified, given their looping and infinitely descending referential chains.) Kripke's (1975) strengthened liar *L* is ground-unspecifiable:

*L* : *L* is not true.

By *TM* and *TC* we must conclude that it's alethically undecidable. We cannot say it is true, false, that it lacks truth-value, or that it's both true and false. Indeed, we cannot say if *L* has any truth-conditions or not.<sup>5</sup> In particular we cannot use the *T*-schema and classical logic to derive a contradiction. That's because the required *T*-sentence is alethically undecidable, because it's undecidable that its constituent sentences, *L* is true and *L* is not true, have truth-conditions, since they are ground-unspecifiable sentences:

*T<sup>L</sup>* : *L* is true iff *L* is not true.

Therefore, we must refrain from asserting *T<sup>L</sup>*.

You might object that this means denying the validity of the *T*-schema. It does not. Logical principles concern sentences that *decidably* have truth-conditions – that accords with *Inference-Rule* above. After all, logic concerns the laws of thought. So, we should think of the *T*-schema's

<sup>5</sup> This last conclusion is in conflict with the theorists like Goldstein (1999) who argue that Liar-sentences express no proposition. But that conclusion is unstable. If *L* expresses no proposition, then a commentary sentence like *C* affirms a truth-about it. But *C* and *L* are grammatically and referentially identical. So why it is that *L* expresses no proposition, whereas *C* does? I don't think Goldstein has a satisfactory answer. Goldstein thinks *L* cannot express a proposition because supposing *L* expresses a proposition generates a contradiction (assuming propositions are either true or false). But his argument ignores the possibility of indeterminacy: that one cannot say either way, whether it does or does not express a proposition.

validity thus: the *T*-schema is valid if and only if for all sentences *S* whose possessing truth-conditions is *decided* the corresponding *T*-sentence is true. On that understanding the *T*-schema is valid, despite  $T^L$ 's not being assertable.

### Incoherence?

Another objection is that the *TM*-solution is incoherent. In order to affirm (C) *One cannot say that L is true*, we are committed to the embedded sentence *L is true* having truth-conditions. But the latter's truth-conditional status is meant to be undecidable. The objection fails. Asserting *C* does not commit us to *L is true* having truth-conditions. (Compare *Step 2*. Given the meanings of '&', '∃*x*' and '¬' we can deduce that the truth-conditions of *step 2* depends on *S is true*'s having truth-conditions – see main text. But there is no comparable argument for *C* and *L is true*.) *C*'s assertability only commits us to *L is true* being meaningful (which it is). The undecidability solution, given *TC*, does not affirm an identity between a sentence's saying something and its having-conditions (the central hypothesis of truth-conditional semantics). In short, there is an implicit challenge to truth-conditional semantics in this solution.

### Revenge

Although standard strengthened liars like *L* won't generate a revenge contradiction for the alethic undecidability approach, other kinds of revenge sentence might appear to threaten it. Consider *R*:

*R* : One cannot assert that *R* is true.

Does *R* have truth-conditions? Although *R* uses an alethic predicate, *true*, *R* does not describe an alethic condition. It's about our capability of asserting that a sentence is true, not about a sentence being true or false, and so on. So, *R* is not a ground-unspecifiable sentence like *L*. However, when we attempt to ascertain whether what *R* says is the case or not, we enter into a *processing loop*. Suppose that we can say *R* has truth-conditions. If *R* is true, it's because the condition  $R^*$  holds:

$R^*$ : *One cannot say that R is true.*

But if the condition  $R^*$  obtains, it does so because the denotation of ' $R$ ' meets one of three conditions: it's false, it lacks truth-conditions, or it's undecidable that it has truth-conditions. To determine if one of these conditions obtains we must return to  $R$  itself, and ask about its truth-conditions. Which brings us back to  $R^*$ . That's the processing loop. This means that an ideal assertor given information  $I$  will not be able to assert that  $R^*$  obtains or that  $R^*$  does not obtain, despite the fact that  $I$  records perfectly non-alethic reality, truth's dependence, and all the facts about  $R$ 's linguistic meaning and referential properties. That's because the ideal speaker will get caught up in the processing loop. We must conclude then:

$M$ : It's undecidable that  $R^*$  obtains or not.

What follows from  $M$ ? Nothing. There's no revenge, just a higher-order indeterminacy. It's undecidable whether the truth of  $R$  is undecidable. We cannot say that it's true that we cannot say that  $R$  is true. We cannot say that  $R$  follows from  $M$ .

### Conclusions

There is a lot more to say here about the nature of alethic undecidability, and its implications for how we understand truth, semantic properties, logic, and semantics. But that's enough for now. What's been shown is that the basic principle of truth-making,  $TM$ , implies a new kind of solution to the semantic paradoxes, one quite distinct from that developed by Kripke's (1975) reflections on truth. Whether or not this form of solution can be generalized to other cases cannot be demonstrated here. My hunch is that it can.

## BIBLIOGRAPHICAL REFERENCES

Armstrong, D. *Truth and Truthmakers*. Cambridge: Cambridge University Press, 2004. Print.

Barker, S. J. "Expressivism about Making and Truth-Making". Schnieder, B. and F. Correia (eds.). *Grounding and Explanation*. Cambridge: Cambridge University Press, 2012. Print.

Barker, S. J. and M. Jago. "Being positive about negative facts". *Philosophy and Phenomenological Research*. 117-138. Print.

Field, H. *Saving truth from Paradox*. Oxford: Oxford University Press, 2008. Print.

Goldstein, L. "A unified solution to some paradoxes". *Proceedings of the Aristotelian Society*. Sep. 2004: 53-74. Print.

Kripke, S. "Outline of a theory of truth". *The Journal of Philosophy*. Feb. 1975: 53-81. Print.

Maudlin, T. "Reducing revenge to discomfort". J. C. Beale (ed.). *Revenge of the Liar*. Oxford: Oxford University Press, 2007. Print.

Priest, G. *In Contradiction*. Dordrecht: Martinus Nijhoff, 1987. Print.

Rodriguez-Pereyra, G. "Why Truthmakers". Dodd, J. and H. Beebe (ed.). *Truthmakers*. Oxford: Oxford University Press, 2005. Print.





# SELLARS AND PRETENSE ON “TRUTH & ‘CORRESPONDENCE’” (WITH A DETOUR THROUGH MEANING ATTRIBUTION)

SELLARS Y LA PRETENSIÓN EN “VERDAD Y ‘CORRESPONDENCIA’”  
(CON UN DESVÍO A TRAVÉS DE LA ATRIBUCIÓN DEL SIGNIFICADO)

BRADLEY ARMOUR-GARB  
University at Albany – SUNY, USA. barmour-garb@albany.edu

JAMES A. WOODBRIDGE  
University of Nevada Las Vegas, USA. woodbri3@unlv.nevada.edu

RECIBIDO EL 23 DE JULIO DE 2012 Y APROBADO EL 24 DE AGOSTO DE 2012

## RESUMEN    ABSTRACT

La explicación de la verdad que Wilfrid Sellars ofrece en su artículo “Verdad y ‘correspondencia’”, nos presenta una aparente tensión *interna* entre dos aspectos supuestamente incompatibles con sus puntos de vista sobre la verdad. Para resolver plenamente esta tensión, es necesario tener en cuenta una explicación de la atribución de significado. Creemos que la explicación que Sellars ofrece, sobre todo en “El significado como clasificación funcional”, incluye los elementos básicos necesarios para resolver la tensión interna en su explicación de la verdad, sin embargo, sus puntos de vista sufren de una aparente tensión *externa*, con respecto a un supuesto conflicto entre su explicación y nuestras prácticas lingüísticas e inferenciales con la expresión ‘significa’.

En este artículo, mostramos cómo la tensión interna en la comprensión de Sellars de la verdad, así como la tensión externa en su explicación de la atribución de significado se puede resolver sin perder de vista el espíritu sellarsiano, apelando a las particulares explicaciones *fictionalistas* de habla sobre la verdad y habla proposicional (incluida la atribución de significado) que hemos desarrollado en otro lugar.

The account of truth that Wilfrid Sellars offers in his paper, “Truth and ‘Correspondence’”, presents us with an apparent *internal* tension between two seemingly incompatible aspects of his views on truth. To fully resolve this tension, it is necessary to factor in an account of meaning attribution. We think that the account Sellars offers, most centrally in “Meaning as Functional Classification”, includes the basic elements required for resolving the internal tension in his account of truth, but here his views suffer from an apparent *external* tension, regarding an apparent conflict between his account and our actual linguistic and inferential practices with the expression ‘means’.

In this paper, we show how the internal tension in Sellars’s understanding of truth, as well as the external tension in his account of meaning attribution, can be resolved while adhering to a Sellarsian spirit, by appealing to the particular *fictionalist* accounts of truth-talk and proposition-talk (including meaning-attribution) that we have developed elsewhere.

## PALABRAS CLAVE    KEY WORDS

deflacionismo, fictionalismo, pretensión, atribución del significado, habla proposicional, habla sobre la verdad.

deflationism, fictionalism, meaning attribution, pretense, proposition-talk, truth-talk.

## Introduction

The account of truth that Wilfrid Sellars offers in his paper, “Truth and ‘Correspondence’”, presents us with a *prima facie* tension (Sellars, *Truth and*).<sup>1</sup> In particular, there is an apparent *internal* tension between two seemingly incompatible aspects of his views on truth. To fully resolve this tension, it is necessary to factor in an account of meaning attribution. We think that the account Sellars offers, most centrally in “Meaning as Functional Classification” (Sellars, *Meaning as*)<sup>2</sup>, includes the basic elements required for resolving the internal tension in his account of truth, but here his views suffer from an apparent *external* tension, regarding an apparent conflict between his account and our actual linguistic and inferential practices with the expression ‘means’.

Our aim here is to show how the internal tension in Sellars’s understanding of truth, as well as the external tension in his account of meaning attribution, can be resolved while adhering to a Sellarsian spirit, by appealing to the particular *fictionalist* accounts of truth-talk and proposition-talk (including meaning-attribution) that we have developed elsewhere.<sup>3</sup> In order to show this, we will first discuss the central aspects of Sellars’s account of truth, along with how the tension we find in it arises. We will then explain how Sellars faces a further problem, in virtue of his views on truth, coming from Hilary Putnam’s model-theoretic argument (Putnam, *Meaning and*) (Putnam, *Models and*). To deal with the challenges Putnam presents, we will introduce the basics of our pretense account of truth-talk, supplement it with our pretense account of proposition-talk (specifically, as it operates in meaning attributions), and explain how this combination resolves the tensions we find in Sellars’s views while avoiding the further problem we think he faces.

One general point we want to make here is that Sellars’s views and pretense views fit together well. Thus, after explaining how we can resolve the aforementioned tensions, we will turn to some important insights that we have gained from reading Sellars. These insights

---

<sup>1</sup> All citations are to the reprint in Sellars *Science*. We recognize that Sellars later offers a somewhat different account of truth (hinted at in the note on page 224 of Sellars (ibid.)), but Lionel Shapiro has convinced us that this shift is highly problematic.

<sup>2</sup> See also Sellars (*Naturalism and*), especially Chapter 4.

<sup>3</sup> We are referring to the pretense-based accounts presented in Armour-Garb and Woodbridge (*Why deflationists*) and (*The story*). Woodbridge (*Truth as*) and (*Propositions as*) offer earlier versions of pretense accounts of truth-talk and proposition-talk, respectively.

help us resolve a *prima facie* challenge that critics sometimes raise against *fictionalist* accounts like ours – accounts that deal with various problematic fragments of discourse by explaining them in terms of *pretense*. Their challenge is that pretense-based accounts fail in the face of what we call the *Engagement Complaint*: the objection that speakers typically do not think of themselves as pretending when they employ these ways of talking. Sellars's careful analysis of different kinds of rule-governed behavior helps us bolster our response to this challenge.

## I

### Sellars, truth, and the deflationism/inflationism distinction

In "Truth and 'Correspondence'", Sellars presents an account of truth that involves two seemingly conflicting views on how to understand truth. In that paper, one of the things that Sellars explains is how one can grant at least some of the importance of Alfred Tarski's semantic theory of truth without closing off further insights regarding additional features that he takes to be essential to understanding truth. He does this by proposing two modes of "correspondence", both of which a proper understanding of truth must recognize, at least for a certain class of truths.

The first mode of correspondence that Sellars acknowledges is the one supposedly captured by the collection of the instances of the truth schema,

(T) 'S' is true iff S,

the entailment of which Tarski took to be a condition of adequacy on any formal theory of truth (Tarski). Unlike Tarski, however, Sellars takes propositional truth to be prior to sentential truth and holds that the latter is derivative of the former. Hence, for Sellars, the relevant, or primary, instances of the equivalence schema for truth are more like those presented in Horwichian Minimalism than those from a Fieldian or Quinean disquotationalism (Horwich) (Field) (Quine). But, actually, even this is not quite right, for, while we might present the necessary instances of the truth schema for propositions by using the standard (material) biconditional sign, viz., as

(PT) <p> is true  $\leftrightarrow$  p,

where 'p' serves as a sentential variable that can be replaced by any declarative (English) sentence, and where the angle brackets, '<' and '>', serve as a device for nominalizing any sentence that goes in for 'p', as Sellars reads '<=>', it is to be understood as capturing logical equivalence or, as Sellars puts it, "reciprocal entailment". Accordingly, a standard instance of the truth schema for propositional truth, such as

(0) <snow is white> is true  $\leftrightarrow$  snow is white,

is to be "put forward" as

(0') <<snow is white> is true> entails, and is entailed by, <snow is white>.

This, he says, governs inferences like the following:

- (1) <snow is white> is true.  
(2) So, Snow is white. (Sellars, *Truth and 206*)<sup>4</sup>

Sellars takes these inferences, the ones codified in the equivalences, which are taken to express the first sense of correspondence, to provide the semantical rules for, and thus the sense and, hence, the meaning of, the word 'true'. According to Sellars, if this is right, then "instead of standing for a relation or a relational property of statements (or, for that matter, thoughts) 'true' is a sign *that something is to be done*—for inferring is a doing". (Ibid)

In addition to this semantical-rules aspect of truth and the kind of correspondence between truth-attributions and claims about the world that it captures, Sellars is also concerned with explaining a *second* sense of correspondence, one that goes beyond that captured by the truth-schema equivalences. He takes this second sense of correspondence to involve a sort of *picturing* relation, holding between certain linguistic objects—names, in effect, as they appear in elementary "matter-of-factual" empirical claims—and the nonlinguistic objects that they name or represent (Ibid. 213-14). This picturing relation is supposed to be an extension of the kind of projection that a map of some geographical area involves—that is, some projection from the pictured objects, in the arrangement that they have, to the elements of the picture, and the arrangement that they have. The rules of projection may be fantastically

<sup>4</sup> We are here (and in (0')) replacing Sellars's use of 'that'-clauses with the use of angle-bracket notation.

complex, but Sellars maintains that they are definable in terms of matter-of-factual (in contrast with conventional) relationships, such as an isomorphism between two systems of objects. It is this second sense of correspondence that Sellars suggests may be “essential to understanding truth”, at least for a particular class of claims.

One central point that it is important to recognize here is that this second sense of correspondence involves a shift in focus from propositional truth (as in the first sense of correspondence) to sentential truth, since the relevant picturing is supposedly something that linguistic objects do. We are now “reflecting on what it means to say of a form of words that it is true” (Ibid 223). Sellars exemplifies this kind of case as follows.

‘Chicago est grande’ (in French) =  
is true.  $\text{Df}(\exists \text{ that-}p)$  ‘Chicago est  
grande’ (in French)  
expresses the sense that- $p$   
and that- $p$  is true. (Ibid 205, 233)

Given that Sellars takes expressing a sense to be the same thing as expressing a proposition, he understands sentential truth as an extension of propositional truth. What the extension involves is attributing the expressing of a proposition or sense to some linguistic item, i.e., a meaning attribution. So, a full explanation of how Sellars’s second sense of correspondence fits into his understanding of truth will require an appeal to an account of meaning attribution. We will return to this below, after explaining the *prima facie* internal tension we find in Sellars’s views.

Although the two Sellarsian modes of “correspondence” just sketched, which he takes to be, in one sense or another, essential to our understanding of truth, both seem coherent enough; we find a tension in his view because of a fairly standard bifurcation among contemporary accounts of truth. As is familiar in contemporary work on truth, there is a split between *deflationary* and *inflationary* accounts of truth.<sup>5</sup> Deflationists about truth-talk (henceforth, *T-deflationists*) see the expression ‘is true’ as a device of semantic descent, whose meaning is exhaustively provided by instances of schemata like (T) or (PT) (Sellars 223).<sup>6</sup> In general, T-deflationists have a lot to say about the *concept* of truth, but they typically chagrin talk about any *property* of truth. It is not

<sup>5</sup> For more on the distinction between deflationism and inflationism about truth, see Armour-Garb.

<sup>6</sup> See also Armour-Garb and Woodbridge (*From mathematical*).

part of their view *qua* T-deflationism that there is no property of truth. Rather than explicitly denying that there is a property of truth, standard T-deflationists often take a more cautious stance and, so, do not accept that there is a substantive property of truth. By contrast, *T-inflationists*, while they must ultimately say something about the concept of truth, also accept that there is a property of truth, and they go to some lengths in trying to explain its nature.

The internal tension that we find – the *apparent* tension, since we think that, ultimately, it can be dissolved (albeit, as we will explain below, not in an entirely worry-free manner) – arises from the fact that, as Sellars contends, one mode of correspondence exhaustively captures the *concept* of truth, while the other mode of correspondence ends up being essential to “the understanding of truth”. Put differently, the tension appears to arise because Sellars seems to be expressing commitments to both a central aspect, or feature, of T-deflationism *and* a central aspect, or feature, of T-inflationism. But T-deflationist and T-inflationist accounts are supposed to be incompatible. Hence, we get what appears to be a tension.

Given this *prima facie* tension, we might ask about Sellars’s motivation for endorsing this apparent aspect of T-inflationism in his understanding of truth. His inferential reading of the standard truth schema comports with his general “use-theoretic” approach in philosophy of language, but his inclusion of an account of truth as a substantive (albeit ultimately matter-of-factual) picturing relation is harder to fit with his general approach.<sup>7</sup>

We think that one clue to Sellars’s motivations on this front might appear in his claim that “... whatever else language does, its central and essential function, the *sine qua non* of all others, is to enable us to *picture* the world in which we live” (Sellars, *Truth and* 213). This point comes a bit after a discussion of how to modify Wittgenstein’s *Tractarian* account of picturing (Wittgenstein, *Tractatus*), in order to avoid a reading of it that amounted to “a thorough-going linguistic idealism,” a view that Sellars considered “a miserable absurdity” (Ibid. 213). Recognizing picturing the world as the central function of language, it is then a short step to

---

<sup>7</sup> The typical T-inflationist motivations of thinking that truth must play some fundamental explanatory role would not do the trick here. Appeals to truth in explanations of successful action or the success of the scientific method have been undercut by T-deflationist replies (Cf. Williams and Leeds), and the remaining appeal that some T-inflationists think must be made, in an account of linguistic meaning and mental content (Cf. Devitt), are entirely out of bounds for Sellars.



understanding the truth of a form of words (at least in an elementary matter-of-factual claim) as involving *correct* picturing of the world.

However, we think that in appealing to a *Tractatus*-inspired isomorphism between the elements of a linguistic picture-object and the objects in the world being pictured, in order to accommodate a “thick”, picturing, correspondence intuition regarding truth (in contrast to the “thin” correspondence intuition that some T-deflationists think is captured in the instances of (T) or (PT) (Williams 233) (Horwich 104-05), Sellars overshoots the mark. His own understanding of what counts as picturing requires no more than that picture-objects get produced via certain interaction and entanglement with objects and events in the world.<sup>8</sup> This is why, in addition to the fully conceptual, natural-linguistic objects produced as part of the linguistic behavior of speakers, even such “non-conceptual” objects as phonographic recordings count as pictures (e.g., of concerts) (O’Shea 147). So, to avoid linguistic idealism and understand language as a means of picturing the world, all that Sellars needs is a means of describing natural-linguistic objects as connected to and entangled with objects and events in the world. This can involve a “fantastically complex system of rules” (Sellars, *Truth and* 215), but it does not require any explicit appeal to any special notion of picturing explained in terms of an isomorphism.<sup>9</sup> We will say more about this below when we turn to the role of meaning attribution in the second sense of correspondence. Before doing so, however, we return to the apparent internal tension in Sellars’s views on truth.

Taking Sellars’s two-part account as given, we think that one way to resolve the apparent tension that arises in it is to recognize that, while Sellars’s views about truth involve both a deflationary component and an inflationary component, the components seem directed at two different sorts of *bearers* of truth, propositions and sentences, respectively. In turn, this suggests that we might see Sellars as endorsing a sort of T-deflationism about propositional truth and a sort of T-inflationism about sentential truth. Rather than seeing Sellars as *both* a T-deflationist *and* a T-inflationist, we might see his view as in some ways deflationary (as regards propositional truth) and in some ways inflationary (as

<sup>8</sup> See Sellars (*Truth and* 219-22), where he describes his hypothetical super-inscriber.

<sup>9</sup> What we find interesting is that, in this same paper (at least the 1963 version); Sellars already presents a means for doing this with his “dot-quote” notation on (Ibid. 204-05, 223). The explanations involving dot-quotes were substituted into the proofs of Sellars (*Science*).

regards sentential truth). This resultant *hybrid view* would dissolve the internal tension that we find in Sellars's view of truth.<sup>10</sup>

Although this hybrid view seems to be a coherent one for a truth-theorist to adopt, one aspect of Sellars's version of it seems to suffer from a very serious worry. We have in mind a potential problem for his preferred account of sentential truth, given that it is grounded in an appeal to the notion of an isomorphism, and given that he has an overarching commitment to nominalism. Here is the problem.

Sellars maintains that the truth of elementary matter-of-factual statements essentially involves correct picturing, where this relies on the notion of a projection from the objects pictured to the names employed in the picture. However, according to Putnam's model-theoretic argument, there could be an infinite number of mappings, i.e., projections, from objects to names, any number of which could count as the relevant picturing relation (Putnam, *Meaning and*) (Putnam, *Models and*). The reason that the model-theoretic argument is especially worrisome for Sellars, given his commitment to nominalism, is because nominalism seems to undermine the most familiar replies to Putnam's challenge. These replies typically involve an appeal to privileged relations or some appeal to privileged classes of particulars,<sup>11</sup> which some have described as "anti-nominalist" (see van Fraassen). If a commitment to nominalism does undercut one's ability to block the conclusions of Putnam's model-theoretic argument, then this is a problem for Sellars's account of sentential truth.<sup>12</sup>

---

<sup>10</sup> In a review of Horwich (*Truth*), Field (*Critical notice*) questions whether a reliance on an equivalence schema like (PT) should be called a "minimalist" account of truth on grounds that even T-inflationists can accept a version of (PT). Be that as it may, given the importance that Sellars attributes to (PT), we find it useful to see Sellars as adopting the sort of hybrid views that we are describing.

<sup>11</sup> See, for example, the responses given in Glymour and Lewis.

<sup>12</sup> Lest anyone think that Sellars has a way out of this dilemma in virtue of his recognition and incorporation of a normative dimension in his account of truth as *correct* picturing according to a "fantastically complex system of rules", first note that the whole point of the model-theoretic argument is that there are too many candidate isomorphisms, with no way to single one out as the one that makes for correct picturing. Moreover, any proposed normative selection one hopes might be performed by the system of rules involved in language picturing the world will also fail to single out a unique candidate isomorphism. This is because any norm that is supposed to do this work must be either explicit in the rules or implicit in the behavior of speakers. But in the former case, the rule, and thus the relevant norm expressed, will fall victim to Putnam's "just more theory" objection (Putnam, *Meaning and* 126-27) and (*Models and*, 18) (Lewis 224-25). And in the latter case, Wittgenstein's rule-following considerations again leave us with too many candidate norms implicit in the behavior, with no fact of the matter about which one is correct ((Wittgenstein, *Philosophical*) and Kripke).



This worry might motivate trying to accommodate Sellars's "thick" correspondence intuition, regarding the truth of elementary matter-of-factual claims, in some way that avoids the challenges of Putnam's model-theoretic argument while still adhering to Sellars's nominalist scruples. We think that bringing in our pretense-based accounts of truth-talk and proposition-talk can accomplish this. We will explain how, beginning with our account of truth-talk.

## II

### A pretense account of truth-talk

The analysis of truth-talk that we have developed employs a Walton-inspired strategy involving *semantic pretense*.<sup>13</sup> The central idea of our account is that our use of truth-locutions, such as 'is true', 'is false', 'is not true', etc. is underwritten by a rule-governed pretense that establishes a systematic dependency between some of what is *to be pretended* and certain real-world conditions outside of the pretense.

This type of pretense is similar to that found in children's games of make-believe. Make-believe involves two kinds of *prescribed* pretenses: those that are stipulated background pretenses that are expressly pretended, and additional pretenses that are systematically *generated from reality* via a game's principles of generation (Cf. Crimmins 5). Consider, for example, two children, Zev and Dex, playing a make-believe game of *Star Trek*. Certain pretenses will be stipulated as the background pretenses for the game, for example,

#### *Star Trek* game pretenses

The props for the game are Dex, Zev, cell phones, flashlights, the living room, and the bathroom. It is to be pretended that

- i) Dex is Captain Kirk,
- ii) Zev is Mr. Spock,
- iii) cell phones are communicators,
- iv) flashlights are phaser pistols,
- v) the living rooms is the bridge of the Starship Enterprise,
- vi) the bathroom is the transporter station...

<sup>13</sup> See Walton. For other Walton-inspired, pretense-based accounts of various fragments of discourse, see Evans; Yablo (*How in*) (*Go figure*), and (*The myth*); Kroon (*Characterizing*) (*Fictionalism and*), and (*Descriptivism*); Crimmins; and Egan.

In addition to these rules regarding various props that are used in the make-believe, the rules of the game also generate additional prescribed pretenses, based on the stipulated pretenses along with how things really are in the world outside of the game. That is, the stipulated background pretenses also underwrite various implicit *principles of generation* – rules of the game that determine what further pretenses are prescribed, depending on the circumstances of the props in the world outside of the game. Thus, given (i)-(vi), the pretenses displayed in

(3) Captain Kirk stole Mr. Spock's phaser pistol and hid it in the transporter station.

are prescribed in the game if, and only if, the real-world conditions specified directly in the following claim obtain.

(4) Dex took a flashlight that belongs to Zev away from him and hid it in the bathroom.

Putting forward (3) as prescribed expresses a commitment to the obtaining of the conditions specified in (4), thereby asserting indirectly what one could assert directly using (4). Thus, although (3) involves pretense, it can be used to make a serious assertion about the world – albeit indirectly. Claims involving make-believe can thus be used to make genuinely true, serious statements indirectly. They are true exactly when the pretenses they display are prescribed. Pretense-employing claims can therefore be used for serious purposes (e.g., providing information about the real world, giving explanations, e.g., of why Zev is scouring the house, etc.). So, pretense-based accounts are not error-theories; there is a difference between *pretense-involving fictionalism* (or PIF) accounts and *error-theoretic fictionalism* (or ETF).<sup>14</sup>

Understanding truth-talk in terms of semantic pretense is to take it as underwritten by a game of make-believe governed, at least in part, by principles of generation like the following.

### **Truth-Talk game pretenses**

The props for the game are the linguistic expressions 'is true', 'is false', 'is not true', and their cognates (e.g., 'is so', 'is right', etc.), as well as the expressions 'truth' and 'falsity'. It is to be pretended that

<sup>14</sup> For further discussion of this distinction, see our (*Linguistic puzzles*), (*Why deflationists*), (*The story*), and (*Pretense and*).

- i) the adjectival expressions 'is true', 'is false', etc. function predicatively, to describe objects as having or lacking certain properties;
- ii) the nominal expression 'truth' picks out the property attributed with the expression 'is true' (and 'falsity' picks out the property attributed with the expression 'is false');
- iii) the most basic objects that directly have or lack the properties that 'is true', etc. attribute are abstract, mind and language-independent entities called "propositions"; other kinds of objects (e.g., linguistic items) can have the properties that 'is true', etc. attribute only derivatively, in virtue of expressing a proposition that has the relevant property.

The game of make-believe based on these stipulated grounding pretenses also includes *explicit* principles of generation, which are rules that serve to determine what else is to be pretended, given how things are outside of the game – i.e., what further pretenses are generated from reality. The basic principles of generation are as follows.

(PGT) The pretenses displayed in an utterance of '(The proposition) that p is true<sup>1</sup>' are prescribed if and only if p.

(PGF) The pretenses displayed in an utterance of '(The proposition) that p is false<sup>1</sup>' are prescribed if and only if  $\neg$  p.

(PGS) If  $S_1$  and  $S_2$  are sentences that are alike except (in some transparent context) one has a subsentence 'p' where the other has '<p> is true', then one can directly infer  $S_1$  from  $S_2$  and  $S_2$  from  $S_1$  (where 'p' serves as a variable that can be replaced by a sentence and '<p>' stands for a nominalization of such a sentence).

The roles and functions of the rules for the pretense should be understood as follows.

The explanation of the rules for the game of make-believe that underwrites truth-talk begins with a specification of certain linguistic expressions as the props for this game. Rules (i)-(iii) then all state stipulated, *expressly* made-believe, background pretenses *about* these props. In particular, Rule (i) specifies and explains what is to be pretended about the *adjectival* props. One consequence of this rule is that all uses of 'true', 'false', 'not true', etc. involve pretense *intrinsically*, which is to say: There are no pretense-free uses of these truth-locutions

because pretense is invoked in their basic functioning. As a consequence, the only *serious* content (about the real world, outside of the make-believe) that an instance of truth-talk has (or: possesses) must come from the operation of the make-believe's *principles of generation* – specifically, rules (PGT) and (PGF).

Principles (PGT) and (PGF) cover what are arguably the most basic cases of truth-talk, what we call “transparent propositional truth-talk” so an account of these cases provides a core for our more general account. These principles of generation make the correctness of a putative attribution of truth or falsity, to some nominalized, sentential content-vehicle, a function (possibly negating) of whether the conditions specified by a use of that content-vehicle obtain. This makes the utterance of an instance of truth-talk an indirect means for specifying those very same conditions, thus determining the serious content of the instances of truth-talk. Since these indirectly specified conditions can actually obtain, this makes it possible for instances of truth-talk to make (what we might, now employing the very pretense being explained, describe as) “genuinely true” claims about the world outside of the pretense.

Rule (PGS) satisfies another important condition of adequacy for any theory of truth-talk, as it provides a version of a rule of *intersubstitution*. Such a rule further captures the sense in which the content of a putative ascription of truth to some content-vehicle just is the content of the content-vehicle itself. The general intersubstitution licensed by this rule is integral to a pretense account yielding the right content for the more interesting cases of truth-talk, viz., and ‘true’-involving generalizations. Since those cases are what make truth-talk so useful, it seems to be a fairly central aspect of any adequate account of this fragment of discourse.

To illustrate how truth-talk functions according to this account, consider a straightforward instance of truth-talk, such as

(5) It is true that crabapples are edible.

For reasons we explain elsewhere (*Why Deflationists*), we take (5) to be more perspicuously rendered by

(5′) That crabapples are edible is true,

where 'that crabapples are edible' is (in the context of the pretense) a referring expression that picks out the proposition that crabapples are edible.<sup>15</sup> Syntactically speaking, the 'that'-clause is a nominalization of the content-vehicle

(6) Crabapples are edible.

When asserted, a 'true'-involving sentence like (5') presents the pretenses it displays as prescribed, where being prescribed is determined by:

a) the particular principle of generation that governs those pretenses (here, the relevant instance of Rule (PGT)),

and

b) whether the conditions, whose obtaining the relevant principle of generation makes prescriptive for the pretenses, actually obtain.

Recall that Rule (PGT) has it that the prescriptive conditions for the pretenses displayed in (5') are those specified by the use of the content-vehicle that is nominalized as the subject expression of (5') – in this case, (6). In short, by presenting the pretenses it displays as prescribed, (5') specifies, indirectly, precisely those conditions that (6) specifies directly.

Our talk of the conditions specified by a claim pertains to a central component of what we mean by the *content* of an utterance. However, because, as explained above, our PIF account of truth-talk views the expression 'truth' as simply a prop for the game of make-believe that is the core of the account, this makes it just a stipulated background pretense of the game that this expression picks out any property. Since this means that, really, there is no property of truth, we do not hold that the conditions specified by a claim can be understood fundamentally as truth-conditions. We call them *M-conditions*. While M-conditions are objective, worldly situations that either obtain or fail to obtain, on our view, truth-conditions have only a thin, derivative status, as conditions for the appropriate use of the pretense-involving truth-predicate. The truth-conditions for a sentence are a *by-product* of its meaning, of which M-conditions are a significant component. This thought is in line with the meaning-to-truth conditional,

<sup>15</sup> Again, we will say more about talk putatively about propositions below. For a fuller account, see our (*The story*).

(MTC) If S means that p, then S is true iff p,

no instance of which we reject.<sup>16</sup>

Now, while some sentences specify M-conditions directly, as is the case with (6), other sentences specify M-conditions only indirectly. Indeed, as should be apparent, one of the consequences of our PIF account of truth-talk is that any specification of M-conditions (that obtain or fail to obtain outside of the pretense) that is accomplished by a 'true'-involving sentence will be accomplished only indirectly, via the operation of the pretense that governs the functioning of the truth-predicate.

The resulting, quasi-anaphoric, identity of content between an instance of transparent propositional truth-talk of the form 'It is true that p' and the content-vehicle nominalized in it (the sentence that goes in for 'p') means that the game of make-believe behind truth-talk generates all the instances of the equivalence schema

(ES) It is true that p iff p.<sup>17</sup>

This is an important result because, as T-deflationists have argued (and Sellars acknowledges), these equivalences are (at least some of) the central principles governing truth-talk. Our PIF account has them follow directly from the functioning that truth-talk is given by the game of make-believe that underwrites it.

Because our PIF view generates every instance of (ES), it thereby captures the first sense of correspondence that Sellars attributes to truth. Capturing the second sense of correspondence will require fleshing out Rule (iii) of the pretense behind truth-talk, by bringing in the relevant parts of our PIF account of proposition-talk – those pertaining to meaning attribution. We maintain that it is a merit of our account – over T-deflationist accounts that do not explicitly recognize an element of fiction operating in truth-talk<sup>18</sup> – that ours also helps explain certain T-inflationist intuitions we might have. In particular, we can also accommodate the kind of "thick" correspondence intuition that Sellars expresses in the second sense of correspondence he attributes to

<sup>16</sup> Note that the instances of (MTC) also comport with Rule (iii) of the pretense.

<sup>17</sup> For present purposes, this is taken to be equivalent to (PT)  $\langle p \rangle$  is true  $\leftrightarrow$  p.

<sup>18</sup> But see our (*Why deflationists*) and (*From mathematical*) for arguments that all T-deflationists should endorse alethic fictionalism.

truth – the one involving correct picturing in virtue of an isomorphism between the elements of natural-linguistic objects and objects and events in the world. This is something that other, thorough-going deflationists cannot do. To show how our view accomplishes this, while maintaining a Sellarsian spirit with respect to nominalism and meaning attribution, we will first explain the basic elements of Sellars's account of meaning attribution and the external tension we find in his view.

### III

#### Sellars, "picturing the world", and non-relational meaning

To account for the second "correct picturing" sense of correspondence that Sellars attributes to truth, we need to understand the picturing aspect of natural-linguistic objects and the functioning of meaning attributions made to them. We will focus on meaning attribution at the sentential level, which is performed by employing a 'that'-clause in a claim of the form,

(MA) S means that p,

where 'S' is filled in with an expression denoting a sentence, and 'p' is filled in with an English sentence-in-use.

We think that Sellars's views about 'that'-clauses and their role in meaning attributions provide a foundation for an account of how language can be understood as picturing the world without the need for any special representational relations, or even any explicit appeal to projections or isomorphisms. As we mentioned above, Sellars is a nominalist and thus no fan of abstract entities. However, his method of avoiding abstract meaning entities (like propositions) is to deny that 'means' is a relational expression, and this is what generates the external tension we find with his account of meaning attribution. On his view, in its "quasi-formal" representation, 'means' ends up as a specialized form of the copula – one that serves to classify lexically specified linguistic expression types in terms of an illustrating functional sortal (Sellars, *Meaning as* 431).<sup>19</sup> The "quasi-formal" notation that Sellars develops to express this kind of functional classification is his well-known "dot-quote" notation.

<sup>19</sup> See also (Sellars, *Naturalism and* 73-4).



So, on Sellars's account, a claim of the form

(MA) S means that p

is to be understood in terms of a claim of the form

(SMA) Ss are ●p●s,

where the expression that goes in for 'S' is what Sellars calls a *distributive singular term*, here specifying a type of linguistic item, and placing the dot-quotes around whatever expression goes in for 'p' forms another distributive singular term, here specifying a functional role, as the one played in our language by the expression thus dot-quoted (Ibid 428).<sup>20</sup> (More specifically, it specifies a broadly inferential role, involving Language Entry Transitions, Intra-linguistic Moves, and Language Departure Transitions). (Ibid 423-24)<sup>21</sup>

There is much we like about Sellars's account: his understanding of meaning in terms of use (to put it very broadly), the essentially social aspect of language, and his analysis of meaning attributions as, in effect, specifying the use-features of an expression, by illustrating them – that is, by, in a sense, putting them on display with dot-quotes. What it is important to recognize about the operation of dot-quotes is that even though this device operates by holding up instances of the meaning attributor's home language, it still serves as a way to indicate connections between linguistic items and non-linguistic items (objects and events that factor into the linguistic functional-role indicated by the use of dot-quotes) and does not simply collapse into translation. So, rather than just stating that some linguistic expression has the same meaning as some home-language expression (as in a translation), a meaning attribution gives or "illustrates" the meaning of the expression (Ibid. 431), and thus specifies something outside, even of the speaker's home language.

It is because Sellars's dot-quotes analysis of meaning attributions explains this type of claim as connecting natural-linguistic objects to something outside of language that we think dot-quotes provide Sellars with enough to avoid linguistic idealism. More specifically, it is because the linguistic functional-roles put on display with dot-quotes include Language Entry Transitions (in which objects in the world affect speakers

<sup>20</sup> See also (Sellars, *Naturalism and* 73-4, 77-8).

<sup>21</sup> See also (Sellars, *Naturalism and* 69).



and lead to the production of natural-linguistic objects) and Language Departure Transitions (in which natural-linguistic objects lead speakers to affect objects in the world through actions) that natural-linguistic objects with the roles put on display can, via a “fantastically complex system of rules”, have the kind of interaction and entanglement with non-linguistic objects and events that is sufficient to count as language picturing the world, on Sellars’s matter-of-factual understanding of picturing. This is why we think, as we mentioned above, that Sellars’s explicit talk of isomorphisms and projections as the basis of picturing is overshooting the mark when it comes to avoiding linguistic idealism.

However, while eschewing talk of isomorphisms and projections – and perhaps any attempts explicitly to *describe* a picturing relations between natural-linguistic items and objects in the world, rather than just *displaying* it with dot-quotes – would allow Sellars to avoid the challenges of Putnam’s model-theoretic argument, while maintaining nominalist principles, we see a problem with Sellars’s account of meaning attributions. This is the aforementioned *external* tension we find in his view, pertaining to a conflict it has with our inferential practices involving the expression ‘means’. This conflict is particularly problematic for Sellars for the following reasons.

In “Language as Thought and as Communication,” Sellars presents an understanding of fully conceptual linguistic behavior as largely a matter of its thoroughly rule-governed nature, with the meanings of expressions arising out of the rules that govern their use – in particular, in inferences (Sellars, *Language as* 510 512-13, 517). In keeping with Sellars’s views, then, we should consider the kinds of inferences that form MA meaning attributions sanction and are sanctioned by – in order to consider what we might want to say about the meaning of ‘means’. One thing that we have emphasized in our recent work is that, Sellarsians or not, we should all recognize that our linguistic and inferential practices involving ‘that’-clauses (including their use in MA claims) treat these expressions as singular terms, putatively denoting mind- and language-independent abstract entities (Armour-Garb and Woodbridge, *The story*).<sup>22</sup>

To see this, consider the following inference examples.

---

<sup>22</sup> See also, Schiffer (*Language-created*) and (*Pleonastic*).

- (I) "Holzäpfel sind eßbar" means that crabapples are edible.  
So, "Holzäpfel sind eßbar" means something.  
(There is something that "Holzäpfel sind eßbar" means).
- (II) "Holzäpfel sind eßbar" means the same thing as "Crabapples are edible".  
"Crabapples are edible" means that crabapples are edible.  
So, "Holzäpfel sind eßbar" means that crabapples are edible.
- (III) "Holzäpfel sind eßbar" means that crabapples are edible.  
Corey believes that crabapples are edible.  
So, "Holzäpfel sind eßbar" means something that Corey believes.

The inferential practices displayed in these arguments are easily regimented into first-order logic, if we take the claims involved to traffic in a domain of non-linguistic entities, of the sort known as *propositions*. Even in meaning attributions, the practices of quantifying into the positions occupied by 'that'-clauses and trading on the repetition of a 'that'-clause as an identity clearly suggest a referential treatment of these expressions. And this, in turn, suggests a relational treatment of the expression 'means' in form MA meaning attributions.

However, we have also explained that this kind of meaning-entity realism faces serious metaphysical and epistemological problems (Armour-Garb and Woodbridge, *The Story*), which suggest that if we can do without a commitment to propositions, we should. Accordingly, we have argued that while our use of 'that'-clauses (what we call *proposition-talk*) appears to traffic in abstract entities of a certain sort, actually, no such entities exist. What we offer, as a resolution of these problems, is a PIF account of proposition-talk. Such an account vindicates our inferential and linguistic practices involving 'that'-clauses and carries with it a *linguistic commitment* to talking *as if* there were such entities, because of the expressive advantages provided by such talk, while at the same time avoiding any worrisome *ontological* commitment. The ontological commitment is bracketed by taking this way of talking to be

underwritten by pretense (Armour-Garb and Woodbridge, *The Story*). Before we turn to the basic of our account, let us return to Sellars's views and explain the external tension we find with them.

Previously, we noted that Sellars's views about 'that'-clauses and their role in meaning attributions can capture language picturing the world without the need for problematic abstract entities, such as propositions. While this is all for the better, what we find worrisome about Sellars's account is how difficult and how unnatural it is for him to wed his views with the inferential and linguistic practices that we engage in with meaning attributions, practices of the sorts exhibited in arguments (I)-(III), presented above.

To see the problem, consider argument (I). We can render the inference made there schematically as

- (A) S means that p.  
 So, S means something.  
 (i.e., There is something that S means).

Taking this inference schema at face value, and treating the 'that'-clause as a referring expression and 'means' as a relational predicate, a formal rendering of this inference schema might be given as

- (B)  $\frac{M(s, \langle p \rangle)}{\exists x M(s, x)}$

where 's' names, in the formal language, whatever sentence 'S' names, and the angle brackets function as described above. Here the inference is rendered into straightforward first-order logic as a clearly valid argument.

Now consider how argument (I) would be recast on Sellars's account, given his commitment against reading 'means' as a relational predicate. Since Sellars is trying to give an account of how speakers can attribute senses to sentences without incurring any ontological commitment to abstract meaning entities, he cannot accept (B) as a correct formal rendering of (A). At the same time, he would not wish to deny the validity of (A); hence, he must offer different forms for both the premise and the conclusion employed in that argument.

We will begin with the premise in (A). Since Sellars takes the application of dot-quotes to represent the form of a sentence like ‘S means that p’, he will likely take something like the following to capture the form of the premise,

$$(SMA_F) \forall y(y \in S \rightarrow y \in \bullet p \bullet).$$

Here, ‘S’ gets filled in by a particular expression for a class of lexical items, and ‘p’ gets filled in by a particular sentence from the speaker’s language, to form an expression specifying a functional class by displaying the functional role the items in the class share.

*Prima facie*, there is nothing the matter with  $(SMA_F)$ . Worries arise, however, when we consider how Sellars will have to represent the conclusion of (A). The idea behind the conclusion is to capture the thought that there is some functional classification that classifies all the Ss. Our conjecture is that, short of wheeling in second-order quantification, which we assume Sellars would not do (since it would commit him to properties), the only way to represent the conclusion, while preserving the validity of the argument and allowing that  $(SMA_F)$  represents the form of the premise, is to represent the conclusion as

$$(C) \exists x \forall y(y \in S \rightarrow \Sigma q(x = \bullet q \bullet \wedge y \in x)),$$

where ‘ $\Sigma$ ’ is the existential-substitutional quantifier, with a substitution class consisting of all the sentences of (all possible extensions of?) our language.

Here is the problem we see with this. Although we do not have a knock down argument against it, we find it implausible to take (C) to represent the *form* of the sentence, ‘There is something that S means’. We recognize that if one accepts the validity of argument (I) and that  $(SMA_F)$  represents the form of the premise, then (C), or something equally complicated, is needed to represent the form of the conclusion. As a result, we think that the implausibility of the claim that (C) represents the form of the conclusion sheds doubt on the claim the  $(SMA_F)$  actually represents the form of the premise. But, this, in turn, raises a question about whether Sellars’s employment of dot-quotes actually represents the form of any sentence with the surface form, ‘S means that p’. Accordingly, if argument (I) is valid and the charge of implausibility regarding (C) is correct, then there is a reason for rejecting, or resisting, Sellars’s non-relational account of form MA meaning attributions.

Now, perhaps Sellars would deny that he is in the business of representing the form of the sentences in such arguments. But if that is not what he is doing, then we are somewhat puzzled about this aspect of his project. At any rate, we do not think that Sellars would deny that he is representing the form of such sentences, and, so, we will say no more about this possible rejoinder. Instead, we claim that if there were *another* way of reading, or representing, the sentences in (A), a way that does not require representing the conclusion with something as protracted as (C), yet which still avoids incurring ontological commitments to meaning entities, etc., then it seems that that might be preferable to Sellars's non-relational analysis. This is precisely what we claim our PIF account of proposition-talk can do. We therefore now turn to a brief sketch of the account to show how.

#### IV

#### Propositions, meaning attribution, and pretense

For present purposes, we will focus on the elements of our PIF account of proposition-talk most relevant to the topics at hand, namely, the use of 'that'-clauses in MA form meaning attributions.<sup>23</sup> So, consider a meaning attribution of this sort, e.g.,

(7) 'Rauchen ist verboten' means that smoking is forbidden.

We explain this claim as operating in virtue of belonging to a pretense that involves stipulated background pretenses like the following.

#### Propositions game pretenses

The props for the pretense behind proposition-talk are 'that'-clauses and other expressions that can be replaced by them. The following are stipulated, background pretenses for the game.

'That'-clauses (and expressions they can be substituted for) are referring expressions.

- i) They denote mind- and language-independent abstract entities, called "propositions".
- ii) These entities are "fine-grained neo-Russellian" entities, i.e., they are structured entities built out of "worldly" elements.
- iii) Any 'that'-clauses whose embedded sentences employ

<sup>23</sup> For more details about the account as a whole, see Armour-Garb and Woodbridge (*The Story*) and (*Pretense and*).

- different expressions, or have different structures, (pretend-) denote distinct propositions.
- iv) The expression ‘means’ attributes a relation (called “meaning”) that certain things (e.g., certain linguistic expressions) can bear to propositions – a relation constituted by a particular isomorphism (“picturing projection”) holding between the elements of the meaningful item and the elements of the proposition.<sup>24</sup>

In addition to these background pretenses, as the propositions game pertains to MA form claims like (7), it also involves the following principle of generation that establishes further pretenses as also prescribed, depending on whether certain real-world conditions obtain.

(PGMA) The pretenses displayed in a utterance of  $\lceil$ (The sentence) S means that p $\rceil$  are prescribed iff the sentence S has a broadly inferential role similar to that of ‘p’-as-the-speaker-actually-uses-it (in the utterance).

Explaining a claim like (7) in terms of these rules for the pretenses that we hold the claim involves, takes the pretenses displayed in (7) as prescribed only under certain real-world conditions – those specified in

(8) ‘Rauchen is verboten’ has a broadly inferential role similar to that of ‘smoking is forbidden’-as-the-speaker-actually-uses-it (in the MA-form utterance).

So, given (PGMA), putting forward the pretenses displayed in (7) as prescribed commits one to the obtaining of the conditions specified in (8); a speaker can therefore use (7) to say indirectly what (8) says directly. Thus, what meaning attributions of form MA accomplish (indirectly) are the specification of a broadly inferential role and the functional classification of some specified linguistic item in terms of that role. This has a Sellarsian ring to it.

But this is not yet the whole story. As with Sellars’s dot-quotes, a major expressive advantage that proposition-talk provides, over an opaque descriptive specification of the broadly inferential role in question as in (8), is something akin to the collapse of the use/mention distinction. While sentences embedded in a ‘that’-clause are nominalized, they are

<sup>24</sup> The pretense about a picturing-projection isomorphism can be restricted to “elementary, matter-of-factual” claims, as Sellars seems to prefer doing.

still *used*. So the meaning of such a sentence is still in play; its broadly inferential role is engaged or active. What the pretense of proposition-talk does is put that role on display via a kind of deferred ostension. Then, through the operation of the pretense of there being a proposition serving as the relatum of the relations 'means', this allows us to pick out that role and attribute it to some linguistic expression. This captures the "illustrating" aspect of the operation of Sellars's dot-quotes. So the pretense account of form MA meaning attributions has much of the same upshot as the Sellarsian non-relational/specialized copula account.

Moreover, any intuition one has that language pictures the world *via* an isomorphism can be written right into the pretense about the expression 'means'. Since the pretense also takes propositions to be neo-Russellian entities, an isomorphism to the elements of a proposition meant is an isomorphism with elements of the world. While this is just a pretense on our account, it still avoids linguistic idealism in the same way that the use of dot-quotes does, but unlike the dot-quotes account, our view treats the logic of meaning attributions just what it appears to be on the surface. This accommodates and vindicates our inferential and linguistic practices with 'that'-clauses and the expression 'means'. One might conclude, therefore, that a pretense account of meaning attributions should be attractive to someone with Sellarsian aspirations.

## V

### Sellars's second sense of correspondence

Now that we have an account of meaning attribution on the table, we can explain how we can accommodate Sellars's intuitions about language picturing the world (avoiding linguistic idealism) and the truth of basic matter-of-factual empirical claims involving correct picturing, understood in terms of an isomorphism, all while avoiding the challenges of Putnam's model-theoretic argument without appealing to any "anti-nominalist" factors.

Sellars's second sense of correspondence — the one involving picturing — seems to be the sort that T-deflationists explicitly reject (or: want nothing to do with) (Horwich, *Truth Second* 116-17). However, the combination of our PIF account of proposition-talk and our PIF account of truth-talk can even vindicate (in a sense) this second, stronger version of the correspondence intuition. This is because the "thick" version of the correspondence intuition — including Sellars's talk of picturing in virtue of



an isomorphism – can effectively be written into the stipulated grounding pretenses of the account (of meaning attribution in particular). Thus, it is just part of the pretense that when some elementary, matter-of-factual statement expresses some (neo-Russellian) proposition, it does so in virtue of (i) a projection from the names employed in the statement to the objects involved in the proposition, and (ii) a particular isomorphism between the linguistic picture and the proposition. This captures a robust intuition about language picturing the world and truth being correct picturing. At the same time, our PIF account of proposition-talk allows us to capture the “thinner” notion of picturing that Sellars’s dot-quotes can provide, but in a way that avoids the external tension confronting Sellars’s views. In this way, our PIF accounts can accommodate the second sense of correspondence Sellars attributes to truth, in addition to accommodating his first sense of correspondence.

Finally, it should be clear that our PIF accounts satisfy Sellars’s nominalist scruples, since the property of truth, the meaning relation, and propositions are all only pretend – none of them really exist on our view. At the same time, however, we can deflect the challenges posed by Putnam’s model-theoretic argument – the challenge of there being “too many isomorphisms”. While the pretense that underwrites meaning attribution does portray there being a particular isomorphism involved in linguistic picturing (so that the correct picturing that makes for the truth of elementary, matter-of-factual statements involves that particular isomorphism rather than any other), this does not saddle us with a problem. On our view, all of this is a stipulated pretense. We can, if need be, accept the model-theoretic argument’s conclusion that really there are too many isomorphisms and no way to single one out as the basis of any linguistic picturing relation. The claim that there is one isomorphism or projection that is special is just a pretense. Since, on our PIF views, there actually is no property of truth and no meaning relation, there can be no requirement of finding a way of picking out one of the isomorphisms as capturing the truth of elementary, matter-of-factual claims to the exclusion of all the others. We can respond to Putnam’s challenges with a shrug.

## VI

### The engagement complaint and a Sellarsian response

We will now consider how Sellars can help us with an objection that is often launched at certain fictionalist accounts – what we call the



*Engagement Complaint* (henceforth *EC*). Although the objection can apply to any non-revolutionary fictional account, we will focus on the objection as it pertains to our PIF account. Here is the objection: It is absurd and implausible in the extreme to suppose that, in general, speakers who make meaning attributions or use proposition-talk ('that'-clauses) are engaged in pretending and are *aware* of any pretense. But PIF accounts of this way of talking seem to require that speakers are aware of and allude to the pretense that such accounts describe. Thus, such accounts are implausible.<sup>25</sup>

While we have argued elsewhere that *EC* undermines a number of PIF accounts,<sup>26</sup> it does not apply to our own favored pretense account. This is so, at least in part, because our pretense analysis is not intended as an account of speakers' attitudes or activities. We do not claim, for example, that when speakers make meaning attributions or otherwise employ 'that'-clauses, they are thinking of themselves as pretending. Nor do we assume that they are alluding to or are even aware of any pretense at all, when they are engaged in proposition-talk. Indeed, such speakers need not be aware of any pretense at all in order to use that talk. On our view, pretense is part of the explanation as to how the talk functions semantically; it does not enter as part of an account of what speakers intend to do.

Why think, then, that *EC* applies to our particular semantic pretense account of proposition-talk (or, for that matter, to any discourse that we provide a semantic pretense account)? We think that *EC* seems applicable because of a misunderstanding about how our pretense account works. More specifically, the *Engagement Complaint* comes in because objectors assume that we accept

(EC1) Competent users of a fragment of discourse the functioning of which involves pretense, are making as if they are engaged in some game of make-believe involving the central locutions of that fragment of discourse.

If we accepted (EC1), the objector would be right. But we don't accept (EC1). Rather, we accept

<sup>25</sup> Versions of this objection appear in Richard and Stanley.

<sup>26</sup> In particular, it undermines the accounts developed in Kroon (*Descriptivism*) and (*Existence in*). See Woodbridge and Armour-Garb (*Linguistic puzzles*) and Armour-Garb and Woodbridge (*The Story*) and (*Pretense and*).

(EC2) *It is as if* competent users of a fragment of the language are engaged in a game of make-believe with respect to a fragment of discourse that we analyze with a PIF account.

As is clear, given (EC2), there is simply no reason for thinking that we are attributing any kind of pretense awareness, let alone engagement, to competent users of the relevant discourse. Given that we accept (EC2) but not (EC1), it is clear, therefore, that the Engagement Complaint is *misdirected*, if aimed at our particular semantic pretense account.

So where does pretense come in on our PIF account? Although speakers need not engage in the pretense operating in some way of talking, in order to use that talk fully competently, certain *theorists*, who are offering an account of that fragment of discourse, will mention pretense, in order to explain which claims about the world its instances make (and how they do this). But no one needs to engage in, or even be aware of, the games of make-believe that figure in the explanation of how pretense-involving ways of talking function in order to be a fully competent speaker with the relevant fragments of discourse. Thus, we accept something a bit more specific than (EC2), *viz.*,

(EC2\*) *From the perspective of a language theorist*, it is as if competent users of a fragment of the language are engaged in a game of make-believe with respect to a fragment of discourse that we analyze with a PIF account.

Of course, a theorist, who is theorizing about the relevant discourse (the “talk”), may talk about pretense linking up utterances with their serious content, but at no point do speakers have to be engaged in any such pretense. Just as speakers need not be aware of deference to experts securing reference for natural kind terms, they do not even have to be aware of such a pretense, nor do they even have to be disposed to acknowledge that they are engaged in a sort of pretense if asked.<sup>27</sup> It is also compatible with our PIF accounts that ordinary speakers do not even take an attitude towards their use of truth-talk or proposition-

---

<sup>27</sup> Nor need speakers intentionally allude to or even be aware of any pretense at work in any talk for which we want to give a PIF account. Still, at the level of *theorizing* about how the talk functions, we should understand these fragments of discourse to be “as if” ways of talking that involve a systematic dependency on how things actually are. And we should explain the kind of “as ifness” involved in terms of pretense, specifically, games of make-believe.

talk, or towards whether anything in proposition-talk commits them to any sort of *as ifness* (or to whether they feel – or even *can* feel – that the propriety of their stance depends on anything approaching a belief that S). As such, our pretense account avoids EC by keeping all reference to pretense within the *theorist's explanation* of the claim's semantics and of how it ends up with the content it has.

Now, while we believe that our reply to EC is sufficient, as we will now show, Sellars's nuanced analysis of linguistic rules can strengthen our reply to the Engagement Complaint. In laying out his distinction between Rules of Criticism ("ought-to-be's" or OTBs) and Rules of Action ("ought-to-do's" or OTDs), Sellars explains how the Rules of Criticism, which govern linguistic behavior, need *not* be acknowledged by the subjects to whom the rules apply (Sellars, *Language as* 509). In fact, even though subjects might not have the conceptual repertoire or recognitional capacities needed to understand the rules, they can still manage to act in accordance with them and to engage in the right sort of rule-governed behavior. A subject who is in this position will not count as full-fledged language-user, but she can still count as a "user-in-training" and she will still be a full member of the linguistic community. Moreover, through the influence of various language-trainers, her linguistic behavior can *become* pattern-governed and conform to the Rules of Criticism (Ibid. 511-13). In this way, Sellars explains how people can act in accordance with certain rules, and have them govern their behavior, even when they have no awareness of those rules.

This insight is helpful, when we think about EC. Sellars notes that language-trainees need not have any conception of, let alone be intentionally following, the semantic rules, i.e., Rules of Criticism, to which their pattern-governed behavior conforms. While we recognize that a *full-fledged language-user* must be aware of, and intentionally follows, the Rules of Criticism that govern a fragment of discourse, on our pretense account, in following those rules, such a language user could, with a complete lack of awareness, also be acting in accord with another set of rules – ones that govern a possible game of make-believe, say – which serve to *explain* why the Rules of Criticism that govern the discourse she uses are the particular rules they are. By engaging in pattern-governed behavior that accords with the rules of a possible game of make-believe, a language-user can implement the resulting semantic mechanisms that explain the Rules of Criticism, again, without awareness of how those "ought-to-be's" result from those semantic

mechanisms. Her lack of awareness of any such explanations would have no impact on her capability to employ the language as a full member of her linguistic community. It is only language theorists who seek an explanation of why the Rules of Criticism that govern the language are what they are, so that is the only level of engagement with the language that would require awareness of any pretense at work in its linguistic functioning. Graphically, we might represent this relation as follows.

Language Trainee-----Language User-----Language Theorist  
(Accords with OTBs) (Follows OTBs (+ OTDs)) (Explains OTBs)

### Concluding remarks

As we have seen, we can helpfully augment a Sellars's understanding of truth, by bringing in our own PIF accounts of truth-talk and meaning attribution. And, looking at the combination from the other direction, we have seen that a Sellarsian analysis of the nature of linguistic rules can help a pretense account deflect a complaint often launched at pretense analyses of not explicitly figurative fragments of discourse. We conclude that Sellarsian views and pretense accounts combine well.

### BIBLIOGRAPHICAL REFERENCES

Armour-Garb, B. "Deflationism (about theories of truth)". *Philosophical Compass*. Apr. 2012: 267-277. Print.

Armour-Garb, B. and J. Woodbridge. "Why deflationists should be pretense theorists (and perhaps already are)". Pedersen, N. and C. Wright (eds.). *New Waves in Truth*. Basingstoke and New York: Palgrave Macmillan, 2010. Print.

---. "The story about propositions". *Noûs*. Dec. 2012: 635-674. Print.

---. "From mathematical fictionalism to truth-theoretic fictionalism". *Philosophy and Phenomenological Research*. Forthcoming.

---. *Pretense and pathology*. Cambridge: Cambridge University Press, 2014. Forthcoming.

Crimmins, M. "Hesperus and Phosphorus: Sense, pretense, and reference". *Philosophical Review*. Jan. 1998: 1-48. Print.

Devitt, M. "The metaphysics of truth". M. Lynch (ed.). *The nature of truth*. Cambridge: MIT Press, 2001. 579-611. Print.

Egan, A. "Pretense for the complete idiom". *Noûs*. Sep. 2008: 381-409. Print.

Evans, G. "Existential statements". J. McDowell (ed.). *Varieties of reference*. Oxford: Clarendon Press, 1982. 343-369. Print.

Field, H. "Critical notice: Paul Horwich's Truth". *Philosophy of Science*. Jun. 1992: 321-330. Print.

---. "Deflationist views of meaning and content". *Mind*. Jul. 1994: 249-285. Print.

Glymour, Clark. "Conceptual scheming or confessions of a metaphysical realist". *Synthese*. May. 1982: 169-180. Print.

Horwich, P. *Truth*. First edition. Oxford: Clarendon Press, 1990. Print.

---. *Truth*. Second edition. Oxford: Clarendon Press, 1998. Print.

Kripke, S. *Wittgenstein on rules and private language*. Cambridge, MA: Harvard University Press, 1982. Print.

Kroon, F. "Characterizing non-existents". *Grazer Philosophische Studien*. Jan. 1996: 163-193. Print.

---. "Fictionalism and the informativeness of identity". *Philosophical Studies*. Dec. 2001: 197-225. Print.

---. "Descriptivism, pretense, and the Frege-Russell problems". *Philosophical Review*. Jan. 2004: 1-30. Print.

---. "Existence in the theory of definite descriptions". *Journal of Philosophy*. Jul. 2009: 365-389. Print.

Leeds, S. "Truth, correspondence, and success". *Philosophical Studies*. Jul. 1995: 1-36. Print.

Lewis, D. "Putnam's paradox". *Australian Journal of Philosophy*. Oct. 1984: 221-236. Print.

O'Shea, J. *Wilfrid Sellars: Naturalism with a normative turn*. Cambridge: Polity Press, 2007. Print.

Putnam, H. *Meaning and the moral sciences*. London: Routledge & Kegan Paul, 1978. Print.

---. "Models and reality". *The Journal of Symbolic Logic*. Sep. 1980: 464-482. Print.

Quine, W. V. O. *Philosophy of logic*. Cambridge: Harvard University Press, 1986. Print.

Richard, M. "Semantic pretense". Everett, A. and T. Hofweber (eds.). *Empty names, fiction, and the puzzles of non-existence*. Stanford: CSLI Publications, 2000. Print.

Schiffer, S. "Language-created, language-independent entities". *Philosophical Topics*. Jan. 1996: 149-167. Print.

---. "Pleonastic fregeanism". A. Kanamori (ed.). *Analytical Philosophy and Logic: The Proceedings of the Twentieth World Congress of Philosophy*, Vol. 6. Ohio: Bowling Green Documentation Center, 2000. 1-15. Print.

Sellars, W. "Truth and 'Correspondence'". *Journal of Philosophy*. Jan. 1962: 29-56. Reprinted in Sellars (1963). 197-224. Print.

---. *Science, perception and reality*. London: Routledge & Kegan Paul, 1963. Print.

---. "Language as thought and as communication". *Philosophy and Phenomenological Research*. Jun. 1969: 506-527. Print.

---. "Meaning as functional classification". *Synthese*. Jul. 1974: 417-437. Print.

---. *Naturalism and ontology*. Atascadero: Ridgeview Publishing Company, 1979. Print.

Stanley, J. "Hermeneutic fictionalism". *Midwest Studies in Philosophy Figurative Language*. Dec. 2002: 36-71. Print.

Tarski, A. "The semantic conception of truth". *Philosophy and Phenomenological Research*. Mar. 1944: 341-376. Print.

van Fraassen, B. "Putnam's paradox: Metaphysical realism revamped and evaded". *Noûs*. Jun. 1997: 17-42. Print.

Walton, K. *Mimesis as make-believe*. Cambridge, Mass.: Harvard University Press, 1990. Print.

Williams, M. "Do We (Epistemologists) need a theory of truth?" *Philosophical Topics*. Jan. 1986: 223-242. Print.

Wittgenstein, L. *Tractatus logico-philosophicus*. London: Routledge & Kegan Paul, 1974. Print.

---. *Philosophical investigations*. 3<sup>rd</sup> edition. Oxford: Basil Blackwell, 1978. Print.

Woodbridge, J. "Truth as a pretense". M. Kalderon (ed.). *Fictionalism in metaphysics*. Oxford: Oxford University Press, 2005. Print.

---. "Propositions as semantic pretense". *Language & Communication*. Jul. 2006: 343-355. Print.

Woodbridge, J. and Armour-Garb, B. "Linguistic puzzles and semantic pretense". S. Sawyer (ed.). *New waves in Philosophy of Language*. Basingstoke and New York: Palgrave Macmillan, 2009. Print.

Yablo, S. "How in the world?" *Philosophical Topics*. Apr. 1996: 255-286. Print.

---. "Go figure: A path through fictionalism". *Midwest Studies in Philosophy XXV: Figurative Language*. Dec. 2001: 72-102. Print.

---. "The myth of the seven". M. Kalderon (ed.). *Fictionalism in metaphysics*. Oxford: Oxford University Press, 2005. Print.





# BETTER TO BE THAN NOT TO BE?

*¿ES MEJOR SER QUE NO SER?*

GUSTAF ARRHENIUS

Stockholm University, Sweden. [gustaf.Arrhenius@philosophy.su.se](mailto:gustaf.Arrhenius@philosophy.su.se)

WŁODEK RABINOWICZ

Lund University, Sweden. [wlodek.Rabinowicz@fil.lu.se](mailto:wlodek.Rabinowicz@fil.lu.se)

RECIBIDO EL 15 DE AGOSTO DE 2012 Y APROBADO EL 17 DE OCTUBRE DE 2012

## RESUMEN    ABSTRACT

¿Puede ser mejor (o peor) para mí existir que no existir? Diversos filósofos lo han negado, en razón de que si lo fuera, entonces, si yo no existiera, habría sido peor (mejor) para mí, lo cual es un absurdo. En este artículo argumentamos que dichos filósofos están equivocados: las afirmaciones sobre el valor o la falta de valor comparativo de la existencia no tienen por que implicar un absurdo. Estas afirmaciones, que son de vital importancia para la ética de las poblaciones, así como para el status de la denominada "restricción de la persona que afecta" pueden racionalizarse si nos adherimos al llamado análisis de valor de las actitudes que se condicionan.

Can it be better (or worse) for me to exist than not to exist? Several philosophers have denied this, on the ground that if it could, then if I didn't exist, this would have been worse (better) for me, which is absurd. In our paper we argue that these philosophers are mistaken: Claims about the comparative value or disvalue of existence need not imply any absurdities. Such claims, which are of central importance for population ethics and for the status of the so-called Person-Affecting Restriction, can be rationalized if one adheres to the so-called fitting-attitudes analysis of value.

## PALABRAS CLAVE    KEY WORDS

relación de mejor condición, persona que afecta la restricción, mejor condición personal, análisis de valor de las actitudes que se condicionan, valor de la existencia.

betterness relation, person-affecting restriction, personal betterness, fitting-attitudes analysis of value, value of existence.

“Not to be born at all is best, far best that can befall.” Sophocles (*Oedipus at Colonus* 1224).<sup>1</sup>

“Gut ist der Schlaf, der Tod ist besser – freilich das beste wäre, nie geboren sein.” Heinrich Heine (*Morphine* 15-16).

“Life is so terrible, it would have been better not to have been born. Who is so lucky? Not one in a hundred thousand!” Old Jewish saying.

## I. Introduction

Can it be better or worse for a person to be than not to be, that is, can it be better or worse for her to exist than not to exist at all? This old and challenging philosophical question, which we can call the existential question, has been raised anew in contemporary moral philosophy. There are roughly two reasons for this renewed interest. Firstly, traditional so-called “impersonal” ethical theories, such as utilitarianism, have paradoxical and very counterintuitive implications in regard to questions concerning procreation and our moral duties to future, not yet existing people. Secondly, it has seemed evident to many that an outcome can only be better than another if it is better for someone, and that only moral theories that are in this sense “person affecting” can be correct. The implications of this so-called Person Affecting Restriction will differ radically, however, depending on which answer one gives to the existential question.

Hence, many of the problems regarding our moral duties to future generations turn around the issue of whether existence can be better or worse for a person than non-existence. Some think so, others adamantly deny it. Sigmund Freud, for instance, described the Jewish saying we have quoted above as a “nonsensical joke”.<sup>2</sup> Others, as illustrated by the quotes from Sophocles and Heine above, seem to have a different view. Thus, for example, Melinda Roberts (“Can it”) and Matthew Adler have defended an affirmative answer to the existential question. Contrariwise,

<sup>1</sup> Translated by F. Storr. (London, Heinemann; New York, Macmillan, 1912-13).

<sup>2</sup> Freud (57) quoted after Benatar (3). Freud tried to account for the nonsensicality of the joke by this observation: “who is not born is not a mortal man at all, and there is no good and no best for him.” (ibid.). His suggestion thus seems to be that the existential question requires a negative answer. However, the joke would still of course be nonsensical (and for that reason funny) even if the existential question were answerable in the affirmative.

Derek Parfit, John Broome (*Ethics out*), and others have worried that if we take a person's life to be better for her than non-existence, then we would have to conclude that it would have been worse for her if she did not exist, which is absurd: Nothing would have been worse or better for a person if she had not existed.

We shall suggest an answer to the existential question to the effect that one can claim that it is better or worse for a person to exist than not to exist, without implying any absurdities. First, however, we shall explain in more detail why this question has again moved to the forefront of moral philosophy. We shall then discuss some of the proposed answers in the literature and our own suggestion. Lastly, we shall consider and rebut some possible objections to our position.

## II

### The 'Person Affecting Restriction' and the 'Existential Question'

The Person Affecting Restriction, put as a slogan, states that an outcome can only be better than another if it is better for someone. The restriction has a strong intuitive appeal and it has been suggested that it is presupposed in many arguments in moral philosophy, political theory, and welfare economics.<sup>3</sup> Moreover, several theorists have argued that the counterintuitive implications in population ethics of so-called "impersonal" welfarist theories could be avoided by adopting the restriction. This applies in particular to the well-known Repugnant Conclusion, which – as has been pointed out by Parfit – is entailed by classical utilitarianism. (Parfit 388)<sup>4</sup>

It is not easy to discern what exactly the distinction between "impersonal" and "person affecting" theories amounts to in the literature, partly

<sup>3</sup> Temkin (*Inequality, "Harmful Goods..."*). The term "Person Affecting Restriction" was introduced by Glover (66), but see also Narveson.

<sup>4</sup> For an overview of these counterintuitive implications, see (Arrhenius et al., *The Repugnant*) and (Arrhenius, *Future generations; Population*). The Repugnant Conclusion is the claim that for any world inhabited by people with very high welfare, there is a possible world in which everyone has a life that is barely worth living which is *better*, other things being equal. Imposing the Person Affecting Restriction can block the derivation of the Repugnant Conclusion only if it is conjoined with a negative answer to the existential question. Then it is arguable that a world in which everyone has a life barely worth living cannot be better than a world consisting of individuals with very high quality of life, since the former is not better for anyone, not even for the people who exist in the former but not in the latter world. Since we are going to argue that the existential question should be answered in the affirmative, however, we are sceptical about this manoeuvre. Making population ethics more 'person affecting', so to speak, does not suffice to save it from counterintuitive implications, see (Arrhenius, "Can the person"; *Population*).

because different authors have had a different take on the distinction and partly because other ideas have been conflated or mixed with the Person Affecting Restriction. As has been shown elsewhere, one can interpret the restriction in a manner which makes it perfectly compatible with impersonal welfarist theories such as classical utilitarianism (Arrhenius, *The Person; Can the Person; Population*). Thus, it could be understood as an idea about what kind of facts moral goodness supervenes on, for example, that goodness is essentially related to the interests of human beings. We are, however, interested in a stronger reading of the restriction which stresses the individualist aspect of value even more by claiming that morality is essentially *person comparative*:

*The Person Affecting Restriction*: If an outcome A is better (worse) than B, then A is better (worse) than B for at least one individual.<sup>5</sup>

This is the principle that we shall henceforth refer to as the Person Affecting Restriction (or “the restriction” for short). In cases involving only the same people in the compared outcomes, this view is quite straightforward and, we surmise, widely accepted.<sup>6</sup>

In comparisons between outcomes involving different people, however, and in particular in cases involving people whose existence is contingent on our choices, the restriction becomes ambiguous. An outcome A is better than B for Peter if Peter has a higher welfare in A as compared to B. We can assume that much. But what if Peter exists in outcome A but not in

---

<sup>5</sup> An interesting question is whether the restriction should be supplemented with a person affecting necessary condition for outcomes being equally good. We would suggest the following condition: If outcome A is equally as good as B, then either A and B are equally as good for at least one individual, or A is better (worse) for at least one individual and B is better (worse) for at least one individual. What if both A and B are empty worlds? We think that it is in the spirit of the person affecting idea that such worlds are not ranked as equally good but rather that they completely lack value from a person affecting perspective.

<sup>6</sup> The term “Person Affecting Restriction” might be misleading, since many theorists would, sensibly we think, weaken the restriction to also include other sentient beings. Cf. Holtug (*In defence*). Notice that since the Person Affecting Restriction is formulated without any *ceteris paribus* clause, value pluralists are not likely to accept it since it leaves little room for other values apart from welfarist ones. For instance, one might believe in some non-welfarist values such as virtue, reward in accordance to desert (Cf. Feldman *Adjusting utility; Utilitarianism*), beauty (cf. Moore, section 50), variety of natural species, or what have you (for a general discussion of value pluralism, see Rabinowicz & Rønnow-Rasmussen, (*The strike*); for a discussion of this issue in connection with the Person Affecting Restriction, see Arrhenius, (*The Person; Can the Person; Population*). Moreover, certain welfarist theories might also be ruled out by the restriction, such as some versions of welfarist egalitarianism (Arrhenius, *Can the Person; Population*). However, we shall only discuss implications of the restriction in cases where one can assume that other values are not at stake. Hence, the arguments below also apply to a *ceteris paribus* version of the restriction.

outcome B? Is A then better than B for Peter? More generally, can existence be better or worse for a person than non-existence? In other words, what is the correct answer to the existential question? Hence, depending on the answer to the existential question we get different versions of the Person Affecting Restriction and very different implications regarding how to morally evaluate different possible futures.

### III

#### Neither Better nor Worse to Be than Not to Be

A popular answer to the existential question is to claim that existence cannot be better or worse than non-existence for a person, nor equally as good for that matter, since existence and non-existence are, in some sense, incomparable in value for a person. David Heyd seems to endorse this position when he argues against the view that existence could be worse than non-existence by claiming that such a view “is inconsistent with a person-affecting theory as it presupposes the comparability of non-existence with life of a certain quality” (Heyd, *Procreation* 161).<sup>7</sup>

In his early pioneering work in population ethics, Narveson seems to share Heyd’s concern, although he formulates it in terms of happiness comparisons rather than comparisons in value:

If you ask, “whose happiness has been increased as a result of his being born?”, the answer is that nobody’s has. [...] Remember that the question we must ask about *him* is not whether he is happy but whether he is happier as a result of being born. And if put this way, we see that again we have a piece of nonsense on our hands if we suppose the answer is either “yes” or “no”. For if it is, then with whom, or with what, are we comparing his new state of bliss? Is the child, perhaps, happier than he used to be before he was born? Or happier, perhaps, than his alter ego? Obviously, there can be no sensible answer here. (Narveson 67)<sup>8</sup>

<sup>7</sup> See also (Heyd, *Genethics* 124-5). Heyd states that his view is “grounded in an ‘anthropocentric’ conception of value according to which value is necessarily related to human interests, welfare, expectations, desires and wishes – that is to say to human volitions” (Heyd, *Procreation* 164).

<sup>8</sup> Cf. Dasgupta (383): “Recall our definition of the zero level of well-being. This isn’t a standard arrived at through a comparison with ‘non-existence’. Such comparisons can’t be made. The ‘unborn’ aren’t a class of people. It makes no sense to attribute a degree of wellbeing, low or high or nil, to the ‘state of not being born’”.

Similarly, Alan Buchanan et al. claim that “when the alternative is nonexistence, there is no individual who is made worse off by being conceived and born” (Buchanan et al. 234)<sup>9</sup> and John Broome states that “...it cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all”. (Broome, *Ethics out* 168)<sup>10</sup>

The negative answer to the existential question in combination with the Person Affecting Restriction has such counterintuitive conclusions that it is hard to believe that anyone would seriously endorse the conjunction of these two views. Consider the Future Bliss or Hell Case:

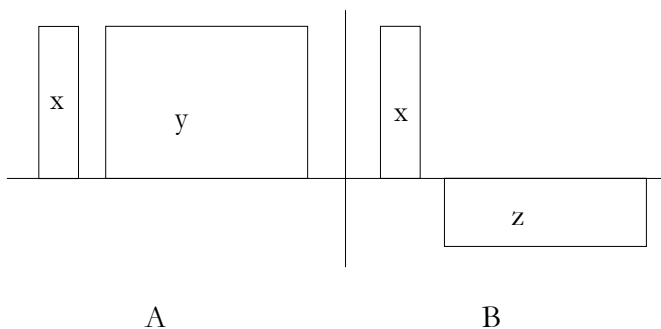


Diagram 1

The blocks in the above diagram represent populations. The width of each block represents the number of people in the corresponding population, whereas the height represents their welfare. Assume that we can either see to it that all the people in the future have excellent lives (the y-people in outcome A) or that they have hellish lives (the z-people in outcome B). Assume further that these two possible futures consist of different but the same number of people and that these two outcomes are equally good for us, the present x-people.

Most evaluators, we presume, would consider outcome A clearly superior to outcome B and agree that we ought to realise A rather than B. However, since the y- and z-people are uniquely realisable (i.e. exist

<sup>9</sup> This passage continues: “Nonexistence is not a condition that is better for an individual only in rare cases like having Lesh-Nyhan or Tay Sachs disease; it is no condition at all, and so it is not better or worse than any other condition.” (Buchanan et al. 234)

<sup>10</sup> See also Parfit (395, 489).

in just one of the two outcomes), the negative answer to the existential question implies that outcome A is neither better nor worse for the y- and z-people as compared to B. Moreover, the two outcomes are equally good for the x-people. Hence, according to the Person Affecting Restriction, A cannot be better than B since it is not better for any individual. Nor is of course B better than A. Consequently, if combined with the negative answer to the existential question, the Person Affecting Restriction ranks these outcomes as either equally good or as incomparable in value.<sup>11</sup> But that is clearly the wrong diagnosis of the Future Bliss or Hell Case.

This and other counterintuitive implications of the Person Affecting Restriction in combination with the negative answer to the existential question have led philosophers to abandon the restriction (the majority) or to accept not only that existence can be better or worse for a person than non-existence but also that a non-existent person has a certain welfare level (namely, zero welfare) and that, consequently, non-existence can be better or worse for that non-existent person than a life at some specified level of welfare. As we shall show, both of these moves are uncalled for.

#### IV

### The Argument from Absurdity

What is the reason behind the negative answer to the existential question? Well, one worry seems to be that if we give an affirmative answer to the existential question, that is, if we take a person's life to be better or worse for her than non-existence, then we would have to conclude that *it would have been worse or better for her if she did not exist*, which is obviously absurd: Nothing would have been worse or better

<sup>11</sup> The Person Affecting Restriction coupled with a negative answer to the existential question yields a position close to what we call *Strict Comparativism*: When comparing outcomes, one should only count the welfare of people who exist in both of the outcomes that are being compared and completely disregard the welfare of people who only exist in one of them. This seems to be, for example, Broome's take on the restriction: "Suppose [an alternative X] contains a certain number of people, and [an alternative Y] contains all the same people, and some more as well ... Then [the person-affecting view] is that [X] is at least as good as [Y] if and only if it is at least as good for the people who exist in both" (Broome, *Counting* 124). Broome rejects the restriction understood in this way, but Heyd seems to accept it, since he argues that "[e]xcluding the welfare and interest of future merely possible persons ... is a necessary consequence of a coherent person-regarding theory of value" (Heyd, *Procreation* 159-61); see also (Heyd, *Genethics* 124-5). See Arrhenius (*The Person; Can the Person; Population*) for a discussion of different versions of Comparativism.



for a person if she had not existed.<sup>12</sup> This argument is eloquently stated by Broome:

...[I]t cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would have been worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her. (*Ethics out* 168)<sup>13</sup>

However, this *Absurd Conclusion* (italicized above) does not follow. A triadic relation consisting in one state (having a certain life) being better for a person *p* than another state (non-existence) cannot hold unless all its three relata exist. Now, the states in question are abstract objects and thus can be assumed to exist even if they do not actually obtain. Consequently, the triadic relation in question can indeed hold as long as *p* exists. However, if a person is a concrete object, which is the received view (and, we surmise, the correct one), then this relation could not hold if *p* didn't exist, since the third relatum, *p*, would then be missing.<sup>14</sup>

---

<sup>12</sup> Rabinowicz (*Broome and* fn. 2), ascribes this worry to Derek Parfit (1984), who writes: "in being caused to exist, someone can be benefited. (...) we need not claim that this outcome is *better* for that person than the alternative. This would imply the implausible claim that, if that person had never existed, it would have been worse for this person". (395, cf. also *ibid.* 189 489) Parfit asserts, however, that causing someone to exist still can be *good* for the person in question. Good, but not better. It seems to us that Parfit needn't have been so cautious. His discussion of the matter contains all that is needed for the bolder betterness claim (see below).

<sup>13</sup> Notice that this argument, if correct, would also work equally well against the idea that existence could be worse for someone than non-existence: If it were worse for a person that she exists than that she should never have existed, then it would have been better for her if she had never existed. But if she had never existed, then there would have been no her for it to be better for, so it could not have been better for her. Thus, it cannot be true that it could be worse for a person to exist than not to exist.

<sup>14</sup> On the other hand, if, contrary to the received view, a person were itself constructed as a collection of (abstract) states of affairs, then it would be correct to say that she would exist, as an abstract object, even if she didn't 'obtain', so to speak. Hence, one might then say that there is nothing absurd in claiming that, if she didn't obtain, this state could have been worse for her than her actual state, since all three relata would then exist as abstract objects. However, this interpretation of persons as abstract objects is a view that few philosophers would be prepared to accept.

It should be noted that in this paper we simply presuppose that relations require the existence of relata. This presupposition could, of course, be questioned. (We are indebted to Staffan Angere for pressing this point.) Thus, for example, it could be argued that intentional attitudes are best interpreted as relations between the subject and the intentional object, which would mean that a relation after all can obtain even when one of the relata (the intentional object) happens not to exist. On our view, such a relational account of intentional attitudes is unsatisfactory. An ontology that allows obtaining of the relations without existing relata leads to a whole host of problems, but the discussion of these issues would take us too far afield.



Consequently, even if it is better for  $p$  to exist than not to exist, assuming she has a life worth living, it doesn't follow that it *would have been worse* for  $p$  if she did not exist, since one of the relata,  $p$ , would then have been absent. What does follow is only that non-existence *is* worse for her than existence (since 'worse' is just the converse of 'better'), but not that it *would have been worse* if she didn't exist. Hence, Broome's argument is a non-sequitur and the Absurd Conclusion doesn't follow from the idea that existence can be better or worse for a person than non-existence.<sup>15</sup>

It might be that Broome assumes that the following general principle is true:

*Subjunctive Connection 1 (SC1):* An outcome A is better (worse, equally as good) for  $p$  than (as) another outcome B if and only if outcome B would be worse (better, equally as good) for  $p$  than (as) A if B came about.<sup>16</sup>

Krister Bykvist has suggested a similar principle which he calls *Accessibility*: "If A is better (worse) for S than B, then A would be better (worse) for S than B even if A obtained." (Bykvist 348)<sup>17</sup> However, as we pointed out above, it doesn't follow logically from "it is better for  $p$  to exist than not to exist" that "it would have been worse for  $p$  if she did not exist" since in the latter case one of the relata,  $p$ , would be absent. Moreover, it seems clear that SC1 is false, *mutatis mutandis*, for related concepts, such as "considered better by" / "would be considered worse by" and "preferred by" / "would be dispreferred by".<sup>18</sup> So it is not clear

<sup>15</sup> Rabinowicz suggested this argument already back in 2000 in personal conversation with Arrhenius, Broome, Bykvist and Erik Carlson at a workshop in Leipzig; and he has briefly presented it in Rabinowicz (*The Size* 29), and in more detail in Rabinowicz (*Broome and 2*). For a similar argument, see Arrhenius (*Population Axiology* 158), who suggests that an affirmative answer to the existential question "only involves a claim that if a person exists, then she can compare the value of her life to her non-existence. A person that will never exist cannot, of course, compare "her" non-existence with her existence. Consequently, one can claim that it is better ... for a person to exist ... than ... not to exist without implying any absurdities". In fact, even though he accepted the negative answer to the existential question (and instead went for the view that it can be good but not better for a person to exist than not to exist), Parfit (1984) came very close to making the same point as we make when he observed that there is nothing problematic in the claim that one can benefit a person by causing her to exist: "In judging that some person's life is worth living, or better than nothing, we need *not* be implying that it would have been worse for this person if he had never existed. --- Since this person *does* exist, we can refer to this person when describing the alternative [i.e. the world in which she wouldn't have existed]. We know who it is who, in this possible alternative, would never have existed" Parfit (487-8), emphasis in original; see also Holtug (*On the Value*), Bykvist and Johansson.

<sup>16</sup> As we have seen, the only-if part of this principle is assumed by Parfit (395, 489).

<sup>17</sup> Symbols have been changed in this quotation, for the sake of consistency.

<sup>18</sup> This is acknowledged by Bykvist (349).

to us why one should go for SC1 rather than for the following connection between “better for” and “would be worse for”:

- (i) *Subjunctive Connection 2 (SC2)*: (i) If a person  $p$  exists in both outcomes A and B, then A is better (worse, equally as good) for  $p$  than (as) B if and only if B would be worse (better, equally as good) for  $p$  than (as) A, if B obtained.
- (ii) If a person  $p$  exists in A but not in B, then A can be better (worse, equally as good) for  $p$  than (as) B although B would not be worse (better, equally as good) for  $p$  than (as) A, if B obtained.

Of course, one might find SC1 more attractive than SC2, perhaps because one finds it more in line with common language use: If we consider one outcome as being better for someone than another outcome, then we are normally prepared to conclude that the other outcome would be worse for that person (and not just that it *is* worse). So we acknowledge that there might be a price to pay here, in terms of departure from common usage, for our preferred answer to the existential question. However, as long as no other reason for SC1 has been brought forward, we find the price worth paying.<sup>19</sup> Moreover, if SC1 is accepted, then one would have to give up the idea that existence can be better or worse than non-existence for a person, since with SC1 as an extra premise, Broome’s argument would be valid and the Absurd Conclusion would follow from an affirmative answer to the existential question.

Notice that our argument is not based on any revision of the logic of “better for” and “worse for”. In one of his earlier contribution to this topic, Nils Holtug seems to suggest such a revision to avoid the Absurd Conclusion:

There is a clear sense in which existence can be better for a person than nonexistence, even if nonexistence is not worse for her (a person can have no properties in a possible world in which she does not exist). (Holtug, *In Defence* 171)

---

<sup>19</sup> Johan Brännmark has suggested (in private communication) that rejecting SC1 is a more serious move than we make it appear. The connection between indicative ascriptions of properties and relations and subjunctive claims about what would happen under hypothetical circumstances is a pervasive feature of our language and our way of thinking about the world. When I say, for example, that the medicine I take decreases blood pressure or that exercise keeps me fit, I imply something about what would be the case if I didn’t take my medicine or if I didn’t exercise. That this kind of subjunctive connection doesn’t hold for better-for on our account of this relation is worrying in his view. We agree, but we don’t think it shows that our account is wrong. SC2 seems sufficient to capture the connection between the indicative ascriptions of better-for relations and subjunctive claims about hypothetical circumstances.

Holtug seems to suggest that one can avoid the Absurd Conclusion by revising the logic of “better for”: One can hold that it can be better for a person to exist than not to exist, but deny that the opposite is worse for her. It is clear that a state X is better than a state Y if and only if state Y is worse than state X (this seems to us to be a conceptual truth, if any). Holtug seems to deny that this logic also holds for “better for”, that is, that a state X is better for a person than another state Y if and only if state Y is worse for the person than state X. His reason is that “better for” and “worse for” are only applicable when a person to which the “for” refers to exists. See also his (1996), p. 77:

When saying that a person has been benefited by coming into existence, I mean that this person is better off than if he had never existed. Of course, normally, if a person is better (worse) off in a situation X than in a situation Y, he is worse (better) off in situation Y. While this is normally true, it is not true when Y involves his nonexistence. And there is a perfectly natural explanation for that. The property of “being worse off”, like other properties, does not apply to people in worlds in which they do not exist. (Holtug, *In Defence* 77)<sup>20</sup>

In our view, there is no need for a revision of the logic of value comparisons. If A is better for *p* than B, then it trivially follows that B is worse for *p* than A. What does not follow is that B *would* be worse for *p* if it obtained, for *p* might then be missing. Revising the logic of “better” and “worse” is unnecessary unless one conflates “worse for” with “would be worse for”.<sup>21</sup>

<sup>20</sup> On another reading of this passage, which is closer to its actual wording, Holtug here only denies that being *better off* in X than in Y entails being *worse off* in Y than in X. However, even this suggestion seems to us unmotivated: It is strictly incorrect to say that a person who in X has a life worth living is better off in that state than she is in the state Y in which she does not exist (it is another matter that X is better for her than Y). Comparisons of how “well off” a person is in two different states do seem to presuppose that she exists in both states that are being compared. See next section for a further discussion of this point.

<sup>21</sup> It should be mentioned that in his (*On the Value*), Holtug gives up on his earlier proposal and instead moves to a position similar to the one defended here, referring to personal communication with Rabinowicz.

## V

### The Argument from Welfare Level Comparisons

To save the Person Affecting Restriction from cases like the Future Bliss or Hell case, Melinda Roberts has suggested that we should accept not only that existence can be better or worse for a person than non-existence, but also the apparently absurd conclusion that, in cases like this, non-existence would be better or worse for a person. The reason is that according to Roberts a non-existing person has a certain welfare level, namely, zero welfare:<sup>22</sup>

...Nora does not have any properties at all at any alternative at which she does not exist and ..., where Nora has no properties at all, all the properties that she does have – that empty set – add up to a zero level of wellbeing. --- It would have been better for Nora not to have any wellbeing at all – to have zero wellbeing – than to have the negative level of wellbeing that she in fact has. It would have been [better] for Nora ... never to have existed at all than it is for Nora to exist. (*Can it Ever* 168-9)<sup>23</sup>

However, in our view it is quite nonsensical to ascribe any wellbeing level at all to a person in a state in which she does not exist. Wellbeing presupposes being. Having a zero degree of wellbeing is a property and property instantiations require the existence of property bearers.<sup>24</sup> Moreover, as we have shown above, one can endorse an affirmative answer to the existential question without being committed to affirming that non-existence could have been better or worse for a person and without assigning any welfare levels to persons who don't exist.

<sup>22</sup> Adler (1506) tentatively embraces a similar position: "Existence can be better or worse for an individual than nonexistence. Nonexistence can be better or worse for an individual than existence. Where an outcome set contains potential nonexistents, their interests should be taken into account by assigning them a utility level of zero in the outcomes where they do not exist".

<sup>23</sup> Moreover, Roberts writes that "I am thus supposing that it is at least possible that *s* has more well-being in a world in which *s* does not exist than *s* actually has. Suppose *s*'s existence in *X* is unavoidably less than one worth living ... and that *s* has, in any world in which *s* does not exist, a zero level of well-being. Under these conditions, *s*'s level of well-being at zero is actually greater than *s*'s well-being in *X*..." (*Child versus* 64) (emphasis in original). On the other hand, she there also claims that "[t]here is no need, ever, on my account of either wrongful life or the non-identity problem to assign a value, even a value of zero, to nonexistence for an individual who never exists in the world that is subject to appraisal" (*Ibid.* 174-5) (emphasis in original).

<sup>24</sup> This is a bit dogmatic, of course. Here, we gloss over the controversial issue of property ascriptions to fictional objects and an even more controversial issue of tropes, i.e., concrete property-like particulars whose existence is supposed to be ontologically independent of the existence of property bearers.

However, one might insist that the suggestion we make still doesn't make sense: that we cannot make sense of one state, *A*, being better for *p* than another state, *B*, if we cannot compare the wellbeing levels of *p* in the two states in question. This might be what Heyd and Narveson have in mind in the quotes above.<sup>25</sup> Likewise, when Bykvist claims that SC1 (his Accessibility principle) is true about any "interpretation of 'better for' that is conceptually linked to well-being..." (Bykvist 348), it seems that his idea is that "better for"-claims are analyzable in terms of comparisons between well-being levels possessed by a given individual in different outcomes. This would entail SC1 given that no individual has any level of wellbeing in an outcome in which she does not exist. The idea is that there is a necessary connection between "better for" and "has a higher welfare than":<sup>26</sup>

*Welfare Level Connection (WLC):* An outcome *A* is better for a person *p* than another outcome *B* if and only if *p* has higher welfare in *A* as compared to *B*.

But again, *p* would not have any welfare level at all in a state in which she would not exist.<sup>27</sup> However, it seems to us that "better for" comparisons can be made without comparisons of welfare levels. Consequently, one should reject the suggested tight connection between "better for" and comparisons of welfare levels as expressed by WLC.

## VI

### Guardian Angels and Fitting Attitudes

Instead of relying on WLC, one might explicate "better for" in terms of what a benevolent impartial observer would choose for a person when she is only considering what is in the interests of the person in question, or - better - in terms of what that person's guardian angel would be

<sup>25</sup> See also the quote from Dasgupta in fn. 9.

<sup>26</sup> Adler (1503) considers a similar conceptual connection, which he presents as the connection between "worse for" and "worse off than". However, unlike Bykvist, Adler's discussion leads him to reject, at least tentatively, this supposed conceptual link. (Ibid. 1505)

<sup>27</sup> This seems to hold even if we were to construct persons as abstract objects that can obtain or not obtain. A specific welfare level is something an abstract person can possess only in a world in which she obtains.

willing to choose.<sup>28</sup> According to this view, an outcome A is better for a person than another outcome B if and only if this is what her guardian angel would choose for her sake. If a person exists in the two compared outcomes, then trivially the guardian angel will choose the state in which her charge has the highest welfare level.

However, if the guardian angel has a choice between bringing her charge into existence with negative welfare or not bringing her into existence at all, she would choose the latter. Moreover, if the guardian angel had the choice between bringing her charge into existence with a positive welfare or not bringing her into existence, she would choose the former. Or so it may seem, at least.

We can think of this idea of a guardian angel as just a *criterion* for the “better for”-relation. On this criterial interpretation, we can try to find out what is better for a person by putting ourselves, in imagination, in her guardian angel’s shoes and then trying to determine what our preferences would be in that hypothetical position. Additionally, on a view that is philosophically more far-reaching and radical, the idea of a guardian angel can also be seen as a metaphor for a certain *analytical* proposal. More precisely, on this reading, we should take it as an application to “better for” of the so-called *fitting-attitudes analysis of value*. Along the lines of this format of analysis, we could say that

A is better for *p* than B if and only if one ought to prefer A to B for *p*’s sake.<sup>29</sup>

---

<sup>28</sup> Rabinowicz suggested the guardian angel approach in 2000 (see fn. 16) and Arrhenius (*The Person*) proposes the benevolent impartial observer approach. See also Bykvist. Broome (*Weighing* 63), credits Rabinowicz with a suggestion that is simpler but less plausible: A history (or a world) X is better for *p* than a history Y if and only if *p* prefers X to Y. As Broome points out: “A person may prefer one history to another even if she does not exist in both of them” (ibid.). Obviously, however, this simple proposal is not satisfactory as it stands.

The advantage of appealing to the preferences of the guardian angel rather than to those of the benevolent impartial observer is that the latter are supposed to track what is impersonally good (good, period) rather than what’s good for the person under consideration. A benevolent impartial observer tracks impersonal goodness even when she only focuses on the interests of that person and of no one else. The task of a guardian angel is different in that respect.

<sup>29</sup> Cf. Darwall for this proposal. As Darwall puts it: “[W]hat it is for something to be good for someone *just is* for it to be something one should desire for him for his sake, that is, insofar as one cares for him” (8). See also Toni Rønnow-Rasmussen, where this fitting-attitudes account of value-for is elaborated and defended. That this account can be used to clarify comparisons between existence and non-existence has been suggested in Rabinowicz (*Broome and fn. 2*).



This analytic proposal could be made to work provided we can make some sense of locutions such as “preferring A to B for  $p$ ’s sake”.<sup>30</sup> Again, it seems reasonable to say that in the choice between bringing  $p$  into existence with negative welfare or not bringing her into existence at all, one ought to prefer the latter for  $p$ ’s sake. Likewise, in the choice between bringing  $p$  into existence with positive welfare or not bringing her into existence at all, one ought to prefer the former for  $p$ ’s sake.

On both these interpretations, the criterial and the analytic one, if a person  $p$  has higher welfare in an outcome A as compared to another outcome B, then A is better for  $p$  than B, but the reverse doesn’t always hold. Hence, there is a connection between “better for” and “has higher welfare than” but this connection isn’t as tight as WLC would have it.

As a matter of fact, the fitting-attitudes analysis of value can be used to define the very notion of wellbeing. To say that  $p$  has a positive (negative, zero) wellbeing in a given outcome A presupposes that  $p$  exists in A and means that to exist in A is better (worse, equally as good) for  $p$  than (as) not to exist at all, with the latter claim interpreted on the lines of the fitting-attitudes account.

This format of analysis makes it also possible to account for interpersonal comparisons of wellbeing, of the form “ $p$ ’s wellbeing in A is greater than  $q$ ’s wellbeing in B.” To do it, though, we first need to re-interpret preference. Instead of treating it as a dyadic comparative attitude we need to think of it as a comparison between monadic attitudes. More

<sup>30</sup> The challenge here is whether the “for  $p$ ’s sake”-locution can be independently understood, without presupposing the notion of “better for” as already given. If preferring something for  $p$ ’s sake just means “preferring it insofar as one only cares for what is better for  $p$ ”, then the analysis becomes circular. On the other hand, if “for  $p$ ’s sake” is given an independent interpretation, then it is not obvious that all that one ought to prefer for  $p$ ’s sake is better for  $p$ . In particular moral considerations (and perfectionist ideas in general) complicate matters at this point. Thus, for example, it might be argued that one ought to desire for  $p$ ’s sake that  $p$  experiences sorrow at the thought of others’ suffering. But it is not obvious that such a set of mind would be good for  $p$ . (We are indebted to Johan Brännmark for pressing this point.) Maybe, therefore, a circular analysis of “sake” would, after all, be preferable! It should be noted that even circular analyses can be instructive to some extent: They can be used to exhibit structural connections between concepts appearing in the analysans and the analysandum. Thereby, they can provide relevant information to those who already possess the concepts involved but are not clear about their mutual relationships. Thus, to take an example, David Wiggins adheres to the sentimentalist version of the fitting-attitudes account even though he explicitly recognizes the charge of circularity. Still, as he argues, the account is informative in its “detour through sentiments” see (Wiggins 189). Cf. Rabinowicz & Rønnow-Rasmussen (*The Strike, Buck-Passing*). The circularity Wiggins has in mind is different from the one mentioned here, though. He thinks that it might be essential to the fitting sentiments with regard to objects that these attitudes themselves already involve evaluations.

precisely, that a person prefers A to B would on this re-interpretation be understood as a claim to the effect that the person in question favours A to a higher degree than she favours B.<sup>31</sup> This would mean that we end up with the following series of analyses:

*A is better than B* = A ought to be preferred to B = It is required that A is favoured to a higher degree than B is favoured.

*A is better för p than B* = A ought to be preferred to B for p's sake = It is required that A is favoured for p's sake to a higher degree than B is favoured for p's sake.

*A is better för p than B is for q* = It is required that A is favoured for p's sake to a higher degree than B is favoured for q's sake.<sup>32</sup>

*p's wellbeing in A is greater than q's wellbeing in B* = p exists in A, q exists in B, and A is better för p than B is for q.

## VII

### Person Affecting Restriction Revisited

As for the connection between “better” and “better for”, the Person Affecting Restriction remains an attractive option. It does seem plausible to claim that, to the extent we focus on welfare, an outcome cannot be better than another outcome without being better for someone. While this restriction would lead to counterintuitive implications if combined with the negative answer to the existential question (see the case of Future Bliss or Hell above), we have argued in this paper that the existential question should be answered in the affirmative.

---

<sup>31</sup> This formulation is simplified and covers only one form of preference: when both alternatives are favoured, though to a different degrees. But preference for A over B might instead be a matter of disfavouring A to a lesser degree than B, or of favouring A but disfavouring or being indifferent to B, or – finally – of being indifferent to A and disfavouring B. Complete analysis of preference needs to cover all these four possible cases.

<sup>32</sup> It is at this point that the re-interpretation of preference as a comparison between monadic attitudes is crucial. That A is favoured for p's sake to a higher degree than B is favoured for p's sake is the thought that could not be expressed in terms of a dyadic comparative attitude with respect to A and B. That dyadic attitudes should be replaced by comparisons between monadic attitudes in the fitting-attitudes account of value relations has also other reasons. On this issue, see Rabinowicz (*Values*).



It should be noted, however, that even coupled with the affirmative answer to the existential question, the Person Affecting Restriction, as we have stated it above, leads to counterintuitive implications, unless it is appropriately weakened. The reason is that the betterness relation between outcomes does not require for its obtaining the actual existence of the affected persons. Persons enter as relata in the triadic 'better for'-relation and therefore must exist for that relation to obtain, but they are not relata in the dyadic betterness relation that obtains between outcomes. This contrast between the triadic and the dyadic relations of betterness explains why the Person Affecting Restriction cannot be correct as it stands.

Thus, to give an example, consider a variant of the Future Bliss or Hell Case above in which only the x-people exist in outcome A (i.e., in this variant, outcome A does not contain any future y-people) while outcome B still in addition contains z-people that lead hellish lives, and suppose that outcome A is the one that actually obtains. The Person Affecting Restriction would imply, counterintuitively, that A is not better than B, since – as things actually are – there exists no one for whom A is better than B: The added people in the hypothetical outcome B, for whom A would have been better, do not actually exist. Intuitively, however, if A would have been better than B if B obtained, then A *is* better than B even if B does not obtain.<sup>33</sup> To solve problems like this, Holtug (*Person-affecting*) has argued that we should replace the restriction with a weaker version, which in our formulation runs as follows:

*The Wide Person Affecting Restriction:* If an outcome A is better than B, then A would be better than B for at least one individual if either A or B would obtain.

<sup>33</sup> This counterfactual invariance of the dyadic betterness relation is possible only because its relata (outcomes) can be assumed to exist even if they do not obtain. By contrast, the triadic relation of 'better for' can only satisfy a weaker condition of counterfactual invariance: If A would have been better for *p* than B if B obtained, then A is better for *p* than B even if B does not obtain, provided that *p* exists.

In the example above, it is the second disjunct of this weaker restriction that is applicable. Clearly, it is only this wide, disjunctive version of the restriction that deserves serious consideration.<sup>34</sup>

Before we finish, we should say more about ordinary language formulations of value comparisons between outcomes. The reader might have got an impression that, on our view, counterfactual claims such as “not to exist would have been better/worse for  $p$  than to exist” are absurd. But are such claims really so implausible? Think of Melinda Roberts’ statement about Nora: “It would have been better for Nora never to have existed at all than it is for Nora to exist.” This doesn’t sound absurd at all. But, if not, then perhaps the Argument from Absurdity doesn’t even get started?

Still, how should a counterfactual statement like the one about Nora be understood? Here is what we’d like to suggest. When we use ordinary language formulations of the form “ $A$  would have been better for  $p$  than  $B$ ,” what we mean is something along these lines: (i) we state that  $A$  is better than  $B$  for  $p$ , i.e., that a certain triadic relation obtains; and (ii) we imply that  $A$  does not obtain.

On this analysis of “ $A$  would have been better for  $p$  than  $B$ ,” when we use such formulations, we don’t take a stand on what relation would obtain between  $A$ ,  $B$  and  $p$  under counterfactual or subjunctive circumstances such as, say, if  $A$  had obtained. Which explains the absence of absurdity in Nora-type statements. However, we do take such a stand when we expressly state that “ $A$  would have been better/worse for  $p$  than  $B$ , if  $A$  had been the case.” Thereby we do state that the betterness/worseness relation would have obtained between  $A$ ,  $B$  and  $p$  under the counterfactual circumstance in which  $A$  would have obtained. Consequently, the following is absurd: “Not to exist would have been

---

<sup>34</sup> In fact, even this disjunctive version of the restriction might be too strong. In principle, it is conceivable that  $A$  is better than  $B$  because it would be better for some individual who only exists in the actual outcome, but does not exist in either  $A$  or  $B$ . Example: Suppose that  $C$  is the actual outcome and  $p$  exists in  $C$  but not in  $A$  or  $B$ . Suppose that  $p$  in  $C$  devotes his whole life to a certain goal  $G$ . For  $p$ , this goal is categorical: he wishes it to be realized even if it weren’t his goal in the first place. While  $p$  would not exist in either  $A$  or  $B$ , let us suppose that  $G$  would be better realized in the former outcome than in the latter. On the assumption that the realization of  $p$ ’s fundamental goals is one of the things that determine how good an outcome is for  $p$ , and that this applies, in case of categorical goals, even to possible outcomes in which  $p$  would never strive for the goals in question, it might well be the case that  $A$  is better for  $p$  than  $B$ , even though – as it happens – there would be no one in either  $A$  or  $B$  for whom the former outcome would be better than the latter.

better/worse for  $p$  than to exist, *if  $p$  had not existed.*" It is this kind of statements that the Argument from Absurdity focuses on. Therefore, if that argument is to be invalidated, it has to be met head on, as we have done, rather than rejected on the grounds that there is no absurdity there to begin with.

## VIII. Summary

We have defended an affirmative answer to the existential question to the effect that one can claim that it is better or worse for a person to exist than not to exist, without implying any absurdities. Hence, not only is your existence, dear reader, better for us than your non-existence; it is also better for you.<sup>35</sup>

### BIBLIOGRAPHICAL REFERENCES

Adler, M. D. "Future generations: A prioritarian view". *The George Washington Law Review*. Sep. 2009: 1478-1520. Print.

Arrhenius, G. *Population axiology*. Toronto: Philosophy University of Toronto, 1999. Print.

---. *Future generations: A challenge for moral theory*. Uppsala: University Printers, 2000. Print.

---. "The Person Affecting Restriction, comparativism, and the moral status of potential people". *Ethical Perspectives*. Sep-Dec. 2003. 185-195. Print.

---. "The repugnant Conclusion". *Stanford Encyclopaedia of Philosophy*, 2006. Online.

---. "Can the Person Affecting Restriction solve the problems in population ethics?" Roberts, M. and D. Wasserman (eds.). *Harming future persons*. Ashgate: Springer, 2009. Print.

---. *Population ethics*. Oxford: Oxford University Press, 2013. Print.

Arrhenius, G., Ryberg, J. and T. Tännsjö. "The repugnant conclusion". *Stanford Encyclopaedia of Philosophy*, 2006. Online.

---

<sup>35</sup> We would like to thank Matthew Adler, Staffan Angere, Johan Brännmark, Ingvar Johansson, Krister Bykvist and Toni Rønnow-Rasmussen for helpful comments. We gratefully acknowledge financial support from the Bank of Sweden Tercentenary Foundation, the Swedish Research Council, and the Swedish Collegium for Advanced Study.

Benatar, D. *Better never to have been: The harm of coming into existence*. Oxford: Oxford University Press, 2006. Print.

Broome, J. *Counting the costs of global warming*. Cambridge: White Horse, 1992. Print.

---. *Ethics out of economics*. Cambridge: Cambridge University Press, 1999. Print.

---. *Weighing lives*. Oxford: Oxford University Press, 2004. Print.

Bykvist, K. "The benefits of coming into existence". *Philosophical Studies*. Jan. 2007: 35-362. Print.

Buchanan, A., Brock D. W., Daniels N. and D. Wikler. *From chance to choice: Genetics and justice*. Cambridge: Cambridge University Press, 2000. Print.

Darwall, S. *Welfare and rational care*. Princeton: Princeton University Press, 2002. Print.

Dasgupta, P. *An inquiry into well-being and destitution*. Oxford: Oxford University Press, 1995. Print.

Feldman, F. "Adjusting utility for justice: A consequentialist reply to the objection from justice". *Philosophy and Phenomenological Research*. Sep. 1995: 567-85. Print.

---. "Justice, desert, and the repugnant conclusion". *Utilitas*. Nov. 1995: 189-206. Print.

---. *Utilitarianism, hedonism, and desert: Essays in moral philosophy*. Cambridge: Cambridge University Press, 1997. Print.

Freud, S. *The standard edition of the complete psychological works of Sigmund Freud*. London: The Hogarth Press, 1960. Print.

Glover, J. *Causing death and saving lives*. London: Penguin Books, 1977. Print.

Heyd, D. "Procreation and value: Can ethics deal with futurity problems?" *Philosophia*. Jul. 1988: 151-170. Print.

---. *Genethics: Moral issues in the creation of people*. Berkeley: University of California Press, 1992. Print.

Holtug, N. "In Defence of the Slogan". W. Rabinowicz (ed.). *Preference and value: Preferentialism in ethics*. Lund: Lund University, 1996. Print.

---. "On the value of coming into existence". *The Journal of Ethics*. Sep. 2001. 361-384. Print.

---. "Person-affecting moralities". Ryberg, J. and T. Tännsjö (eds.). *The Repugnant conclusion. Essays on population ethics*. Dordrecht: Kluwer, 2004. Print.

Johansson, J. "Being and betterness". *Utilitas*. Sep. 2010: 285-302. Print.

Narveson, J. "Utilitarianism and new generations". *Mind*. Jan. 1967: 62-72. Print.

Moore, G. E. *Principia ethica*. Cambridge: Cambridge University Press, 1903. Print.

Parfit, D. *Reasons and persons*. Oxford: Clarendon Press, 1984. Print.

Rabinowicz, W. "The size of inequality and its badness: Some reflections around Temkin's inequality". *Theoria*. Aug. 2003: 60-84. Print.

---. "Broome and the intuition of neutrality". *Philosophical Issues*. Oct. 2009: 389 - 411. Print.

---. "Values compared". *Polish Journal of Philosophy*. Jan. 2009: 73-96. Print.

Rabinowicz, W. and T. Rønnow-Rasmussen. "The strike of the demon: On fitting pro-attitudes and value". *Ethics*. Apr. 2004: 391-423. Print.

---. "Buck-passing and the right kind of reasons". *Philosophical Quarterly*. Jan. 2006: 114-120. Print.

Roberts, M. A. *Child versus Childmaker: Future persons and present duties in ethics and the Law*. Lanham: Rowman and Littlefield, 1998. Print.

---. "Can it ever Be Better Never to have existed at all? Person-based consequentialism and a new repugnant conclusion". *Journal of Applied Philosophy*. Oct. 2003: 159-185. Print.

Rønnow-Rasmussen, T. "Analyzing personal values". *The Journal of Ethics*. Jan. 2007: 405-435. Print.

Temkin, L. S. *Inequality*. Oxford: Oxford University Press, 1993. Print.

---. "Harmful goods, harmless bads". Frey, R. G. and C. W. Morris (eds.). *Value, welfare, and morality*. Cambridge: Cambridge University Press, 1993. Print.

Wiggins, D. *Needs, values, truth: Essays in the philosophy of value*. Oxford: Blackwell, 1987. Print.



# NOMINALIST ANALYSES OF AN ENTITY BEING CHARACTERED

*ANÁLISIS NOMINALISTA DE UNA ENTIDAD QUE ESTÁ SIENDO CARACTERIZADA*

MICHAEL ISTVAN

Texas A&M University, USA. m-istvan@philosophy.tamu.edu

RECIBIDO EL 17 DE SEPTIEMBRE DE 2012 Y APROBADO EL 17 DE OCTUBRE DE 2012

## RESUMEN    ABSTRACT

Este artículo está concebido principalmente como una herramienta de referencia para quienes participan en el debate entre realismo y nominalismo sobre los universales. A su vez, ofrece un catálogo exhaustivo de los análisis básicos de una entidad que está siendo caracterizada y que los nominalistas pueden emplear tanto en una ontología no-constitutiva, como en una constitutiva.

This paper is intended primarily as a reference tool for participants in the debate between realism and nominalism concerning universals. It provides an exhaustive catalogue of the basic analyses of an entity being characterized that nominalists can employ in both a constituent and nonconstituent ontology.

## PALABRAS CLAVE    KEY WORDS

ontología constitutiva, nominalismo, ontología no-constitutiva, particulares, problema de los universales, propiedad, universales.

constituent ontology, nominalism, non-constituent ontology, particular, problem of universals, property, universals.

## Introductory remarks

Nominalism (that is, antirealism) concerning universals is the view that it is impossible for something to have the intrinsic capacity to be wholly present through multiple entities at one and the same time. It is the view, in other words, that there can be no strict identity had among members of a multiplicity and thus that the agreement or sameness between things is never grounded in identity between those things. The purpose of this paper is to delineate the three basic and exhaustive nominalist analyses of an entity being charactered. I intend this brief taxonomy to serve as a resource, a reference tool, for thinking about the controversy between realism and nominalism concerning universals.

There can be only the following two exhaustive analyses of an entity being charactered: (1) a constituent analysis, according to which what makes it correct to predicate P of entity o is some property had by o, and (2) a nonconstituent analysis, according to which it is not the case that what makes it correct to predicate P of entity o is some property had by o. In light of this coarse division, there can be only the following three exhaustive categories of nominalism: relational and austere nominalism (which are the two exclusive and exhaustive forms of nonconstituent nominalism), and constituent nominalism<sup>1</sup>. In this section, I will outline these three views (all of which have been occupied in the history of philosophy), and then conclude with a note about how the infamous medieval theory of conceptualism is supposed to fit on this map.

### I

#### Nonconstituent nominalism

Nonconstituent nominalism, the orthodox form of nominalism, takes individuals to be the only sorts of entities possible, where by “individual” it is usually meant a non-property item (a non-property item that is usually going to be charactered), and takes these individuals to be particulars, where by “particular” I do not mean *specific* (as in the colloquial sense of the term), but rather that which cannot appear in,

---

<sup>1</sup> Some expressions of relational nominalism can be found in Roscelin and Ockham (Loux 63). An expression of austere nominalism can be found in Ockham (Loux 83fn21). A “common intellectual currency” according to D. C. Williams (106), some expressions (although not necessarily advocacy) of constituent nominalism (that is, trope theory) can be found in Plato, Aristotle, Boëthius, Avicenna, Averroës, Aquinas, Ockham, Scotus, Buridan, Suárez, and Leibniz (Mertz ch. 4) (Loux 73). Williams (107) finds trope theory prevalent in Descartes and Spinoza, and so does Stout (9) and Seargent (13).



be a constituent of, multiple entities at the same time. Nonconstituent nominalism, in other words, (1) denies that individuals (usually taken to be non-property items) in themselves have any characteristics and (2) defines individuals as particulars (as non-universal items). There are only two nonconstituent nominalist analyses possible: relational and austere nominalism.

### **Relational nominalism**

Relational nominalism holds that a particular individual's being characterized in a certain way is due merely to that individual's relation to some other individual that is also a particular. For example, according to predicate nominalism, a subjectivist form of relational nominalism that traditionally is all that nominalism has been thought to be, this particular individual apple is green if and only if it falls under the predicate 'green', such that there is nothing like greenness that the green apple has and if there were no predicate term 'green' – or at least no possibility of the predicate term 'green' – the apple would not in fact be green. All the analyses of relational nominalism have the following reductive form, then: to say that particular individual *o* is *P* is merely to say that *o* has a relation to some particular individual *x* (such that there is nothing like *P*ness that a *P* thing has).

### **Austere nominalism**

As with relational nominalism, austere nominalism (1) denies that individuals (usually taken to be non-property items) intrinsically have any characteristics and (2) defines individuals as particulars (as non-universals). Although both in effect view individuals in themselves as ontologically unstructured simples, and thus are extreme insofar as they deny any reality to properties, austere nominalism refuses to give an account of what it means to say that a particular individual is a certain way, has a certain character.<sup>2</sup> The truthmaker, the ontological ground, for attributing *P* of particular individual *o* is nothing more and nothing less than the ontologically unstructured individual that is *o*, in which case the sweetness of the bonbon is nothing more, nothing less, than the ontologically (although not mereologically) unstructured bonbon itself. The austere nominalist "analysis" has the following form, then: particular individual *o* is *P* just means that *o* is *P* – nothing more than that can be said.

<sup>2</sup> Quine is taken to be the father of this view. There are more recent defenders, such as Devitt and Parsons.

As with relational nominalism, austere nominalism does not deny that Socrates is characterized in many ways and can be correctly described in various ways. But whereas relational nominalism explains why this does not mean that the individual has intrinsic features, austere nominalism does not provide an explanation (even though it continues to say such things as “this figure is round” and “both of these are cars”). To put this important difference another way, relational nominalists think that the resources for explaining why an individual is characterized and yet devoid of properties cannot just be the thing itself and for this reason they bring in other individuals besides the thing in question to help in the explanation. Austere nominalists, on the contrary, refuse to expand their explanatory resources beyond the thing in question. Given that the only resource they have for explaining the individual is the individual itself, when asked why *o* is *P* the best they can do is point to *o* (and then, as the quip against them goes, stick their head in the sand, which is why they are sometimes called “ostrich nominalists”). In effect, austere nominalism offers a pretty thin analysis of a thing having a property, it just says that *o* is *P* if and only if *o* is *P* (such that there is nothing like *P*ness that a *P* thing has).

## II Constituent nominalism

In contrast to relational and austere nominalism, the two basic and exhaustive forms of nonconstituent nominalism, constituent nominalism is, in a sense, much more moderate in that it does not deny the reality of properties. Called “trope theory” in recent literature, this form holds that there really are properties constituting characterized items—items commonly understood as individuals by trope theorists even when trope theorists view such items as nothing but a bundle of properties (as they commonly do). So, in contrast to relational nominalism, the state of affairs of particular individual *o* having property *P* is not parasitic upon *o* being in relation to some entity. Rather, it is just is a matter of *o* having a *P* property as a constituent. Unlike the other views, then, a *P* thing possesses *P*ness: *o* is *P* if and only if *o* has *P*ness. What makes this view nominalist, however, is that properties, such as this *P*ness, are taken to be particulars rather than universals, which means that these properties are intrinsically unable to be wholly present in more than one (nonconcurrent) entity at one time, and thus cannot serve as the respect of similarity between two or more entities. In effect, if there are

two exactly similar qualities on this view, they will not be identical (for to say that they are is to accept realism concerning universals). Rather, they would be merely perfectly resembling, merely indiscernible (where the term “merely” is meant to stress that they are not thereby identical in any way).<sup>3</sup> So although constituent nominalism is moderate in the sense that it preserves the everyday belief in qualities, it is extreme in the sense that it denies that objective indiscernibility between qualities means identity, something that runs against everyday intuition (and is presumably why proponents of the other two forms of nominalism feel they must reject qualities had by an individual in order to reject realism).

### III

#### What about conceptualism?

If one is not a realist concerning universals, then one must fall within one of the above three nominalist categories. First, there is no other choice but to adopt a constituent or nonconstituent ontology; this is an exhaustive taxonomic division. Second, if one adopts a constituent ontology as a nominalist, one can only be saying that there are properties that are non-universals (trope theory). Third, if one adopts a nonconstituent ontology as a nominalist, in which case one rejects the reality of properties had by an individual (an individual typically construed as itself a non-property), then one can account for something being characterized in some way either by saying that it is in relation to some other entity (the relational explanation of relational nominalism) or else by saying simply that that something is characterized in that way (the non-relational explanation of austere nominalism). Now, historians of philosophy generally will be aware of the medieval position known as conceptualism, which is often packaged as a middle path between realism and nominalism. In order to obviate the response that I have not been sensitive to all the options, I will close this section by explaining that conceptualism, if it is antirealism, does not fall in any way outside of the parameters that I have laid out above.

Medieval conceptualism is the view, generally put, that any identity among the members of diversity is merely in the mind. Either this means

---

<sup>3</sup> See Istvan. Trope theory is perhaps the most popular form of nominalism today. Here are some of the popular contemporary advocates of the view: G. F. Stout (1936); D.C. Williams (1966); Keith Campbell (1990); John Bacon (1995); and Anna-Sofia Maurin (2002). Trope view is, however, ancient. For more on the ancient legacy of tropes, see Mertz (1996).

that there is nothing physical or mental (or so on) that has the capability of being wholly present in multiple physical or mental entities at one and the same time, or else it means that that which has the capability of being present in multiple mental entities can only be mental. In the first case, we are just dealing with nominalism (and thus the worldview that there can be no strict identity – however partial – among many things).<sup>4</sup> Most likely this view then will be classified as what is now called “concept nominalism,” which is a relational form of nominalism that analyzes a particular individual being characterized in the following way: o is P just means that o falls under the concept P (such that there is nothing like Pness that a P thing has). In the second case, we are just dealing with realism (and thus the worldview that there can be strict identity – however partial – among many things); it is just that, according to this particular brand of realism, only mental items can exemplify universals serving as the respects in which these items are similar.

#### BIBLIOGRAPHICAL REFERENCES

Bacon, J. *Universals and property instances: The Alphabet of Being*. Oxford: Blackwell, 1995. Print.

Campbell, K. *Abstract particulars*. Oxford: Blackwell, 1990. Print.

Devitt, M. “Ostrich nominalism or ‘mirage realism’?” *Pacific Philosophical Quarterly*. Aug. 1980: 433-49.

Istvan, Michael Jr. “On the possibility of exactly similar tropes”. *Abstracta*. Jun. 2011: 158-177. Print.

Loux, M. *Metaphysics: A contemporary introduction*. New York: Routledge, 2006. Print.

Maurin, A-S. *If Tropes*. Dordrecht: Kluwer Academic Publishers, 2002. Print.

---

<sup>4</sup> The phrase “partial strict identity” may have a ring of self-contradiction, but it is not upon consideration. I say that two purported entities have *full* strict identity if everything about the one is strictly – as opposed to loosely and popularly – identical to everything about the other. I say, therefore, that two purported entities have partial strict identity if something, but not everything, about them is strictly identical. Here is an example. Assume that A is nothing more than x, y, and z. A would have partial strict identity with B if, for instance, (1) B was x, g, and r (this covers the talk of “partial” since A and B merely have x in common) and (2) A’s x is identical – rather than merely similar or exactly similar – to B’s (this covers the talk of “strict”).

Mertz, D. W. *Moderate realism and its logic*. New Haven: Yale University Press, 1996. Print.

Parsons, J. "There is no 'truthmaker' argument against nominalism". *Australasian Journal of Philosophy*. Jun. 1999: 325-34. Print.

Quine, W. V. O. "On what there is". *From a Logical Point of View*. Cambridge: Harvard University Press, 1954. Print.

Seargent, D. A. J. *Plurality and continuity: An Essay in G. F. Stout's Theory of universals*. Dordrecht: Martinus Nijhoff, 1985. Print.

Stout, G. F. "Universals again". *Proceedings of the Aristotelian Society*. Suppl. 15. 1936: 1-15. Print.

Williams, D. C. "The elements of being". *Principles of empirical Realism: Philosophical Essays*. Springfield: Charles C. Thomas, 1966. Print.



# THE LIFE OF THE PHILOSOPHER: TESTIMONY OF PLUTARCH AND PORPHYRY

*LA VIDA DEL FILÓSOFO: TESTIMONIO DE PLUTARCO Y PORFIRIO*

ISHA GAMLATH

University of Kelaniya, Sri Lanka. isha@kln.ac.lk

RECIBIDO EL 15 DE AGOSTO DE 2012 Y APROBADO EL 17 DE OCTUBRE DE 2012

## RESUMEN      ABSTRACT

Este artículo explora en los testimonios de Plutarco y Porfirio, el rol y función del filósofo cuyo modo de vida negocia la liberación del rango de los sentidos. Esto incluye una investigación en su dieta y sacrificio cuya culminación es la re-evaluación de su demarcación de las masas.

The paper explores on the testimony of Plutarch and Porphyry, the role and function of the philosopher whose mode of life negotiates liberation from the range of the senses. This includes an investigation into his diet and sacrifice, the culmination of which is the re-assessment of his demarcation from the masses.

## PALABRAS CLAVE      KEY WORDS

Plutarco, Porfirio, vegetarianismo, sacrificio sin sangre, la buena vida.

Plutarch, Porphyry, vegetarianism, bloodless sacrifice, the good life.

The pagan holy man's way of living demarcates him from the laity in respect of his conduct and vision.<sup>1</sup> Insight in to the discipline of philosophy within perspectives pertaining to a way of life, not necessary restricted within limits of personal holiness, has driven Pierre Hadot to survey the role and function of the sage by appreciating him as a living concrete model.<sup>2</sup> Hadot's focus is on Socrates whose representation in Greek philosophy inspires the application of spiritual exercises for the perfection of intellectual and moral training throwing much light on his assertion that "philosophy was a method of spiritual progress which demanded a radical conversion and transformation of the individual's way of living." (Hadot 265)<sup>3</sup>

The logic behind Hadot's position is that philosophy is a way of living that separates a person from others but it does not fulfill the requirements in full length for the understanding of the philosopher's life within the wider frame of a *bios*. This void will be filled with the testimony of Plutarch and Porphyry.

The study of the wise man in the form of the combination of complete philosopher and active citizen projects in Plutarch's demonstration of Socrates who in Riley's view stands midway between a philosophical topic typical of the 'Moralia' and an active historical narrative from the 'Lives'. (Riley 273) Regrettably, the focus of current Plutarchean scholarship is more on character portrayal of significant social figures where the character of Socrates is often ranked superior among others who dominate the scene and Plutarch is driven to emphasize this position when he declares that "abstention from pleasure in what is allowed is a training of the soul to resist what is forbidden". (Ibid. 15.585 a-b=7.579c-d 15.584 d-e 15.585 b-c c-d) (Duff 159) (Mounard 339)

Socrates is identified as the ideal of the philosophic life and accommodates a domain illuminated by the emanations of the daemons (Ibid. 20.589b-

---

<sup>1</sup> Brown (80) traces the holy man "popularity as a product of the oppression and conflict that the social historian often ends to see as a blatant feature of east Roman society"; also his power in varied proportions (81-7), his role as peace maker (89-0), athlete (94), social status (91), intimacy with god (94); his rise is identified with what Plutarch writes of the decline of oracles. (99-00)

<sup>2</sup> Davidson notices Hadot's study to be of existential value not only a moral one (476) and that: "what Hadot has done beyond his influence on any particular thinkers is to open up dimensions of ancient philosophy we have typically overlooked or forgotten. Thus he has re-discovered and re-conceptualized the significance of ancient philosophy for our present moment in philosophy's unfolding history". (482)

<sup>3</sup> Hadot draws from standard academia for his treatment of Socrates: Nietzsche, Kierkegaard, Gaiser, Bohm, Bertram, Hildbrant, Spiegelberg, Lausberg (171-73); (Davidson 476).



c), the existence and assistance of the personal daemon (Ibid. 24.593 d-f) where every soul understands the irrational and the unintelligent. (Ibid. 22.591d-e e-f 22.592b-d) (Opsomer 197-203) (Jones) (Brenk 27-49)

For Porphyry Socrates represents the same tenor in *De Abstinencia* (3.1) but he cannot be identified as completely 'good'. (3.8) He is neither a vegetarian (1.15 3.26) nor does he possess human consideration for animals (3.22) although he does display some concern for them as he used to swear by animals. (3.16) As such, Socrates does not represent habits, which in the opinion of Porphyry and for that matter what Plutarch strives in the 'Moralia' to demonstrate, as central for the good life – the practice of vegetarianism and the possession of humane consideration for all life forms – although Plutarch traces absence of sexual lust in his response to Alcibiades' infatuation and the appreciation of a life free from indulgence in material pleasures as outstanding. (Duff 40) (Nikolaidis 275-88) Applied to Porphyry's ascetic mold it loses substance. (Clark, *Augustine's Porphyry* 127) (Clark, *Fattening the*)

The Socratic ideal of renunciation in *De Genio*, central to virtuous conduct expands from its Middle Platonic context in to a different frame in Neoplatonism when Porphyry's description of his mentor, Plotinus, whose personal details bear a degree of living in accordance to a program of discarding the corporeal. For example, his birthday and ancestry (*Vitae Plotini*, 2) except for a few childhood memories (3) are concealed under cover of 'ashamed of being in the body' (1) which belongs to a broader program of "striving to give back to the Divine in himself to the Divine in All." (2) The culmination of his habitual indulgence towards approaching the Supreme (9) though he is under the impression that "it is for those Beings (ancestors or spirits) to come (to him) not for him to go to them," a claim that may have risen with his possession of a God for a personal daemon. (10) His desire for the re-establishment of the Platonic tradition in the form of the Platonopolis (12) despite his love of following his own tradition (12) though subject to failure demarcates him from the host of contemporaries driven to seek the beauty of philosophical endeavor (3 7 9 14-5 20) inspires Porphyry himself to look upon him as a model (20-1) even at the point when he attempts suicide. (11) The representation of a separate way of living reaches its zenith at the moment of his death, even with an unusual occurrence (2) followed by the declaration of Apollo:

Spirit man once, but now the diviner lot of a spirit as the bond of human necessity has been loosed for you and strong in heart, you swam swiftly from the roaring surge of the body to that coast where the stream flows strong, far apart from the crowd of the wicked, there to set your step firm in the easy path of the pure soul, where the splendor of God shines around you and the divine law abides in purity far from lawlessness, wickedness, O blessed one, you have borne so many contests and now move among the holy spirits crowned with mighty life. (22)

Porphyry represents Plotinus as one who 'refuses' medicaments containing substances from wild beasts and reptiles and did not approve the consumption of meat of animals reared for the table. (2) His non-violence is questionable, given the fact that he may not have practiced what Porphyry defines in *De Abstinencia* as total vegetarianism except perhaps for hygienic purposes. (2) There is discrepancy in Plotinus' type of non-violence since once he encourages torture of a man who stole a necklace even though he sensed that he was the thief. (11) But could one who loved all encourage violence to anybody? Still, Porphyry ensures that he never had an enemy except Olympus whose plans against Plotinus were based on jealousy. (10) True enough, he lives for others (as well as for himself) (8) enjoying the reputation of being guardian of the young as well their progenitor's consolation of their being in 'holy hands'. (9) Where his association with women is concerned he has no discomfort in living in the same house with such women whose intentions are philosophic, as Gemina and Amphiclea (9) although his conduct seems to be a little awkward when his blushful reaction at the sight of Origen is covered by his declaration that "the zest dies down when the speaker feels that his hearers have nothing to learn from him". (14) His reaction to Diophanes' claim of sexual intimacy between master and pupil is expressed in his exclamation to Porphyry: "so strike and be a light to men," a position similar to his reaction to Porphyry's essay on Sacred Marriage. (15) There is nothing to confirm apart from this incidence that sexual attraction has any significant influence on him. It may have attributed to Porphyry's type of celibacy which drives him to a ten month old marriage with the Jewish widow, Marcella, which to him is a 'gift from heaven' (*Ad Marcellam*, 3), his reasons for marriage being the propitiation of gods of generation (2) and protection of Marcella who conceived an attraction suitable for 'true philosophy'. (3)

Separation from externals, this being the way Porphyry addresses this issue in *De Abstinētia* (external: 1.28, 2.52; exoteric discipline: 1.30; external tumult: 1.34; external concerns: 1.54; external calamity: 1.33) (*Ad Marcellam*, 5 6-8 14 28-0 33) is a process that inspires the philosophic life when men could become 'truly rich' (1.54); 'live truly' (1.31); 'return to what is truly theirs' (ours) and to the 'truly existing thing' (1.29-0) with the love of the true being (1.33) which is literally the only kind of 'true salvation' known to Porphyry forming the true philosopher (*Ad Marcellam*, 27) who himself honors truth and the true good. (5) He is after all the priest of the God and loved by Him (2.49), constantly accompanied by Him (19-0) and in this sense, he finally might as well become himself a God. (15) (*De Abstinētia*, 1.54) He toils to accomplish virtues like freedom from sloth and torpor which cause idle speech and lies. (1.49 = 1.27-8) His one concern is attainment of virtue. (*Ad Marcellam*, 12) The soul's ascent depends on virtue (16) but 'neglecting virtue and wisdom and mere reasoning faith without right living does not make it possible to attain to God'. (23) One ought to live according to divine laws (26), eliminate wrong conceptions of God (21-3), cultivate self-mastery (29), cultivate faith, truth, love, hope (24) and eventually casting away the body. (34) Non-use of wrong words and non-engagement in wrong deeds (8 12 14-5 17) following Plato's track to facilitate the journey from the sensibles to the intelligible (10) is no easy venture but in reality a contest. (5-8 9 = 1.56) The education of the man who abides by such a way of living consists not in absorbing knowledge but in casting off affections of the soul (9) the culmination of which will be its application to those who desire such knowledge. (2.61) Education, for Plutarch, however assumes the form of adoption of Hellenic culture which cast major Roman figures in to this mold who are not necessarily versed in philosophy but whose social program is in harmony with the welfare of the Roman public in mind. (Gill 469-87) (Bebekar 193-203) (Swain 127-8)

The training of the philosopher within the mold of asceticism begins with discarding sensory pleasures, a project devoted to the elimination of the energization of the irrational part of the soul. It gives rise to a condition exemplified by the Essenes who are "averse to pleasures conceiving them to be vicious but they are of opinion that continence and not yielding to the passions constitute virtue." (4.11) Deliverance from corruption, in Porphyry's *schema* is fundamental for the philosophic type of life (*Ad Marcellam*, 33) and if possible casting away the body not just in parts but perhaps even go to the extent of dying for God. (34) Cultivation of all that is good is a vital means of being good to all men. (35 = 3.20)

Purity for Plutarch from the point of view of righteousness is to be self-sufficient and this is not a 'quibble of Orpheus'. (*Septem Sapientiam*, 16.159 b-c c-d) This self-sufficiency dominates Plutarch's arguments regarding maintenance of purity free from gluttony (*De Tuenda*, 6.124 e-f) more so, gluttony and the lust to kill (*De Esu*, 2.997a-d) which is totally deranging, disturbing and foreign to nature. (*De Tuenda*, 7.125 c-d e-f 18.131f-132b) What Plutarch laments is the transformation from previously 'lawful desires' when men depended on vegetable produce (*De Esu*, 1.2.993f-994b) to 'unnatural and anti-social pleasures' (*De Esu*, 1.2.993 d-e = *Vice and*, 2.101a-b) which are not in tune with human nature (*De Esu*, 1.5.995 a-b b-c 1.6.995 d-e e-f 7.996a-b) leading to the conclusion that "he who tortures a living creature is no worse than he who slaughters it outright." (1.7.996 b-c) Dismissing the mythical crimes of cannibalism (1.7.996 b-c 2.2.997 e-f) but accepting the program initiated by Pythagoras and the Pythagorean Empedocles whose precepts were 'law' for the ancient Greeks (2.3.998f 4.999a) Plutarch is more concerned with who exactly was responsible for this drastic social transition. It was the tyrants who began by killing the 'worst of sychophants' like Niceratus, Theramena, Polemarchus (2.4.998 b-c) which gradually cleared ground to kill wild and harmful animals like bird and fish and expanded to the "laboring ox, well-behaved sheep and house-warding cock" the reason for which was entertainment of guests, celebration of marriage, consorts with friends and war. (c-d)

Neither Plutarch nor Porphyry specifically claims a total abstinence from killing animals or eating their flesh where conditions require it. But meat consumption obstructs what in their opinion is viewed as the good life and the soul's eventual return to its lawful origin. They both address the issue of philosophy as the prime motive in this program which is virtually the primary concern of the man whose way of living transcends that of the masses. Philosophy is more a passion than anything else (*Vitae Pythagorae*, 46 = *Vitae Plotini*, 3) where stress is laid on the fulfillment of the acquisition of truth. (*Vitae Pythagorae*, 41 = *Vitae Plotini*, 13-8)

Plutarch professes knowledge of both a tradition of bloodless sacrifice (*Apophthegmata Romana*, 5.15.267 c-d) and a tradition which did not approve the violation of a maiden (at the shrine of Diana). (3.264 c-d) He responds to the popularity of blood sacrifice (3.52.217 a-b 3.68.280 b-c) and the use of wine during sacrifice (3.45.275 e-f) when meat consumption was so common that even a public meat market was built by Marcellus. (3.54.e-f) Plutarch's defense of vegetarianism, however,

is grounded on a wider and broader frame where animal intelligence is appreciated as far better than that of the humans (*Chance*, 3.98c-d) *Bruta Animalia* is replete. (4.987 d-e 5.989 a-b 4.987 e-f 4.988 b-c 6.998 e-f 7.990 f-8.991b 9.992 d-e 6.989 c-f) Their 'native virtue' makes it possible to train them for the sake of utility. (9.992b-c c-d) In contrast, men are gluttons (8.991 b-c) for they eat simply everything (8.991c-d) and commit sexual acts with even animals (8.990 e-f) so that men are the ones who ought to be called beasts not animals because their cruelty exceeds that of savage serpents, panther and lion. (8.99b-c) Some animals are revered. (*Isis et Osiris*, 5.70.379 d-e) Plutarch downright denies that gods are metamorphosized as animals. (5.71.379f-380b)

Porphyry's quest for a universal mode of salvation was foremost in his philosophical carrier for which he applied his knowledge of Indians and Chaldeans (*De Civitate*, 10.32) and interaction with Iamblichus in the form of the *Letter to Anebo*. The eminent polymath declares his preference for true philosophy (*Ad Marcellam*, 3-8) which embodies the 'proper mode of life' and 'guidance' (3) being the 'only true refuge'. (5) Access to the ascent to the gods remains in philosophy yet this ascent is 'steep and rough' (fr. 323 = *Ad Marcellam*, 5-12) (*Cult Images*, fr.10 = *Preparatio Evangelica*, 3.11.45-3.20) the culmination of which encapsulates the last words of Plotinus: "try to bring back the god in you to the Divine in the All." (*Vitae Plotini*, 2)

The 'true philosopher' following nature and not vain opinions is self-sufficient in all things (*Ad Marcellam*, 27) extends in his 'naked' emptiness all throughout his life (29-31) and finally approaches Him in that condition (*De Abstinencia*, 1.31-33) and thus assumes the condition of "teacher, savior, guardian, leader, nurse (*Ad Marcellam*, 26 = *Vitae Plotini*, 9 = *Vitae Pythagorae*, 18 19 33 = *De Abstinencia*, 2.61) clearing ground for the love of mankind (*Ad Marcellam*, 35 = *De Abstinencia*, 3.26-27 = *Vitae Pythagorae*, 30 = *Vitae Plotini*, 9) removal of ignorance, wickedness and passions (*Ad Marcellam*, 13-4 25 27-9 34) and liberation from the state of being bound in chains (*Ad Marcellam*, 33 = fr.297-298 (*De Regressu*) ascribe to the good life (*Ad Marcellam*, 14-7) for then only will he who practices philosophy will be as worthy as a god the representation of whom is idealized in Plotinus who rises above humankind after death. (*Vitae Plotini*, 22-3 = *Ad Marcellam*, 15-6 24 32) The link between right action and Porphyry's conception of salvation is clear. Elizabeth DePalma's views are truly enlightening in this regard:



Our Protestantized society tends to view souls as being either deserving or unworthy of salvation, and all-or-nothing status. Third-century Platonists, however, thought that a range of activities positioned the soul, depending on its condition, at different levels within the celestial spheres, even to union with transcendent divinity, or *henosis theōi*. Philosophers were especially interested in divine union which, according to Porphyry, might allow their souls to break out of the cycle of metempsychosis. Activities were salvific if they improved the condition of the soul; for ordinary persons certain activities would perhaps allow their souls to achieve *henosis theōi* in a future life. (522-3)

#### BIBLIOGRAPHICAL REFERENCES

Baltes, M. *Der Platonismus in der Antike, Bd. IV*. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1996. Print.

---. *Der Platonismus in der Antike, Bd. V*. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1998. Print.

Bebekar, J. "Drunken violence and the transition of power in Plutarch's". Ribeiro Ferreira, J., Delfim L., Troster, M. and Paula Barata Dias (eds.). *Symposion and Philanthropia in Plutarch*. Coimbra: Classica Digitalia, CECH, 2009. Print.

Bland, M. "The eschatological aspects of Porphyry's Anti-Christian Polemics in a Chaldaean - Neoplatonic Context". *C&M*. Sep. 2001: 193-215. Print.

Brenk, F. E. "Plutarch's Middle-Platonic god: About to enter (or remake) the Academy". Hirsch-Luitpold, R. (ed.). *Gott und die Götter bei Plutarch*. Berlin-New York: Götterbild-Gottesbilder-Weltbilder, 2005. Print.

Brown, P. "The rise and function of the holy man in late antiquity". *Journal of Roman Studies*. Jul. 1971: 80-101. Print.

Clark, G. "Fattening the soul: Christian asceticism and Porphyry on abstinence". *Studia Patristica*. Jul. 2000: 39-43. Print.

---. "Augustine's Porphyry and the Universal Way of Salvation". Karamanolis, G. and Anne Sheppard (ed.). *In Studies on Porphyry*. London: Institute of Classical Studies, 2007. Print.

---. "Philosophic lives and the philosophical life: Porphyry and Iamblichus in Greek Biography and Panegyric in late antiquity". Hagg, T. and Philip Rousseau (eds.). California: Berkeley University Press, 2000. Print.

Christ, W. *Plutarch's dialog vom daemon des Socrates*. Munchen: Loeb Classical Library. 1901. Print.

Davidson, A.I. "Spiritual exercises and ancient philosophy: An introduction to Pierre Hadot". *Critical Inquiry*. Sep. 1990: 475-482. Print.

De Palma, D. "Porphyry, Lactantius and Path to God". *Studia Patristica*. Jul. 2001: 522-533. Print.

---. "The power of religious rituals: A philosophical quarrel on the Eve of the Great Persecution". Cain, A. and Noel Lenski (ed.). *In The Power of Religion in Late Antiquity*, Colorado: University of Colorado, 2010. Print.

Duff, T. "The opening of Plutarch's life of Themistokles". *Greek, Roman and Byzantine Studies*. Sep. 2008: 159-179. Online.

---. "Plato's Symposium and Plutarch's Alcibiades". Ribeiro Ferreira, J., Delfim L., Troster, M. and Paula Barata Dias (eds.). *Symposium and Philanthropia in Plutarch*. Coimbra: Classica Digitalia, CECH, 2009. Print.

Edwards, M. J. "Two episodes in Porphyry's Life of Plotinus". *Historia*. Aug. 1991: 456-464. Print.

Fowden, G. "The pagan holy man in late antique society". *Journal of Hellenic Studies*. Jul. 1982: 33-59. Print.

Gill, C. "The question of character development: Plutarch and Tacitus". *Classical Quarterly*. Dec. 1983: 469-487. Print.

---. *The Character-Personality in CBR Pelling eds Characterization and Individuality in Greek literature*. Oxford: Oxford University Press, 1990. Print.

---. *Personality and Greek epic, tragedy and philosophy: The self in dialogue*. Oxford: Oxford University Press, 1990. Print.

Hadot, P. *Philosophy as a way of life: Spiritual exercises from Socrates to Foucault*. London: Blackwell, 1995. Print.

Jones, R. M. *The Platonism of Plutarch*. Wisconsin: Menasha, 1916. Print.

Johnson, A. "Arbiter of the oracular: Reading religion in Porphyry of Tyre". Cain, A. and Noel Lenski (ed.). *In The Power of Religion in Late Antiquity*, Colorado: University of Colorado, 2010. Print.

Männlein-Robert, I. "Biographie, Hagiographie, Autobiographie-Die Vita Plotini des Porphyrios". Kobusch, T. and Michael Erler (ed.). *Metaphysik und Religion: Zur Signatur des spätantiken Denkens*. Munich-Leipzig: K. G. Saur, 2002. Print.

Mounard, H. *La psychologie de Plutarque*. Paris: Moxon, 1960. Print.

Nikolaidis, A. "Philanthropia as sociability and Plutarch's unsociable heroes". Ribeiro Ferreira, J., Delfim L., Troster, M. and Paula Barata Dias (eds.). *Symposion and Philanthropia in Plutarch*. Coimbra: Classica Digitalia, CECH, 2009. Print.

O'Meara, J. J. *Porphyry's philosophy from Oracles in Augustine*. Paris: Etudes Augustiniennes, 1959. Print.

Opsomer, J. *In search of truth: Academic tendencies in Middle Platonism*. Brussels: Koninklijke Academie voor Wetenschappen, Letteren en SchoneKunsten van Belgie, 1998. Print.

---. "Demiurges in early imperial Platonism". Hirsch-Luitpold, R. (ed.). *Gott und die Götter bei Plutarch*. Berlin-New York: Götterbild-Gottesbilder-Weltbilder, 2005. Print.

Pelling, C. *Plutarch and history*. London: Eighteen studies, 2002. Print.

Riley, M. "Purpose and unity of Plutarch's De Genio Socratis". *Greek, Roman and Byzantian Studies*. Mar. 1977: 257-274. Online.

Sandbach, F. H. "Plutarch and Aristotle". *Illinois Classical Studies*. Sep. 1982: 207-232. Print.

Schoppe, Ch. *Plutarchs Interpretation der Ideenlehre Platons*. Hamburg: Münster, 1994. Print.

Smith, A. "Porphyrian studies since 1913". *ANRW*. Sep. 1987: 717-773. Print.

---. *Porphyry's Place in the Neoplatonic Tradition: A Study in Post-Plotinian Neoplatonism*. The Hague: Martinus Nijhoff Places, 1974. Print.

Sorabji, R. *Time, creation & the continuum. Theories in Antiquity and the early Middle Ages*. Chicago: University of Chicago Press, 2006. Print.

Teodorsson, S-T. "The education of rulers in theory (Mor.) and Practice (Vitae)". Nikolaidis, A. (ed.). *The unity of Plutarch's work*. Berlin-New York: Walter de Gruyter, 2008. Print.

Witt, R. E. *Albinus and the history of Middle Platonism*: Cambridge: Cambridge University Press, 1937. Print.



# A CONCEPÇÃO DO VALOR: UMA ANÁLISE ACERCA DA OBJETIVIDADE DOS VALORES PROPOSTA POR P. GRICE

*THE CONCEPTION OF VALUE: AN ANALYSIS ABOUT THE OBJECTIVITY  
OF VALUES PROPOSED BY P. GRICE*

LAUREN DE LACERDA NUNES\*

Universidade Federal do Pampa (UNIPAMPA-RS), Brasil. laurenlacerdanunes@gmail.com

GABRIEL GARMENDIA DA TRINDADE\*\*

Universidade Federal de Santa Maria (UFSM-RS), Brasil. garmendia\_gabriel@hotmail.com

RECIBIDO EL 23 DE NOVIEMBRE DE 2012 Y APROBADO EL 14 DE DICIEMBRE DE 2012

## RESUMO ABSTRACT

Este artigo pretende analisar tópicos específicos do primeiro capítulo intitulado "Value and objectivity" do livro *Conception of value* de P. Grice. Através desta análise pretende-se realizar uma discussão metaética acerca da possibilidade de uma objetividade ou ceticismo com relação aos valores morais. Os referidos tópicos são: 1) a abordagem de P. Grice acerca da Teoria do Erro de J. Mackie; 2) a contribuição de R. M. Hare à discussão da Teoria do Erro de J. Mackie; 3) as críticas dirigidas por Grice à abordagem de R. M. Hare e 4) a análise de P. Grice acerca do anti-objetivismo radical de J. Mackie contraposto à sua visão objetivista. Conclui-se a discussão questionando-se a cética Teoria do Erro de Mackie que afirma serem os valores morais falsos quando na realidade Mackie os teria tratado metodologicamente como "sem sentido". Com isso, o ceticismo de valores proposto por Mackie se desqualifica frente à possibilidade do objetivismo dos valores proposto por P. Grice.

This article aims at analyzing specific topics of the first chapter "Value and objectivity" of P Grice's book *Conception of value*. Through this analysis we intend to do a metaethical discussion about the possibility of an objectivity or skepticism concerning moral values. The referred topics are: 1) P. Grice's approach of J. Mackie's Error Theory; 2) the contribution of R. M. Hare to the discussion of J. Mackie's Error Theory; 3) P. Grice's criticisms to R. M. Hare's approach and 4) P. Grice's analysis of the radical anti-objectivism of J. Mackie while opposed to his own objectivist vision. The discussion is concluded by disagreeing with J. Mackie's Error Theory, which asserts moral values as false, when in fact Mackie would have treated it methodologically as "no sense". With it, the skepticism of values proposed by J. Mackie is disqualified when it faces the possibility of an objectivism of values proposed by P. Grice.

## PALAVRAS CHAVE KEY WORDS

objetivismo, Grice, Teoria do Erro.

objectivism, P. Grice, Error Theory.

\* Doutoranda pelo Programa de Pós-Graduação em Filosofia - PPGF da Universidade Federal de Santa Maria (UFSM-RS), Brasil. Professora assistente na área de humanidades da Universidade Federal do Pampa (UNIPAMPA-RS), Campus São Borja.

\*\* Mestrando pelo Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Maria (UFSM-RS), Brasil. Licenciado em Filosofia pela Universidade Federal de Santa Maria (UFSM-RS). Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## Introdução

O presente trabalho<sup>1</sup> pretende expor em linhas gerais alguns tópicos do primeiro capítulo intitulado “Value and objectivity” do livro *Conception of value* de P. Grice. Neste capítulo Grice pretende elucidar o que seria propriamente uma concepção acerca dos valores morais. É importante ressaltar que, para Grice, o que importa é oferecer uma noção geral de valor, ou seja, o lugar desse conceito no mapa de estudos da filosofia e a sua conexão com outras questões também representadas nesse mapa.

Entretanto, Grice não pretende com isso oferecer uma leitura transcendental ou metafísica dos valores, apesar de ressaltar que deve-se estar preparado para oferecer respostas com adequado suporte metafísico, se isso for possível, quando o assunto é a concepção dos valores. Ao se buscar esse adequado suporte metafísico Grice defende que isso fornecerá uma clara configuração da natureza da ética teórica e o seu lugar na taxonomia das investigações que constituem a filosofia. Pois, a teorização em ética não é isolada: reclama o mesmo método das outras temáticas filosóficas.

Oferecidas tais considerações preliminares, pode-se passar ao que Grice realmente tem como metas em seu capítulo “Value and objectivity”. Primeiramente, pretende tratar problemas fundamentais acerca da natureza do valor, que poderiam ser divididas em: (i) Questões sobre a objetividade dos valores e (ii) Questões sobre a possibilidade de defender ou refutar o ceticismo sobre os valores. É importante notar que aquilo que subjaz tanto a primeira quanto à segunda questão é o fato de que Grice busca, através de sua investigação sobre objetividade, responder à questão elementar de se há uma “entidade” que corresponda àquilo que comumente chama-se de “valor”. Seria o ceticismo a favor ou contra o objetivismo? Essas parecem ser questões provocativas, no sentido em que Grice afirma que pretende “descobrir” em sua análise quem exatamente seria “o cético”, se o objetivista com relação aos valores ou o subjetivista.

Dessa forma, a escrita de Grice volta-se nesse primeiro capítulo a esclarecer questões sobre a objetividade dos valores relacionando-as o tempo todo com o ceticismo moral. Na medida em que Grice já escreveu em outras obras em defesa de uma espécie de objetivismo na moralidade,

---

<sup>1</sup> As obras em inglês utilizadas no presente trabalho são uma tradução livre realizada pelos autores.

ele deseja determinar se tal linha de raciocínio é defensável ou não. Por isso, um de seus objetivos primordiais no presente capítulo é abordar duas concepções anti-objetivistas: J. L. Mackie e P. Foot. Este trabalho se focará especialmente na interpretação de Grice acerca de J. L. Mackie e sua Teoria do Erro.

## I

### Grice e uma exposição preliminar da Teoria do Erro de Mackie

Grice inicia sua exposição acerca da Teoria do Erro de Mackie com a célebre frase deste autor que afirma “Não há valores morais objetivos” (Mackie 15). Quais teriam o sido os passos de Mackie para chegar a tal afirmação radical acerca dos valores morais? De acordo com Grice, Mackie possui uma tese central a respeito do status da ética. Essa tese diria respeito ao fato de Mackie não considerar os “valores” como uma exclusividade da ética: outras áreas do saber também os possuiriam e também nessas áreas a objetividade de tais valores seria de vital importância. Um exemplo seriam os valores estéticos. Pois, como afirmou Mackie, uma vez que a objetividade é considerada, as mesmas observações valem tanto para valores estéticos quanto para valores morais, pois ambos teriam o mesmo status. Uma discussão séria sobre a objetividade dos valores estéticos só não ocorre porque estes trazem uma problemática “menos urgente” (Ibid. 43).

Dito isso, Grice retorna à frase célebre de Mackie de que “não há valores morais objetivos” e afirma que geralmente existem três reações típicas a ela:

- 1) Alguns autores a veem como falsa, perniciosa e uma ameaça à moralidade;
- 2) Alguns autores a veem como uma verdade trivial que não vale a pena ser discutida;
- 3) Alguns autores a consideram sem significado ou vazia, não trazendo nenhuma questão real. (Grice 26)

Grice opta por tratar da terceira questão, associada por Mackie com uma reação típica do pensamento de Hare. Dessa forma, é iniciado outro processo no texto de Grice que prioriza a análise da terceira questão supracitada e as reações de Mackie com relação às colocações de Hare.

## II

### Grice, Hare, Mackie e a objetividade dos valores

Grice inicia sua exposição afirmando que sequer possui significado para Hare uma afirmação acerca da objetividade dos valores. Pois, a atitude de pensar que “alguém agiu errado” e censurar essa pessoa por isso, é comum tanto a objetivistas quanto a subjetivistas. Hare afirma que para os subjetivistas trata-se de uma atitude de desaprovação e para os objetivistas trata-se de uma intuição moral. Seriam apenas nomes diferentes para a mesma coisa e nenhum poderia ser preferível ao outro.

Importante ser dito que tal concepção de Hare reflete muito de duas ideias expostas em *A Linguagem da moral*. Nesta obra, Hare realiza um estudo acerca das inferências morais e em especial acerca da natureza dos imperativos. Acaba concluindo que seria possível separar a parte “frástica” da parte “nêustica” de um imperativo e através de tais distinções estabeleceu aquilo que haveria de comum tanto a imperativos quanto a indicativos. Evidentemente, esse tópico mereceria ser mais aprofundado, e como esse trabalho não pretende expor o pensamento de Hare de maneira exaustiva, basta apenas mencionar que esta é uma temática recorrente em suas obras.

Dando continuidade à exposição de Grice acerca do pensamento de Hare, pode se passar à exposição de Grice do argumento deste último. O mencionado argumento é o de que não haveria diferença substancial entre subjetivistas e objetivistas no tocante à concepção dos valores morais, como foi dito.

Suponha-se que A expressa ou reporta a presença de desaprovação em comer carne em si mesmo e que B expressa sua falta de desaprovação em comer carne em si mesmo. Passe-se à análise do argumento com relação a A e B de acordo com Hare (25-7):

A pensa que comer carne é errado

B pensa que comer carne não é errado, que ele mesmo (não o subjetivista).

está em posição de afirmar que um deles está errado.

A isso, o subjetivista de Hare responderia:

1. Se alguém (x) pensa que A julga erroneamente que comer carne é errado.

≡ x desaprova o julgamento de A de que comer carne é errado.

≡ x desaprova a desaprovação de A sobre comer carne.

≡ x não desaprova comer carne.

Então:

2. Se alguém (x) pensa que B julga erroneamente que comer carne não é errado.

≡ x desaprova o julgamento de B de que comer carne não é errado.

≡ x desaprova a não-desaprovação de B sobre comer carne.

≡ x desaprova comer carne.

Então:

3) Qualquer pessoa x deve desaprovar ou não-desaprovar comer carne (desaprovação pode estar ou não presente nesta pessoa).

Logo,

4) Qualquer pessoa x deve julgar que ou A ou B julgaram erroneamente com relação a comer ou não comer carne (3º excluído).

É importante ressaltar que de acordo com o argumento exposto, “julgar erroneamente” é o mesmo que desaprovar. Além de sustentar tal argumento, Hare em uma citação realizada por Mackie chega a supor a existência de dois mundos: um no qual os valores foram construídos objetivamente e outro no qual foram aniquilados. Nos dois mundos as pessoas estarão preocupadas com as mesmas coisas, afirma Hare, não haverá diferença na preocupação “subjetiva” que as pessoas terão pelas coisas, apenas no seu valor “objetivo”. Qual a diferença entre os estados de coisas nesses dois mundos? “Nenhuma” é a única resposta fornecida por Hare.

Partindo dessa ideia de Hare, Mackie decide examinar a indistinguibilidade dos dois mundos propostos por Hare e estabelece três pontos contra esta tese: 1) Hare não consegue provar o que pretende com os seus dois mundos, apenas abre a questão de que é preciso pensar em uma ética de primeira e segunda ordem; 2) os julgamentos feitos em uma ética

de primeira ordem devem ser totalmente independentes de qualquer julgamento contra ou a favor da objetividade dos valores, que cairão em uma ética de segunda ordem. Tendo isso exposto, Mackie afirma que isso tudo não mostra, como Hare gostaria, o vazio, ou a indecisão sobre tais questões e objetividade. Tais questões não são vazias como Hare pensa como Mackie mostra em seus dois comentários posteriores.

O primeiro dos comentários de Mackie diz respeito ao fato de que se forem admitidas objetivamente crenças de valores, elas se proverão de um apoio justificativo para as valorações em geral, o que de outra forma, não ocorreria. O segundo comentário de Mackie diz o seguinte: se o mundo for organizado em torno de valores objetivos, como Hare supõe, estaria disponível então, uma maneira simples de modificar ou adquirir as direções de interesse de cada um.

Afinal, se poderiam deixar as realidades do reino dos valores influenciarem as atitudes, “deixando o pensamento ser controlado pela maneira como as coisas fossem” (Mackie 22).

Para Mackie, Hare expõe a sua teoria a uma crítica positivista. Seria até mesmo possível vislumbrar um “positivismo hareano”, comparável com o berkeliano, que insiste que as aparências sejam exatamente aquilo que são, exista ou não um mundo material por detrás delas. Nesta parte do raciocínio de Mackie, Grice intervém ao afirmar não estar totalmente satisfeito com as críticas de Mackie dirigidas a Hare. De acordo com ele, Mackie não teria visto outros pontos vitais na teoria de Hare, os quais poderiam mudar a interpretação de alguns fatores. A próxima sessão se ocupará, portanto, das críticas de Grice endereçadas a Hare, considerando ainda, a Teoria do Erro de Mackie como objeto de análise.

### III

#### As críticas de Grice a Hare: objetivismo x subjetivismo

Grice considera errônea a tendência de Mackie a apostar toda a sua crítica a Hare na distinção entre uma ética de primeira e segunda ordem. De acordo com Grice a real fraqueza do argumento de Hare reside na conclusão do argumento anteriormente exposto:

- 3) Qualquer pessoa x deve desaprovar ou não-desaprovar comer carne (desaprovação pode estar ou não presente

nesta pessoa). Logo, 4) Qualquer pessoa  $x$  deve julgar que ou  $A$  ou  $B$  julgam erroneamente com relação a comer ou não comer carne. (3º excluído)

A conclusão do argumento repousa essencialmente na ideia de que para qualquer conteúdo  $\phi$ , qualquer pessoa deve desaprovar  $\phi$  ou não desaprovar  $\phi$ . Hare afirma que o raciocínio trata-se de uma tautologia, mas Grice diz que isso não fornece ou implica a premissa vital que ele precisa para o seu argumento ir adiante. Qual seja: “para qualquer  $\phi$ , qualquer pessoa tem ou uma atitude de desaprovação com relação à  $\phi$  ou uma atitude de não desaprovação com relação à  $\phi$ ”.

Logo, para Grice o raciocínio exposto não se trata de uma tautologia, uma vez que a ausência de desaprovação apenas conduz a uma atitude de não desaprovação se alguma condição adicional for também preenchida: a pessoa em questão deve ter conhecimento mínimo sobre o assunto. Dessa forma, tudo isso mostra que o argumento de Hare não serve formalmente para dissolver a diferença entre objetivismo e subjetivismo, como ele pretendeu.

Apesar de considerar que Mackie não dirigiu suas críticas ao ponto certo no pensamento de Hare, Grice afirma estar inclinado a considerar que, ao final da presente discussão, Mackie estaria certo.

Da mesma forma que Mackie, Grice não se sente apto a concordar com Hare na afirmação de que não haja diferença entre subjetivistas e objetivistas. E dada à falha apontada no argumento de Hare, Grice toma para si, a partir de então, a tarefa de identificar mais precisamente a tese sobre a qual objetivistas e subjetivistas discordariam entre si obrigatoriamente. Dessa forma, a partir da consideração de que Mackie é um anti-objetivista Grice vai tentar elucidar qual tese, claramente, Mackie sustenta.

#### IV

#### A análise de Grice acerca do anti-objetivismo da Teoria do Erro de Mackie

Em primeiro lugar, Grice julga adequado que se analise o que são os tópicos de primeira e segunda ordem em ética a que Mackie se refere, localizados nas distinções de mundos proposta por Hare. A explicação vem na forma de exemplos. Julgamentos éticos de primeira ordem



consistem, de acordo com Mackie, em comentários avaliativos sobre ações particulares e princípios amplos e gerais como, por exemplo, “todos devem lutar pela felicidade geral”. Por sua vez, julgamentos éticos de segunda ordem dizem respeito ao que ocorre quando alguém realiza um julgamento ético de primeira ordem: se tal julgamento expressa uma descoberta ou uma decisão; ou se gera um ponto interessante para se pensar sobre a ética; ou aborda uma visão sobre o significado dos termos éticos. Dito isso, Grice enumera alguns aspectos salientes na teoria de Mackie.

O primeiro aspecto saliente para Grice é que na tese anti-objetivista de Mackie, questões sobre o status da ética são de segunda ordem, bem como sua teoria como um todo. Além disso, para Grice, Mackie afirma haver considerável medida de independência entre os reinos de primeira e segunda ordem. Isso dito menciona-se o exemplo do ceticismo moral. De acordo com Mackie alguém pode ser um cético moral de segunda ordem sem ser um cético moral de primeira ordem e vice-versa. Por exemplo, uma pessoa pode sustentar fortes visões morais de conteúdo bastante convencional, ao mesmo tempo em que acredita que tais visões sejam simples atitudes e condutas que condigam com o comportamento da maioria. Da mesma forma, alguém pode rejeitar toda a moral estabelecida ao mesmo tempo em que acredita ser uma verdade objetiva que essa moralidade seja má e corrupta.

O segundo aspecto que Grice considera digno de nota na teoria de Mackie é que a sua versão de um anti-objetivismo ou ceticismo moral é uma tese negativa. Para esclarecer esse ponto é importante elucidar o entendimento do próprio Grice acerca do objetivismo. Em primeiro lugar, ele aborda a visão de Mackie. De acordo com Mackie, uma visão anti-objetivista supostamente afirma que não existem entidades ou relações de certo tipo, valores e demandas objetivos, que muitos pensam existir. Do ponto de vista de Grice, uma visão objetivista, apesar de sua aparência positiva, também se torna inteligível enquanto negação de alguma posição que carregue o rótulo “subjetivista”. Ou seja, como a negação da disputa de que as sentenças morais são redutíveis ou realmente correspondem à expressão de certas atitudes como aprovação ou desaprovação. Para Grice, nesse sentido, o subjetivismo ou o não-objetivismo, apesar de sua aparência negativa, seria o termo positivo: afinal, os valores postulados poderiam ser reduzidos a expressões subjetivas de aprovação e desaprovação, algo que o objetivista condena. Mas, para Mackie, afirma Grice, “objetivista” não é um termo cripto-



negativo, ou seja, que esconda uma faceta positiva. Grice considera que Mackie avaliou erroneamente a sua própria visão do objetivismo ao afirmar isso.

O terceiro aspecto saliente da teoria de Mackie mencionado por Grice é que a negação do objetivismo efetuada por Mackie não pode ser tomada como uma tese ética de segunda ordem, meramente semântica (sobre o significado dos termos valorativos). Nem como uma tese lógica (sobre a estrutura e os tipos de determinados argumentos). Para Grice a tese de Mackie seria ontológica: aceitaria ou negaria a existência de certos itens na realidade, e por isso, só poderia se tratar de uma tese de primeira ordem.

O quarto e último aspecto saliente da teoria de Mackie para Grice diz respeito à afirmação polêmica de Mackie (que caracteriza toda a sua teoria) de que as valorações ditas ou pensadas por pessoas ordinárias são sistematicamente, compreensivelmente e abrangentemente falsas. Em suma, os valores morais (e não apenas eles) seriam falsos. De acordo com Grice, o ceticismo moral extremo de Mackie é tomado pelo autor não como uma tese do significado daquilo que os julgamentos morais afirmam, mas sim, sobre a não presença de certos itens no mundo real. Por isso, parece adequado para Mackie afirmar que a existência de valores reais é sustentada por aquilo que as pessoas ordinárias pensam e dizem. Contudo, isso não seria, de fato, um aspecto do mundo, o que gera a consequência de que as valorações proferidas pelas pessoas seriam falsas.

Para tornar o ponto anterior mais inteligível, Grice expõe a comparação efetuada por Mackie acerca de sua Teoria do Erro com a teoria das cores de Boyle (Mackie 20) e Locke (Ibid. 20). Para esses autores trata-se de uma crença falsa aquela de que as coisas no mundo real possuem qualidades como a cor. As coisas reais possuiriam apenas disposições que dariam a sensação da cor, bem como qualidades primárias (forma, tamanho), que seriam fundamento de tais disposições. Mackie salienta que nenhum desses itens que explicam a sensação da cor deve ser identificado com as cores particulares, pois nada pode ser identificado com uma cor. Para Mackie a situação dos valores seria análoga à das cores, considerando-se o pensamento de Boyle e Locke.

De acordo com Grice o raciocínio de Mackie até agora exposto leva a duas questões: 1) Por que Mackie sustenta que demandas por objetividade

seriam incorporadas em julgamentos de valores ordinários? E 2) Por que Mackie acredita que tais demandas sejam falsas?

Alguns pontos preliminares antes de se passar à resposta da primeira questão. Para Grice, Mackie sustenta que uma demanda por objetividade esteja incorporada em um julgamento de valor ordinário. Tal demanda seria, conseqüentemente, parte do significado de tais julgamentos de valor (ou as sentenças nas quais são expressos). Isso é algo que não parece ser considerado por Mackie: tal inclusão de significado. Logo, para Grice, Mackie não pode afirmar consistentemente que o seu anti-objetivismo não é uma tese sobre significado de afirmações de valores. A única coisa que Mackie poderia afirmar é que sua tese contém uma tese sobre significado dos valores, mas, não estaria restrita a isso.

Para Grice, a visão de Mackie de que uma demanda moral por objetividade seria incorporada a um julgamento de valor ordinário parece repousar, de forma insegura, na seguinte sugestão. A de que há duas alternativas à sua suposição de que julgamentos de valor ordinários teriam a função de introduzir valores objetivos no discurso sobre a conduta e a ação dos agentes: o não-cognitivismo e o naturalismo.

Mackie caracteriza o não-cognitivismo como a visão moral que defende que os juízos morais ou valores não seriam asserções, no sentido de serem acessíveis ao entendimento, mas sim expressões de sentimentos e desejos do agente. O naturalismo, por seu turno, seria a visão moral onde os juízos morais corresponderiam ou seriam redutíveis a objetos naturais. Mackie lembra, de maneira perspicaz, que qualquer uma dessas visões torna inoperante a aparente autoridade da ética.

A referida autoridade da ética repousaria, de acordo com Mackie, nos sentimentos do homem ordinário. Parece que este homem sente certo desconforto com relação à moralidade, que somente seria aliviado quando ele consegue lidar com questões como “tal curso de ação é errado em si mesmo ou certo em si mesmo”. Inevitavelmente isso leva Mackie a mostrar que esse sentimento nada mais é do que uma espécie de “objetivismo cotidiano”, que afirmaria que as qualidades não-naturais dos juízos morais seriam apenas uma espécie de reconstrução “de filósofo”. Por isso, Mackie defende a tese de que os elementos objetivistas estariam embutidos em julgamentos morais ordinários e que estes, seriam falsos. Para tornar ainda mais claro seu ponto de vista, Mackie ainda apresenta mais dois argumentos: o argumento da

relatividade e o argumento da estranheza. Passe-se à análise de cada um desses argumentos.

O argumento da relatividade de Mackie baseia-se na premissa de que há diferentes códigos morais entre diferentes sociedades, em diferentes períodos entre grupos ou classes mais complexas e etc. A um primeiro olhar, tais diferenças seriam assunto da antropologia e por si só, não sustentariam nenhuma conclusão ética de primeira ou de segunda ordem. Porém, podem interferir de maneira indireta.

Mackie supõe que é mais plausível pensar que crenças morais refletem modos de vida mais do que qualquer outra coisa é capaz de fazer. Por exemplo, as pessoas que aprovam a monogamia, vivem monogamicamente, ao invés de viverem monogamicamente porque aprovam a monogamia. Tais constatações tornam, na visão de Mackie, muito mais fácil explicar as divergências de valores morais, pois se podem encarar tais divergências como produtos de diferentes modos de vida e não somente como percepções distorcidas de valores morais supostamente objetivos.

O argumento da relatividade proposto por Mackie apresenta-se como uma das suas principais arma contra o objetivismo. Mas Grice pontua que o objetivista poderia contra-argumentar afirmando que as crenças morais divergentes seriam derivadas do resultado da operação de um único conjunto já escolhido de princípios gerais em circunstâncias diversas de aplicação. Contudo, a contra-argumentação do objetivista tende a ruir, pois como lembra Grice, muitas vezes parece que crenças morais divergentes não necessariamente surgiram por derivação de princípios gerais, mas sim, de algo como “senso moral” ou “intuição moral”.

O argumento da estranheza como proposto por Mackie é um dos mais importantes para sua teoria. De acordo com esse argumento, para sustentar sua posição, o objetivista postula entidades e aspectos valorativos de tipos completamente diferentes daqueles a que estamos familiarizados. E como não há nenhuma faculdade cognitiva capaz de percebê-los, o objetivista se vê, na imensa maioria dos casos, obrigado a afirmar que possuímos algo como uma “intuição moral” ou “senso moral” capaz de apreender tais entidades, que seriam “estranhas” (Mackie 37-8).

Realizadas todas essas colocações acerca da teoria de Mackie, Grice ocupa-se de refletir acerca de alguns pontos principais, que de acordo com ele, demonstram a fragilidade da Teoria do Erro. Em especial, Grice retoma alguns pontos acerca do objetivismo e em que sentido exatamente Mackie teria tomado o termo, bem como o argumento da “estranheza” e o argumento da relatividade de Mackie, supramencionados.

Há, para Grice, como foi mostrada, uma versão positiva e uma versão negativa do objetivismo. A versão positiva afirma que atribuir objetividade a algo é proclamá-lo como pertencente ao “esquema básico” da realidade. De acordo com Grice, essa parece ser a versão defendida por Mackie. A versão negativa do objetivismo seria, afirma Grice, atribuir objetividade a algo negando que sentenças sobre este objeto sejam redutíveis ou elimináveis (de segunda ordem). Para Grice, Mackie vacilou entre essas duas versões de objetivismo ao formular o seu célebre anti-objetivismo.

Além disso, em uma análise dos argumentos da “estranheza” e da relatividade, Grice afirma não estar convencido de que não existam valores morais objetivos. Com relação ao primeiro argumento, Grice afirma haver dois tipos de “estranheza” não mencionados por Mackie. O primeiro tipo seria dizer que certas propriedades não-naturais seriam consideradas valores que de alguma maneira “misteriosa” se colocariam acima de aspectos naturais e familiares. O segundo tipo de “estranheza” estaria na suposição de que o reconhecimento da presença dessas propriedades não-naturais motivaria ou poderia motivar os agentes sem auxílio de nenhum interesse ou desejo que estes poderiam ter.

Ainda, as “estranhezas” a que Mackie denuncia não seriam, de acordo com Grice “esqueletos nos armários” dos objetivistas, como ele teria proclamado. Estas seriam apenas alguns dos desafios que os objetivistas têm de resolver e até mesmo o próprio Mackie, mesmo que este já tenha se adiantado na resolução dos mesmos. A verdade é que tais desafios ainda não foram esgotados e por isso, salienta Grice, não podem ser jogados como “tijolos contra toda a construção objetivista” (Grice 45).

Para Grice é como se Mackie estivesse dizendo que duvidasse sinceramente da aritmética, pois se tal coisa fosse possível, teria de ser sobre números, e números são coisas muito estranhas, completamente inacessíveis a qualquer observação. Ou ainda, como se Mackie estivesse a pensar algo como “não vejo como o matrimônio possa existir”, pois se

existisse as pessoas deveriam ser amarrada uma na outra no casamento, e tudo que o que se vê na vida real e no cinema sugere que a única maneira de as pessoas estarem amarradas umas às outras é através de cordas, complementa Grice. Em suma, Grice desqualifica os principais argumentos utilizados por Mackie em sua visão anti-objetivista.

### Conclusão

Pode-se concluir a presente análise do primeiro capítulo de Grice “Value and objectivity” de seu livro *Conception of value* elucidando alguns pontos dignos de nota em seu texto. O primeiro ponto é tornar evidente a escolha de Grice por Mackie para se analisar uma visão anti-objetivista. A opção de Grice por Mackie revela a intenção do primeiro por realmente pôr em cheque a sua própria visão acerca de um objetivismo dos valores. Pois, Mackie é um dos autores mais radicais em ética quando se trata de negar que exista algo como uma moralidade possivelmente objetiva. Mackie sequer admite os valores morais, chegando ao extremo de afirmar não apenas que estes não existiriam, mas que seriam falsos.

Também foram abordados pontos do pensamento de Hare, outro autor chamado à discussão por Grice e citado na obra de Mackie. Pois, se Mackie é radical ao afirmar que os valores morais seriam falsos, Hare também causa certo estranhamento ao afirmar que não haveria diferença substancial nas visões objetivista e subjetivista da moralidade. Grice analisou os argumentos de cada um dos autores (priorizando, evidentemente, Mackie) e mostrou porque não estaria satisfeito com nenhuma das leituras.

Sobre Hare, Grice expõe a análise realizada por Mackie a respeito das ideias do autor e de sua possível resposta ao ceticismo com relação aos valores levantado por Mackie. Mackie apenas expõe em linhas gerais a resposta fornecida por Hare a respeito da objetividade dos valores: a de que não haveria diferença substancial entre objetivistas e subjetivistas. Grice analisa de forma profunda tal colocação de Hare ao reconstruir seu argumento e provar que o mesmo não se sustenta.

A respeito da Teoria do Erro de Mackie, Grice torna evidente seu descontentamento com a argumentação deste. Para Grice, parece que Mackie atribui falsidade a algo sem sentido. Mackie nega o objetivismo, mas precisa do valor objetivamente compreendido para chamá-lo

de falso. Talvez, se Mackie tivesse optado por considerar os valores morais como “sem sentido”, pudesse escapar de tal crítica, pois poderia argumentar que os objetivistas atribuem sentido a algo que não o possui. Ainda assim, Mackie teria permanecido na esfera de uma teoria do erro.

Além disso, Grice coloca que os argumentos centrais de Mackie, a saber, o argumento da “estranheza” e da relatividade dos valores morais, é na realidade bastante comum e podem ser considerados desafios aos objetivistas e para o próprio Mackie, e não elementos capazes de pôr totalmente em cheque toda uma concepção objetivista da moralidade.

#### REFERÊNCIAS BIBLIOGRÁFICAS

Grice, P. “Value and objectivity”. *The Conception of value*. Oxford: Oxford University Press, 2001, p. 23-45. Print.

Hare, R. M. “Some confusions about subjectivism”. *Essays in ethical theory*. London: London Paperbacks, 1993, p. 15-32. Print.

---. *A Linguagem da Moral*. São Paulo: Martins Fontes, 1996. Print.

Mackie, J. L. *Ethics: Inventing right and wrong*. London: Penguin Philosophy, 1977. Print.

# O MODELO ONTOLÓGICO ESTRATIFICADO NO NATURALISMO BIOLÓGICO DE JOHN SEARLE: UMA CONTROVÉRSIA COM JAEGWON KIM\*

*THE ONTOLOGICAL LAYERED MODEL IN JOHN SEARLE'S BIOLOGICAL NATURALISM:  
A CONTROVERSY WITH JAEGWON KIM*

TÁRIK DE ATHAYDE PRATA\*\*

Universidade Federal de Pernambuco (UFPE), Brasil. tarikbilden@yahoo.de

RECIBIDO EL 1 DE OCTUBRE DE 2012 Y APROBADO EL 23 DE NOVIEMBRE DE 2012

## RESUMO ABSTRACT

O artigo discute a controvérsia entre Jaegwon Kim (1934-) e John Searle (1932-) a respeito do modo como este último utiliza o modelo ontológico estratificado em sua filosofia da mente. Após uma exposição e discussão geral desse modelo (seção 2), são discutidas as respostas de Searle a duas das críticas de Kim ao naturalismo biológico (seção 3). Apesar de usar o modelo estratificado e tentar construir um fisicalismo não redutivo, Searle recai em um dualismo de propriedades (seção 4).

This paper discusses the controversy between Jaegwon Kim (1934-) and John Searle (1932-) about the way the latter uses the ontological layered model in his philosophy of mind. After an exposition and a general discussion of such model (section 2), Searle's answers to some of Kim's criticisms to biological naturalism are discussed (section 3). Even though Searle uses a layered model and tries to construct a nonreductive physicalism, he ends up recalling a property dualism.

## PALAVRAS CHAVE KEY WORDS

Jaegwon Kim, John Searle, níveis de complexidade, ontologia, problema mente-corpo.

Jaegwon Kim, John Searle, levels of complexity, ontology, mind-body problem.

---

\* O presente trabalho foi produzido no âmbito do projeto de pesquisa "O naturalismo biológico de John Searle e o fisicalismo não-redutivo", registrado no Departamento de Filosofia da Universidade Federal de Pernambuco (UFPE).

\*\* Doutor em Filosofia pela Ruprecht-Karl Universität Heidelberg (Alemanha). Professor adjunto do Departamento de Filosofia da Universidade Federal de Pernambuco (UFPE), Recife - Brasil. Possui experiência de pesquisa nas áreas de filosofia da mente e filosofia da psicologia.



## I Introdução

O debate a respeito do estatuto ontológico dos fenômenos mentais e de sua relação com os fenômenos físicos se divide entre diversas posições possíveis, sendo algumas extremamente difundidas, enquanto outras possuem pouquíssimos defensores. Em termos extremamente gerais, o debate mente-corpo pode ser dividido entre os defensores do *monismo* (só há um domínio de fenômenos) e do *dualismo* (existem dois domínios ontologicamente diversos), porém, essas posições básicas podem ser concebidas de maneiras muito diferentes. O dualismo, na sua forma mais tradicional, foi concebido como um dualismo entre *substâncias*, no sentido ontológico do termo (Imaguire 272-83), (Heil 19), (Maslin 40) e (Cunningham 3), porém, mais recentemente, tornou-se predominante um dualismo entre *propriedades* mentais e físicas. Já o monismo pode ser articulado em termos *mentalistas* (tudo o que existe é mental), em termos *fiscalistas* (tudo o que existe é físico), ou *neutros* (defendendo-se que “o mesmo ‘material’ primitivo, ordenado de diferentes maneiras, constitui, de um lado, o mundo mental e, de outro, o mundo físico” (Blackburn 75)). Porém, no cenário atual da filosofia da mente, o verdadeiro embate é travado entre o dualismo *de propriedades* e o monismo *fiscalista*.

John Rogers Searle (1932-), professor da Universidade da Califórnia (Berkeley), é um dos filósofos mais influentes de nossa época<sup>1</sup>, e ao longo dos últimos cinquenta anos procurou articular uma explicação abrangente de como diferentes aspectos da realidade humana (tais como consciência, intencionalidade, linguagem, sociabilidade, liberdade, etc.) se integram ao mundo tal como estudado pelas ciências da natureza (em especial a física, a química e a biologia) (Searle, *Consciousness and 1*) ([IX-X])<sup>2</sup>. Nesse vasto empreendimento filosófico, a reflexão a respeito da natureza da mente e sua relação com o mundo físico desempenha um papel importante, já que os fenômenos mentais (em especial a consciência e a intencionalidade) são vistos por Searle como o fundamento da linguagem (Searle, *Intentionality* vii) ([VII]), da sociabilidade (Searle, *Mind language* 121) ([113]) e de quaisquer fenômenos humanos.

<sup>1</sup> “John R. Searle, (...) é um dos mais famosos e influentes filósofos da atualidade” (Grewendorf Meggle vii). “John Searle tem sido uma figura dominante e altamente influente entre os filósofos contemporâneos” (Smith i).

<sup>2</sup> O primeiro número é o da página na edição original. O número entre colchetes é o da paginação da tradução para o português, quando disponível. Nos casos onde não há número entre colchetes, a citação foi traduzida pelo autor do presente artigo.



Com sua teoria a respeito da mente, o “naturalismo biológico”, Searle pretende *superar* as posições tradicionais no debate mente-corpo (Prata, *Irredutibilidade ontológica* 110), que ou compreendem a mente como uma realidade essencialmente diversa dos fenômenos físicos (dualismo), ou a compreendem como parte da realidade física (fiscalismo). Essa teoria se baseia, por um lado, na ideia de que os fenômenos mentais são *causados* por processos cerebrais no nível micro, ao mesmo tempo em que são *propriedades do sistema cerebral* no nível macro -- (Searle, *The Rediscovery* 1) ([7]); (Searle *John* 545); (*Consciousness and* 9) ([4]) -- (de modo que a eficácia causal da mente sobre fenômenos do mundo físico não seria problemática). Mas, por outro lado, Searle defende com veemência que os fenômenos mentais são *ontologicamente irredutíveis* a qualquer fenômeno objetivo, em virtude do fato de eles *existirem de modo subjetivo* (Searle, *The Rediscovery* 117) ([170]; (*Mind a* 119)<sup>3</sup>. Ou seja, ao mesmo tempo em que concebe a mente como um fenômeno biológico, causado e realizado no cérebro, Searle defende a irredutibilidade *ontológica* das propriedades mentais. Desse modo, ele acredita ter articulado uma concepção que conserva as vantagens dessas duas posições, deixando de lado suas desvantagens (Searle, *Mind a* 126). Ao defender seu ponto de vista a respeito da mente, ele afirma de modo direto:

Aquilo em que quero insistir sem cessar é que podemos aceitar fatos óbvios da física – por exemplo, que o mundo é constituído inteiramente de partículas físicas em campos de força – sem, ao mesmo tempo, negar os fatos óbvios de nossas próprias experiências – por exemplo, que somos todos conscientes e que nossos estados de consciência têm propriedades fenomenológicas *irredutíveis* bastante específicas. (Searle, *The Rediscovery* 28) ([44-5])

Essa adesão à irredutibilidade da mente indica que Searle não é um fiscalista reducionista (nem eliminativista), de modo que só parecem restar duas opções: ou o naturalismo biológico é um dualismo de propriedades (uma posição problemática e pouco atraente), ou é um

<sup>3</sup> Em formulações recentes da teoria (Searle, *Mind a* 113-4); (*Dualism revisited* 170-1), ele resume o naturalismo biológico nas quatro teses seguintes: (1) A consciência é ontologicamente irredutível; (2) A consciência é causada por processos cerebrais; (3) A consciência é uma característica de nível superior do cérebro; (4) A consciência é causalmente eficaz.

fisicalismo não-redutivo<sup>4</sup> (essas concepções serão discutidas em maiores detalhes na próxima seção do presente artigo).

É importante ter em vista que há um indício muito forte de que Searle defende uma forma de *fisicalismo não reducionista*: em sua tentativa de construir uma explicação abrangente de como os fenômenos psicológicos e sociais se articulam com os fenômenos físicos, químicos e biológicos, Searle recorre a uma influente concepção, segundo a qual esses diferentes tipos de fenômenos se distribuem ao longo de determinados *níveis de complexidade* do mundo que nos cerca<sup>5</sup>.

Searle considera como uma parte fundamental da teoria atômica da matéria, a ideia de que tudo o que existe em nosso universo (sejam coisas de grandes dimensões como galáxias e planetas, ou coisas de dimensões menores como carros ou organismos vivos) são, em última instância, compostas por entidades extremamente pequenas que se costuma denominar “partículas”. Todas as entidades de maior porte

são constituídas de entidades menores que são, por sua vez, feitas de entidades ainda menores, até que finalmente atingimos o nível das moléculas, elas mesmas compostas de átomos, eles mesmos compostos de partículas subatômicas. (Searle, *The Rediscovery* 86) ([128])

E, uma vez que sistemas maiores são constituídos por entidades cada vez menores, é uma ideia fundamental da teoria atômica da matéria que muitos aspectos desses sistemas podem ser *explicados* em termos do comportamento de seus constituintes. Disso se segue que,

haverá diferentes níveis de explanação do mesmo fenômeno, dependendo de se vamos da esquerda para a direita de macro para macro ou de micro para micro,

---

4 Jaegwon Kim (*Supervenience and* 267), recusa o fisicalismo não redutivo como uma opção viável a respeito da mente, defendendo que as únicas opções reais são (a) o dualismo anti-fisicalista, (b) o reducionismo e (c) o eliminativismo. De minha parte, considero o fisicalismo não-redutivo como a opção mais desejável, e entendo que existem bons argumentos para defendê-lo (Pereboom Kornblith 723-4).

<sup>5</sup> De acordo com Kim: “O emergentismo, que floresceu durante a primeira metade do século XX, foi a primeira formulação sistemática do fisicalismo não-redutivo bem como do modelo multi-estratificado do mundo. Seus principais proponentes incluem não só filósofos acadêmicos como Samuel Alexander, C. D. Broad e A. O. Lovejoy, mas também cientistas de formação como C. Lloyd Morgan” (*Philosophy of* 226). A proximidade de Searle como emergentismo indica que a sua adesão ao modelo estratificado o aproxima do fisicalismo não-reducionista.

ou de baixo para cima de micro para macro. (Searle, *The Rediscovery* 87) ([129])<sup>6</sup>

Essas colocações evidenciam o compromisso de Searle com aquilo que o importante filósofo Jaegwon Kim denomina o modelo ontológico estratificado (Kim, *Supervenience and* 337), que remonta, pelo menos, aos emergentistas britânicos, que atuaram no início do século XX. Em uma discussão crítica a respeito do naturalismo biológico, Kim expressa a sua análise da maneira como o modelo estratificado é aplicado por Searle, adotando uma postura bastante negativa em relação à teoria deste último – “pode haver realmente uma solução simples para o problema mente-corpo, mas eu não creio que ela se encontre no naturalismo biológico de Searle” (Kim, 1995, p. 189).

De acordo com Kim (*Mental causation* 190), Searle *substitui* a antiquada ontologia cartesiana do mental e do físico pelo modelo multi-estratificado, modelo no qual, segundo Kim, a ordem entre os diferentes níveis é gerada pela relação de *ser parte de*, ideia que ele identifica na teoria de Searle, tomando o exemplo da passagem citada acima (Searle, *The Rediscovery* 86) ([128]), onde é afirmado que entidades maiores *são feitas de* [*made up of*] entidades cada vez menores.

Além disso, Kim ressalta que o principal significado do modelo estratificado para a metafísica e a metodologia da ciência seria *o tipo de relação* vigente entre as propriedades típicas de cada nível, e discute a relação de *superveniência causal* que Searle acredita existir entre processos cerebrais e estados conscientes (Kim, *Mental causation* 191). Ao mesmo tempo em que defende essa forma de superveniência, Searle rejeita o dualismo de propriedades. Entretanto, por outro lado, ele rejeita também a *redução* das propriedades conscientes aos processos cerebrais, o que o leva, na opinião de Kim, a recair no dualismo de propriedades por ele rejeitado (Kim, *Mental causation* 192).

Outra dificuldade da posição de Searle é que, apesar de ele próprio utilizar o já antigo e até hoje influente modelo estratificado, que (de acordo com Kim) seria uma *alternativa* ao modelo cartesiano, ele insiste com veemência que os debates atuais na filosofia da mente estão presos

<sup>6</sup> Searle explicita que a metáfora de “direita” e “esquerda” é baseada na convenção das línguas européias, segundo a qual se escreve nesse sentido (Searle, *The Rediscovery* 251) ([361]), nota 2, ao capítulo 4), e essa metáfora busca exprimir que se trata de eventos cronologicamente anteriores explicando eventos cronologicamente posteriores (Ibid. 87) ([129]).

ao modelo bifurcado de Descartes (que faz a cisão entre o mental e o físico), e que essa influência cartesiana é a fonte de todas as dificuldades que acometem as teorias da mente (Ibid. 192). Kim não consegue entender como Searle usa o modelo estratificado e, mesmo assim, não percebe que visões alternativas ao modelo cartesiano são desenvolvidas na filosofia atual<sup>7</sup>.

Mas em sua réplica ao texto de Kim, Searle discorda de vários pontos dessa análise (aqui apenas resumida)<sup>8</sup>. Na opinião de Searle, Kim se equivoca ao afirmar que “o modelo de diferentes níveis, ou estratos, de descrição de um sistema (...) é gerado pela relação de ‘ser uma parte’” (Searle, *Consciousness* 217), pois propriedades de sistemas físicos (em níveis de maior complexidade) *não são partes* de moléculas ou dos sistemas formados por moléculas. A análise de Kim se mostra equivocada, pois:

se a tese dos níveis fosse equivalente à relação parte-todo, então ela seria supérflua. É precisamente porque existem *características* sistêmicas de nível superior que não são partes dos sistemas que nós precisamos dos níveis em primeiro lugar. (Searle, *Consciousness* 217)

Além disso, Kim estaria enganado ao achar que a concepção dos níveis é uma alternativa ao cartesianismo, pois o modelo estratificado seria perfeitamente consistente com o dualismo de Decartes, já que cada uma dessas concepções é pensada para responder a questões bastante diferentes.

De acordo com Searle:

O cartesianismo é uma resposta à pergunta ‘quantas categorias metafísicas existem?’ A tese dos níveis é uma resposta à pergunta ‘como a realidade física se organiza?’ (Ibid. 218)

<sup>7</sup> “Os emergentistas britânicos foram provavelmente os primeiros, no início deste século, a dar uma articulação sistemática [ao modelo estratificado] e ele é o esquema metafísico que tem formado o pano de fundo da maior parte da discussão contemporânea sobre o problema mente-corpo. Assim, é um mistério porque Searle afirma, com alguma veemência, que os debates atuais sobre os problemas mente-corpo estão ainda inextricavelmente enredados com o modelo cartesiano de um mundo bifurcado, e que isso é a fonte de todos os males e deficiências que, na visão dele, tem assolado a atual filosofia da mente”. (Kim, *Mental Causation* 192)

<sup>8</sup> Por motivos de concentração temática, deixamos de lado as considerações de Kim a respeito da causalidade mente-corpo e a respeito do problema da sobredeterminação causal. Para uma discussão sobre esse problema (Prata, *Dificuldades da*).

Portanto, ao contrário do que afirma Kim (*Mental Causation* 192), não haveria nenhum “mistério” no fato de Searle aceitar o modelo estratificado e, ainda assim, considerar que o cartesianismo infesta a atual filosofia da mente (mesmo que de modo oculto)<sup>9</sup>.

Mas então se colocam as seguintes perguntas: as críticas de Kim são, de fato, equivocadas? Searle consegue, em suas respostas, articular uma concepção viável a respeito da tese dos níveis de complexidade? Responder essas perguntas seria um passo importante para se poder avaliar se o modo como Searle aplica o modelo ontológico estratificado é capaz de fundamentar sua teoria da mente. O presente artigo pretende responder essas perguntas, e esclarecer o modo como o modelo estratificado organiza a teoria de Searle sobre a relação mente-corpo. Para isso, será feito o seguinte percurso: em primeiro lugar, será feita uma exposição e discussão geral do modelo ontológico estratificado (seção 2). Em seguida serão avaliadas as réplicas de Searle resumidas acima, em especial a respeito das considerações de Kim sobre a *relação parte-todo* e sobre o modelo estratificado como *alternativa ao cartesianismo* (seção 3). Nas considerações finais, será discutida a insuficiência da réplica de Searle à acusação de uma recaída do naturalismo biológico em um dualismo de propriedades (seção 4).

## II

### O modelo ontológico estratificado

As posições filosóficas a respeito da ontologia da mente podem ser formuladas com base em um par de conceitos básicos: os conceitos de *coisa* e *propriedade*. É uma idéia bastante intuitiva (isto é, bastante de acordo com nosso modo pré-teórico de pensar) que as coisas ao nosso redor, como mesas, cadeiras, carros e pessoas, ao mesmo tempo em que são únicas, partilham muitas características em comum: diversas mesas podem ter, por exemplo, o mesmo peso, diversas cadeiras podem ser feitas do mesmo material, diversos carros podem ter a mesma cor, etc. Sendo assim, é usual dizer, por exemplo, que uma *coisa* (como uma cadeira) pesa, digamos, vinte quilos (ou tem a *propriedade* de pesar vinte quilos).

<sup>9</sup> Conforme será discutido a seguir (seções 3 e 4), Searle não responde detalhadamente a objeção de que sua teoria, por negar a redução da consciência, recai no *dualismo de propriedades*. A análise dessa deficiência nas réplicas de Searle será importante para nossas considerações finais (seção 4).

Coisas e propriedades podem ser concebidas de muitas maneiras. Um dos debates mais importantes da história da filosofia trata da questão sobre se as propriedades possuem um estatuto ontológico (isto é, um *modo de existência*) radicalmente diferente do das coisas particulares (de modo que elas poderiam existir em diversos lugares e tempos) ou se tudo o que existe é, tal como as coisas, restrito a porções determinadas do espaço e a períodos determinados do tempo (trata-se da “controvérsia dos universais” (Loux 18)). Por outro lado, existem diversas maneiras como se pode conceber as coisas: como formadas por um *substrato* incognoscível que “suporta” as propriedades (Imaguire 273), como um *feixe de propriedades* (desprovido de qualquer substrato) mantidas juntas pela relação de co-presença (Garrett 55-6) ou como aglomerados de propriedades particularizadas, os chamados *tropos* (Imaguire 286-7); (Garrett 59-60). Essas várias, e altamente controversas, concepções ontológicas sobre as coisas e as propriedades não serão abordadas aqui. Para os objetivos do presente trabalho, basta se ter em mente a distinção entre coisas (sejam elas formadas por um substrato particular, por um feixe de universais etc.) e propriedades (sejam elas universais ou tão particulares quanto as coisas), pois a problemática ontológica que será relevante aqui não diz respeito diretamente a esse debate sobre categorias<sup>10</sup>.

A problemática ontológica aqui enfocada diz respeito a certas distinções entre *domínios* do ser, distinções estas que, justamente, são fundamentais para que se compreenda a relação da mente com o mundo físico. A esse respeito foi especialmente influente a distinção operada por Descartes entre a alma e o corpo. Ele considerava que o nada não pode ter nenhuma propriedade (Beckermann 46), donde se segue que, para que uma propriedade exista, é necessário existir *algo* que a possui, de modo que as propriedades necessitam, sempre, de um portador, no qual elas inerem<sup>11</sup>. Assim, Descartes se integra à antiga tradição da ontologia da *substância*, que remonta a Aristóteles. O ponto é que ele defendia a existência de *dois* tipos de substâncias, diferenciados por certas propriedades essenciais mutuamente excludentes. Por um lado,

<sup>10</sup> Uma ontologia pode ser entendida como a lista das categorias do que *existe* (p. ex. substâncias, atributos, eventos, estados de coisas etc.). Sobre o debate a respeito das categorias ontológicas, cf. o excelente livro de Michael Loux (2006).

<sup>11</sup> “Uma das noções comuns é que o nada não pode ter nenhum atributo, propriedade ou qualidade. Por essa razão, logo que encontramos algum atributo podemos concluir que é o atributo de alguma substância, e que tal substância existe”. (Descartes 46 §52)



existiriam as substâncias cuja propriedade essencial é o *pensamento* (*res cogitans*), por outro, as substâncias cuja propriedade essencial é a *extensão espacial* (*res extensa*).

Para essa distinção, Descartes ofereceu uma série de argumentos que, por motivos de espaço, não poderemos discutir aqui<sup>12</sup>. O fato é que tal divisão entre domínios ontológicos acabou desacreditada. Os progressos da ciência natural desde o século XVII mostram como implausível a idéia de que existiriam particulares não físicos, o que levou a uma modificação significativa da concepção ontológica.

Nas palavras de Jaegwon Kim:

A imagem ontológica que tem dominado o pensamento contemporâneo sobre o problema mente-corpo é nitidamente diferente da imagem cartesiana. O modelo cartesiano de um mundo *bifurcado* foi substituído por um mundo *estratificado* [*layered*], uma estrutura hierarquicamente estratificada de 'níveis' ou 'ordens' de entidades e suas propriedades características. (Kim, *Supervenience and* 337)

Uma importante formulação contemporânea desse modelo se encontra no clássico artigo de Paul Oppenheim e Hilary Putnam, intitulado "A unidade da ciência como hipótese de trabalho" (1958). Nesse texto, os autores propõem um modelo para se alcançar a unidade das ciências (uma meta que eles reconhecem como ainda distante) baseada na idéia de *micro-redução*, isto é, a redução de um ramo da ciência  $B_2$  a outro ramo  $B_1$ , quando:

os objetos no universo de discurso de  $B_2$  são totalidades que podem ser decompostas em partes próprias que pertencem ao universo de discurso de  $B_1$ . (Oppenheim Putnam 6)

Tomemos o exemplo de um ramo da ciência  $B_2$  cujo universo de discurso é formado por seres vivos pluricelulares. Esse ramo seria redutível a um ramo  $B_1$ , cujo universo de discurso é formado por células, quando certas condições fossem satisfeitas (Ibid. 5), e já que os objetos estudados por  $B_1$  são partes dos objetos estudados por  $B_2$  essa seria uma *micro-redução*.

Para oferecer um padrão de referencia para as micro-reduções, os autores

<sup>12</sup> Uma apresentação detalhada dos argumentos de Descartes para o dualismo mente-corpo se encontra em Maslin (53-64).

articulam um modelo de ordenação dos ramos da ciência, no qual cada ramo, com seu universo de discurso, será sempre uma potencial base de redução para os ramos superiores (se existirem). O modelo se baseia nas premissas de que existem vários níveis, em numero finito, com um único nível fundamental, sendo que os objetos de cada nível (exceto o fundamental) têm de ser decomponíveis em objetos do nível imediatamente inferior, e nunca podem ter partes pertencentes a um nível superior. Os níveis foram selecionados de um modo justificável do ponto de vista da ciência empírica contemporânea. O modelo é o seguinte (Oppenheim Putnam 9):

Figura 1: o modelo estratificado na versão de Oppenheim & Putnam

6. Grupos sociais
5. Organismos pluricelulares
4. Células
3. Moléculas
2. Átomos
1. Partículas elementares

A idéia é que cada um dos tipos de particulares citados na lista, que povoam os diferentes níveis, possuem propriedades típicas, que não se encontram nos objetos dos outros níveis, sejam inferiores (mais elementares) ou superiores (mais complexos). Esse modelo envolve a idéia que *todos* os particulares existentes no universo<sup>13</sup> são decomponíveis, em ultima instância em partículas físicas, isto é, a idéia de que todos os particulares são *físicos*, o que já evidencia a ruptura com o dualismo cartesiano de substâncias (isto é, de particulares). Mas essa concepção a respeito dos particulares deixa em aberto uma questão decisiva.

Nas palavras de Jaegwon Kim:

Como as *propriedades* características de cada nível estão relacionadas com as propriedades dos níveis adjacentes? Dado que as entidades dos diferentes níveis são ordenadas pela relação parte-todo, é o caso que as propriedades

---

<sup>13</sup> Evidentemente, objetos extremamente grandes, como montanhas, planetas, estrelas ou galáxias, não encontram lugar no modelo de Oppenheim & Putnam, mas isso certamente se deve ao fato de os autores estarem concentrados num modelo de unificação que abranja ciências naturais e humanas. Isso não é um problema, pois particulares como planetas e galáxias certamente possuem partes que podem ser decompostas em outras partes, que possam por sua vez ser decompostas em moléculas, átomos e partículas subatômicas.



associadas com os diferentes níveis também são ordenadas por alguma relação distintiva e significativa? (Kim, *Supervenience and* 338)

Oppenheim e Putnam defendiam uma teoria reducionista a esse respeito, mas o modelo estratificado e essa questão levantada por Kim na citação acima são compatíveis com teorias contrárias ao reducionismo, que (apesar de aceitarem um fisicalismo a respeito dos *particulares*) defendem um dualismo a respeito das *propriedades*<sup>14</sup>. Se, na filosofia contemporânea, poucos são os que vêem a mente como uma coisa não-física, um particular imaterial que detém as propriedades mentais, ainda existem muitos filósofos que (embora aceitem que todos os particulares são físicos) consideram essas propriedades como parte de um domínio separado. Temos, assim, a contraposição entre monismo fisicalista (que defende o caráter físico tanto dos particulares quanto das propriedades) e dualismo de propriedades (que faz a distinção entre propriedades físicas e mentais, mas defende que os dois tipos de propriedades pertencem a particulares físicos), que é um dualismo metafísico um pouco mais afinado com os resultados das ciências da natureza, embora, mesmo assim, seja normalmente considerado incompatível com os conhecimentos trazidos por essas ciências.

Porém, existem aqueles que pensam que certo dualismo de propriedades (mais moderado) é compatível com os resultados das ciências da natureza. Muitos pensadores procuram preservar o caráter (em certo sentido) físico dos fenômenos mentais e, ao mesmo tempo, preservar sua autonomia e peculiaridade próprias, rejeitando sua redução a algo outro. Tratam-se dos defensores do *fisicalismo não-reduutivo*, uma concepção a respeito da mente que exerceu forte influência desde o final da década de 1960 (Kim, *Supervenience and* 311); (Schlosser 73). Os fisicalistas não-reducionistas defendem, simultaneamente, a natureza física de todos os objetos existentes no universo (em porções determinadas do espaço e em períodos determinados do tempo)<sup>15</sup> e a irredutibilidade das propriedades mentais. Mas a irredutibilidade é uma tese negativa (Propriedades mentais *não são* propriedades físicas), e os fisicalistas

<sup>14</sup> Esse ponto será de grande importância para que se possam avaliar umas das réplicas de Searle às críticas de Kim. Confira a seção 3 abaixo.

<sup>15</sup> “A tese ontológica básica do fisicalismo não-reduutivo confere ao físico um certo tipo de primazia: todos os existentes concretos são físicos – não existem particulares não-físicos, não existem almas, substâncias mentais cartesianas, ‘princípios vitais’ ou ‘enteléquias’. Formulada como uma tese sobre propriedades, a primazia do físico nesse sentido resulta no seguinte: todas as propriedades mentais são instanciadas por particulares físicos”. (Kim, *Supervenience and* 339-40).

não-redutivos pretendem fornecer uma concepção positiva sobre a relação entre propriedades físicas e mentais, e o título de “fiscalismo” atribuído a esse ponto de vista provém exatamente da suposta *primazia* das propriedades físicas em relação às mentais. Nas palavras de Kim:

A maioria dos fiscalistas não-redutivos quer ir além da afirmação de que as propriedades mentais são instanciadas por sistemas físicos; eles querem defender uma tese de *primazia*, (...), das propriedades físicas em relação às propriedades mentais. A idéia principal aqui é que, a despeito de sua irreduzibilidade, as propriedades mentais são em algum sentido forte *dependentes de* ou *determinadas por* propriedades físico-biológicas. (*Supervenience and* 340)

Ou seja, além das posições filosóficas que afirmam (reduccionismo) ou negam (dualismo) a identidade entre as propriedades mentais e físicas, há uma posição intermediária que nega essa identidade (admite uma diferença entre elas), mas afirma que as propriedades físicas são ontologicamente mais fundamentais. Podemos, então, representar as três posições filosóficas sobre a relação mente-corpo discutidas acima através da seguinte figura:

Figura 2: três concepções sobre a mente

**Fiscalismo reducionista**

(Propriedades mentais = propriedades físicas)

**Dualismo de propriedades**

(Propriedades mentais ≠ propriedades físicas)

**Fiscalismo não-redutivo**

(Propriedades mentais ≠ propriedades físicas) &

(Propriedades mentais *são dependentes* das propriedades físicas).

As reflexões expostas acima fornecem meios extremamente úteis para que se possa avaliar a ontologia da mente proposta por Searle. Não é difícil perceber que ele formula sua concepção ontológica de um modo extremamente afinado com a versão anti-reducionista do fiscalismo, e diversos interpretes já chamaram a atenção para isso (Cunningham 34); (Chalmers 164); (Maslin 170). Se por um lado ele considera os fenômenos mentais como *ontologicamente irreduzíveis*, por outro ele afirma que tais fenômenos são dependentes de sistemas físicos, no sentido de que eles

são *propriedades de nível superior* desses sistemas, propriedades que são *causadas* pelos processos que ocorrem no nível inferior. Discutindo sua concepção sobre o reducionismo, Searle (*Consciousness* 229) afirma que ele não usa o termo “unidade da ciência”, nem acredita que algo assim faça sentido, pois, apesar de que nós vivemos em *um único mundo* (o que indica certo monismo ontológico), este mundo é abordado pelas diferentes ciências com diferentes interesses e a partir de diferentes perspectivas, coisa que decorre do fato de que,

no único mundo existem várias características [*features*] diferentes: biológicas, elétricas, econômicas, geográficas, mentais, morais, estéticas, etc. Em algum sentido, todas elas são ‘físicas’, porque o mundo consiste inteiramente de entidades físicas. (Ibid. 229)<sup>16</sup>

Formulada nesses termos, a concepção de Searle é, no que diz respeito aos particulares, *fisicalista*, uma vez que ele concebe todas as entidades *concretas* como sistemas físicos. Conforme exposto acima, isso ainda é compatível com o dualismo de *propriedades*. Em minha opinião, da forma como a teoria de Searle está, hoje, articulada, ela se mostra incapaz de sustentar um fisicalismo verdadeiro, pois a maneira como Searle concebe a irredutibilidade ontológica da consciência faz da sua teoria, como afirma Kim (*Mental causation* 192) uma forma de dualismo de propriedades<sup>17</sup>. Porém, agora, devemos, por enquanto, deixar esse problema do dualismo de lado, pois nosso *principal objetivo* aqui é determinar se Searle consegue rebater as críticas de Kim e articular uma concepção viável do modelo ontológico estratificado, de modo a esclarecer a maneira como esse modelo *estrutura* (organiza) o naturalismo biológico. Cabe agora examinar as réplicas de Searle contra Kim.

### III

#### Avaliando as réplicas de Searle contra Kim

Conforme exposto anteriormente, em sua resposta às críticas de Jaegwon Kim, Searle afirma que seu oponente se equivoca ao pensar que o

<sup>16</sup> Discutindo sua pretensa síntese entre dualismo e fisicalismo, ele afirma que este “tenta dizer com razão que o universo é inteiramente feito de partículas físicas que existem em campos de força e são freqüentemente organizadas em sistemas”. (Searle, *Mind* a 126)

<sup>17</sup> De acordo com Prata (*É incoerente* 572), uma vez que Searle entende a subjetividade e a objetividade como diferentes *modos de ser*, e uma vez que ele critica a fisicalismo por ser, supostamente, incapaz de dar conta das *características essenciais* da mente, sua teoria se mostra como um dualismo entre propriedades ontologicamente subjetivas e ontologicamente objetivas.

modelo dos níveis é gerado pela relação de *ser uma parte*, pois, como Kim certamente sabe,

solidez e liquidez são características de sistemas físicos em um nível superior ao das moléculas, mas solidez e liquidez não são partes das moléculas nem são partes do sistema composto de moléculas. (Searle, *Consciousness* 217)

A discussão feita na seção anterior é mais do que suficiente para mostrar que a resposta de Searle é completamente inadequada, pois ele está ignorando a diferença entre *concretos particulares* (os sistemas), por um lado, e suas *propriedades*, por outro lado (cf. a nota de rodapé n. 16, acima).

É evidente que propriedades não podem ser partes dos sistemas que as exemplificam, pois as propriedades, por um lado, e os sistemas, por outro, pertencem a categorias ontológicas diferentes. E Kim está plenamente ciente dessa diferença categorial, o que fica bastante claro quando ele afirma:

dado que as *entidades* [particulares] dos diferentes níveis são ordenadas pela relação parte-todo, é o caso que as *propriedades* associadas com os diferentes níveis também são ordenadas por alguma relação distintiva e significativa? (Kim, *Supervenience and* 338) (grifos meus)

O que ele está afirmando é que os *particulares* são ordenados pela relação parte-todo, e ele questiona, justamente, se há *outra relação* ordenando as propriedades em cada nível de complexidade (como a solidez ou a liquidez, citadas por Searle).

Já no caso da réplica de Searle à crítica de que ele emprega o modelo estratificado, mas, mesmo assim, não percebe que existem alternativas ao modelo bifurcado (de Descartes), penso que ele é mais bem sucedido. De fato, como coloca Searle, o modelo dos níveis de complexidade (nível subatômico, nível atômico, nível molecular, nível celular, etc.) se presta a esclarecer “como a realidade física se organiza”, de modo que ele é logicamente compatível com o dualismo ontológico, seja de substâncias, seja, *principalmente*, de propriedades.

Considerando que o modelo estratificado organiza os *particulares* através da relação parte-todo (partículas subatômicas formam átomos, átomos forma moléculas, moléculas formam células, etc.), fica em aberto a relação

entre as *propriedades* de cada nível, como por exemplo, as propriedades e capacidades mentais manifestadas por certos organismos (como os seres humanos) e as propriedades de certas estruturas biológicas em nível celular. Sendo assim, permanece a possibilidade de que tais propriedades mentais sejam *ontologicamente distintas* das propriedades biológicas, constituindo um dualismo de propriedades perfeitamente afinado com o modelo dos níveis. Nesse ponto, a resposta de Searle me parece adequada.

Figura 3: dualismo entre propriedades físicas (F) e mentais (F)

(Organismo pluricelular)	$F_1, F_2, (...), F_N [F_1, F_2, (...), F_n].$	(propriedades)
	$X''$	(particular)
(Células)	$f'_1, f'_2, f'_3, (...), f'_n.$	(propriedades)
	$x'$	(particular)
(Moléculas)	$f_1, f_2, f_3, (...), f_n.$	(propriedades)
	$x$	(particular)

Entretanto, como foi mencionado acima (cf. a nota de rodapé n. 11), Searle não responde de modo detalhado a uma importante crítica colocada por Kim: a acusação de que, ao rejeitar a redução da consciência a propriedades objetivas do cérebro, o naturalismo biológico se identifica com o dualismo de propriedades. Na tentativa de rechaçar essa interpretação Searle se limita<sup>18</sup> a dizer duas coisas:

Primeiro, o fracasso da redução da consciência não é tanto um problema sobre a consciência quanto é um problema sobre as ambigüidades na noção de redução. A consciência é "irreduzível" por razões triviais que tem a ver com nossas práticas definicionais. Segundo, um ponto relacionado, precisamos distinguir entre aquelas formas de redução que são eliminativas, que mostram que não havia nada em primeiro lugar, e aquelas que não são eliminativas, que "reduzem" uma propriedade "emergente", como a solidez, às suas bases causais por redefinição. (Searle, *Consciousness* 220)

<sup>18</sup> "Eu dediquei um capítulo inteiro de *A redescoberta da mente* a esse assunto, e não vou repetir os argumentos, exceto para dizer duas coisas". (Searle, *Consciousness* 220)

As reduções eliminativas seriam impossíveis para coisas que *realmente* existem, como a consciência. Já a redução de propriedades “emergentes”, como é (de acordo com Searle) a consciência, seria possível, mas no caso particular da consciência, seria inadequada, pois nos faria perder aquilo que nos faz precisar da noção de consciência em primeiro lugar: a *subjetividade* da nossa experiência consciente.

Sobre a ambigüidade do termo “redução”; penso que Searle tem certa razão, na medida em que esse termo pode ser usado para designar procedimentos muito diferentes (Searle, *The Rediscovery* 113-4) ([164-6]). Todavia, permaneço bastante cético diante da idéia de que o problema está, antes, *nas nossas práticas* e nos nossos interesses ao efetuarmos uma redução do que *no próprio fenômeno* da consciência. No livro *A redescoberta da mente*, Searle afirma que, de um modo geral, as reduções visam eliminar experiências subjetivas de modo a excluir essas experiências da definição de fenômenos reais (se, no senso comum, definimos “calor” em termos de certa sensação subjetiva, o conhecimento físico e a *redução* da termodinâmica à mecânica estatística nos permitem definir “calor” em termos de energia cinética média, uma propriedade objetiva) e esses fenômenos reais passam a ser definidos em termos das características que mais nos interessam (nesse caso, as características objetivas estudadas pela física). Mas, “nos casos em que os fenômenos que mais nos interessam são as próprias experiências subjetivas, não há como eliminar nada” (Searle, *The Rediscovery* 121) ([176]).

Com essa teoria, Searle sugere que a redução ontológica da consciência *seria possível*<sup>19</sup>, mas que ela não nos interessa. É isso que ele quer dizer quando afirma que a irredutibilidade da consciência tem a ver com nossas práticas *definicionais*, e não com a essência do fenômeno. Entretanto, essa idéia não me parece coerente com as afirmações que Searle faz a respeito da subjetividade, que seria uma *característica essencial* da consciência, que a faz diferente de tudo o mais no universo<sup>20</sup>. A insistência dele na

<sup>19</sup> “Por que a consciência não pode ser redefinida em termos dos processos neurofisiológicos da mesma maneira que redefinimos o calor em termos de processos físicos subjacentes? Bem, logicamente, poderíamos fazer a redefinição caso insistíssemos nisso. Poderíamos simplesmente definir, por exemplo, ‘dor’ como padrões de atividade neurônica que causam sensações subjetivas de dor. e, se tal redefinição ocorresse, teríamos chegado à mesma espécie de redução para a dor à qual chegamos para o calor”. (Searle, *The Rediscovery* 121) ([175])

<sup>20</sup> “Estados e processos conscientes possuem uma característica especial não possuída por outros fenômenos naturais, a saber, a subjetividade” (Ibid. 93) ([138]). Em um texto mais recente ele, novamente, afirma uma diferença tão forte que sugere um dualismo entre a consciência e a realidade natural: “A consciência tem três aspectos que a fazem diferente de outros fenômenos biológicos e, na verdade, *diferente de outros fenômenos no mundo natural*” (Searle, *Consciousness and* 39) ([56]). (grifo meu).



tese de que a subjetividade constitui um *modo de existência* (ontologia) radicalmente diferente da objetividade, parece tornar inevitável a conclusão de que sua recusa da redução da consciência resulta em um dualismo, como pensa Kim.

#### IV Considerações finais

A discussão apresentada acima permite concluir que, apesar de pecar por certa falta de clareza a respeito das *propriedades* típicas de cada nível de complexidade, Searle articula a sua visão a respeito da ontologia da mente segundo o modelo estratificado, de um modo que é relativamente bem sucedido. Em sua resposta à crítica de Kim segundo a qual ele, apesar de usar esse modelo, não percebe que existem alternativas ao modelo bifurcado (cartesiano), Searle demonstra ter uma visão lúcida a respeito do modelo dos níveis e de sua compatibilidade com o dualismo ontológico. Sua resposta é que o modelo estratificado se apresenta como uma solução para o problema de como a realidade física se organiza, o que deixa aberta a possibilidade da existência de uma realidade *não-física*.

Todavia, o naturalismo biológico de Searle padece da dificuldade de recusar a redução da consciência de uma maneira tão forte que resulta em um dualismo de propriedades. Isso é problemático porque, apesar de o modelo estratificado ser *compatível* com o dualismo (inclusive com o dualismo entre propriedades subjetivas e objetivas), o grande sentido de se empregar esse modelo ontológico é se aproximar de alguma forma de *fisicalismo* (como fica claro no fato de o modelo estratificado ter sido usado por *reducionistas*, como Oppenheim e Putnam, e por *fisicalistas não-redutivos*, como os emergentistas britânicos).

A maneira como Searle tenta defender a tese de que o seu anti-reducionismo não é extremo, isto é, não conduz a um dualismo de propriedades (não tem “consequências metafísicas profundas”), não se sustenta diante da diferença ontológica radical que ele atribui à *subjetividade* da nossa experiência consciente. Fenômenos subjetivos, como sensações, percepções e desejos, na medida em que são dotados de um *aspecto qualitativo*, possuem características peculiares que jamais poderiam ser descritas através de um vocabulário objetivo (referente a entidades como fendas sinápticas, neurotransmissores ou potenciais de ação), o que evidencia a *diferença* entre essas propriedades subjetivas e quaisquer propriedades objetivas (Searle, *The Rediscovery* 117) ([170]).



A adesão de Searle a esse dualismo entre subjetividade e objetividade afasta o seu naturalismo biológico do fisicalismo tradicionalmente associado ao modelo ontológico dos níveis de complexidade.

#### REFERÊNCIAS BIBLIOGRÁFICAS

Beckermann, A. *Descartes' metaphysischer Beweis für den Dualismus*. Freiburg; Munique: Alber, 1986. Drucken.

Blackburn, S. "Metafísica". Bunnin, N. and E. P. Tsui-James (ed.s). *Compêndio de filosofia*. 1ª edição. São Paulo: Loyola, 2002. Impresso.

Chalmers, D. *The Conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press, 1996. Print.

Cunningham, S. *What is a mind? An integrative introduction to the philosophy of mind*. Indianápolis: Hackett Publishing, 2000. Print.

Descartes, R. *Princípios da filosofia*. Lisboa: edições 70, 1997. Impresso.

Garrett, B. *Metafísica: conceitos-chave em filosofia*. Porto Alegre: Artmed, 2008. Impresso.

Grewendorf, G. and G. Meggle (eds.). *Speech Acts, mind and social reality: Discussions with John R. Searle*. Dordrecht, Boston, London: Kluwer Academic Publishers, 2002. Print.

Heil, J. *Philosophy of mind: A contemporary introduction*. Londres: Routledge, 1998. Print.

Imaguire, G. "A substância e suas alternativas: feixes e tropos". Imaguire, G., Almeida, C. L. S. and M. A. Oliveira (org.). *Metafísica contemporânea*. Petrópolis: Vozes, 2007. Impresso.

Kim, J. *Supervenience and mind: Selected philosophical essays*. Cambridge: Cambridge University Press, 1993. Print.

---. "Mental causation in Searle's 'Biological Naturalism'". *Philosophy and Phenomenological Research*. Mar. 1995: 189-194. Print.

---. *Philosophy of mind*. Oxford, Boulder: Westview Press, 1996. Print.

Loux, M. *Metaphysics: A contemporary introduction*. 3ª ed. London-New York: Routledge, 2006. Print.

Maslin, K. T. *Filosofia da mente*. 2ª edição, Porto Alegre: Artmed, 2009. Impresso.

Oppenheim, P. and H. Putnam. "The unity of science as working hypothesis". Feigl, H., Scriven, M. and G. Maxwell (eds.). *Minnesota*

*Studies in the Philosophy of Science*, Vol. 2. Minneapolis: University of Minnesota Press, 1958. Print.

Pereboom, D. and H. Kornblith. "The metaphysics of irreducibility". Heil, John. *Philosophy of mind: A guide and anthology*. Oxford: Oxford University Press, 2004. Print.

Prata, T. A. "Dificuldades da concepção de John Searle sobre a redução da consciência: o problema das capacidades causais". *Princípios*. Jul. 2008: 5-29. Impresso.

---. "Irreducibilidade ontológica versus identidade: John Searle entre o dualismo e o materialismo". *O que nos faz Pensar*. Ago. 2009: 107-124. Impresso.

---. "É incoerente a concepção de Searle sobre a consciência?". *Manuscrito*. Jul. 2011: 557-578. Impresso.

Schlosser, M. E. "Nonreductive physicalism, mental causation, and the nature of actions". Hieke, A. and H. Leitgeb (eds.). *Reduction: Between the mind and the brain*. Frankfurt-Paris-Lancaster-New Brunswick: Ontos Verlag, 2009. Print.

Searle, J. R. *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press, 1983. Print.

---. *The Rediscovery of the mind*. Cambridge Mass.-London: MIT Press, 1992. Print.

---. "Searle, John". Guttenplan, S. (ed.). *A companion to the philosophy of mind*. Oxford-Cambridge MA: Basil Blackwell, 1994. Print.

---. "Consciousness, the brain and the connection principle: A reply". *Philosophy and Phenomenological Research*. Mar. 1995: 217-232. Print.

---. *Mind, language, and society: Doing philosophy in the real world*. London: Weinfeld & Nicolson, 1999. Print.

---. *Consciousness and language*. Cambridge: Cambridge University Press, 2002. Print.

---. "Why I am not a property dualist". *Journal of Consciousness Studies*. Dec. 2002: 57-64. Print.

---. *Mind: A brief introduction*. Oxford: Oxford University Press, 2004. Print.

---. "Dualism revisited". *Journal of Physiology*. Jul. 2007: 169-178. Print.

Smith, B. (ed.). *John Searle*. Cambridge: Cambridge University Press, 2003. Print.



# LA REALIDAD BÁSICA Y LA REALIDAD HUMANA: CAPÍTULO INTRODUCTORIO PARA EL VOLUMEN DE MÜNSTER

*THE BASIC REALITY AND THE HUMAN REALITY:  
INTRODUCTORY CHAPTER TO THE MÜNSTER'S VOLUME*

JOHN R. SEARLE  
Universidad de Berkeley, USA. j.searle@berkeley.edu

RECIBIDO EL 28 DE NOVIEMBRE DE 2012 Y APROBADO EL 14 DE DICIEMBRE DE 2012

En la conferencia que dicté en la ciudad de Münster, al comienzo del congreso sobre el cual está basado este volumen, discutí mi trabajo en ontología social, y específicamente, la naturaleza de la realidad social institucional, la realidad del dinero, la propiedad, el gobierno, el matrimonio, entre otros. Después de haber dictado dicha conferencia publiqué el libro *Haciendo el mundo social*, en el cual se discuten estos temas mucho más extensamente. En lugar de repetir el material que ya he publicado, se me ocurrió, después de escuchar y leer las diferentes contribuciones a este volumen, que podría ser una buena idea formular mi posición filosófica general, el enfoque global en el cual está basada mi obra y

varios aspectos a los que están dedicados los artículos de este volumen. En vez de repetir lo que ya dije en Münster, voy a discutir los fundamentos de mi empresa filosófica en general.

Me parece que en la actualidad hay un problema prioritario en filosofía y que, después de haber estado haciendo filosofía literalmente por décadas, éste era el problema que yo estaba tratando. Como una formulación preliminar, podríamos expresarlo diciendo que la cuestión es reconciliar una cierta concepción que tenemos sobre nosotros mismos y nuestra realidad humana con lo que de hecho sabemos acerca de cómo es el mundo a partir de los avances

---

\* Agradecimientos al profesor John Searle por la autorización para la traducción al español de este artículo. De la misma manera, agradezco especialmente al profesor Raúl Meléndez por sus correcciones al mismo.

\*\* Traducción al español realizada por Laura Giraldo. Docente Departamento de Filosofía de la Corporación Universitaria Minuto de Dios. Auxiliar de investigación, grupo de investigación "De Interpretatione: filosofía y ciencia la interpretación", de la pontificia Universidad Javeriana de Bogotá. Email: lauragiraldoceballos@hotmail.com.

en el conocimiento que han ocurrido en los últimos tres siglos. Para decirlo en términos más detallados, el problema es el siguiente: sabemos por disciplinas como la física y la química que el mundo está compuesto enteramente por entidades que consideramos conveniente, sino completamente adecuado, llamar "partículas físicas". No pretendo respaldar una fase particular de la historia de la física y la química, la cual, sin duda, cambiará y mejorará. Sin embargo, la idea básica de que las cosas grandes están hechas de cosas pequeñas está tan bien establecida que la daré por sentada. Sabemos que esas "partículas" existen en campos de fuerza y que están organizadas en sistemas. Los límites de esos sistemas se establecen por relaciones causales. Moléculas, galaxias y seres vivos son ejemplos de esos sistemas. Nuestra imagen de la realidad la cual sabemos que es exacta, tanto como sabemos cualquier cosa es que, en últimas, el mundo real consta de entidades físicas carentes de mente y carentes de significado. Pero tenemos una concepción de nosotros mismos que es, a primera vista, difícil de conciliar con eso. Pensamos de nosotros mismos que somos ante todo conscientes. Y pensamos de nosotros mismos que junto con la consciencia tenemos intencionalidad, y junto con consciencia e intencionalidad tenemos un montón de otras características notables. Entre ellas, el lenguaje, la racionalidad, el libre albedrío, la sociedad, la ética, la estética, las obligaciones políticas y las instituciones sociales. Nuestra cuestión en su forma más general puede formularse así: ¿cómo podemos conciliar la concepción de la realidad, que sabemos, que de hecho es correcta (la realidad básica como es descrita por la física, la química y las otras ciencias naturales) con la concepción de la realidad con la que vivimos cada día, con la realidad humana? Sólo para tener una forma abreviada de plantear la cuestión, diré que se trata de reconciliar la *realidad básica* con la *realidad humana*. Y nuestro problema no es sólo hacer que los hechos humanos sean *consistentes* con los hechos básicos, sino mostrar cómo ellos son un *desarrollo natural* a partir de los hechos básicos.

Dados los hechos básicos de la física y la química, debemos mostrar cómo es posible e incluso inevitable que haya consciencia, sociedad, racionalidad, lenguaje, moral, obligaciones y placeres estéticos. Muchas cuestiones discutidas dentro de esta problemática son asuntos tradicionales de la filosofía tales como el problema mente-cuerpo; otros, como el de la naturaleza de la ontología social, son mucho menos discutidos en la filosofía tradicional. El enfoque completo pone a las cuestiones filosóficas ya sean tradicionales o contemporáneas bajo una luz un poco diferente, pues ya no estamos obsesionados como lo estaban

nuestros predecesores filosóficos con la búsqueda de fundamentos epistémicos, con la superación del escepticismo, con la prueba de la posibilidad misma del conocimiento.

Podemos ahora dar por sentado el conocimiento de la estructura básica del universo y preguntar cómo acomodar, si es que podemos, nuestra concepción de nosotros mismos dentro de ese marco. Ésta es la tarea actual de la filosofía tal y como yo la veo. Quizá, finalmente, tendremos que renunciar a algunos de nuestros más apreciados supuestos como por ejemplo, el libre albedrío. Tal vez no podamos hacer que nuestra concepción general de la realidad sea consistente con nuestra creencia en que tenemos una genuina libertad de la voluntad. Este proyecto filosófico como lo he descrito tomaría más de una vida. Pero puede valer la pena en esta ocasión resumir en dónde nos encontramos en el desarrollo de dicho proyecto. Algunas partes, pienso yo, son bastante fáciles por lo menos en lo que tiene que ver con la filosofía, aunque puedan tener asociados problemas científicos difíciles. Algunas partes son bien difíciles. Tal vez al final se mostrarán como imposibles de resolver. En cualquier caso, este es el proyecto y, ahora voy a esbozar donde pienso que estoy ubicado al trabajar en él.

Aquí estamos en nuestro gran universo, compuesto principalmente de espacio vacío, pero que contiene astronómicamente un gran número de moléculas, ellas mismas compuestas de átomos y ellos mismos compuestos de partículas subatómicas, y muchas de esas moléculas están organizadas dentro de sistemas más grandes. Entonces en nuestro pequeño planeta y con toda probabilidad, en planetas situados en otros sistemas solares, ocurrió algo asombroso aunque no necesariamente al mismo tiempo en cada planeta. Comenzó la vida. Aquí viene el primer vacío en nuestro entendimiento. No comprendemos el origen de la vida. No sabemos cómo o dónde o cuándo comenzó la vida en nuestro planeta. Pero vamos a dar por sentado que ella comenzó y continuó. Una vez más limitándonos a nuestro pequeño planeta, durante un periodo de tres a cinco billones de años, la vida se desarrolla en todas sus formas actuales. Así que no tenemos sólo partículas físicas, organizadas en sistemas, sino que algunos de esos sistemas están vivos. Ellos están compuestos en buena medida de grandes moléculas basadas en carbono, combinadas con grandes cantidades de nitrógeno, hidrógeno y oxígeno. Supondremos, como sabemos que es efectivamente el caso, que estas formas de vida siguieron evolucionando hacia las actuales especies vegetales y animales. También ocurre otro desarrollo que suscita la

primera cuestión filosófica seria en nuestro inventario de problemas filosóficos, el comienzo de la *realidad mental*.

## I

### El problema mente-cuerpo: consciencia e intencionalidad

Con la evolución de las formas de vida animal llega la evolución de órganos dentro de esas formas, los cuales contienen grandes cantidades de un tipo inusual de célula llamada neurona. Una neurona es una célula como cualquier otra. Ésta tiene un núcleo celular, un cuerpo celular y una pared celular. Sin embargo, la neurona es inusual puesto que tiene estructuras anatómicas peculiares en forma de axones y dendritas y un sistema peculiar de relaciones. Por medio de él las neuronas se comunican con otras enviando señales electro-químicas a través de un pequeño agujero en un punto donde la neurona entra en contacto con otra neurona, la hendidura sináptica. Específicamente, un impulso eléctrico pasa por el axón donde desencadena la secreción de neurotransmisores en la hendidura sináptica y éstos inician entonces un impulso eléctrico en el lado post-sináptico o dendrítico. Así que la secuencia es eléctrica-química-eléctrica. Estoy dejando de lado un montón de detalles que son desesperadamente importantes para nuestras vidas porque quiero llegar a los temas principales de este capítulo. Con el desarrollo de sistemas neuronales vienen dos procesos lógicamente relacionados los cuales son absolutamente impresionantes y son la base para el resto de historias filosóficas de las que estaremos hablando. Con la evolución del sistema neuronal viene la evolución de la *consciencia* y la *intencionalidad*. Éstas requieren una discusión especial. A primera vista, consciencia e intencionalidad parecen ser independientes la una de la otra. La consciencia consiste en aquellos estados de sentimientos o de sensibilidad, o de darse cuenta, que los organismos tienen mientras están despiertos. Sin embargo, cuando ellos están dormidos algunas veces tenemos consciencia en la forma de sueños. Intencionalidad, es esa capacidad que tienen los estados mentales de ser acerca de o de referirse a objetos o estados de cosas en el mundo. Es habitual usar la palabra "representación" en conexión con la palabra intencionalidad. La intencionalidad consiste en representaciones mentales de objetos y estados de cosas en el mundo.

Hasta el momento todo va bien, pero esto puede ser engañoso porque la manera en que las creencias y deseos representan es totalmente diferente



de la manera en que las percepciones y las acciones intencionales representan. Consciencia e intencionalidad marcan una ruptura con las formas de la realidad que hemos estado describiendo anteriormente, porque tienen esencialmente lo que podríamos llamar una ontología subjetiva o de primera persona. La consciencia sólo existe como experimentada por humanos o por animales. La intencionalidad sólo existe en la medida en que es experimentada conscientemente o es por lo menos el tipo de cosa que, aunque con frecuencia inconsciente, podría ser experimentada conscientemente. Hay muchas cosas que están pasando en mi cerebro cuando tengo estados intencionales, como por ejemplo, descargas neuronales en la sinapsis. Pero éstas no tienen una realidad mental porque no son el tipo de cosas que son estados conscientes o que podrían ser estados conscientes. En este punto, llegamos al primero de los difíciles problemas filosóficos tradicionales que tenemos que abordar y resolver, el llamado problema “mente-cuerpo”. Sin embargo, a partir de la explicación que hemos venido dando éste parece tener una solución más bien simple. Todos nuestros estados conscientes son causados por procesos neuronales. No sabemos los detalles de las estructuras neuronales que causan la consciencia, pero sabemos que las neuronas y los sistemas neuronales están esencialmente involucrados. Y los estados conscientes causados por el comportamiento de las neuronas están ellos mismos realizados en, es decir, existen en, los sistemas neuronales, como características de ellos. Sabemos que los estados conscientes funcionan causalmente. Por ejemplo, tengo la intención-en-acción de (intento) levantar mi brazo y he aquí que el brazo se levanta. Y al igual que cualquier rasgo de nivel superior como la solidez de las piedras o la liquidez del agua, los estados conscientes sólo pueden funcionar causalmente porque se basan en un sistema compuesto de elementos de nivel inferior. Así, para decirlo explícitamente, nuestra solución al problema mente-cuerpo, en la medida en que implica la consciencia, involucra cuatro puntos:

- (i) La consciencia es real. Puesto que tiene una ontología de primera persona no puede reducirse a algo que sólo tiene una ontología de tercera persona tal como el comportamiento o la neurofisiología.
- (ii) Es causada por el comportamiento de elementos neuronales.
- (iii) Se vuelve realidad en sistemas neuronales.
- (iv) Funciona causalmente.

Tenemos un problema ligeramente más difícil en la relación entre consciencia e intencionalidad, pero éste también tiene una solución

más bien fácil. Muchos estados de la consciencia no son intencionales, por ejemplo, los sentimientos de ansiedad no dirigida. Y la mayoría de estados intencionales son inconscientes la mayor parte del tiempo, por ejemplo, mis creencias y deseos no cesan mientras duermo; ellas permanecen en un estado inconsciente. De todos modos, existe una conexión conceptual entre consciencia e intencionalidad. Un estado que es un estado intencional, pero a veces no consciente, es por lo menos el tipo de cosa que podría, en principio, volverse consciente. ¿Por qué? Vimos que la intencionalidad consiste en representaciones, pero las representaciones se dan siempre bajo algún aspecto u otro. Por ejemplo, podría desear agua bajo el aspecto "agua" pero no bajo el aspecto "H<sub>2</sub>O". Mi deseo de agua es diferente a mi deseo de H<sub>2</sub>O, aunque ambos serán satisfechos por la misma sustancia en el mundo. Pero el representarse algo explícitamente bajo un aspecto es algo que sólo puede ser hecho por un ser consciente. En lo concerniente a la ausencia de consciencia, no hay diferencia entre guardar una relación con el agua y una relación con H<sub>2</sub>O, porque el agua es H<sub>2</sub>O. Cuando se trata de intencionalidad, hay una diferencia crucial porque la forma aspectual del estado intencional tiene que tener una realidad y la única realidad cuando el sistema es inconsciente es la de una serie de fenómenos ontológicos de tercera persona tales como estructuras neuronales y descargas neuronales. Lo que vuelve real a la forma aspectual cuando es inconsciente es su manifestación potencial como un estado consciente.

## II

### La estructura de la intencionalidad

Hasta el momento hemos hecho un progreso real. En nuestro universo de hechos básicos hemos dado cuenta de seres conscientes que tienen intencionalidad tanto consciente, como inconsciente. Pero aún así, no hemos llegado muy lejos en nuestro intento de solucionar nuestro inventario de problemas filosóficos.

Hemos respondido la pregunta, ¿cómo es posible la consciencia? La respuesta es: ésta se hace posible por la actividad de ciertos sistemas neuronales. El comportamiento neuronal causa la consciencia. La forma básica de explicar cómo es posible la intencionalidad, es explicar cómo son posibles las formas conscientes de la intencionalidad y esto hace que nuestra explicación neurobiológica de la intencionalidad dependa de nuestra explicación de la consciencia. Sin embargo, hay muchas formas

específicas diferentes que requerirán explicaciones separadas. Así, por ejemplo, la percepción es diferente de la acción intencional, aunque están, por supuesto, entrelazadas de muchas maneras en nuestra vida. Ambas, tanto la acción intencional, como la percepción requerirán, presumiblemente, apelar a tipos de mecanismos explicativos diferentes a los que se usan en el caso, por ejemplo, de procesos de pensamientos y emociones. Todas estas son cuestiones empíricas que dejo a la neurobiología, la neuropsicología y a la ciencia cognitiva en general. El punto filosófico es que la relación general entre la realidad humana y la realidad básica es clara. Los detalles tienen que ser elaborados por las ciencias especiales. En general, los estados intencionales vienen con una estructura que consiste en el tipo de estado que son, junto con el contenido que tienen. Así pues, yo puedo creer que está lloviendo y desear que llueva. Las estructuras son:

Creencia: (Está lloviendo).

Deseo: (Está lloviendo).

Más aún, correspondiendo a la distinción en el sistema nervioso entre el sistema nervioso sensorial y el sistema nervioso motor, los estados intencionales tienen diferentes relaciones con la realidad. Se supone que las percepciones y creencias representan cómo son las cosas en el mundo. Se supone que los deseos e intenciones-en-acción representan no cómo son las cosas, sino cómo nos gustaría que fueran las cosas o cómo vamos a tratar que sean. Correspondiendo a estas distinciones, introduzco la terminología dirección de ajuste “mente-a-mundo” característica de las creencias y percepciones, y la dirección de ajuste “mundo-a-mente” característica de los deseos e intenciones. Estoy dejando de lado los detalles, pero estas dos formas que la mente tiene de relacionarse con la realidad serán cruciales para la discusión que viene y más importante aún, para nuestras vidas en esta tierra. Sólo para tener una ayuda mnemotécnica (mnemonic). Represento la dirección mente-a-mundo con una flecha hacia abajo ↓ y con una flecha hacia arriba la dirección de ajuste mundo-a-mente ↑. He estado apelando implícitamente a la idea de que los fenómenos intencionales de la mente representan cómo son las cosas o cómo nos gustaría que fueran, ya sea con éxito o sin él, así que necesitamos introducir una noción que concuerde con el éxito o con el fracaso. Para ello, introduzco la noción “condiciones de satisfacción”. Lo que hace que una creencia sea verdadera es lo que hace que los deseos sean satisfechos, es lo que hace que las intenciones sean llevadas a cabo, es lo que hace que las percepciones sean verídicas. En todos estos

casos, diremos que los estados intencionales establecen las *condiciones de satisfacción* y el estado tiene éxito o fracasa dependiendo de si las condiciones son satisfechas o no. Las creencias tienen éxito o fracaso dependiendo de si ellas son verdaderas o falsas. Los deseos tienen éxito o fracaso dependiendo de si son o no satisfechos y así sucesivamente para muchos otros casos.

Dado que el organismo está en contacto causal constante con el medio ambiente, nos damos cuenta que muchos estados intencionales tienen un componente causal integrado en sus condiciones de satisfacción. Si veo un objeto en frente de mí, yo solo veré ese objeto, como suele decirse, “verídicamente” si la presencia y los rasgos del objeto causan la experiencia consciente de verlo. De manera similar, si trato de levantar mi brazo, entonces mi intención-en-acción de levantar mi brazo será satisfecha sólo si la intención misma causa el resto de sus condiciones de satisfacción, es decir, ella causa que el brazo se levante. Describo ambos casos diciendo que los fenómenos intencionales en cuestión son “causalmente auto-reflexivos”. El estado intencional mismo requiere una relación causal entre el estado y las condiciones de satisfacción en el mundo. La percepción será satisfecha sólo si el objeto percibido causa la percepción del mismo. La intención-en-acción será llevada a cabo sólo si la intención-en-acción causa los movimientos corporales que el agente está tratando de lograr. La intención-en-acción es sólo el intentar.

Una última distinción; algunos de nuestros fenómenos intencionales están tan directamente relacionados con el medio ambiente que resulta engañoso caracterizarlos como representaciones. Ellos son “presentaciones” directas de sus condiciones de satisfacción. Así que, cuando veo una silla en frente de mí, mi experiencia visual consciente, no sólo representa, sino que de hecho *presenta* la silla misma. Todos los fenómenos que he descrito pueden ser resumidos en una gráfica, la adjunté con la esperanza de que el lector haga un estudio cuidadoso de ella. Si usted comprende los detalles de esta gráfica comprenderá la intencionalidad mejor que un gran número de filósofos profesionales.

	Cognición			Volición		
	Percepción	Memoria	Creencia	Intención-en-acción	Intención previa	Deseo
Dirección de ajuste	↓	↓	↓	↑	↑	↑

Dirección de causalidad	↑	↑	N/A*	↓	↓	N/A
¿Causalmente auto-reflexivo?	Sí	Sí	Sí	Sí	Sí	No

\*N/A: No Aplica.

Entonces ya arrancamos y estamos en nuestro camino. Tenemos consciencia, agentes intencionales con estados intencionales y la clave para entender la intencionalidad son las condiciones de satisfacción. Los estados intencionales son representaciones (o presentaciones) de sus condiciones de satisfacción y algunas veces hay un componente causal, construido dentro de la relación intencional.

Tenemos todavía que añadir algo no trivial a esta explicación de la intencionalidad. Tenemos que señalar que los estados intencionales, aunque siempre están en la mente de los individuos pueden, algunas veces, ser *colectivos*. Cuando estoy haciendo algo juntos —estamos limpiando la casa juntos— digamos, y no es sólo que yo estoy limpiando la casa y usted esta limpiando la casa, sino que yo estoy haciendo lo que estoy haciendo como parte de nuestro hacer lo que estamos haciendo; la intencionalidad colectiva va a ser crucial tanto para el lenguaje, como para la creación de la sociedad humana y, por tanto, de la civilización humana.

Obsérvese que en nuestra solución del problema mente-cuerpo y en nuestra discusión de la naturaleza de la intencionalidad, ni siquiera surgió ninguno de los grandes debates filosóficos acerca de cómo la mente es posible en un mundo físico. No le damos ningún uso a las categorías tradicionales de lo “mental” y lo “físico” sólo estábamos hablando acerca de cómo funciona el mundo.

### III

#### **Relatividad al observador e independencia del observador. Objetividad y subjetividad**

Hay algunas distinciones que haremos y que son esenciales para comprender la vida y la civilización humana. Una vez un conjunto de animales tienen intencionalidad, y en particular, intencionalidad colectiva, entonces ellos son capaces de crear porciones de realidad que sólo existen relativas a su intencionalidad. A ellas las llamo “dependientes del observador”, pero podría igualmente ser llamado “dependientes

de la intencionalidad” de las características del mundo. Así el dinero, las propiedades y el matrimonio son todas características del mundo real, pero todas son rasgos reales del mundo. Pero ellos son creaciones humanas. Los seres humanos, junto con otros animales, tienen también la capacidad de asignar funciones a objetos naturales. Pero la asignación de las funciones existentes sólo es relativa a la intencionalidad del agente. Por supuesto, la mayor parte de la realidad existe independientemente de lo que cualquiera piense. Montañas, moléculas y placas tectónicas son todas independientes del observador. Pero algunas otras partes, de gran interés para nosotros, tienen una existencia que es, como diré, relativa-al-observador o relativa a la intencionalidad y éstas son cosas como los Estados nacionales, los sistemas éticos y la propiedad privada.

Relacionada a esta distinción está la distinción entre lo objetivo y lo subjetivo. Esta última distinción es importante para nuestra cultura intelectual, pero pocas distinciones son tan completamente confusas. Básicamente, la confusión se deriva del hecho de que hay tanto un sentido epistémico, como uno ontológico de esta distinción. Epistémicamente, la distinción entre objetividad y subjetividad es una distinción entre tipos de afirmación. Por ejemplo, la afirmación según la cual Botticelli vivió en Florencia es, como se dice, objetiva. Se trata de una cuestión de hecho objetivo. El hecho de que Botticelli fue mejor pintor que su profesor Lippi es, como se dice, una cuestión de opinión subjetiva. Pero el sentido ontológico de la distinción objetivo-subjetivo es una distinción en modos de existencia. Así, montañas y moléculas tienen un modo objetivo de existencia en el sentido de que su existencia no depende de ser experimentados por algún agente. No obstante, los dolores, las cosquillas y las picaduras tienen una ontología subjetiva, un modo subjetivo de existencia porque su misma existencia depende de ser experimentado por un sujeto consciente. Estas distinciones son cruciales para nuestra explicación de la civilización humana y la racionalidad por la siguiente razón, entre otras: todo lo que tiene una ontología relativa al observador tiene un elemento de subjetividad ontológica en su modo de existencia. Pero y este es el punto crucial para nuestra investigación actual, la subjetividad ontológica de un dominio, como son nuestros dolores, cosquillas y picaduras no implica necesariamente subjetividad epistémica en nuestras afirmaciones sobre ese dominio. Podemos hacer afirmaciones epistémicamente objetivas sobre, por ejemplo, los dolores o la economía, a pesar de que los propios dominios tienen un modo de existencia ontológicamente subjetiva. En una palabra, la subjetividad ontológica del dominio no excluye la objetividad epistémica de las declaraciones sobre el dominio.

## IV Lenguaje

Nuestro siguiente tema es el lenguaje. ¿Cómo es posible que animales como nosotros corriendo por la superficie de la tierra teniendo intencionalidad tanto consciente, como inconsciente y relacionándose unos con otros con intencionalidad colectiva, tengan un lenguaje? ¿Qué tenemos que hacer para tener un lenguaje? Mi objetivo al hacer esta pregunta no es especular acerca de los orígenes reales del lenguaje humano, sino de responder a la cuestión filosófica: ¿cuáles son los componentes del lenguaje, los recursos además de la intencionalidad pre-lingüística? Hay tres componentes del lenguaje que explicaremos a continuación. El primero, es el *significado* y la *comunicación del significado*. ¿Qué quiere decir que un hablante diga algo y signifique algo con lo que dice, y cómo comunica ese significado a un oyente?

El segundo componente es la *convención*. Los elementos actuales del lenguaje humano natural tales como las palabras, son convencionales. ¿Qué significa esto y qué tienen los hablantes que imaginamos cuando tienen convenciones? El tercer componente, que es un desarrollo impresionante, es la *sintaxis* interna de las oraciones. Los seres humanos desarrollan más que dispositivos simples y no divididos para comunicarse entre sí. Tienen estructuras que se dividen en frases nominales y frases verbales. Descritas funcionalmente, éstas pueden ser sujetos y predicados. Este también es un desarrollo importante, como lo veremos más adelante.

*Significado del hablante*: la manera más simple de explicar la naturaleza del *significado del hablante* es describir casos en los que las personas tienen de hecho un lenguaje. Luego veremos que el desarrollo de un lenguaje, depende, justamente, de la posibilidad del *significado del hablante*, pero ahora sólo estamos tratando de explicar la noción de *significado del hablante*. ¿Cuál es la diferencia entre decir algo significándolo y decir algo sin significarlo? Wittgenstein frecuentemente planteaba preguntas como estas tratando de hacernos ver que el “significado” no es el nombre de un *proceso* mental, como tener una imagen residual o sentir un dolor. De todos modos hay una diferencia entre decir algo significándolo y decir algo sin significarlo. Debemos estar en condiciones de decir exactamente cuál es esta diferencia. Imaginemos que estoy practicando el alemán y que digo una y otra vez “Es regnet”, “Es regnet” (está lloviendo); aquí estoy diciendo algo sin significarlo. Ahora, imaginemos



que efectivamente quiero decirle a un hablante alemán cómo está el clima y digo “Es regnet”. En ambos casos, dije las mismas palabras, pero la intencionalidad fue diferente. En el primer caso mi intención fue sólo pronunciar correctamente una oración en alemán, pronunciar correctamente una secuencia de palabras. Pero en el segundo caso mi intención es que la pronunciación de la secuencia, que era ella misma una condición de satisfacción de la intención de pronunciar esa secuencia, debería ella misma tener condiciones de satisfacción. En este caso, las condiciones de verdad.

Es importante formular este punto con cuidado. La importante noción de “condiciones de satisfacción” tiene la ambigüedad usual proceso-producto entre el requerimiento y la cosa requerida. La condición de satisfacción de la creencia de que está lloviendo en el sentido del requerimiento es: para que ella sea verdadera, tiene que estar lloviendo. La condición de satisfacción en el sentido de la cosa requerida es el estado de cosas efectivo de que está lloviendo en el mundo. En estos ejemplos, el requerimiento está en la mente; la cosa requerida está en el mundo. En el caso de practicar la pronunciación del alemán, la condición de satisfacción de mi intención fue sólo producir una secuencia en alemán. En el segundo caso, donde realmente significo lo que digo, también tenía ésta como una condición de satisfacción, pero ella no era la condición total de satisfacción. La condición total de satisfacción requiere que yo produzca la secuencia y la secuencia debe ella misma tener condiciones adicionales de satisfacción. A saber, que debería ser el caso que esté lloviendo. La hipótesis general que esto sugiere, y que yo creo que es cierta, es que la esencia del significado del hablante es la imposición intencional de condiciones de satisfacción (en el sentido del requerimiento) sobre las condiciones de satisfacción (en el sentido de la cosa requerida). La intención es que *el requerimiento* en el primer caso, de que yo produzca una cierta enunciación, es que en la producción de la enunciación, la cosa requerida, debe ella misma tener condiciones de satisfacción.

Dicho brevemente, la creación del significado del hablante en un enunciado es la imposición intencional de condiciones de satisfacción (en el sentido del requerimiento) sobre las condiciones de satisfacción (en el sentido de la cosa requerida). Es este doble nivel de intencionalidad —tener la intención de producir un enunciado y tener la intención de que el enunciado mismo tenga significado en el sentido de condiciones

de satisfacción— es éste doble nivel de intencionalidad el que indica que los hablantes hacen algo y significan algo con ello.

*La comunicación del significado:* el siguiente paso es la comunicación. En el caso de la comunicación, el hablante no sólo dice algo y significa algo con ello, sino que también pretende que su oyente entienda esto y que el oyente entienda el enunciado, es decir, que simplemente él entienda su significado. Y nosotros sabemos que para entender el significado del hablante el oyente debe entender la intención del significado del hablante. Así que la intención total involucrada en decir algo, significarlo y comunicarlo, es que el hablante pretende producir un enunciado que es tal que la enunciación es una condición de satisfacción de la intención de producirlo y el enunciado mismo tiene condiciones adicionales de satisfacción y se pretende que el oyente reconozca el enunciado y, también reconozca que dicho enunciado tiene esas condiciones de satisfacción. En una filosofía del lenguaje completa, tendríamos que hacer un análisis un poco más complicado, pero esta es la idea básica subyacente. Cuando un hablante dice algo y significa algo con ello, él tiene que: (A) tener la intención de producir un enunciado; (B) tener la intención de que ese enunciado que él produce tenga ciertas condiciones de satisfacción; y (C) tener la intención que su oyente reconozca las intenciones A y B (y que el hablante pretende que él reconozca que se pretende que él haga estos reconocimientos).

*Convención:* el siguiente paso es llegar a tener dispositivos que el oyente puede utilizar de manera estándar para significar algo y comunicar lo que significa. Asumimos que en el grupo que incluye a los hablantes y los oyentes, se han desarrollado ciertos procedimientos por medio de los cuales usted puede dar a entender de manera estándar cosas como: “comida”, “peligro”, “fuego”, “lluvia”. Cada uno de estos procedimientos pueden ser convencionalizados en el sentido de que existe un procedimiento estándar para hacer algo y dicho procedimiento tiene las dos siguientes características de las convenciones. Primero, las convenciones son más o menos arbitrarias en el sentido de que otros procedimientos hubieran podido funcionar igualmente bien. Pero, segundo, una vez el procedimiento es adoptado, es decir, una vez el procedimiento es reconocido generalmente, luego los otros tienen el derecho de esperar que cuando el hablante invoque la convención de un enunciado está siguiendo las reglas vinculadas a la convención. Si el hablante emite algo que significa “lluvia” entonces él está comprometido con la opinión de que está lloviendo.

*Sintaxis*: el tercer e igualmente notable paso es la introducción de la estructura sintáctica. Obsérvese que lo que hemos hecho hasta ahora es describir cómo las formas pre-lingüísticas de la intencionalidad pueden evolucionar en el lenguaje, en el sentido lingüístico. Lo que es más notable acerca de los lenguajes humanos es que no tienen simplemente oraciones sin estructura, sino más bien que las oraciones típicamente tienen una estructura sintáctica interna de cosas tales como frases nominales y frases verbales. Este es un avance intelectual notable porque esto le da al hablante la capacidad de manipular libremente los elementos sintácticos, de una manera que los animales, que carecen de lenguaje, pero sólo tienen estados intencionales, no pueden manipular libremente los componentes de su estado intencional. Con la introducción de la estructura sintáctica, el hablante está en posición de crear arbitrariamente un número indefinidamente grande de representaciones lingüísticas según sea su voluntad. Mi perro, por ejemplo, puede creer que alguien está tocando a la puerta. Pero como carece de sintaxis, no puede hacer una distinción entre esta creencia y la creencia de que la puerta está siendo tocada por alguien. Mas aún, el perro no puede fantasear sobre estados de cosas como: “deseo que miles de personas se acerquen a la puerta”; “deseo que la gente se acerque a la puerta todo el tiempo”. Una vez que se tienen manipulados elementos sintácticos, se puede separar la intencionalidad de sus causas inmediatas, en forma de percepciones de una manera tal que no es posible hacer separaciones de elementos representacionales estructurados asintácticamente.

Así que entonces tenemos una comunidad de humanos que tienen la completa gama biológica de formas pre-lingüísticas de intencionalidad y ellos construyen formas lingüísticas de intencionalidad sobre esto. Ahora tienen significado lingüístico y comunicación y notablemente ellos incluso se comunican con oraciones que tienen una estructura sintáctica interna. Con todo este aparato, se encuentran en una posición mucho más poderosa que cualquier otra forma de vida animal pre-lingüística que conozco.

## V Racionalidad

Con la estructura que contiene consciencia, intencionalidad (incluyendo intencionalidad colectiva) y lenguaje estamos ahora en posición de introducir una de los rasgos humanos más distintivos: la racionalidad.

Es un error pensar, como Aristóteles aparentemente lo hizo, que la racionalidad es parte de la definición de la humanidad de manera tal que nos distingue de todos los otros animales. No es el caso que los humanos sean los únicos “animales racionales”, pero tienen formas de racionalidad que no tienen los animales que carecen del tipo de estructuras lingüísticas que tienen los humanos. Cualquier intencionalidad, sin importar cual sea, creencias, intenciones, deseos está sujeta a las restricciones de la racionalidad. Si, por ejemplo, un animal tiene una creencia y después da cuenta de que su creencia es falsa, entonces la racionalidad exige que modifique la creencia inicial. Ésta es una restricción mínima a la racionalidad.

El punto para la presente discusión es que una vez que los seres humanos tienen un lenguaje, entonces la racionalidad impone restricciones más ricas y tenemos ahora un completo conjunto de restricciones al pensamiento y al habla las cuales son constitutivas de la racionalidad humana, acerca de las cuales hemos desarrollado teorías bastante elaboradas en la forma de la lógica, incluyendo la lógica inductiva, diversas lógicas probabilísticas y modales, entre otras. El punto, pues, es este —y este es un punto crucial sobre el cual la tradición filosófica está equivocada— la racionalidad no es una *facultad separada*. Una vez que se tiene el pensamiento y el lenguaje, una vez que tenemos intencionalidad y formas lingüísticas de expresión, entonces ya hemos incorporado las restricciones de la racionalidad. La racionalidad no es una *adición* a las formas humanas de representación; ella es una característica estructural de esas formas de la representación.

Hay otro punto crucial sobre el que debemos llamar la atención ahora. Es un error pensar que la naturaleza no contiene normas, que la naturaleza es, de alguna manera, esencialmente no-normativa. Lo que hemos visto es que la consciencia, la intencionalidad y la racionalidad son partes de la naturaleza. Y en efecto, a pesar de que es convencional en sus formas específicas como el francés o el alemán, el lenguaje es esencialmente parte de la naturaleza, de una manera que he tratado de describir y es esencialmente normativo. Si, por ejemplo, hacemos una afirmación, la norma requiere que, si la afirmación es exitosa, tiene que ser verdadera, tenemos que tener evidencias de que es así y tenemos que creerla, de lo contrario no estaríamos siendo sinceros. Todas esas normas están incorporadas dentro de la estructura misma del acto de habla de la aserción, así como la aserción está incorporada dentro de una intencionalidad humana desarrollada, así también la intencionalidad

humana es parte de la naturaleza. Es un error, un profundo error, contrastar naturaleza y normatividad. La naturaleza no sólo contiene la *base* de la normatividad, como veremos cuando discutamos acerca de la ética, sino que ésta tiene incorporadas restricciones normativas de racionalidad.

## VI Las posibilidades del lenguaje

Hemos visto antes que hay dos maneras básicas en que la mente se relaciona con la realidad, la dirección de ajuste cuesta abajo o mente-a-mundo y la dirección de ajuste cuesta arriba o mundo-a-mente. Una vez que se tiene el lenguaje, se obtienen posibilidades interesantes de variación en estas dos direcciones de ajuste. En el lenguaje, hay cinco y sólo cinco tipos posibles de actos básicos de habla como las afirmaciones, ordenes, promesas y preguntas. Austin los bautizó como "actos ilocucionarios". Primero, existen casos en los que el punto del acto ilocucionario es representar cómo son las cosas. Yo los llamo *asertivos*. La principal característica de éstos es que son representaciones de cómo son las cosas en el mundo. Por tanto, ellos tienen la dirección de ajuste de palabra-a-mundo. Ejemplos de casos asertivos son los enunciados, afirmaciones y descripciones. La segunda es la de los *imperativos*. El propósito que define a los imperativos es tratar de lograr que el oyente haga algo. Ellos tienen la dirección de ajuste mundo-a-palabra, algunos ejemplos son órdenes, comandos y peticiones. La tercera clase son los *compromisorios* y el punto que los define es comprometer al hablante para que haga algo y, como los imperativos, tienen la dirección de ajuste palabra-a-mundo. El ejemplo favorito del filósofo de un compromisorio es la promesa, pero otros ejemplos son votos, amenazas, contratos y garantías. La cuarta clase es algo peculiar. A los actos que componen esta clase los llamo *expresivos*. Esta clase se da cuando el hablante expresa algunos sentimientos o actitudes acerca del estado de cosas que presupone que existen, antes de ser emitidas, por lo que cuentan con lo que yo llamo la dirección de ajuste "nula", algunos ejemplos son apologías, agradecimientos y felicitaciones. El punto de los expresivos es expresar algún sentimiento o emoción acerca del estado de cosas cuya existencia se da por sentada. La última clase es especialmente interesante para nuestra presente discusión. A los actos de esta clase los llamo *declaraciones*. Ellos son casos en los que el hablante realiza una emisión cuyo propósito es producir cambiar la realidad para

que se ajuste al contenido proposicional y este cambio se produce representando la realidad como habiendo sido cambiada así. Algunos ejemplos favoritos de los filósofos son los que comúnmente se describen como “performativos” donde, por ejemplo, un hablante debidamente autorizado levanta la sesión diciendo simplemente “se levanta la sesión” o cuando alguien pueda hacer una promesa diciendo “me comprometo a venir a verte”. En estos casos, el hablante realiza el acto nombrado por el verbo, levantar o prometer, pero él lleva a cabo el acto representándose a sí mismo como realizándolo. Así mismo, hacemos que algo sea el caso diciendo que es el caso y por eso tenemos que decir que estos actos de habla tienen ambas direcciones de ajuste a la vez. Ambos representan cómo son las cosas, como en el caso de los *asertivos*, pero ellos también cambian la realidad, como en el caso de los *directivos* y los *compromisorios*: pero y esto es lo distintivo de ellos, las *declaraciones* no contienen dos actos de habla separados, sino más bien un acto de habla singular con ambas direcciones de ajuste simultáneamente, de modo que uno crea un cambio en la realidad al representarla como cambiada.

Hasta el momento tenemos un rico conjunto de recursos humanos, todos los cuales son desarrollos naturales de la realidad básica subyacente y todos ellos productos de la evolución biológica y cultural. Específicamente tenemos consciencia, intencionalidad (incluyendo la intencionalidad colectiva con todo el aparato que ello implica). Y entonces, tenemos este hecho notable: los lenguajes humanos con significado, convenciones y estructura sintáctica al interior de las oraciones. La oración es la unidad mínima en un lenguaje porque es la mínima expresión de un estado intencional y, por tanto, dado que cada acto de habla es la expresión de un estado intencional, es la unidad mínima del acto de habla. Es difícil exagerar el creciente poder incrementado que los seres humanos adquieren con el lenguaje y es un punto familiar decir que ahora tienen la capacidad de transmitir información en una forma más rica que cualquier otro animal. Particularmente cuando se agrega al lenguaje la capacidad de preservar registros de los actos de habla en forma escrita, entonces se tiene una capacidad indefinidamente grande para crear y almacenar información. Pero hay una función aún más importante del lenguaje como veremos en la siguiente sección.

## VII Realidad social

Los humanos usan el lenguaje para crear una realidad que sólo existe por acuerdos humanos o aceptación. La realidad así creada □ familia, propiedad privada, dinero, gobierno, universidad, entre otros □ distingue a la especie humana de cualquier otra especie que conozco. No es una exageración decir que los humanos usan el lenguaje para crear una civilización humana distintiva. Explicaré ahora esta noción. Los humanos tienen la capacidad, que comparten con algunos otros animales, de imponer funciones a los objetos. De modo que, por ejemplo, el objeto pueda ser usado como una herramienta. Las funciones en este sentido son siempre relativas al observador. La función de las cosas tales como herramientas, casas, botes y armas, son funciones que están, o pueden realizarse, en virtud de la estructura física de la entidad a la que se le impone la función. Pero, dado el lenguaje, los humanos tienen la capacidad de imponer un determinado tipo de función sobre los objetos y otras personas que se diferencia de las funciones que se realizan en virtud de la estructura física. Muchas de las funciones que la gente realiza tales como la función del presidente de los Estados Unidos o la función de un objeto de servir como dinero, son funciones que no se pueden realizar en virtud de la estructura física de la entidad o al menos, no en virtud de la estructura física por sí sola. Así, que no hay nada en la hoja de papel como tal que la haga valiosa como dinero, nada en el pedazo de metal que hace que sea una moneda de valor. Tiene valor, para decirlo crudamente, porque los seres humanos piensan que tienen un valor. Y la forma que adopta esa asignación del valor del dinero es la de asignar al pedazo de papel o a la moneda, un cierto estatus y con ese estatus una función que sólo puede llevarse a cabo en virtud de la aceptación colectiva de dicho status. A éstas las llamo “funciones de status”. Ellas son el pegamento que mantiene a la sociedad unida y, en algún sentido, son un elemento esencial de la civilización humana. La función de status incluye dinero, propiedades, gobierno, matrimonio y también universidades, abogados, médicos, vacaciones de verano y cocteles. Todos estos son casos de “función de status”, en los que una entidad, persona, grupo de personas o un proceso tienen un status y con ese status una función que se realiza en virtud de la aceptación colectiva de ese status. ¿Cómo se crean y mantienen las funciones de status? La respuesta breve es que ellas son creadas y mantenidas por representaciones lingüísticas y aunque éstas no siempre están en la forma de actos de habla explícitos, siempre tienen la misma estructura



lógica subyacente. Todas ellas son declaraciones en el sentido que he explicado anteriormente. Así que, cuando hacemos que alguien sea presidente o le damos a alguien un título de universidad, o tratamos a algo como dinero, hacemos que sea el caso, que sea un presidente o un título universitario, o dinero, al representárnoslo como siendo el caso. Llamaré a estas declaraciones de funciones de status y ellas tienen la característica que tienen todas las *declaraciones*, de tener doble dirección de ajuste. Ellas hacen que algo sea el caso al representárselo como siendo el caso.

El efecto de crear funciones de status es crear facultades. En virtud de tener dinero, estructuras políticas o propiedad privada incrementamos nuestras facultades. Los juegos humanos son modelos simples para estudiar las funciones de status, porque puede verse de forma aislada del resto de la vida humana. Por ejemplo, en el beisbol el lanzador, el bateador y el corredor de base tienen ciertas facultades que no tendrían si no fueran parte del juego. Y ¿qué tipo de cosas son esas facultades? Éstas son un conjunto muy peculiar de facultades, marcadas por algunas palabras como “derechos”, “funciones”, “obligaciones”, “requisitos”, “autorizaciones”, “permisos”. Sólo para tener una etiqueta general las llamo “facultades deónticas”. A los que encarnan estas facultades deónticas, como el hecho de que alguien es un corredor de base o que alguien es el presidente de los Estados Unidos, o que un pedazo de papel es un billete de veinte dólares yo los llamo “hechos institucionales”. Podemos ahora establecer una serie de relaciones dejando de lado distintas cualificaciones de modo que se pueda ver la estructura de los huesos desnudos de la civilización humana. La realidad institucional en la forma de hechos institucionales es creada por medio de declaraciones de funciones de status. Todas ellas crean y luego mantienen las funciones de status y las funciones de status son sin excepción facultades deónticas.

¿Por qué son tan importantes las facultades deónticas? Ellas son el pegamento que mantiene unida la sociedad humana. ¿Cuál es el poder de ese pegamento? La respuesta es que en la medida en que la gente reconoce la validez de las funciones de status, ellos las reconocen como teniendo un status deóntico y, por ésta razón, las reconocen como dando razones para la acción que son *independientes de sus inclinaciones inmediatas*. Abreviaré esta idea diciendo que *las funciones de status proveen razones para la acción que son independientes del deseo*. Los ejemplos son bastante obvios. Si hago la promesa a los organizadores en Münster de producir mis contribuciones para un volumen, entonces, reconozco

que tengo una *obligación de hacer eso* y a su vez independientemente de mis otras inclinaciones, sin importar o no si quiero hacerlo, llega un momento en el que reconozco que tengo una razón para hacer algo que es independiente de mis inclinaciones. Hasta donde sé, ningún animal tiene algo como eso. Cuando entreno a mi perro, lo entreno para tener inclinaciones que se acomoden a mis deseos. Lo entreno de manera tal que cuando me oiga llamarlo, sentirá una inclinación a hacer lo que yo quiero que haga, venir en la dirección desde la que lo estoy llamando. Lo que no hago, o no puedo hacer, pero sí puedo hacer por los seres humanos, es darle un sentido de obligación. Si alguien promete venir y verme en un momento determinado entonces, a diferencia de mi perro, ese alguien tiene una razón para hacerlo, que es independiente de sus inclinaciones. ¿Por qué no puede el perro actuar de acuerdo a obligaciones? La respuesta es que para actuar de acuerdo a obligaciones, uno tiene que estar en la capacidad de razonar a partir de obligaciones, y para razonar a partir de obligaciones uno tiene que tener el concepto de obligación y para tener ese concepto uno tiene que disponer de medios lingüísticos para expresar dicho concepto. Hay razones complejas de esto, pero la idea básica es que la noción de obligación se encuentra demasiado lejos de la experiencia sensorial inmediata para poder ser poseída por animales que no tienen algunos medios lingüísticos para expresar ese concepto.

La discusión hecha hasta aquí plantea un interesante enigma. Dado que todas las acciones intencionales, conscientes y voluntarias son expresiones de deseos para realizar esa acción aquí y ahora ¿cómo puede ser el caso que hayan razones independientes de los deseos para realizar acciones? ¿Cómo pueden las razones —independientemente de los deseos— motivar acciones que expresan deseos? La respuesta es un poco complicada, pero la idea básica es esta: en la medida en que se reconoce la *validez* de una razón para la acción independiente del deseo, se reconoce que se tiene un *fundamento* o un *motivo* para realizar la acción y en esa medida se tiene una base para *desear* realizar la acción. Esto en principio no es más misterioso que el hecho de que uno pueda formar deseos para hacer cosas que son desagradables como un medio para hacer algo que es agradable. Así, por ejemplo, no es divertido para mí ir al dentista, pero deseo ir al dentista, porque hay algo que quiero, a saber, buenos dientes. Mi deseo de ir al dentista es un deseo *derivado*, derivado de otro deseo. Mi deseo de mantener mi promesa es también un deseo derivado, pero éste no es derivado de otro deseo, sino más bien se deriva del reconocimiento de la validez de un hecho deóntico — el hecho que estoy bajo una obligación.

Si parece misterioso que alguien alguna vez pueda tener un deseo sobre la base de una razón independiente del deseo, considérese la razón teórica, y se verá que es obvio que eso pasa. Hay cosas que yo encuentro que preferiría no creer. De todos modos, si descubro que algo es verdadero entonces tengo una razón independiente del deseo para creerlo. Tengo una razón para creer en eso, incluso si no quiero creerlo. Esto no es más misterioso en el caso de la razón práctica. Puedo tener un motivo para hacer algo que, normalmente, no me agradaría hacer, y ese motivo, cuando reconozco su validez, puede proporcionar una base para *querer* hacerlo a pesar de que el motivo por sí mismo no sea un deseo.

Nótese que yo digo que *puede* darme el motivo. A menudo las personas reconocen que tienen razones independientes del deseo para hacer algo y luego ellos simplemente no hacen aquello para lo cual tienen una razón. El reconocimiento de las razones para la acción independientes del deseo no siempre produce un deseo para realizar la acción, incluso después de que uno ha reconocido la validez de la razón para hacerlo. De todos modos tales reconocimientos pueden proporcionar deseos y, por tanto, pueden motivar racionalmente las acciones.

Así obtenemos un conjunto de ecuaciones y derivaciones. Podemos resumir esta parte de la discusión con el siguiente sumario del conjunto de relaciones:

Hechos institucionales = función de status → facultades deónticas  
→ razones para la acción independientes del deseo → posibles motivaciones para la acción.

En buen castellano, todos los hechos institucionales son funciones de status, las funciones de status contienen facultades deónticas y las facultades deónticas, cuando se reconoce su validez, proporcionan razones para la acción independientes del deseo y a su vez proporcionan posibles motivaciones para actuar.

## VIII Ética

Hasta ahora hemos considerado la manera de conciliar la realidad humana con la realidad básica donde la reconciliación tiene que

mostrar no solo que la realidad humana es coherente con la realidad básica, sino que, en sus diversos rasgos, es la consecuencia natural de la realidad básica. Hemos considerado la consciencia, la intencionalidad, la racionalidad, el lenguaje y la sociedad. Estoy suponiendo que todas estas son, en algún sentido crucial, fenómenos naturales, en tanto en cada paso podemos ver el desarrollo de cada etapa como una consecuencia natural de la etapa anterior y, si el carácter natural es transitivo, entonces todos esos fenómenos son consecuencias naturales de la realidad básica como es descrita por la física y la química. El siguiente paso es la ética. ¿Podremos mostrar de manera similar que la ética es, en el sentido que hemos estado mostrando, un fenómeno natural?

Para responder esta pregunta primero tenemos que explicar qué entendemos por "ética". Por ética entiendo el conjunto de normas, principios, criterios y prácticas que regulan el comportamiento de los miembros de las comunidades, sobre las cuestiones que la gente en las comunidades considera importantes, de una manera especial. Tenemos que decir "importante" para distinguir la ética de otras formas de principios tales como la manera de atarse los zapatos o la manera de peinarse el cabello, las cuales no son cuestiones de ética a menos que la comunidad las considere importantes en este sentido moral especial. Es importante también ver que la ética es una cuestión de regular el comportamiento de los miembros en comunidades. Es posible, por supuesto, que una persona tenga obligaciones éticas consigo mismo, pero éste es un caso límite. En el caso normal, la ética es una cuestión de regular la conducta entre personas, regular el comportamiento social. Creo que usando el aparato desarrollado hasta aquí podemos mostrar que la ética, en este sentido, es un fenómeno natural y yo describiré la forma de hacerlo.

En primer lugar, tengo que señalar que mientras escribo esto, se hacen varios esfuerzos para explicar la ética por medio de un argumento genético-evolutivo que dice que hay normas éticas universales y que son resultado de la selección natural darwiniana y que están codificados en el cerebro de todos los seres humanos normales. Se supone que esto es análogo a la manera en que los lenguajes humanos naturales están codificados en los cerebros humanos en lo que solía ser llamado "gramática universal". Se supone que análogamente hay una "gramática ética universal" en nuestros cerebros como resultado de procesos evolutivos de selección. Considero este tema como una cuestión empírica que, nosotros probablemente, no estamos todavía en condiciones de

responder. Pero, en cualquier caso, no es la cuestión que estoy tratando de responder. El problema que estoy tratando de resolver no es un problema de la base genética para el buen comportamiento ético, sino más bien el desarrollo de la ética como una práctica humana independientemente de la medida en que esté determinada genéticamente. Me parece que hay algunas características de la ética que la señalan como un fenómeno natural en el sentido que he venido explicando. En primer lugar, todo tipo de conflictos surgen inevitablemente entre los miembros de la comunidad en su lucha por cosas tales como bienes escasos, acceso a parejas sexuales, límites territoriales, y es esencial para la supervivencia y el florecimiento de la comunidad que ellos tengan alguna manera de regular su comportamiento sobre estas y otras cuestiones. Hasta ahora, sin embargo, esto no nos da como resultado la ética. En alguna medida, esos conflictos se dan entre animales no-humanos. Los principios, que son específicos de la ética humana y sobre los que quiero llamar la atención son:

- 1) La ética, en el sentido que he explicado, requiere de un *lenguaje*. ¿Por qué? La ética en el sentido humano requiere que las personas sean capaces de razonar sobre la base de sus obligaciones, deberes y requisitos éticos en general. Tal razonamiento requiere un modo lingüístico de representación. ¿Por qué? Porque para razonar a partir de obligaciones e imperativos éticos de distintos tipos, uno tiene que tener los conceptos relevantes y esos conceptos no pueden basarse solamente en, por ejemplo, la percepción. Uno tiene que tener el modo lingüístico de percepción de modo que uno pueda manipular efectivamente los conceptos. Sólo las criaturas que tienen un lenguaje pueden crear y actuar sobre la base de los tipos de razones que son específicamente éticas. Esto me lleva a mi siguiente punto.
- 2) La ética está esencialmente interesada en las *razones para la acción independientes del deseo*. Vimos en nuestra discusión sobre realidad institucional que ella funciona en la sociedad humana para crear facultades deónticas y que ellas son esencialmente cuestiones de *razones para la acción independientes del deseo*. No todas las razones éticas son institucionales y claramente no todas las razones institucionales son éticamente relevantes, pero es evidente que hay una superposición, y más importante para esta discusión, el aparato creado por el uso del lenguaje en la realidad institucional es esencial, es una condición necesaria para el funcionamiento de la ética. Kant estaba en lo cierto al ver que la ética se ocupaba esencialmente

de razones para la acción independientes del deseo. Si todo el mundo actuara solamente según sus inclinaciones y por algún milagro esas inclinaciones armonizaran de tal manera que todo el mundo se beneficiara con ese comportamiento egoísta, entonces la ética, en el sentido que he tratado de describir, no existiría y nadie la necesitaría. El punto en el que estoy haciendo énfasis no pretende ser un punto sólo acerca de la palabra "ética". Si alguien quiere decir que las aves que se sacrifican por sus hijos se comportan éticamente, eso está bien para mí. Pero el punto en el que estoy haciendo énfasis es que es un rasgo distintivo de la ética humana es que ella se desarrolle en torno a razones para la acción que no siempre concuerdan con nuestra inclinación pre-ética. A menudo, y precisamente, en una situación ética, se nos pide anular nuestras inclinaciones naturales haciendo prevaler las cuestiones éticas.

Hay todo tipo de razones éticas para la acción que no son cuestiones de la operación de hechos institucionales, como las promesas. Así, por ejemplo, la mayoría de las comunidades reconocen que sus padres tienen una obligación no-institucional y, sin embargo, ética de cuidar a sus hijos. Vamos a suponer que ésta es una norma ética universal. Su universalidad, para nuestro propósito, no importa. Lo que importa es que es un tipo de razón para la acción que se extiende más allá de la inclinación biológica normal de los padres a cuidar a sus hijos, porque requiere que los padres estén preparados para actuar en beneficio de sus hijos, incluso en casos en que los padres no están inclinados a hacerlo. Y este es un rasgo que define a la ética.

- 3) Una tercera característica de la ética, que ya está integrada en la estructura del lenguaje es la *universalidad*. En la medida que uno reconoce que otros están bajo la obligación ética de ayudarlo a uno, precisamente en esa medida, la formulación lingüística de la razón para la acción requiere *lógicamente* que uno mismo está bajo una obligación de ayudar a otros cuando estén en situación similar. La independencia del deseo y la universalidad están, por tanto, integradas en la naturaleza lógica de la ética. Puede verse esto si se consideran algunos ejemplos. Si usted se lesiona y tiene un dolor, y cree que eso es una razón moral para que otras personas lo ayuden, entonces la forma lógica de su enunciado lo compromete con el punto de vista de que en una situación similar □ cuando alguien más esté lesionado y con dolor □ usted estaría en la obligación moral de ayudarlo. La sola universalidad del lenguaje requiere que casos

similares se describan de manera similar y esto implica que una razón que usted se da a sí mismo en virtud de una característica que usted instancia, se aplicará en general a cualquier persona que instancie esa característica.

Tal vez sea importante enfatizar que esta explicación de la ética no pretende darnos un código moralmente aceptable de ética. Estoy identificando ciertos rasgos formales, pero es perfectamente posible tener un código ético que satisfaga estas condiciones y que sea bastante perverso. No es mi interés *prescribir* una forma específica de código ético, sino más bien describir ciertas características formales que todos los códigos de ética tienen en común y que yo pienso son derivaciones naturales de la explicación de la vida humana y la sociedad que he estado dando. Si “son derivaciones naturales”, designa una noción transitiva, entonces la ética entendida como un conjunto de normas para regular el comportamiento social sobre asuntos importantes, es una consecuencia natural del comportamiento de las partículas físicas.

## IX Estética

El objetivo de esta sección es mostrar como la realidad humana no sólo es consistente con la realidad básica, sino que es un desarrollo natural de la misma. Un buen caso que lo confirma es la *estética* y en un nivel de explicación de la posibilidad de la estética es un caso más bien fácil dado todo lo que hemos dicho hasta ahora: todas las experiencias conscientes tienen una dimensión estética. Cada experiencia consciente es tal que cae dentro de un ámbito que admite términos evaluativos como “agradable” o “desagradable”, “bello” o “feo”, “placentero” o “no placentero”. Entonces, por supuesto, a medida que lleguemos a casos más elaborados, aparecen cosas tales como “vulgar” o “sublime”, “kitsch” o “sofisticado”. Pero estos se encuentran en el nivel superior de la evaluación estética.

En los planes de estudios universitarios, ética y estética se encuentran a menudo juntas como si hicieran parte de un tema más amplio. Pero, de hecho, la estética es totalmente diferente de la ética tanto en la materia, como en sus requisitos formales. Un requisito formal inmediato en el que difieren es este: si hago una crítica ética a alguna persona o a alguna acción, entonces estoy lógicamente obligado a decir lo que la



persona debió haber hecho en lugar de hacer lo que de hecho hizo. No estoy diciendo que hay un sentido ético del “deber” que implique “poder”. Estoy diciendo que es un requisito lógico de juicios negativos éticos que quien hace el juicio, se compromete a ser capaz de decir qué debió haber sido hecho, en lugar de la forma éticamente inaceptable de conducta. Esta restricción no se traslada a la estética. Si yo critico una obra de arte, no estoy comprometido a decir cómo el autor debió haberla hecho. Si yo critico, por ejemplo, el último acto de *La Traviata* sobre la base de que la muerte de *Violetta* es exagerada hasta rayar en la vulgaridad, no estoy comprometido a decir cómo Verdi y su libretista debieron haberla hecho. Cualquier otra manera de hacerlo hubiera podido hacer que la opera fuera todavía peor. Sin embargo, uno podría criticar la opera por estas razones. La crítica estética no requiere la especificación de una mejor alternativa, mientras que la crítica ética sí. La diferencia se deriva de la observación que hice anteriormente de que la ética tiene que ver con la regulación de la conducta humana en asuntos de gran importancia. ¿De qué se ocupa la crítica estética? En un sentido, es fácil mostrar como la estética es un fenómeno natural porque *toda la consciencia humana tiene una dimensión estética*. Para cada experiencia consciente, es siempre apropiado preguntar: ¿fue agradable?, ¿placentero?, ¿desagradable?, entre otros. Y más importante aún, para el objeto de la experiencia consciente, la escena que está usted viendo, los sonidos que está escuchando, la comida que está degustando, hay toda una gama de dimensiones estéticas de evaluación. ¿Es la escena bella o fea?, ¿es agradable u opresiva?, ¿es liberadora o limitante? Esto es aún más fuerte en el caso de las experiencias del gusto. El rango de la experiencia del gusto, desde lo delicioso hasta lo repugnante, es casi enteramente una cuestión de hacer juicios estéticos acerca de la calidad de la experiencia. Por tanto, la estética como un dominio de la experiencia humana es absolutamente ubicua. No hay manera de que se pueda tener una experiencia consciente sin que se caiga en el dominio de la apreciación estética.

Por ésta razón, tenemos que hacer una distinción entre la estética y el arte. Toda la vida humana contiene una dimensión estética mientras que el arte, en el sentido en el que se ha desarrollado en la sociedad occidental, desde los tiempos de los griegos es una institución humana muy especial. Ahora quiero identificar algunas de sus características. Es poco probable que todas las culturas tengan algo parecido a nuestra institución del arte. Sin embargo, nuestra institución ha evolucionado y tiene por lo menos las siguientes características formales:

- 1) La *intencionalidad* de la obra de arte. Un objeto es para nosotros una obra de arte si es objeto de una evaluación estética como obra de arte, sólo en la medida en que consideramos que es un producto de la intencionalidad humana. El caso límite es cuando alguien, por ejemplo, encuentra una piedra que uno trata como un *objet d'art trouvé*. Esto es perfectamente aceptable y yo mismo tengo tales piedras. Pero es importante ver que la elección de la piedra y su distinción de otros objetos es en sí misma una forma límite de la intencionalidad. Esto conduce a la segunda característica.
- 2) La obra de arte requiere algún tipo de *límite* entre ella y el resto del mundo. La pintura tiene un marco que la rodea. La ópera tiene un comienzo, una continuación, una parte media y un final. La novela se encuentra, típicamente, en un volumen separado. Las obras de arte son auto-contenidas. La separación física de la obra de arte de su entorno se relaciona con una tercera característica.
- 3) Las obras de arte suelen *estar aparte* del resto de nuestras actividades diarias. Si considero mis zapatos o mi carro como obras de arte, como de hecho yo lo hago en el caso de mi carro, estoy ubicando su dimensión estética como algo especial acerca de su eficacia funcional. Esto nos lleva a la cuarta característica.
- 4) Kant dijo que nuestro interés en una obra de arte es algo desinteresado, en la medida en que la obra de arte es, ella misma, un vehículo de poder o un objeto útil. En la medida en que la consideramos bajo este aspecto, no estamos considerándola como una obra de arte. Creo que hay algo correcto en esto, pero por supuesto, muchas obras de arte son ellas mismas obras de propaganda. Piénsese en la *Guernica* de Picasso como un ejemplo famoso.

En la medida en que la obra de arte es un objeto públicamente aceptado como tal, entonces satisface nuestra definición de un hecho institucional. Hay una deontología que se aplica al tratamiento de algo como una obra de arte. Ciertos tipos de actitudes y comportamientos, se supone, que son apropiados. Por eso hubo tanto debate en la época de la revolución en el arte moderno entorno a la cuestión: “¿pero eso es realmente arte?”. La idea no era solamente cuestión de aplicar una etiqueta, sino más bien se trataba de determinar cuál es la reacción y la actitud apropiadas y si las características del objeto en cuestión merecían esa reacción y esa actitud. Pero en la medida en que la obra de arte es un hecho institucional,

estamos dejando de lado algo que es crucial para nuestra concepción occidental del arte, el aspecto de la creatividad y la originalidad del artista. En cuanto el artista actúa como artista, él precisamente no piensa en sí mismo como creando simplemente un hecho institucional de la misma manera en que alguien imprimiendo dinero o designando un nuevo departamento de gobierno piensa de sí mismo que está creando un hecho institucional. Él piensa que su obra de arte es una creación de algo que tiene un tipo especial de mérito, como un acto de originalidad y creatividad humana. Así que hay una interesante tensión en el arte, que no existe en otras formas de hecho institucional, entre el aspecto de la obra de arte como teniendo ella misma un status deóntico y mereciendo ciertos tipos de facultades deónticas y el énfasis en la creatividad original del artista quien precisamente no piensa en sí mismo como creando simplemente otro hecho institucional rutinario, sino como haciendo alguna contribución única a la experiencia estética humana.

# LA FILOSOFÍA DE LA HISTORIA DE LA CIENCIA DE THOMAS KUHN\*

*THE PHILOSOPHY OF HISTORY OF SCIENCE OF THOMAS KUHN*

ALEXANDER BIRD

University of Bristol, UK. alexander.bird@bristol.ac.uk

RECIBIDO EL 19 DE OCTUBRE DE 2012 Y APROBADO EL 23 DE NOVIEMBRE DE 2012

## RESUMEN ABSTRACT

En el presente artículo defiendo que Kuhn fue un historicista en dos aspectos: primero, fue un conservador en el sentido de Mannheim -la tradición es importante para comprender el cambio científico y la evaluación de una idea científica es relativa al contexto histórico-; segundo, Kuhn respaldó el determinismo -hay un modelo de cambio científico, similar a las leyes del desarrollo científico-. Planteo que el determinismo de Kuhn requiere que él sea un internalista acerca de las causas del cambio científico; y que su internalismo contrasta con el externalismo que caracteriza buena parte de los estudios de la ciencia post-kuhnianos. Concluyo considerando cómo se relaciona el historicismo de Kuhn con los propósitos filosóficos de la historia de la ciencia kuhniana.

In this article, I argue that Kuhn was a historicist in two respects. First, he was a conservative in Mannheim's sense – tradition is important for understanding scientific change, and the evaluation of a scientific idea is relative to historical context. Secondly, Kuhn embraced determinism – there is a pattern to scientific change, akin to laws of scientific development. I show that Kuhn's determinism requires that he is an internalist about the causes of scientific change; Kuhn's internalism contrasts with the externalism that characterizes much post-Kuhnian science studies. I conclude by considering how Kuhn's historicism relates to the philosophical purposes of Kuhn's history of science.

## PALABRAS CLAVE KEY WORDS

determinismo, externalismo, historicismo, internalismo.

determinism, externalism, historicism, internalism.

---

\* Traducido por Juan Carlos Aguirre García. Docente Asociado del Departamento de Filosofía, Universidad del Cauca.

## Introducción

Thomas Kuhn fue un historiador de la ciencia; aunque, en realidad, nunca fue miembro de la amplia comunidad de historiadores. Pese a que había trabajado en el Departamento de Historia de la Universidad de Berkeley, es claro que ese ambiente no fue grato para él, pues sentía que tenía muchas más cosas en común con los filósofos de la ciencia. Los objetivos de Kuhn difieren de los de otros historiadores toda vez que tuvo en su trabajo histórico objetivos filosóficos; además, Kuhn pensaba que la naturaleza misma de la historia de la ciencia la hacía diferente de la historia general<sup>1</sup>. En este artículo revisaré las dos visiones historiográficas de Kuhn y defenderé que era un historicista en ambos sentidos; posteriormente, mostraré que Kuhn era un internalista con relación a la explicación del cambio científico. Finalizaré retornando a aquellos objetivos filosóficos de su historia de la ciencia.

### El historicismo de Kuhn

Comenzaré afirmando que Kuhn era un historicista. El significado de la etiqueta 'historicista' es en sí mismo problemático. Con el fin de superar tales debates articularé dos características que, por lo menos, algunos han asociado con el historicismo y que considero son importantes para comprender el trabajo de Kuhn.

La primera característica es la línea *conservadora* del historicismo<sup>2</sup>. En términos muy amplios, afirma que hay una relación íntima entre la evaluación de una idea (o mejor, de cualquier otro producto humano) y su contexto histórico. Así expresada, la afirmación del historicismo conservador es intencionalmente vaga –las distintas especies e intensidades de historicismo pueden diferir en lo que creen que es la naturaleza y cuáles son las consecuencias de esa relación íntima–. Asumida de un modo débil, la afirmación es casi trivialmente verdadera; tomada en un sentido fuerte y aplicada a una ciencia en particular es altamente polémica. Llamo a esta línea 'conservadora', no porque uno tenga que pertenecer a algún tipo de conservadurismo político para respaldarla –aunque podría ser útil–, sino más bien para aprovechar el

---

<sup>1</sup> Ver, especialmente, el ensayo de Kuhn: "The relations between history and the history of science" (1971), que examina en detalle "las barreras que dividen la historia y la historia de la ciencia".

<sup>2</sup> Esta línea del historicismo puede rastrear sus orígenes en Vico y puede también encontrarse en Herder y Hegel.

sentido en el que, de acuerdo con Karl Mannheim, un pensador puede ser un conservador. Mannheim (1953) contrasta ese pensador con uno que asuma una 'ideología de la ley natural' ilustrada. En consonancia con esta ideología, juzgamos las ideas contra normas generales y sempiternas. La razón, por ejemplo, es la aplicación de ciertas leyes inmutables de la lógica. El conservador, al contrario, dice que para entender e incluso para evaluar una idea debemos apelar a factores históricos y locales.

La segunda característica, en ocasiones etiquetada como 'historicista', es la línea *determinista*<sup>3</sup>. De acuerdo con el historicismo determinista hay leyes del desarrollo histórico. Incluso si los detalles se dejan al azar, las estructuras comprensivas de la historia son inevitables o altamente probables. El historiador no se limita a describir y explicar eventos particulares, sino que también puede esperar ver un patrón subyacente en muchos eventos particulares. En este punto, la historia tiene afinidades con las ciencias y el historiador genial como un gran científico, no solo identificará tales patrones, sino que también puede buscar el modo de explicarlos haciendo referencia a un mecanismo subyacente.

La aproximación de Kuhn a la historia de la ciencia ejemplifica ambas características<sup>4</sup>. Desde la perspectiva del filósofo de la ciencia, la primera característica es la más llamativa. Puesto que el empirista lógico y la filosofía positivista de la ciencia que preceden a Kuhn sostienen que la evaluación de una teoría científica es *sub specie aeternitatis*, la evaluación de una teoría es cuestión de aplicar las leyes de la lógica inductiva a la teoría y a la evidencia total disponible<sup>5</sup>. Esas leyes dirán cuán bien respaldada está una teoría por esa evidencia; y aquellas leyes, como las de la lógica deductiva, se asumen por todas las personas en todos los tiempos. El distanciamiento radical de Kuhn con relación a esta aplicación de la 'ideología de la ley natural' a la ciencia, consistió en sugerir que la evaluación de una teoría es relativa a una tradición de resolver problemas. Para Kuhn (*The Structure* 174-75), el término 'paradigma' tiene dos sentidos: el sentido más amplio, abarca los compromisos compartidos por una comunidad científica, para lo cual él también usa el término 'matriz disciplinar' (Ibíd. 182); el segundo

<sup>3</sup> Hegel y Marx son ejemplares preeminentes del historicismo determinista; este es el historicismo atacado por Popper (1957). Otro ejemplo podría ser Augusto Comte.

<sup>4</sup> Reynolds (1999), también menciona a Kuhn en conexión con distintas especies de historicismo.

<sup>5</sup> Aunque Popper rechazó la inducción, este punto también aplica para él, porque exigió el uso de la lógica deductiva al falsar las teorías.

sentido es más restringido: se refiere a los compromisos más centrales de la comunidad, sus *ejemplares* (Ibíd. 187). Los ejemplares son las soluciones que sirven de ejemplo a la comunidad a la hora de enfrentar los problemas científicos y establecen los estándares para la ciencia posterior en ese dominio. Una parte de la buena ciencia, una solución satisfactoria propuesta para resolver un problema, parecerán la solución de enigmas ejemplares. Por tanto, la evaluación de una teoría no es independiente del contexto, sino que es relativa a una tradición de resolver problemas; más aún, los problemas mismos emergen de la tradición de resolver problemas. Los problemas más valiosos son los que parecen existir o emerger de las fisuras de la tradición existente; así pues, la tradición newtoniana establece los problemas que surgen al tratar de reconciliar las observaciones de planetas, satélites y cometas con las leyes de Newton y de medir el valor de la constante gravitacional  $G$  (entre muchas otras). La importancia de la tradición, también se muestra en sí misma en el fenómeno de la inconmensurabilidad, pues Kuhn (Ibíd. 149-0) sostiene que, una teoría científica puede que incluso no tenga sentido para alguien que trabaje por fuera de la tradición en la que se origina; allí puede haber comprensión incompleta porque la evaluación no está anclada solo a los ejemplares de la tradición, sino porque también son importantes los elementos de los significados de los términos científicos.

El que pueda haber patrones y leyes en la historia de la ciencia no es filosóficamente problemático en sí mismo, pues los empiristas lógicos podrían haber esperado que tales patrones reflejaran las leyes de la lógica; se podría hacer un patrón a partir de la aplicación repetitiva del método científico. Pero cuando se considera el asunto más plenamente, se tiene que la expectativa en el patrón no fue especialmente interesante: sería, en primera instancia, la acumulación de más conocimiento, como la montaña de arena que se va agrandando al interior de un inagotable reloj de arena. El modelo de Kuhn es mucho más interesante que éste, con sus fases alternadas de ciencia normal y extraordinaria (revolucionaria). Y Kuhn agrega un detalle: la transición de la ciencia normal a la revolucionaria procede a través de una crisis. No toda crisis culmina en una revolución, pues algunas podrían resolverse antes de que ocurra. Las revoluciones son asuntos complejos en los que entra en disputa el reemplazo del viejo paradigma por el nuevo o la victoria del sucesor sobre sus competidores entra en disputa.



Los dos elementos del historicismo de Kuhn están unidos. Como se mencionó, los científicos pueden identificar patrones en los fenómenos; también pueden querer explicar aquellos patrones haciendo referencia a mecanismos subyacentes o a leyes más generales. Kepler identificó la naturaleza elíptica de las órbitas de los planetas, así como otros patrones; Newton explicó esto mediante su teoría de la gravitación. Mendeleiev descubrió el patrón periódico entre los elementos; esto se explicó por la teoría atómica desarrollada por Rutherford, Bohr, Thompson y Chadwick. Ver un patrón en la historia de la ciencia es una cosa, explicarlo es otra; aunque realmente estos dos procesos difícilmente se pueden separar. La línea determinista en el pensamiento de Kuhn nos lleva a creer que hay un patrón cíclico en la historia de la ciencia; la línea conservadora explica este patrón. Lo que explica el modelo consiste en que la ciencia se desarrolla al explotar una tradición basada en un paradigma de resolución de problemas. La ciencia normal existe porque un campo científico está dominado por un conjunto de ejemplares. Como se mencionó antes, estos ejemplares establecen la agenda para la investigación posterior, del mismo modo, que muestran cómo todos los objetos del sistema solar se ajustan a las leyes de Newton (en el paradigma newtoniano). Estos principios se hicieron relevantes no solo debido a los *Principia* de Newton, sino porque dieron después los medios para resolver aquellos problemas sobre todo a través de ejemplos del uso de la teoría para resolver problemas de este tipo. Esto explica la existencia de la ciencia normal. No toda la ciencia normal que resuelve problemas lo hace de modo riguroso; por ejemplo, los astrónomos matemáticos del siglo XVIII encontraron difícil reconciliar la órbita observada de la Luna con la teoría de Newton. Clairaut y d'Alembert calcularon el valor para el período de revolución del perigeo de la Luna, que es el punto de la órbita de la Luna que está más cerca de la Tierra. Para encontrar esto, se demoraron dieciocho años, aunque por observación se sabía hacía nueve. Estos conflictos aparentes entre los fenómenos observados y la teoría u otros casos en los que los científicos fallan al resolver problemas son denominados *anomalías*. Kuhn (*The Structure* 80) explica que las anomalías no son consideradas contra-evidencia para la teoría que está en el centro del paradigma. Durante la ciencia normal, el fracaso al resolver problemas se atribuye a los científicos o a la comunidad científica; sin embargo, si las anomalías acumuladas son particularmente significantes y recalcitrantes, entonces, por culpa de las anomalías se inicia un cambio de los científicos al paradigma. Esto es lo que ocurre durante los períodos de crisis. El movimiento anómalo de la Luna fue tan suficientemente serio para Leonhard Euler como para

sugerir que la ley de la gravitación de Newton necesitaba un ajuste, hasta que Clairaut mostró que la anomalía se debía, principalmente, a las aproximaciones inexactas que se usaban. Esto podría pensarse como una mini-crisis que fue resuelta exitosamente al interior del paradigma. Más seria fue aquella crisis que surgió a finales del siglo XIX, derivada de la precesión anómala del perihelio de Mercurio, reportada por Le Verrier, en 1859, y del nulo resultado del experimento de Michelson-Morley. Puesto que, de acuerdo con el conservadurismo de Kuhn, la ciencia normal requiere una tradición establecida con un paradigma creíble y las crisis deben resolverse. Si no se resuelven en el paradigma existente, entonces ese paradigma se debe reemplazar; en particular, se debe reemplazar por un paradigma que pueda respaldar una tradición de resolución de problemas. De este modo, tenemos revoluciones científicas; de esta manera, la línea conservadora en el historicismo de Kuhn (el énfasis en una tradición de resolución de problemas) explica la línea determinista (el modelo cíclico legaliforme de cambio científico).

### El internalismo de Kuhn

El trabajo de Kuhn fue un estímulo para un amplio rango de estudios de la ciencia que van desde la historia de la ciencia hasta la sociología de la ciencia; muchos de los trabajos en estos campos se ven a sí mismos como, de manera laxa, herederos de un legado kuhniano. El libro de Barry Barnes *T. S. Kuhn and social science* (1982), es uno de los ejemplos más prominentes de esto. Sin embargo, el mismo Kuhn (1992) repudió en los términos más fuertes, la más importante (y la más sofisticada filosóficamente) escuela de los estudios de la ciencia, el programa fuerte en la Sociología del Conocimiento Científico (SCC), de la que Barnes fue un brillante líder. Por otra parte, el alcance de las críticas de Kuhn incluyó implícitamente un espectro más amplio que el movimiento de los estudios de la ciencia de la Escuela de Edimburgo.

En esta sección deseo explicar de qué manera Kuhn rechazó el constructivismo social que se encuentra en gran parte en SCC y por qué está bien que lo hiciera así, a la luz de su compromiso con el historicismo. El constructivismo social que Kuhn encontró antitético con sus propias ideas sostiene que los factores principales que determinan los resultados de un episodio científico tales como una crisis, son factores sociales y políticos que se originan por fuera de la ciencia. Por ejemplo, el triunfo de las teorías de Pasteur, que rechazan la generación espontánea, no es el resultado del poder probatorio de sus experimentos con frascos de

cuello de cisne; más bien ese éxito se puede atribuir al hecho de que sus ideas estaban en mejor sintonía con las visiones de la jerarquía conservadora católica, que gobernó en la Francia de Luis Napoleón (Farley Geison); (Farley). El éxito del darwinismo no es una consecuencia de los argumentos expuestos en *El origen de las especies*, sino que puede explicarse mejor gracias a la simpatía natural del libre mercado Inglés con la idea de que el desarrollo es la meta de la libre competencia (Young). Estos son ejemplos de explicaciones externas del cambio científico, en contraste con las explicaciones internalistas que se refieren solamente a los objetivos, valores, prácticas y creencias que se originan en la ciencia.

La propia explicación de la ciencia dada por Kuhn da pie a tales influencias externas, ciertamente no las suficientes como para que sean las determinantes principales de los resultados de los debates científicos. Consideremos primero la ciencia normal: como se explicó antes, el progreso durante la ciencia normal es guiado por el paradigma, por el conjunto de soluciones ejemplares a los problemas que definen una tradición de resolución de problemas. Esto establece tanto la agenda –define qué tipo de problemas vale la pena afrontar–, como los estándares con los que se evalúan las soluciones propuestas a aquellos problemas. La explicación de Kuhn no deja espacio, por ninguna parte, a las influencias externas. En tanto Kuhn enfatiza que el grueso de la actividad científica corresponde a la ciencia normal, de esto se sigue que, al menos, la mayor parte del cambio científico está gobernada por factores internos a la ciencia.

Kuhn respalda explícitamente un enfoque predominantemente (si no exclusivamente) internalista. Kuhn nos dice que,

[...] el ambiente intelectual reacciona a la estructura teórica de una ciencia solo en la medida en que puede ser relevante para los problemas técnicos concretos de los que se ocupan los practicantes del campo. (*The relations* 137-38)

Criticando a los historiadores (desde fuera de la historia de la ciencia) que ignoran este hecho. Kuhn reconoce que la (por lo regular antigua) historia de la ciencia podría limitarse debido al internalismo exclusivo; sin embargo, afirma que:

esa limitación no siempre debe ser un defecto, pues las ciencias maduras por lo regular están más aisladas del

ambiente externo, al menos de las ideas, que los demás campos creativos. (Ibíd. 148-48)

Kuhn plantea dos excepciones al aislamiento general que las ideas científicas hacen de las influencias externas: mientras el desarrollo de una tradición es conducido internamente, no es necesario que lo sean los *orígenes* de esa tradición:

al inicio, en el desarrollo de un nuevo campo..., las necesidades sociales y los valores son los mayores determinantes de los problemas en los que se concentran los practicantes. (*The history* 118)

Kuhn contrasta esto con la evolución posterior de una especialidad científica:

los problemas en los que trabajan los especialistas ya no se presentan a la sociedad externa sino que son un reto interno para incrementar el alcance y la precisión del encuadre entre la teoría existente y la naturaleza. (Ibíd. 119)

La segunda excepción de Kuhn tiene que ver con la *velocidad* con la que la ciencia se desarrolla. Kuhn (*The history* 119) nos dice que el ritmo de un avance científico puede estar condicionado por factores externos. Debe ser cierto que mediante las condiciones económicas prevalentes se puede determinar la cantidad de recursos destinados a la investigación. Kuhn también sugiere que debido a la interacción entre las distintas disciplinas científicas puede haber un efecto acumulativo de factores externos en la evolución de la ciencia. Los avances en la tecnología hacen claramente una diferencia en la capacidad de la ciencia para progresar.

Es importante notar que ninguna de las dos excepciones sugiere que las influencias externas repercuten regularmente en los *resultados* de una investigación o debate científico. En SCC puede distinguirse un programa *débil* que estudia el amplio ambiente social y político, así como sus efectos sobre, digamos, la existencia de instituciones científicas tal y como fue ejemplificado por Merton en *Science, technology and society in seventeenth century England* (1938); también puede distinguirse un programa *fuerte* véase por ejemplo: *Leviathan and the Air-Pump* (1985) de Shapin y Schaffer; de acuerdo con dicha obra, el contenido de los resultados aceptados por la ciencia, así como el propio discurso de los términos científicos, están influenciados por factores sociales y

políticos. La mayor parte del trabajo de Kuhn da un respaldo parcial al programa débil. Por ejemplo, en relación con la crisis en la astronomía ptolemaica que precede a la revolución copernicana, Kuhn nos dice que un ingrediente es “la presión social para la reforma del calendario, una presión que plantea, como particularmente urgente, el problema de la precesión” (*The Structure* 69).

Él va a decirnos, con respecto a una crisis científica, que:

en una ciencia madura –y la astronomía ha llegado a ser tal en la antigüedad– los factores externos como los citados anteriormente son significativos, especialmente, para determinar la efectividad de la ruptura, la facilidad con la que se reconoce y en la cual, debido a su atención particular dada, primero ocurre la ruptura. (*The Structure*)

Mientras se reconoce qué tales factores pueden ser importantes, Kuhn enfatiza en que “la ruptura técnica aún podría permanecer en el centro de la crisis”. De ésta manera, mientras los factores externos pueden influir en el modo en que sucede este episodio, permanece la cuestión acerca de los factores internos que explican por qué eso podría ocurrir<sup>6</sup>.

Si la ciencia normal e incluso la crisis pueden ser explicadas sobre fundamentos puramente internos, ¿podríamos esperar, quizás, que el externalismo sea probablemente más verdadero, de acuerdo con la explicación de Kuhn de la ciencia revolucionaria? El propio Kuhn escribe:

Los científicos individuales asumen un nuevo paradigma por todo tipo de razones y, frecuentemente, por muchas a la vez. Algunas de estas razones, por ejemplo, la adoración al sol que ayudó a que Kepler fuera un copernicano, quedan fuera de la esfera aparente de la ciencia plena; otras pueden depender de idiosincrasias autobiográficas o de personalidad. Aún la nacionalidad o la reputación del innovador y sus profesores pueden, algunas veces, jugar un papel importante. (Ibíd. 152-53)

En este pasaje no debería sobrevalorarse el externalismo. Como Kuhn dice, *algunas* de las razones que tiene un individuo para adoptar un paradigma pueden quedar por fuera de la ciencia, y solo da un ejemplo,

<sup>6</sup> Para más detalles acerca del internalismo de Kuhn en relación con SCC, ver Bird (*Kuhn, naturalism*).

la adoración al Sol de Kepler. Por implicación, él cree que las 'otras' que menciona no quedan por fuera de la ciencia. Claramente, la reputación es interna a la ciencia; es cierto que las diferencias personales podrían hacer que los científicos difieran en el grado en el que se disponen a adoptar ideas radicales; o la apertura de un científico a una idea podría estar influenciada por el hecho de que se formó en un laboratorio en donde tales ideas se desarrollaron o debido a que el trabajo en esa teoría ofrece unas mejores perspectivas laborales. Sin embargo, una vez más, no es claro que estas sean consideraciones externas a la práctica de la ciencia, al menos, no de un modo que amenace las afirmaciones centrales del internalista que son importantes para Kuhn: que es el requisito de resolver problemas lo que determina irremediabilmente cuáles ideas se desarrollan y adoptan. Como dice Kuhn, inmediatamente después del pasaje citado:

probablemente la única afirmación más relevante anticipada por quienes proponen un nuevo paradigma es que pueden resolver problemas que habían llevado al paradigma antiguo a una crisis. Cuando legítimamente puede hacerse esto, con frecuencia esta afirmación es la más efectiva posible. (*The Structure* 152-53)

El modo más efectivo de anticipar un nuevo paradigma es mostrar que resuelve los problemas que dejaron al otro en una crisis. Kuhn luego va a señalar que esto no es siempre posible; en efecto, el nuevo candidato a paradigma no puede ayudar a evocar todos los problemas de la crisis. En ese caso, las predicciones novedosas, las predicciones de fenómenos que serían completamente insospechadas en el viejo paradigma pueden ser persuasivas (tales como la predicción de las fases de Venus por la teoría de Copérnico). Después Kuhn menciona el papel de las consideraciones estéticas. También, discute ampliamente las características de las revoluciones que que posteriormente hemos llamado: "la pérdida de Kuhn"<sup>7\*</sup> y la importancia de un nuevo paradigma como una base fructífera para la investigación en la solución de nuevos problemas. Al evaluar si Kuhn dio estímulo directo al estudio externalista de la ciencia, debemos confrontar el pequeño pasaje de la cita con las seis páginas que siguen, en las cuales él enfatiza detalladamente la importancia de la tradición de resolver problemas para determinar su propio desarrollo.

---

<sup>7</sup> La expresión "pérdida de Kuhn" (*Kuhn-loss*) hace referencia a la pérdida de capacidad que sufre un nuevo paradigma o una teoría con respecto a un paradigma o teoría precedente a la hora de ofrecer explicaciones en un cierto dominio de solución de problemas; es decir, un período posterior de la ciencia puede encontrarse sin poder explicar un fenómeno que en un período anterior fue explicado exitosamente. (Nota del traductor).



Sin embargo, ¿no ocurre que durante la ciencia revolucionaria no hay tradición de resolución de problemas que juegue este papel determinante? Y si la elección de una teoría es indeterminada, ¿eso no da oportunidad para que fuerzas extracientíficas determinen el resultado, como defiende Barnes (*On the 'hows'; T. S. Kuhn and*)? Creo que esta es una mala interpretación de Kuhn, basada en la idea que las revoluciones son rupturas totales y radicales con el pasado. Mientras es cierto que Kuhn pudo haber sobreestimado la diferencia entre ciencia normal y revolucionaria, también es cierto que Kuhn da gran énfasis al progreso y la continuidad a través de las revoluciones. En efecto, *The Structure of scientific revolutions* tiene un capítulo con el título "Progreso a través de las revoluciones", mientras el capítulo precedente: "La resolución de las revoluciones", describe las restricciones impuestas al nuevo paradigma por el éxito continuado de su predecesor en la resolución de problemas. Tales restricciones significan que hay una continuidad relevante en la ciencia revolucionaria<sup>8</sup>. Hay más cosas en común que diferencias entre la ciencia normal y la revolucionaria. Tanto en la ciencia normal, como en la revolucionaria, lo que conduce el progreso es, principalmente, resolver problemas científicos; durante la ciencia normal, la necesidad de resolver problemas es satisfecha por el paradigma; durante la ciencia extraordinaria la necesidad perdura, pero ahora debe satisfacerse encontrando un paradigma que lo reemplace. Lo que determina el resultado será, sobre todo, el poder de un paradigma propuesto para resolver problemas. Puede que eso no determine el resultado de modo único y sin ambigüedad -Kuhn enfatiza que hay lugar para el desacuerdo racional acerca del poder relativo para resolver problemas de un nuevo paradigma propuesto, comparado con el viejo paradigma o con un competidor-; sin embargo, el hecho de que la disputa es acerca del poder de resolver problemas científicos restringe las elecciones disponibles. Los participantes en el debate deben ser racionalmente capaces de creer que su solución preferida da más y mejores soluciones a problemas que sus competidoras. En particular, quienes apoyan un nuevo paradigma deben, en la mayoría de los casos, ser capaces de mostrar que resuelve una porción considerable de las anomalías más significativas dejadas por el viejo paradigma, mientras que preservan, a la vez, el grueso del poder de resolver problemas que tiene su predecesor.

<sup>8</sup> Kuhn vuelve sobre este tema en varios de sus escritos posteriores; por ejemplo, en su ensayo: "Objectivity, value judgment, and theory choice", en el que articula los cinco valores científicos (precisión, consistencia, amplitud de alcance, simplicidad, fecundidad). Toda su preocupación radica en rechazar las acusaciones de subjetividad al preferir una teoría a la vez que da espacio para el disenso razonable entre paradigmas.



Aunque encontrar una solución innovadora que logre esto no es fácil, la mayoría de episodios en ciencia revolucionaria darán muy pocas opciones, típicamente habrá solo una propuesta revolucionaria que rete el viejo paradigma. Dado el rango infinito de creencias que un científico pueda tener acerca de un asunto dado, todas, excepto un puñado de ellas, son excluidas inmediatamente por factores internos a la ciencia incluso durante la ciencia extraordinaria.

En efecto, parece que esto todavía da lugar a que influyan factores externos en los resultados de una revolución científica; sin embargo, no creo que Kuhn haya considerado que tales factores jugaran un papel determinante. El que se dé lugar a desacuerdos racionales no significa que haya algo más que inflencie la idea de algún científico individual, de dejar solo las opiniones de la comunidad en su conjunto. Lo que esto significa es que la resolución de una revolución será un asunto mucho más prolongado. En la ciencia normal puede haber disputas, pero por lo regular se pueden resolver usando el recurso del paradigma. Las causas del SIDA fueron disputadas al comienzo; no obstante, las técnicas estándar identificaron un virus particular como la causa; de tal manera, que esto quedó más allá de cualquier disputa racional. En tal sentido, no hay una “pérdida de Kuhn” –no se necesitan creencias, ni compromisos preexistentes a los cuales renunciar–; el éxito de la explicación viral es claro gracias a los estándares establecidos y se hacen evidentes las oportunidades para investigar (lo mismo que los beneficios científicos) ofrecidas por el nuevo descubrimiento. Por otra parte, en la ciencia revolucionaria hay “pérdida de Kuhn” que se confronta con las pretensiones de la solución de problemas; hay, por lo menos, algunos conflictos con los estándares (al menos con las creencias) existentes y debido a esto, es confuso el potencial del pretendido nuevo descubrimiento para respaldar un programa de investigación fructífero (futura solución a un problema) sobre todo cuando tenemos que abandonar una tradición existente. Cuando Barry Marshall y Robin Warren propusieron que la causa principal de la úlcera gástrica era una bacteria en vez de, como se había creído, exceso de ácido producido por factores tales como el estrés, un sub-campo completo de investigación (así como de tratamiento) estaba bajo amenaza; consiguientemente, era confuso saber en ese momento si, en términos de solución de problemas, la nueva propuesta sería más productiva que la idea establecida. No estamos comparando de igual a igual estos casos, ya que comparamos una trayectoria existente con una promesa futura. Así pues, se da plena cabida a la diferencia de opinión como también, a si la nueva visión

debería o no adoptarse. Los factores biográficos, como sostiene Kuhn, pueden jugar cierto rol a la hora de determinar cómo responden los investigadores individuales. Los científicos antiguos invertirán prestigio, experiencia, programas de investigación y laboratorios en el enfoque establecido; por su parte, los investigadores jóvenes verán que la idea más reciente les ofrece oportunidades para progresar de una manera más rápida que la que hubieran tenido de otro modo. Pero el dar cabida a las diferencias de opinión y a la influencia de consideraciones profesionales no puede persistir para siempre. Conforme pasa el tiempo, debido al poder que tiene la nueva idea para resolver problemas, cambiará de la potencialidad a la actualidad y será posible una comparación más directa entre lo viejo y lo nuevo. En el caso de Marshall y Warren, después de la resistencia inicial, la opinión de la comunidad coincidió de un modo razonablemente rápido. Aunque no había un punto definitivo al cual una u otra opinión pudiera apegarse, esto no significó que fuera razonable respaldar indefinidamente una u otra visión. Pese a que puede encontrarse científicos que continuaron creyendo, en 1920, en el éter electro-magnético, la mayoría de los matemáticos y físicos teóricos aceptaron la teoría (especial) de la relatividad de Einstein, originada en 1905, antes de la irrupción de la Primera Guerra Mundial. A su vez, las ventajas que pueden atraer hacia un nuevo campo a un joven científico ambicioso pronto se empañarían si no están a la altura de lo que prometen como vehículo para una tradición productiva en la solución de problemas; la fusión fría es un ejemplo de ello. Así pues, la diferencia entre la ciencia normal y la extraordinaria no se da entre fases en las que los factores internos o externos sean decisivos; sino que es una diferencia entre la rapidez y el modo en que los factores internos, sin ninguna ayuda, alcanzan su conclusión.

Una razón distinta para pensar que los factores externos deben ser importantes es creer que las preguntas sobre las que trabajan los científicos están determinadas frecuentemente por las necesidades materiales de la sociedad en general. Bacon, en el *Novum Organum*, imaginó una ciencia que conduce a la prosperidad económica y, en tal sentido, uno esperaría que esa ciencia se preocupara por cuestiones que estuvieran conectadas directamente con los problemas que surgen en las esferas social y económica. De este modo, podría pensarse en los esfuerzos hechos por los astrónomos para resolver el problema de la longitud. Dada la insistencia en el aislamiento de la ciencia madura de las fuentes externas de problemas, no es sorprendente que Kuhn estableciese una diferenciación explícita entre la ciencia y la tecnología:

como primera aproximación, el historiador del desarrollo socio-económico haría bien en tratar la ciencia y la tecnología como empresas radicalmente distintas, con la misma diferencia que hay entre las ciencias y las artes. (*The relations* 143)

La tecnología responde a demandas externas, pero la ciencia no lo hace. Ciertamente, esa puede ser una mera distinción por definición; sin embargo, Kuhn (Ibíd. 142-7) nos da razón para creer que no es así. Él sostiene que la ciencia y la tecnología han sido esferas de acción históricamente distintas; fue solo en el segundo tercio del siglo XIX, que la visión de Bacon comenzó a alcanzarse y el conocimiento generado por los científicos empezó a configurar una diferencia tecnológica para la sociedad, primero, a través de las materias colorantes y después, mediante los instrumentos y las técnicas eléctricas tales como la pasteurización de la cerveza, el vino y la leche. En efecto, el aislamiento de la ciencia de la tecnología no está garantizado y uno podría maravillarse si la ciencia moderna está en una posición distinta. Ciertamente, los requisitos que hacen los gobiernos a la investigación para que respondan a los retos generados externamente pueden enturbiar la distinción entre la ciencia y la tecnología al punto de erradicarla.

### **Por qué el historicismo de Kuhn requiere el internalismo**

He dicho que Kuhn fue un historicista y un internalista. Esto plantea, *prima facie*, una tensión: la línea conservadora del historicismo de Kuhn implica el relativismo; por su parte, el rechazo del externalismo se asocia con el objetivismo acerca del conocimiento científico; sin embargo, esta tensión es solo aparente. En síntesis, el externalismo conduce al relativismo (o al escepticismo), pero su contrario no es el caso: el relativismo no conduce, necesariamente, al externalismo. Es cierto que los objetivistas, aquellos que creen que la ciencia tiene un éxito razonable al descubrir hechos acerca de un mundo independiente, serán internalistas; pero de esto no se sigue que todos los internalistas deban ser objetivistas. El internalismo da cabida a objetivistas y relativistas que crean que los determinantes del cambio científico están encapsulados en la ciencia misma. Ese es el tipo de internalista que, considero, era Kuhn.

Sin duda, Kuhn tiene que pertenecer a esta clase de internalista, siempre y cuando sea correcta la línea determinista de su historicismo. Si el

externalismo fuera verdadero, aquellos factores que se originan por fuera de la ciencia serían los principales conductores del cambio científico; en consecuencia, no habría razón para suponer que habría patrones en la historia de ese cambio. En vez de eso, uno esperaría que la historia de la ciencia demuestre el mismo caos y contingencia que encontramos en otros asuntos humanos. Tomemos los dos ejemplos de la biología del siglo XIX que ya he mencionado, el rechazo de la generación espontánea hecho por Pasteur y la explicación darwiniana de la evolución a través de la selección natural: si los externalistas están en lo cierto, el origen y éxito de estas dos teorías son resultado de fuerzas socio-políticas que tienen lugar apenas se oponen naturalezas que concurren al mismo tiempo en diferentes países: el clero conservador en Francia y el liberalismo económico en Gran Bretaña (nótese que una de las pretendidas ventajas políticas de los resultados de Pasteur es que amenaza el darwinismo ateo, que muchos tomaban como requisito para asumir alguna forma de generación espontánea). Debido a que estas diferentes fuerzas sociales son producto de diferentes secuencias de eventos históricos en los dos países, es difícil ver cómo la totalidad de estas fuerzas podría conspirar para producir el ciclo ordenado que Kuhn ve en la historia de la ciencia. Para usar una analogía mecánica: Kepler fue capaz de discernir el orden elíptico del sistema solar debido a que el sistema solar es un sistema simple y aislado; si estuviera frecuentemente perturbado por grandes objetos inter-estelares que pasan a través o cerca de él, entonces Kepler no tendría órbitas elípticas por descubrir. Así mismo, una condición necesaria de la verdad de la teoría de Kuhn es que los conductores del (contenido del) desarrollo científico son locales a la ciencia, por lo que se aíslan, en gran medida, de la influencia de otros desarrollos históricos.

El externalismo y el historicismo determinista en ciencia podrían reconciliarse si los patrones en la historia de la ciencia reflejaran patrones históricos más amplios, de modo que permeen lo social y lo político, así como lo científico. Las leyes del desarrollo científico serían manifestaciones de una verdad historicista más amplia. Sin embargo, esto parece implausible por dos razones: en primer lugar, tal determinismo histórico global tiene poco crédito. Los grandes esquemas de Hegel y Comte son creídos por pocos, si no por ninguno, en la actualidad; y ni siquiera todos los marxistas aceptan un determinismo estricto similar al de, por ejemplo, la versión leninista del materialismo dialéctico. En segundo lugar, tal determinismo histórico externo debe *explicar* el modelo cíclico kuhniano, pero no se ha hecho ningún esfuerzo por mostrar cómo procedería tal explicación; ciertamente, parece implausible

que pueda haber dicha explicación. Los deterministas históricos tienden a ver la historia como si mostrara etapas a gran escala (Comte, Marx), pero no podrían explicar el modelo cíclico de Kuhn (especialmente cuando no hay tal modelo, sino muchos; cuando no es necesario que el modelo para un campo de la ciencia coincida con el modelo para otro campo). Además, los deterministas históricos tienden a creer que la historia tiene una dirección incluso una meta (Hegel). Kuhn niega que la historia de la ciencia tenga una dirección, pero, además, se niega a aceptar que podría permitirse un tipo de historia *whig*<sup>9\*</sup> de la ciencia que *The Structure of scientific revolutions* rechaza (nótese que el *whigismo* es una característica de los historiadores marxistas). Así pues, incluso si un determinismo externo explicara los modelos kuhnianos (algo que es implausible), sería antitético con las clases de externalismo promovidos por muchos practicantes de los estudios de la ciencia.

### Los propósitos filosóficos de Kuhn para la historia de la ciencia

Kuhn dice que su interés en la historia de la ciencia estuvo motivado, en gran medida, por preocupaciones filosóficas<sup>10</sup>. Para concluir, defiende que el historicismo de Kuhn es una parte importante para encontrar ciertos objetivos filosóficos. El blanco filosófico de Kuhn fue el empirismo lógico; los empiristas lógicos, vistos de un modo lo suficientemente amplio como para incluir a Popper, estuvieron preocupados por dar explicaciones normativas del cambio teórico. La historia de la ciencia puede usarse para examinar estas explicaciones normativas –bajo el supuesto que los científicos razonan normalmente como deberían razonar–. Este último supuesto es importante, pues sin él la teoría normativa sería una teoría acerca de cómo los científicos deberían cambiar sus hábitos de razonamiento con el fin de mejorarlos. La filosofía de la ciencia de Bacon puede verse en esta dirección. En términos generales, los empiristas lógicos creían que los científicos razonan correctamente; por consiguiente, sus teorías se dirigieron a articular cómo los científicos razonan de hecho. Popper no solo creía que los científicos deberían rechazar las teorías que fueran falsadas, sino que también planteaba que ellos, de hecho, rechazan tales teorías. Así pues, la perspectiva de Popper enfrentaría en sí misma la falsación si la

<sup>9</sup> La historia o la historiografía *whig* es el enfoque historiográfico que presenta el pasado como un progreso inevitable hacia mayor libertad e ilustración, culminando en formas modernas de democracia liberal y monarquía constitucional. En el ámbito de la historia de la ciencia, la historiografía *whig* se centra en la cadena de teorías y experimentos exitosos que conducen a la ciencia del día, ignorando las teorías que fracasan y los callejones sin salida. (Nota del traductor).

<sup>10</sup> Por ejemplo, en el prefacio a *The Structure*.

historia de la ciencia muestra que los científicos se apoyan regularmente en teorías que enfrentan evidencia aparentemente contradictoria. Ciertamente, esto es lo que Kuhn pretende mostrar con el componente conservador de su historicismo, según el cual la ciencia normal está gobernada por una tradición de resolución de problemas. Como hemos discutido, en la perspectiva de Kuhn, los científicos no abandonan una tradición cuando enfrentan una anomalía; más bien, una anomalía será, a menudo, solo otro rompecabezas que resolver. Si un científico aborda ese rompecabezas, pero falla al resolverlo, ese fracaso se atribuye a las limitaciones del científico, no de la tradición. De este modo, la misma existencia de la ciencia normal es un reto mayor al falsacionismo de Popper.

Las cosas son un tanto diferentes con respecto a la línea inductivista más central del empirismo lógico. Aquí el blanco de Kuhn es la concepción de ciencia como una acumulación de creencias verdaderas adquirida por la aplicación repetida del método científico (por ejemplo: alguna forma de lógica inductiva). Tal perspectiva es consistente con la existencia de la ciencia normal. Es la ciencia revolucionaria la que crea el problema para el empirismo lógico, pues se trata de episodios en los que se rechazan las creencias bien establecidas. Sin embargo, ya que tales episodios son, en la terminología de Kuhn, 'extraordinarios', hay cabida al debate con respecto a su valor de evidencia contra la imagen del empirista lógico, toda vez que su relativa rareza permitirá al empirista lógico relacionarlos con excepciones ocasionales, en algunos casos episodios patológicos (o correcciones a la ciencia patológica) o características de la ciencia inmadura, entre otros. Aquí es donde se vuelve relevante la línea determinista del historicismo de Kuhn, ya que Kuhn puede mostrar que el cambio científico tiene una estructura, la estructura cíclica de ciencia normal-crisis-revolución-ciencia normal; entonces, los episodios no acumulativos, las revoluciones, no pueden ser descartadas, son parte inevitable del proceso científico.

El propio Kuhn se dirige hacia una revolución mayor en la filosofía de la ciencia. Rechaza los supuestos comunes de los empiristas lógicos que el objetivo de la ciencia es la verdad y que la racionalidad científica consiste en aplicar algún tipo de lógica a las relaciones entre una teoría y afirmaciones directas relacionadas con la experiencia de los científicos. El paradigma alternativo de Kuhn tiene la intención de ser uno en el que el objetivo de la ciencia sea resolver problemas y la racionalidad científica consista en ajustar la solución de problemas propuesta con las



soluciones de problemas ejemplares<sup>11</sup>. La perspectiva de Kuhn parece que no necesita ser muy radical cuando consideramos que gran parte de la cognición humana ocurre a través del reconocimiento del modelo (piénsese en el reconocimiento de rostros); sin embargo, en su propio contexto histórico, la propuesta era radical y fue percibida como más extrema de lo que debería haber sido. Debido a esto fue tomada como una forma de irracionalismo acerca de la ciencia. Una vez se percibió desde esta luz, no es sorprendente que los detractores y admiradores de Kuhn, lo tomaran como aquél que articula una visión de ciencia en la que los científicos y sus ideas, sin las restricciones de la racionalidad, están sujetos a fuerzas sociales.

He argumentado que la perspectiva kuhniana de la ciencia *no* es ésta. Kuhn desea defender su nueva comprensión de la racionalidad científica, justamente al señalar el modelo que percibe en la historia de la ciencia; dicho modelo se explica mejor a partir de esta nueva concepción de la racionalidad, en comparación con la defensa que hizo la vieja concepción. De este modo, es la línea conservadora de su historicismo la que respalda esa nueva concepción de la racionalidad; a su vez, la racionalidad de la ciencia, independiente de cualquier concepción, requiere el determinismo. Un componente significativamente externalista de la ciencia quebranta la línea determinista del historicismo de Kuhn y, en consecuencia, es antitético con los propósitos filosóficos de Kuhn para la historia de la ciencia. Espero que pensar a Kuhn como un historicista con respecto a la historia de la ciencia nos permitirá repensar su comprensión acerca de ese tema y alcanzar una mejor perspectiva de sus propósitos filosóficos en torno a la naturaleza y la racionalidad de la ciencia.

---

<sup>11</sup> Para más detalles, ver Bird (*Naturalizing Kuhn*). Esta visión es polémica, pues Kuhn no se proclamó a sí mismo, como quien buscaba revisar nuestra noción de racionalidad; esto se debe a que la misma noción de racionalidad es cercana a la idea de seguir las reglas de la razón. Kuhn mostró que la ciencia otorgó reglas pero en vez de eso, empleó el razonamiento en analogía con los ejemplares. Esto último, que se esforzó por enfatizar, no es en modo alguno irracional.



## REFERENCIAS BIBLIOGRÁFICAS

Barnes, B. "On the 'hows' and 'whys' of cultural change (response to Woolgar)". *Social Studies of Science*. Nov. 1981: 481-498. Print.

---. *T. S. Kuhn and social science*. London: Macmillan, 1982. Print.

Bird, A. "Naturalizing Kuhn". *Proceedings of the Aristotelian Society*. Jun. 2005: 109-127. Print.

---. "Kuhn, naturalism, and the social study of science". Kindi, V. and T. Arabatzis (eds.). *Kuhn's The Structure of scientific revolutions revisited*. Abingdon: Routledge, 2012. Print.

Farley, J. "The social, political, and religious background to the work of Louis Pasteur". *Annual Reviews in Microbiology*. Oct. 1978: 143-154. Print.

Farley, J. and G. Geison. "Science, politics and spontaneous generation in nineteenth-century France: The Pasteur-Pouchet debate". *Bulletin of the History of Medicine*. Jul. 1974: 161-98. Print.

Kuhn, T. S. *The Structure of scientific revolutions*. Chicago: University of Chicago Press, 1962. Print.

---. "The history of science". *International Encyclopedia of the Social Sciences*. New York: Crowell Collier and Macmillan, 1968. Print.

---. "The relations between history and the history of science". *Daedalus*. Jan. 1971: 271-304. Print.

---. *The Essential tension*. Chicago: University of Chicago Press, 1977. Print.

Mannheim, K. *Essays on sociology and social psychology*. London: Routledge and Kegan Paul, 1953. Print.

Popper, K. *The Poverty of historicism*. London: Routledge and Kegan Paul, 1957. Print.

Reynolds, A. "What is historicism?" *International Studies in the Philosophy of Science*. Oct. 1999: 275-287.

Shapin, S. and S. Schaffer. *Leviathan and the Air-Pump: Hobbes, Boyle, and the experimental life*. Princeton: Princeton University Press, 1985. Print.

Young, R. "Malthus and the evolutionists: The common context of biological and social theory". *Past and Present*. May. 1960: 109-145. Print.



# LOS PROBLEMAS DEL SENTIDO-REFERENCIA EN LA SEMÁNTICA FILOSÓFICA CLÁSICA: DOS GRANDES CONCEPCIONES EN LAS TEORÍAS REFERENCIALISTAS DEL SIGNIFICADO

*THE PROBLEMS OF SENSE-REFERENCE IN THE CLASSIC PHILOSOPHICAL SEMANTICS:  
TWO GREAT CONCEPTIONS IN REFERENCIALIST THEORIES OF MEANING*

JUAN MANUEL JARAMILLO URIBE  
Universidad del Valle, Colombia. jaramillo.juanmanuel@gmail.com

RECIBIDO EL 12 DE NOVIEMBRE DE 2012 Y APROBADO EL 14 DE DICIEMBRE DE 2012

## RESUMEN    ABSTRACT

La introducción en la filosofía del neologismo “ontoepistemosemántica” por parte de C. U. Moulines busca dar cuenta de los estrechos vínculos existentes entre las reflexiones ontológicas, epistemológicas y semánticas. El primer filósofo en percatarse de esta imbricación fue Frege, dando una especial preeminencia metodológica a la última. En este artículo, se pretenden mostrar y contraponer dos grandes teorizaciones de lo que, en sus comienzos, fue la reflexión ontoepistemosemántica sobre el significado-referencia de los términos singulares: la teoría de la *referencia indirecta* de Frege y la teoría de la *referencia directa* de Russell, al menos para lo que este último denomina “nombres lógicamente propios”. El propósito es mostrar que estas pioneras reflexiones llevadas a cabo a finales del siglo XIX y comienzos del XX, hoy en día, son un referente obligado en cualquier investigación sobre temas óntico-ontológicos, epistemológicos y semánticos.

The introduction of the neologism: ‘ontoepistemosemantics’, by C. U. Moulines seeks to account for the close links among the ontological, epistemological and semantic reflections. The first philosopher in noting this imbrication was Frege, who gave a special methodological prominence to the latter. In this article, I intend to show and contrast two major theorizations of what, in its beginning, was the ‘ontoepistemosemantic’ reflection about the meaning-reference of singular terms: the theory of *indirect reference* of Frege and the theory of *direct reference* of Russell, at least for what he calls “logically proper names”. The purpose is to show that these reflections, undertaken in the late nineteenth and the early twentieth centuries, nowadays are an essential referent in any investigation about ontic-ontological, epistemological and semantic issues.

## PALABRAS CLAVE    KEY WORDS

Frege, ontoepistemosemántica, sentido (*Sinn*), referencia (*Bedeutung*), Russell, teoría de las descripciones.

Frege, ontoepistemosemantics, sense (*Sinn*), denotation (*Bedeutung*), Russell, theory of descriptions.

---

\* Prof. Jubilado U. del Valle.

“Si (como el griego afirma en el Crátilo) el nombre es arquetipo de la cosa en las letras de *rosa* está la rosa y todo el Nilo en la palabra *Nilo*”  
(J. L. Borges, *El Golem*).

## Introducción

El problema *óntico-ontológico* de la realidad y, ligado a él, el problema *epistemológico* de su conocimiento y de modo especial, el problema *epistemológico-alético* de su conocimiento *qua* conocimiento verdadero o aproximadamente verdadero, tradicionalmente han sido tratados en forma relativamente independiente por la metafísica y la teoría del conocimiento. Sin embargo, desde finales del siglo XIX y comienzos del XX y sobre todo, gracias a los trabajos de Frege<sup>1</sup>, las cuestiones puramente metafísicas y epistemológicas se han visto desplazadas hacia cuestiones relativas al análisis lógico-semántico. Frege no sólo es el fundador de la *filosofía analítica*, sino también de la lógica moderna, en el período que va desde su tiempo hasta nuestros días y con el “giro lingüístico” que le introdujo a la filosofía, desarrolló y abrió espacios de interés para el análisis del lenguaje ordinario y para el análisis del lenguaje de las teorías científicas como es el caso de los términos o predicados que se usan en ellas. A partir de Frege — como lo destaca Moulines —,

la tendencia general [es] a reformular cuestiones básicas de ontología y teoría del conocimiento en problemas ontosemánticos o epistemo-semánticos, con el fin de hacerlas más claras y precisas. (Moulines, *Pluralidad y recursión* 130)

Es así como Moulines, acuñó el término “*ontoespistemosemántica*” para destacar las estrechas relaciones entre la *metafísica*, la *epistemología* y la *semántica*, pues considera que para responder a la pregunta acerca de “lo que hay” (*problema óntico*) tenemos que “conocer lo que hay” (*problema epistemológico*) y para conocer lo que hay y conocerlo

<sup>1</sup> Para M. Dummett la producción intelectual de Frege puede dividirse en cinco períodos bien delimitados. En el primero, se destaca la publicación de *Begriffsschriften* (*Conceptografía*) (1879); en el segundo, *Die Grundalgen der Arithmetik* (*Los fundamentos de la aritmética*) (1884); en el tercero, que comienza en 1891 y llega hasta la publicación del segundo volumen de *Die Grundgesetze der Arithmetik* (*Los fundamentos de la aritmética*) desarrolla sus reflexiones ontoepistemosemánticas en artículos como “Funktion und Begriff” (“Función y concepto”), “Über Sinn und Bedeutung” (“Sobre sentido y referencia”) y “Über Begriff und Gegenstand” (“Sobre concepto y objeto”). Así como *Begriffsschriften* lo convierten en el fundador de la lógica moderna, estos artículos lo convierten en el fundador de lo que algunos autores han dado en llamar la “lógica filosófica”, la “semántica lógica” o la “filosofía del lenguaje”.

verdaderamente, así sea en forma aproximada (*problema epistemológico-alético*), tenemos que disponer de antemano de esquemas conceptuales que nos permitan saber qué “significa lo que hay” (*problema semántico*). Más aún, si hacemos eco del aforismo wittgensteniano de que “en un gran número de casos el significado de una expresión es su uso en el lenguaje”, diríamos que también presupone la pragmática y en consecuencia hablaríamos de “*onto-epistemo-semántica-pragmática*”. Pero como dice el mismo Moulines, “aparte de acuñar este trabalenguas sería una ofensa para nuestro interlocutor (basta grave es ya el término “*ontoepistemosemántica*”)” (Moulines, *Ontoepistemosemántica en 20*).

En este artículo me concentraré en la presentación de dos de las grandes tendencias de lo que, en sus orígenes, fue la reflexión ontoepistemosemántica en relación con los problemas del significado-referencia de los términos singulares<sup>2</sup>:

(i) *La teoría de la referencia indirecta* formulada por Frege y que años más tarde John Searle (n. 1932) y otros adoptan con algunas modificaciones y;

(ii) *la teoría de la referencia directa* de Bertrand Russell (1872-1970) y que posteriormente adoptarán, también con modificaciones, Saul Kripke (n. 1940) y Hillary Putnam (n. 1926), principalmente<sup>3</sup>.

En estos últimos, se apela a la noción de causalidad, mas no para dar cuenta de la relación epistémica entre las ideas (representaciones) y los objetos, sino entre las palabras y los objetos, destacándose la “teoría causal de la referencia” en la que sus defensores asumen un compromiso ontoepistemosemántico con el realismo y, en particular, con el llamado “realismo científico”, defendiendo con su noción modal de “designador rígido” un tipo de invariancia referencial tanto de los nombres propios del lenguaje cotidiano, como de los términos de clase del lenguaje científico, cuya significación resultará decisiva en el debate realismo-antirrealismo.

<sup>2</sup> Los términos singulares como categoría general incluyen una serie de subcategorías: los nombres propios, los deícticos (“yo”, “tú”, “ese”, “allí”, “aquí”, entre otros) y las descripciones, de modo especial las descripciones definidas de la forma “el [la] tal y tal” (“*the so-and-so*”) como: “El actual rey de Francia”, “el norte magnético de la Tierra”, entre otras.

<sup>3</sup> No se desconoce que fue John Stuart Mill, quien en su *A system of logic* (1843), planteó por primera vez y en forma relativamente sistemática una teoría de la referencia directa, en la que sostiene que los nombres propios tienen denotación, pero no connotación.

## Dos grandes concepciones acerca de la referencia

### Frege y la teoría de la referencia indirecta<sup>4</sup>

Los antecedentes de las investigaciones ontoepistemosemánticas tienen que ver con la fundación por parte de Frege de la lógica moderna y con ella de su propuesta logicista (en la que tuvo el acompañamiento de Russell), de reducir la matemática (en particular la aritmética) a la lógica<sup>5</sup> y por esa vía, superar la confusión e imprecisión en nociones claves relacionadas con el análisis semántico del lenguaje aritmético y por extensión del lenguaje en general como era el caso de la noción lógico-matemática de “igualdad” o “identidad” que desde los griegos había sido objeto de numerosos y en muchos casos, infructuosos análisis.

En el escrito “Über sinn und Bedeutung” (“Sobre sentido y referencia”) publicado en 1892, aborda desde el comienzo el problema de la igualdad y se pregunta si se trata de una relación, de una relación entre objetos o entre nombres y signos de objetos, inclinándose por esta última como ya lo había avizorado en su libro *Begriffsschrift (Conceptografía)* de 1879. Las razones para esta decisión son las siguientes:  $a = a$  y  $a = b$  son enunciados de distinto valor cognitivo (epistémico), pues mientras la verdad del primero la conocemos *a priori* (Kant la consideraría una verdad analítica), la segunda, la conocemos *a posteriori* (Kant la consideraría una verdad sintética). La igualdad  $a = a$  es una verdad tautológica, si se quiere, trivial y sólo expresa la relación de una cosa consigo misma, mientras que la igualdad  $a = b$  no lo es y sólo resulta verdadera si los signos “*a*” y “*b*”, no obstante, ser distintos, se refieren a *lo mismo*, de tal modo, que la relación de igualdad “ya no sería concerniente a la cosa misma, sino a nuestro modo de designación” (Frege, *Sobre sentido* 85).

Pero lo que la teoría semántica de Frege busca destacar no es la distinción entre enunciados conocidos *a priori* y enunciados conocidos *a posteriori*, ni la distinción entre enunciados *analíticos* y enunciados *sintéticos*, sino el hecho de que el valor cognoscitivo de la igualdad tiene que ver con las

---

<sup>4</sup> En este trabajo seguiremos la versión castellana de los textos de Frege “Los fundamentos de la geometría”, “Estudios sobre semántica” y “Sobre los fundamentos de la geometría”, que aparecen compilados en el libro *Escritos filosóficos* de Jesús Mosterín.

<sup>5</sup> Este programa lo anuncia Frege en el prólogo a *Begriffsschrift* (1879), donde se pregunta si los teoremas de la aritmética son deducibles a partir únicamente de las leyes lógicas o si hay que introducir hechos empíricos para su demostración. Su formalización de la lógica llevada a cabo en esta obra, le permite afirmar que con los solos medios del cálculo lógico es posible deducir formalmente los teoremas de la aritmética.

referencias de los signos, de modo que el enunciado  $a = b$  es verdadero *syss.* los signos  $a$  y  $b$  tienen la misma referencia, no obstante, tener sentidos distintos o como dice Frege, ser “modos distintos de presentación de lo referido”. De ésta manera, las nociones de *sentido-referencia* pasan a convertirse en conceptos básicos de la semántica fregeana, al igual que lo serán las categorías de *objeto-función* para su ontología, como lo veremos más adelante, si bien ambos pares de conceptos, aunque distintos, no se pueden separar radicalmente, pues, como vimos, la semántica, la ontología y la epistemología no son disciplinas aisladas. Así, si profiero la oración: “el autor de *Conceptografía* nació en Wismar” en un ciclo de conferencias sobre Gottlob Frege, la referencia de “el autor de *Conceptografía*” es un *objeto* en el sentido fregeano y la del predicado “...nació en Wismar” una *función* que, para Frege, es una *función* de objetos en valores de verdad, *i.e.*, una función que asigna como referente *la verdad* a Frege y *la falsedad* a Russell, Kripke, Putnam y, en general, cualquier otro individuo que no sea Frege.

Como ya lo destacamos, una de las tesis centrales de la ontosemántica fregeana es que sólo podemos tener *conocimiento de la referencia* a través del *conocimiento de su sentido*, *i.e.*, en forma indirecta y sólo es pertinente afirmar el enunciado de identidad  $a = b$  si “ $a$ ” y “ $b$ ”, no obstante tener sentidos distintos, *i.e.*, valores cognoscitivos distintos, se refieren a *uno* y el *mismo* objeto. Esta diferencia de sentido hace que  $a = b$  no sea un enunciado trivial, como sí lo es el enunciado tautológico  $a = a$ .

El que un mismo objeto pueda ser identificado mediante dos conjuntos distintos de características (sentidos), no contradice el *principio de indiscernibilidad (indistinguibilidad) de los idénticos* o de *sustitutibilidad* de Leibniz y no debe conducir a pensar que en realidad se trata de dos objetos y no de *uno*, y el *mismo* objeto. Sucede lo contrario. Como dice García-Carpintero:

una manifestación de *objetividad* de una entidad es que pueda ser identificada a través de características distintas a aquellas a que se recurre inicialmente para pensar en ella, o designarla. (García-Carpintero 211)

$2 + 2 = 4$  y  $2 \times 2 = 4$ , expresan sentidos distintos (su contenido proposicional es distinto), pero se puede sustituir  $2 + 2$  por  $2 \times 2$  *salva veritate*, porque ambos signos tienen el mismo referente: 4. Lo importante es entender que la noción de “sentido” (“connotación” la llama J. S. Mill) es diferente



de la noción de referencia (“denotación” la llama Mill), pero como lo destaca Dummett, uno de los mejores conocedores y analistas de la obra de Frege, “la noción de sentido, tal como Frege la entendió, está relacionada con nuestro conocimiento de la referencia” (Dummett 589) y en consecuencia, constituye un componente epistémico importante para el conocimiento de los objetos, que resultará fundamental en la discusión realismo-antirrealismo epistémicos, como es el caso de lo que Davidson llamó el “tercer dogma del empirismo”: el dualismo *esquema conceptual-contenido empírico*.

Si se acepta la tesis referencialista de Mill, Russell y otros, un nombre propio como “Venus” simplemente refiere a Venus, pero no nos proporciona ninguna información respecto de él, salvo la de llamarse “Venus”. Pero las descripciones definidas “el lucero de la mañana” o “lucero de la tarde” (modos de presentación del objeto), no sólo nombran el mismo objeto, Venus, sino que además, nos dan cierta información como la de ser el último astro cuyo brillo visible desaparece del cielo visible por la mañana (lucero de la mañana) o la de ser el primer astro cuya luz visible aparece en el cielo visible por la tarde (lucero de la tarde). Esta diferencia de información no se resuelve apelando sólo al *objeto*, sino considerando los *modos de presentación* del objeto, *i.e.*, aquella parte del significado de una expresión que es distinta de la referencia: el sentido. Esto es algo que las teorías *referencialistas* no consideran, pues en ellas el sentido y la referencia son equivalentes. Para Frege, por el contrario,

es natural considerar entonces que un signo (nombre, unión de palabras, signo escrito) además de lo designado, que podría llamarse la referencia del signo, va unido a lo que quisiera denominar el sentido del signo, en el cual se halla contenido el modo de darse [...] La referencia de “lucero vespertino” y de “lucero matutino” es la misma, pero el sentido no sería el mismo. (Frege, *Escritos filosóficos* 173)

Y más adelante expresa:

Un nombre propio (palabra, signo, fila de signos, expresión) expresa su sentido, se refiere a su referencia o la designa. Con un signo expresamos su sentido y designamos su referencia. (Ibíd. 177)

El sentido (*Sinn*) de un nombre propio es, para Frege, idéntico al sentido de alguna descripción definida que asociemos con él, pero si bien cada

descripción definida determina unívocamente *un* referente, a cada referente le pueden corresponder *n*-descripciones definidas distintas, *i.e.*, *n*-sentidos distintos. La relación referencial  $p \subseteq S \times B$  que asocia cada sentido  $s$  con un solo referente  $b$ , donde “ $s$ ” es un elemento del conjunto sentido  $S$  ( $s \in S$ ) y “ $b$ ” un único elemento del conjunto  $B$  ( $b \in B$ ) es, por tanto, una función, pero una función no inyectiva, pues a un sólo y mismo elemento “ $b$ ” del recorrido le pueden corresponder dos o más elementos “ $s$ ” distintos del dominio, como en el caso mencionado por Frege de “el lucero matutino” y el “lucero vespertino” respecto de Venus<sup>6</sup>.

Reconocer como lo hace Frege que es a través del sentido como un signo determina su referencia es importante, pues es esto lo que distingue su *teoría de la referencia indirecta* (mediada por el sentido), de las *teorías de la referencia directa*, en las que el papel lógico de las expresiones lingüísticas (*v. gr.* los nombres propios) es puramente referencial, *i.e.*, donde el significado de esas expresiones y su acción, el significar, se agotan en el referir (estar en lugar de otra cosa) de suerte que el nombre funciona simplemente como una “etiqueta” o “marca” del objeto referido y su significado (sentido) consiste exclusivamente en el objeto referido.

Uno de los aportes importantes de Frege es el de haber propuesto un mecanismo para explicar cómo un nombre propio refiere a un objeto, algo que los defensores de las *teorías de la referencia directa* (*referencialistas*) no logran explicar, quizás, porque intuitivamente su teoría les parece muy obvia. La relación de referencia indirecta entre el signo y su referente mediada por el sentido, formalmente, la podemos representar como una función producto o composición, pues si la función  $\tau$  hace corresponder cada expresión  $e$  del lenguaje  $L$  ( $e \in L$ ) con un sentido  $s$  del conjunto  $S$  ( $\tau: L \rightarrow S$ ) y la función  $\rho$  cada elemento  $s$  del conjunto  $S$  con un elemento  $b$  del conjunto  $B$  ( $\rho: S \rightarrow B$ ), entonces la relación referencial indirecta  $\sigma$  es una función producto o composición de las funciones “ $\rho$ ” y “ $\tau$ ”, así:  $\sigma = \rho \circ \tau$ .

En “Funktion und Begriff” (“Función y concepto”) de 1982 y “Über Begriff und Gegenstand” (“Sobre concepto y objeto”) de 1892, Frege se ocupa del análisis de otro tipo de expresiones distintas de los nombres propios: los conceptos (*términos generales*). En ellos destaca la naturaleza funcional (no objetual) de los conceptos en el sentido particular que él da en su

<sup>6</sup> En el lenguaje natural encontramos signos que carecen de referencia, pero para Frege en un lenguaje “lógicamente perfecto” deben ser excluidos.

ontología a las nociones de *objeto* y *función* realizando una contrastación con los nombres propios (nombres de objeto como también los llama) al considerar que éstos resultan inadecuados para ser predicados gramaticales. A diferencia de los nombres propios, expresiones saturadas o completas, *i.e.*, *objetos*, los conceptos no son entidades saturadas o completas, sino entidades incompletas o insaturadas, *i.e.*, *funciones*, aunque – como lo advierte el mismo Frege – funciones de una clase especial: funciones de *un* argumento (monarias). Así, un enunciado afirmativo como “César conquistó las Galias” puede descomponerse en “César”, que es la parte saturada (*objeto*) y el predicado “conquistó las Galias”, que es la parte no-saturada (*función*). Como función lleva consigo un lugar vacío (función de un argumento). Se trata, por tanto, de una función  $F$  de objetos  $O$  en valores veritativos  $\{V, F\}$ , *i.e.*,  $F : O \rightarrow \{V, F\}$ . Así, la función  $F$  aplica *la falsedad* a Carlos Magno, Bolívar y San Martín, y *la verdad* a César. Confirmar el enunciado como verdadero es establecer que el objeto o entidad que se toma como argumento de la función (predicado-concepto) cae bajo el concepto. En suma, los conceptos son *funciones* que tienen como recorrido *valores de verdad* y como dominio *objetos*.

Para diferenciar los *objetos* de los *conceptos* Frege propone una regla sencilla: “[...]en singular, el artículo determinado siempre indica un objeto, mientras que el indeterminado acompaña un término conceptual” (Frege, *Escritos filosóficos* 210). Si decimos: “el caballo es un animal cuadrúpedo”, la expresión “el caballo” es un nombre propio y como tal designa (refiere) una entidad saturada (*objeto*), mientras que la expresión “es un animal cuadrúpedo” es una entidad insaturada (*función*) y como predicado que es, refiere a un concepto. Al respecto, escribe Frege:

El concepto es la referencia de un predicado, pero el objeto es lo que nunca puede ser la referencia de un predicado, aunque puede ser la referencia de un sujeto. (Ibíd. 213)

Pero Frege, no se limita al estudio del sentido-referencia de los términos singulares y de los términos generales (conceptos y/o predicados), sino que extiende sus análisis a los enunciados<sup>7</sup> reconociendo, en virtud del principio de composicionalidad que el sentido-referencia de los enunciados como expresiones lingüísticas complejas que son, depende

<sup>7</sup> Este análisis del significado (sentido-referencia) de los enunciados lo desarrolla en “Ausführungen über Sinn und Bedeutung” (“Consideraciones sobre sentido y referencia”) del período 1982-1895 el cual permaneció inédito hasta 1969.

del significado de sus expresiones simples componentes. Sin embargo, en el análisis de los enunciados centra su atención en los nombres propios para mostrar que, si en un enunciado sustituimos un nombre propio por otro con la misma referencia, así tenga diferente sentido, la referencia del enunciado, *i.e.*, su valor de verdad, permanece inmodificado como lo explica correctamente Leibniz con su principio de sustituibilidad *salva veritate* y del que Frege se pregunta:

¿qué otra cosa, sino el valor veritativo, podría encontrarse que pertenezca con toda generalidad a cada enunciado en el que interese la referencia de las partes componentes, y que permanezca inmodificado en una sustitución del tipo mencionado? (Frege, *Escritos filosóficos* 181)

Así como la referencia (*Bedeutung*) del enunciado es su *valor de verdad* (*Wahrheitswert*), vale decir, la circunstancia de que sea verdadero o de que sea falso, el *sentido* (*Sinn*) del enunciado es, para Frege, el *pensamiento* (*Gedanke*)<sup>8</sup>. Para la identificación del pensamiento (que se expresa mediante oraciones) tenemos que identificar las partes del enunciado, así como sus categorías, *i.e.*, el sentido o modo de presentación correspondiente a los términos singulares (nombres propios) y el sentido o modo de presentación correspondiente a los términos generales (conceptos) que son representaciones generales bajo las que (al menos lógicamente) podría caer una entidad concreta (un objeto). Pero como el aspecto fundamental del significado de los enunciados son sus condiciones de verdad, entonces, hay que establecer las referencias de sus componentes que contribuyen a establecer su valor de verdad. Este se ve más claro en el caso de la ciencia, pues si consideramos que su objetivo es la búsqueda de la verdad, es esta búsqueda la que, según Frege, “nos incita a “avanzar del sentido a la referencia” (Ibíd. 180).

Ahora bien, como la referencia de un enunciado es su valor veritativo, entonces todos los enunciados verdaderos tienen la misma referencia (*la verdad*) y lo mismo vale para los enunciados falsos (*la falsedad*). Así, “ $2 + 2 = 4$ ” y “García Márquez es colombiano” denotan el mismo objeto: la Verdad. Igualmente, “el Sol gira alrededor de la Tierra” y “Borges es

<sup>8</sup> Por pensamiento dice Frege: “no entiendo la actividad subjetiva del pensar (imagen o representación), sino su contenido objetivo, que es apto para ser propiedad común de muchos” (*Escritos filosóficos* 178 n5). En otras palabras, aunque la primera noción de sentido que propone Frege está ligada a la de “valor cognitivo” a pesar de su propósito de separar la lógica de la psicología, sin embargo, precisa que el sentido es algo público, objetivo, común para todos los hablantes del lenguaje.

el autor de *Cien años de soledad*” denotan el mismo objeto: la Falsedad. Como es fácil advertirlo, esta forma de concebir la referencia de los enunciados resulta bastante problemática como Alonso Church y Kurt Gödel supieron advertirlo, pues todos los enunciados verdaderos y todos los enunciados falsos tendrían una *misma* referencia, así se trate de enunciados con contenidos distintos donde las condiciones para su verdad, *i.e.*, aquello de cuyo darse o no darse depende su verdad, serían distintos. Asimismo,, afirmar que los enunciados no referenciales (como los de ficción o los de la mitología) carecen de valor de verdad resulta problemático, salvo que nos movamos en ámbitos como el de la literatura donde la pregunta por la verdad y no por el sentido, nos llevaría a abandonar el goce estético que le es inherente. Cabe acotar que Frege no examina los enunciados existenciales negativos como “el rey de Francia no existe”, que justamente son verdaderos *syss*. el sujeto gramatical (nombre propio) carece de referencia y que refutarían su tesis de que cuando el constituyente de un enunciado carece de referencia, el enunciado no tiene valor veritativo (Moro Simpson 218).

Finalmente, hay que señalar que aunque en Frege cada signo expresa un sentido y cada sentido determina un referente, sin embargo, existen casos anómalos en los que el signo refiere su sentido, como es el caso de los contextos indirectos u oblicuos, *i.e.*, aquellos en los que se expresan actitudes proposicionales (deseos, creencias, opiniones, entre otros) o que incluyen modalidades<sup>9</sup>.

Si decimos –como lo formula Russell–, que “George IV quiso saber [actitud proposicional] si Scott era *el autor de Waverley*”, la referencia de “*el autor de Waverley*” es indirecta (no directa) y, por tanto, no denota a Scott, sino al sentido expresado normalmente por “*el autor de Waverley*”. Si fuese directa, entonces podríamos sustituir “*el autor de Waverley*” por Scott, y el enunciado “George IV quiso saber...” quedaría “George IV quiso saber si Scott era Scott”, algo que, a todas luces es falso, pues aunque George IV, sin ser un genio, sabía que Scott era Scott, desconocía si había escrito *Waverley* o como irónicamente dice Russell, “difícilmente podría atribuirse al primer caballero de Europa un interés por el principio de identidad” (Russell, *Obras completas* 38). En consecuencia, cuando se trata de contextos indirectos u oblicuos, no es lícito sustituir la descripción definida por un individuo (objeto) que, se supone, es su referente como es el caso de Scott en el enunciado “George IV...”.

---

<sup>9</sup> Este ejemplo es tomado de Valdivia (68).

Todo lo anteriormente expuesto respecto de la ontoepistemosemántica de Frege lo podemos resumir en el siguiente cuadro:

Expresión	Sentido ( <i>Sinn</i> )	Referencia ( <i>Bedeutung</i> )
Término singular (nombre propio).	Significado del término singular (descripción definida).	Objeto (entidad saturada).
Predicado y/o concepto.	Significado de la expresión predicativa.	? El concepto que es el referente lo representamos por una interrogación, pues el concepto es una entidad no saturada, i.e., una función con un lugar vacío que puede ser llenado con un nombre propio, pero que, como función, no posee un referente específico, sino indeterminado.
Enunciado.	Pensamiento ( <i>Gedanke</i> ).	Valor de verdad (V o F).

### Russell y la teoría de la referencia directa: la teoría de las descripciones

En contra de la posición defendida por Frege y haciendo eco de la propuesta milliana expuesta en su libro *A system of logic* (1867, Libro I, Cap. II), Russell defendió en *Principles of Mathematics* (1903), una teoría de la referencia directa según la cual los términos singulares (los nombres propios y quizás los indécicos) son puramente denotativos (referenciales), i.e., términos en los que la relación de significar consiste exclusivamente en denotar (referir), de tal manera, que su significado lo constituye exclusivamente el individuo denotado (referido) y, por tanto, entre el nombre y el objeto denotado no existe ninguna instancia intermedia, pues el primero, refiere directamente al segundo. En otras palabras, Russell no acepta que la distinción fregeana sentido-referencia tenga aplicación en el caso de los nombres lógicamente propios. Como veremos, sólo la tiene en el caso de de los nombres ordinarios y de las descripciones.

Para esta teoría de la referencia directa (en adelante referencialista), los nombres lógicamente propios son el paradigma de este tipo de



expresiones en las que su significado se agota en la entidad denotada. Si esto es así, entonces, cómo identificar el objeto denotado y distinguirlo de otros objetos, algo que Frege resolvió apelando al sentido como una instancia intermedia entre el signo y el referente del signo y que, fundamentalmente, corresponde a las distintas descripciones definidas de la forma “el tal y tal”, pues con ellas identificamos un objeto al proporcionar, a través de ellas, información acerca de él, como en el caso del nombre propio “Aristóteles”, con las descripciones definidas: “el discípulo de Platón”, “el estagirita”, “el maestro de Alejandro Magno”, entre otras. Más aún, si se apelara únicamente al objeto denotado sin tener en cuenta el sentido como algo diferente de dicho objeto, surge el problema que de un mismo objeto se podrían dar varias descripciones distintas, como en el caso mencionado por Frege de “el lucero matutino” y “el lucero vespertino” y ellas, aunque con diferente sentido se refieren a uno y mismo objeto: Venus. Si esto es así, las descripciones serían un contraejemplo para la teoría referencialista y las expresiones lingüísticas significativas como los nombres propios no serían puramente denotativos y su significado no se agotaría sólo en el objeto denotado. Sin embargo, en estos casos, Russell apela un criterio epistémico para la identificación de los objetos, *i.e.*, a su *conocimiento directo* (*knowledge by acquaintance*), distinto del criterio semántico aducido por Frege.

Russell, unos meses antes de publicar sus *Principles of mathematics* (1903), había aceptado que si bien en el caso de los nombres propios la distinción sentido-referencia no aplica, sin embargo, consideraba que en el caso de las descripciones definidas tal distinción resulta importante. Esta nueva idea la desarrolla en su artículo “On denoting” (1905), en *Principia mathematica* (1910, Ch. 3) escrito junto con Whitehead y en su libro *Introduction to mathematical philosophy* (1919), en los cuales formula su teoría de las descripciones.

La argumentación de Russell es que las descripciones tal como las presenta Frege no son “expresiones completas” o “unidades de significación” si se las somete a un riguroso análisis lógico-sintáctico. En un análisis superficial, como es el análisis gramatical, enunciados como “Frege es calvo”, “el actual rey de Francia es calvo”, “algún filósofo es calvo”, “los filósofos son calvos” constan de dos partes o unidades de significación: un sujeto y un predicado, *i.e.*, una entidad o sujeto (Frege; el actual rey de Francia; algún filósofo o los filósofos) de la que, en cada caso, se predica la calvicie, *i.e.*, se trata de enunciados con la misma forma o estructura sintáctica. Sin embargo, Russell muestra en un análisis más profundo (lógico-sintáctico) que esos mismos enunciados constan



de más de dos componentes y, por tanto, las unidades semánticas que para análisis superficial parecen *expresiones completas*, en un análisis lógico se revelan como *expresiones incompletas*, algo que el mismo Frege en su *Begriffsschrift (Conceptografía)* de 1879, ya había advertido para los enunciados particulares como “algún filósofo es calvo” y universales como “todos los filósofos son calvos”. El problema es que Frege no aplica este mismo análisis para enunciados singulares como “Frege es calvo” o para descripciones definidas como “el actual rey de Francia es calvo”. Lo que muestra el análisis lógico-sintáctico russelliano es que también estos enunciados son expresiones incompletas y si en un análisis superficial parecen tener la misma estructura sintáctica sujeto-predicado, en un análisis profundo (lógico), sujeto-predicado no son componentes o piezas del enunciado; dicho análisis muestra que el sujeto gramatical no es idéntico al sujeto lógico.

Justamente, la intuición central de “On denoting” (1905), es una intuición acerca de la forma lógica de las descripciones definidas y, en especial, de aquellas descripciones que no denotan como “el actual rey de Francia”, pues lo que para el análisis superficial aparece como una oración singular con un nombre sin referencia (el sujeto) para el análisis lógico es una proposición existencial en la que la descripción definida “el actual de Francia” no aparece. Advertir como lo hace Russell que las descripciones definidas son expresiones incompletas y, por tanto, que no son expresiones lingüísticas significativas genuinas, hace que ellas no constituyan un contrajemplo para su teoría referencialista, aunque la teoría de las descripciones de Russell no se limita exclusivamente a éste punto, sino a resolver o mejor disolver algunas de las paradojas a los que conduce el análisis superficial de las descripciones como expresiones lingüísticas genuinas de la forma gramatical sujeto-predicado. Russell considera que una teoría lógica como su teoría de las descripciones se pone a prueba por su capacidad de resolver paradojas o rompecabezas (*puzzles*) y para ello propone tres paradojas o rompecabezas que dicha teoría debe ser capaz de resolver:

(i) la primera paradoja tiene que ver con el principio de sustituibilidad *salva veritate* en todos los contextos (directos e indirectos), *i.e.*, con la posibilidad de sustitución de términos o expresiones co-referenciales o co-designativas en todos ellos. Como ya vimos, en los contextos indirectos u oblicuos, como son los contextos intensionales de creencia, la sustituibilidad de términos correferenciales *salva veritate* no se cumple. En consecuencia, hay que afirmar que la teoría referencialista

que identifica el significado no es válida o, al menos, no lo es en todos los contextos.

(ii) La segunda paradoja se relaciona con la anterior y tiene que ver con la violación del principio del tercero excluido o, en el caso de la verdad, del principio de bivalencia, pues si una descripción definida se interpreta como un enunciado particular, a saber *como* “el objeto  $x$  tiene la propiedad  $P$ ”, este enunciado particular de la forma  $P(x)$  es verdadero si  $x$  tiene la propiedad  $P$  y falso si no la tiene, por tanto, debe ser verdadero o bien que “el actual rey de Francia es calvo” o bien que “el actual rey de Francia no es calvo”, pero para que “el actual rey de Francia no sea calvo” sea verdadero, “el actual rey de Francia es calvo” debe ser falso, entonces, ni “el actual rey de Francia es calvo” es verdadero, ni “no es el caso que el actual rey de Francia es calvo” es verdadero, lo que constituye una violación del principio del tercero excluido. Como dice jocosamente Russell refiriéndose a este caso:

si hacemos una enumeración de las cosas que son calvas y de las que no lo son, no encontramos el actual rey de Francia. Quizá los hegelianos que aman la síntesis, concluirán que usa peluca. (Russell 879)

(iii) La tercera y última paradoja tiene que ver con las entidades inexistentes. Una oración de la forma lógica  $P(a)$  es verdadera o falsa *syss*.  $a$  es  $P$  y análogamente  $\text{no-}P(a)$  es verdadera *syss*.  $a$  no es  $P$ . Como las entidades ficticias no existen, entonces su forma lógica sería  $\text{no-}P(a)$  y esta sería verdadera si  $a$  no es  $P$ , *i.e.*, si  $a$ , la entidad ficticia, tiene la “propiedad” de la no-existencia. Pero para que un objeto  $x$  tenga propiedades es necesario que exista y dado que  $a$ , la entidad ficticia, tiene la propiedad de la no-existencia, entonces  $a$  existe. Por tanto,  $a$ , la entidad ficticia, existe y no existe.

Russell disuelve estas paradojas mostrando que hay diferencia entre la forma de los nombres, de los nombres propios lógicos y la de los nombres, de las descripciones definidas. Tal distinción tiene que ver con otra distinción epistémica, también russelliana, entre conocimiento directo (*knowledge by acquaintance*), *i.e.*, sin mediación alguna y conocimiento indirecto o por descripción (*knowledge by description*).

En su versión del atomismo lógico, Russell distingue una categoría especial de los nombres propios: los *nombres lógicamente propios*. Estos son nombres que necesariamente no pueden dejar de tener referencia,

como es el caso de los deícticos “este”, “ese”, entre otros. El que Russell haya privilegiado únicamente a los deícticos como nombres *lógicamente propios* obedece a razones epistémicas, pues si la función de los nombres *lógicamente* es denotar un objeto, este objeto tiene que ser un objeto simple y no un objeto complejo como personas o lugares, pues *stricto sensu* no conocemos directamente (*by acquaintance*) los objetos completos, sino sólo los datos sensoriales (*sense data*)<sup>10</sup>. Es por ello, que las únicas expresiones que Russell admite como “nombres *lógicamente propios*” son las expresiones deícticas (“eso”, “este”, “ahí”, entre otras), pues ellas se refieren a puntos-superficies de nuestro espacio perceptivo, como cuando digo “*esto* cuadrado y rojo está a la derecha de *eso* triangular y verde”, pues se trata de un enunciado básico que expresa una proposición sobre *sense data* que están a la base de nuestro conocimiento. Sin embargo, en ocasiones Russell habla de conocimiento directo de personas o lugares que realmente hemos conocido o visitado (*knowledge by acquaintance*) y no simplemente de personas o lugares que hemos conocido porque hemos oído hablar (*knowledge by description*), pero como lo precisa Haack,

este es un uso relajado, y [...] la teoría estricta según la cual ningún nombre corriente es lógicamente propio es la única [teoría] que debe tomarse en serio. (Haack 84)

Esta limitación hace que para Russell los nombres propios *corrientes* se asemejen más a “descripciones (definidas) disfrazadas” que a verdaderos nombres o nombres *lógicamente propios*; estas descripciones definidas son el significado del nombre corriente y, como en caso de Frege, permiten identificar el objeto denotado o referido por el nombre corriente al especificar ciertas propiedades de él. Sin embargo, Russell se aparta de Frege al proponer una explicación diferente de las descripciones como ya lo hemos insinuado.

Para Russell una descripción definida como “el  $x$  es  $F$ ” equivale lógicamente a la conjunción de los siguientes tres enunciados:

<sup>10</sup> Entre los objetos de conocimiento directo Russell menciona al menos cuatro: los *sense data* de los sentidos externos; los *sense data* de los sentidos internos (introspección); los *sense data* de la memoria y el “yo”, aunque vacila respecto de este último. De estos *sense data* y de las propiedades de y relaciones entre estos *sense data* podemos tener conocimiento directo, pero no de objetos externos, propiedades de y relaciones entre objetos externos (macroscópicos) al no ser objetos de conocimiento directo.

- a) Existe una  $x$ .
- b) Existe *a lo sumo* (exactamente) una  $x$ .
- c) Algo que es  $x$  es  $F$ .

Para Russell la afirmación (b) es muy importante, pues el artículo “el” de la descripción definida supone *unicidad* cuando se lo usa adecuadamente, *i.e.*, no simplemente afirma que “existe algo  $x$  que tiene determinada propiedad  $F$ ”, sino que “existe *a lo sumo* un  $x$  que tiene la propiedad  $F$ ”, *i.e.*, que “existe una y *sólo una* entidad  $x$  que es  $F$ ”.

La conjunción de las tres afirmaciones formalmente se expresa así:

$$(\exists x) [Fx \wedge (\forall y) (Fy \rightarrow y \equiv x) \wedge Gx]$$

En el caso de la descripción definida “el actual rey de Francia es calvo” el análisis lógico muestra que tal descripción equivale a la conjunción de tres enunciados:

- a) Existe un  $x$  [sujeto lógico] tal que:  $x$  es el actual rey de Francia.
- b) Para cualquier  $y$ , si  $y$  es el actual rey de Francia entonces  $y = x$ .
- c) “ $x$  es calvo”.

Lo que el análisis muestra es que esta forma lógica de la descripción definida (y en general todas las descripciones definidas) no es igual a la de la forma lógica gramatical sujeto-predicado, pues las descripciones definidas son expresiones complejas cuantificadas (sin variables libres) que expresan condiciones de verdad generales (no singulares) en la medida en que no involucran ningún nombre propio, ni se refieren, por tanto, a ningún objeto particular. En ellas, sólo aparecen términos generales (predicados como “ser rey de Francia” y “ser calvo”) y no términos singulares (nombres propios), con lo cual se evita el problema de nombres propios que no denotan para una teoría referencialista. Se trata de funciones proposicionales en las cuales las variables ligadas pueden ser sustituidas por términos singulares (nombres propios) para de esa forma establecer el valor de verdad de la función.

Frege era consciente de las dificultades que entrañan los enunciados particulares de la forma “ $P(x)$ ” cuando el objeto  $x$  no existe, pues si no existe, no es posible realizar ninguna predicación particular de él. Aunque reconoce que esta es una situación que idealmente (en un lenguaje lógicamente perfecto) no debe ocurrir, en el evento de que

ocurra se le puede asignar una referencia arbitraria: el conjunto vacío. Esta solución Russell no la acepta, pues además de arbitraria, tiene consecuencias inaceptables, ya que en la teoría de conjuntos en un dominio  $A$ , un conjunto nulo es un subconjunto de todos los conjuntos contenidos en  $A$  y si se acepta el axioma de regularidad, el conjunto nulo es idéntico a  $A$ . Como dice Díez (s.f.):

si a toda descripción impropia se le asigna el conjunto vacío, entonces “el actual rey de Francia es subconjunto de todo conjunto” sería verdadero.

Russell también rechaza la teoría de Meinong quien considera que si el enunciado “el actual rey de Francia es calvo” es significativo, entonces el enunciado debe ser acerca de “el actual rey de Francia” y si el rey de Francia no existe, la sentencia no es acerca de nada, de suerte que debe existir (subsistir) algún sentido de la palabra “el actual rey de Francia”. Meinong no sólo se refiere a este caso, pues además de admitir la existencia de objetos reales, espacio temporales, como los objetos físicos, postula la existencia (subsistencia) de otros objetos no-espacio temporales como los números, las propiedades e incluso, de objetos contradictorios como “el cuadrado redondo”. El filósofo inglés no sólo niega que puedan existir o *subsistir* como dice Meinong objetos irreales o ficticios, sino que señala que cuando se hace un análisis adecuado (lógico) de las descripciones definidas, en la representación formal no aparece ningún término singular, sino sólo variables ligadas, predicados (términos generales) e identidad.

Con su teoría de las descripciones, Russell se libera de asumir compromisos ónticos con un dominio de entidades irreales que son denotación de nombres que no denotan. En la teoría russelliana, los nombres propios han sido sustituidos por descripciones generales en cuya forma lógica sólo aparecen términos generales (predicados) como “ser el actual presidente de Francia” o “ser calvo”. El compromiso óntico russelliano sólo es con los objetos de sustitución de las variables individuales que hacen verdadera la función, algo similar a lo que planteará Quine (1953) con su eslogan “ser es ser el valor de una variable ligada” con el que introduce su criterio de compromiso ontológico. Con la teoría de las descripciones Russell, en contra de Meinong e influenciado por la defensa que Whitehead hace de la navaja de Ockham de no proliferar los entes sin razón, plantea una forma de restricción y de compromiso ontológico, al establecer como existentes sólo aquellos

objetos o entidades de sustitución de las variables ligadas que hacen que la proposición sea verdadera. Tal es el caso mencionado por el mismo Russell de la proposición “la Tierra gira alrededor del Sol” que para él, siendo verdadera, denota una entidad y no la de “el Sol gira alrededor de la Tierra que, siendo falsa, no denota entidad alguna” (Russell 45)<sup>11</sup>.

Finalmente, mientras para Frege una oración como “el actual rey de Francia es calvo” *presupone* que hay en la actualidad un rey de Francia y es calvo, para Russell la oración en cuestión *implica* que en la actualidad existe uno y sólo un individuo que es rey de Francia y es calvo, pues lo que hay es parte de lo que dice. Sin embargo, aunque en Frege no existe un tratamiento formal adecuado de la noción lógica de *presuposición* como relación entre proposiciones en el contexto de una lógica no-bivalente, Strawson adopta la misma noción de Frege y, como él, afirma que oraciones como “el actual rey de Francia es calvo” no son ni verdaderas, ni falsas, *i.e.*, carecen de valor de verdad, pues mientras en la implicación lógica (como la defendida por Russell) si  $S_1$  *implica*  $S_2$  es verdadero y  $S_2$  es falso, entonces,  $S_1$  es falso (validez del *modus tollens*), en la *presuposición* lógica, *i.e.*, como relación entre enunciados, si  $S_1$  *presupone*  $S_2$  es verdadero y  $S_2$  es falso, entonces,  $S_1$  no es ni verdadero ni falso, salvo que  $S_2$  sea verdadero (Haack 89) y (van Fraassen 133ss.).

#### REFERENCIAS BIBLIOGRÁFICAS

Díez, J. A. “Tema 4. La teoría de descripciones de B. Russell”. Web. <http://www.ub.edu/filosofia-del-llenguatge/dossiers/TEMA04.pdf>. Online.

---. “Tema 5. B. Russell: Semántica y epistemología”. Web. <http://www.ub.edu/filosofia-del-lenguaje/dossiers/TEMA05.pdf>. Online.

Dummett, M. *Frege: Philosophy of language*. London: Gerarld Duckworth & Co., 1973. Print.

Frege, G. *Conceptografía. Los fundamentos de la aritmética, otros estudios filosóficos*. México: U.N.A.M, 1972. Impreso.

---. “Sobre sentido y referencia”. *Estudios sobre semántica*. Barcelona: Ariel, 1973. Impreso.

---

<sup>11</sup> Este mismo procedimiento lo extiende Russell a lo que él denomina “el reino de las no-entidades”, “cuadrado redondo”, “el número primo par distinto de 2”, “Apolo”, “Hamlet”, entre otros, cuyas proposiciones no denotan nada (Russell).

---. *Escritos filosóficos*. Barcelona: Crítica, 1996. Impreso.

García-Carpintero, M. *Las palabras, las ideas y las cosas. Una presentación de la filosofía del lenguaje*. Barcelona: Ariel, 1996. Impreso.

Haack, S. *Filosofía de las lógicas*. Madrid: Cátedra, 1978. Impreso.

Moro Simpson, T. *Formas lógicas, realidad y significado*. Buenos Aires: Editorial Universitaria de Buenos Aires, 1964. Impreso.

Moulines, C. U. *Pluralidad y recursión. Estudios epistemológicos*. Madrid: Alianza, 1991. Impreso.

---. "Ontoepistemosemántica en perspectiva estructuralista". Peris-Viñé, L. M. (ed.). *Filosofía de la ciencia en Iberoamérica: metateoría estructural*. Madrid: Tecnos, 2012. Impreso.

Russell, B. *Obras completas, Tomo II. Ciencia y filosofía 1897-1919*. Madrid: Aguilar, 1973. Impreso.

van Fraassen, B. *Semántica formal y lógica formal*. México: U.N.A.M., 1987. Impreso.

Valdivia, L. "Teorías de la referencia". Acero, J. J. (ed.). *Filosofía del lenguaje I. Semántica*. Madrid: Trotta, 1998. Impreso.





# PROPUESTAS Y DIFERENCIAS PRAGMÁTICAS EN TORNO DEL LENGUAJE COMO INSTITUCIÓN: WITTGENSTEIN Y HABERMAS

PROPOSALS AND PRAGMATIC DIFFERENCES ON LANGUAGE AS INSTITUTION:  
WITTGENSTEIN AND HABERMAS

JAVIER ALEGRE\*

Universidad Nacional del Nordeste (UNNE), Argentina. pillancho@yahoo.com.ar

RECIBIDO EL 4 DE OCTUBRE DE 2012 Y APROBADO EL 23 DE NOVIEMBRE DE 2012

## RESUMEN      ABSTRACT

El presente artículo analiza concepciones *institucionalistas* del lenguaje de cuño pragmático dadas en el seno de la filosofía contemporánea, centrándonos específicamente en las elaboraciones señeras hechas por Wittgenstein y la reapropiación y propuestas realizadas por Habermas al respecto. A partir de precisar los núcleos centrales del *giro pragmático*, abordamos los ejes conceptuales e implicancias de la teoría de los *juegos de lenguaje* del segundo período wittgensteiniano y su recepción y modificación en el marco de la *teoría de la acción comunicativa* habermasiana. Este abordaje relacional tiene por objetivo analizar los lineamientos y consecuencias que se desprenden de sendas teorías y, en particular, establecer si la reapropiación de Habermas está en condiciones de retomar la potencia heurística y teórica propia de la concepción pragmática institucionalista del lenguaje iniciada por Wittgenstein.

This article analyzes *institutional* conceptions of language of a pragmatic derivation within contemporary philosophy, focusing mainly in the works of Wittgenstein and in its re-appropriation and proposals by Habermas. After specifying the central tenets of the *pragmatic turn* we will tackle the conceptual axis and implications of the *language games* theory by the second Wittgenstein, as well as its reception and modification in the framework of Habermas's *theory of communicative action*. This comparative approach aims to analyze guidelines and consequences that result from both theories and, in a particular way, determine whether Habermas re-appropriation is able to retake the heuristic and theoretical power common to the institutionalist pragmatic conception of language initiated by Wittgenstein.

## PALABRAS CLAVE      KEY WORDS

acción comunicativa, juegos de lenguaje, instituciones, lenguaje, pragmatismo lingüístico.

communicative action, language games, institutions, language, linguistic pragmatism.

---

\* Doctor en Filosofía. Docente-investigador en el área de Filosofía Contemporánea - Universidad Nacional del Nordeste (UNNE), Chaco - Argentina.

El ámbito temático de este artículo proviene de las concepciones *institucionalistas* del lenguaje generadas a partir del *giro pragmático* contemporáneo, por lo que nuestro interés gira en torno del lenguaje en tanto institución que tiene un rol decisivo en la conformación de las demás instituciones y del carácter que puede llegar a tomar aquel en relación con las propiedades conferidas por las prácticas institucionales. Los ejes cardinales de nuestra perspectiva tienen que ver, por un lado, con que el lenguaje mismo es una institución, debido a que está constituido por determinadas reglas, convenciones, prácticas, entre otras, que posibilitan y regulan su funcionamiento y que todo hablante debe manejar para poder desempeñarse dentro de su universo; y, por otro lado, con que no hay instituciones extra-lingüísticas, lo cual no quiere decir que todas las instituciones existen *en* el lenguaje o que lo hagan gracias a él solamente, sino que todas ellas se conforman a través suyo, necesitan de la participación del lenguaje –y de varios elementos más según sea el caso de cada una de ellas– para constituirse como tales.

Esta participación del lenguaje en la constitución de las prácticas sociales y su relación con las instituciones han sido enfocadas desde diversas perspectivas; la base teórica del artículo está provista por el pragmatismo lingüístico contemporáneo, más precisamente utilizamos como *plataforma móvil* los aportes teóricos inaugurales en esta temática generados por la segunda etapa filosófica de Wittgenstein y tiene por objetivo particular analizar la reelaboración que el legado wittgensteiniano encuentra en los planteos de la *teoría de la acción comunicativa* de Jürgen Habermas. Teniendo en cuenta esto, hemos organizado el escrito en cuatro partes: (i) iniciamos con una breve exposición del abordaje del lenguaje en el *giro pragmático*; (ii) proseguimos con el tratamiento de la visión de Wittgenstein del lenguaje como institución; (iii) luego continuamos con la reapropiación y la propuesta hechas por Habermas al respecto; y (iv) cerramos con las consecuencias y conclusiones que se pueden obtener del análisis hecho, haciendo especial hincapié en la concepción habermasiana.

## I

### La perspectiva pragmática del lenguaje contemporánea

El *giro lingüístico*, acaecido en el siglo XX, no es algo homogéneo, ni propio de un solo movimiento contemporáneo, bajo este apelativo genérico se ha congregado *a posteriori* (Rorty) diferentes vertientes y momentos de la reflexión filosófica de fines del segundo milenio, dentro de las cuales sobresalen con nitidez, por la envergadura de

sus discusiones, el *giro lingüístico analítico*, el *giro hermenéutico* y el *giro pragmático*. Aquí nos detendremos únicamente en este último.

El *giro pragmático* se centra en el análisis pragmático de las estructuras lingüísticas, considera que la base semántico-sintáctica de los estudios lingüísticos debe ser ampliada y complementada con el punto de vista pragmático y surge como resultado de la revisión y complementación o desestimación de determinados fundamentos y objetivos incumplidos del análisis lógico –idealismo lógico, concepción empirista y atomista del conocimiento, pretensiones de objetividad, claridad absoluta y no circularidad, entre otros– (Cabanchik, Penelas y Tozzi). Es por ello que reemplaza la semántica fregeana por la pragmática wittgensteiniana como base, abandona la perspectiva referencialista del significado y la idea de la construcción de lenguajes lógicos e ideales, aborda la investigación de los actos lingüísticos antes que las proposiciones y se interesa por el lenguaje cotidiano, los usos comunes y los factores sociales más que por el lenguaje ideal, las formulaciones científicas y los aspectos formales. El abandono de la semántica realista implica rebatir la noción de que haya entidades definidas que precedan objetivamente a las palabras y constituyan el significado de éstas, sino que el significado tiene realización sólo en las manifestaciones lingüísticas, es inmanente al lenguaje y al contexto pragmático en que es empleado –que incluye instituciones, hábitos, patrones culturales, entre otros–, por lo que presta renovada atención a la relación entre lenguaje, prácticas, convenciones y decisiones humanas y a la posibilidad de realizar muy diferentes funciones mediante las emisiones lingüísticas.

La visión del lenguaje propia del pragmatismo lingüístico está vinculada con la convicción de que las producciones lingüísticas tienen la capacidad de ser el horizonte y marco regulativo que conforman nuestras pautas de pensamiento, sentimiento y acción, condición que hace imposible constituir nuestra subjetividad independientemente de los procesos simbólicos que pueblan la realidad cultural e institucional en la que nos desenvolvemos, a la vez que esta realidad social no puede pensarse, ni realizarse, sino es a través de la participación constitutiva del lenguaje como práctica colectiva. Con ello, se resaltan las funciones pragmáticas y realizativas del lenguaje, por sobre las descriptivas-denotativas-representativas en que se centró clásicamente la tradición filosófica. La fecundidad del enfoque pragmático reside, precisamente, en que se encuentra en mejores condiciones para dar cuenta de la complejidad de los fenómenos humanos debido a que incluye y pone en relación a un número más amplio de variables y a que sitúa en el centro

al componente principal de nuestra instalación en el mundo: la acción, acción que siempre es entendida como simbólica y colectiva (Naishtat).

Nuestra adopción de la perspectiva pragmática está en directa dependencia de la mayor fecundidad que encontramos en ella, dado que contempla un mayor número de prácticas y funciones simbólicas y tiende a buscar las relaciones existentes entre los diferentes niveles de acción, por lo que no realiza recortes abruptos o elimina elementos o variables en pos de establecer la hegemonía de uno solo de ellos (como fuera el enaltecimiento de la función representativa del lenguaje en la tradición filosófica). El enfoque pragmático resalta, por un lado, el carácter público, intersubjetivo y situado social e históricamente de todas las actividades lingüísticas y, por el otro, que cualquier acción, institución u organización social se constituye en parte por la participación del lenguaje como práctica colectiva; de aquí que se centre en el funcionamiento y las interrelaciones de las diversas dimensiones participantes en las acciones lingüísticas. Por eso, la perspectiva pragmática permite integrar ciertos aportes de las aproximaciones semánticas y sintácticas, no vuelve necesaria su eliminación, razón por la cual consideramos que es más abarcativa y brinda mejores herramientas para la comprensión de las prácticas y funciones lingüísticas.

Dentro de esta perspectiva son clave los senderos abiertos por las reflexiones de Wittgenstein y Austin a partir de la década del 50', más precisamente con la publicación de las *Philosophische Untersuchungen* de Wittgenstein en 1953 y las conferencias dictadas por Austin en 1955 y que fueran editadas bajo el título de *How to do Things with Words* en 1962<sup>1</sup>. La especial importancia de Wittgenstein para la filosofía del pragmatismo lingüístico radica en que subraya el papel determinante que tienen los contextos de uso en el establecimiento del significado de las palabras y que pone al lenguaje en relación ineludible con las *formas de vida*, hábitos e instituciones de los colectivos de hablantes a través de lo cual dota de todo un nuevo sustento a las aproximaciones teóricas al lenguaje y los múltiples procesos conexos. Las reflexiones de Wittgenstein proveyeron el sustrato teórico específico para los posteriores análisis institucionalistas del lenguaje y, en particular, fueron reasumidas y utilizadas, en conexión con otras tradiciones, por Jürgen Habermas en su *teoría de la acción comunicativa*.

---

<sup>1</sup> Inicios al que también mucho aportaron, aunque consideramos que secundariamente respecto de la repercusión lograda por Wittgenstein y Austin, los artículos "Two dogmas of Empiricism" de Willard Quine (1951) y "On Referring" de Peter Strawson (1950).

## II

## Lenguaje como institución en Wittgenstein: *formas de vida y juegos de lenguaje*

La teoría del lenguaje de la segunda etapa wittgensteiniana es una crítica radical a la conciencia como ámbito del sentido y como punto de partida del filosofar, la concepción basada en los *juegos de lenguaje* (*Sprachspiele*) y las *formas de vida* (*Lebensformen*) se opone a la extensa tradición que entiende el lenguaje como expresión de contenidos mentales y del significado como representación que yace en la conciencia. La crítica de Wittgenstein se dirige expresamente a la concepción referencialista del lenguaje, la que expone a través de la visión agustiniana, pero que recibe su elaboración más acabada en la tradición cartesiana e implícitamente abarca casi la totalidad de la tradición filosófica occidental (incluido el *Tractatus*), para la cual el lenguaje posee una esencia caracterizada por la capacidad de representar la realidad y por el hecho de que cada palabra tiene un significado y que éste, entendido como entidad mental, es a su vez el que remite a un determinado objeto. Esta concepción es primitiva, parcial y, por tanto, errónea para Wittgenstein; la relación entre palabras y objetos es variada y no se da objetivamente, no está impuesta de modo lógico o natural por ninguna cualidad representativa, ni por condición alguna establecida con anterioridad al empleo contextualizado del lenguaje, sino que es convencional y depende del modo en que se usa y se instruye para dicho uso: el significado depende de las prácticas instituidas dentro de cada comunidad y el lenguaje es él mismo una institución.

El lenguaje posee un carácter complejo e inacabado producto de sus múltiples interrelaciones con distintas esferas y tareas y de la dinámica propia de su devenir histórico –lo que vuelve inútil cualquier intento de búsqueda de esencias, hipostatización o idealización de una determinada función o estructura–, Wittgenstein lo grafica en las *Investigaciones filosóficas* con la siguiente analogía en que compara al lenguaje con una ciudad:

nuestro lenguaje puede verse como una vieja ciudad: una maraña de callejas y plazas, de viejas y nuevas casas, y de casas con anexos de diversos períodos; y esto rodeado de un conjunto de barrios nuevos con calles rectas y regulares y con casas uniformes. (*Investigaciones* §18)<sup>2</sup>

<sup>2</sup> Esta potente metáfora, además, da cuenta del modo histórico en que se constituyen los lenguajes: cualquier intento de *urbanización* lógica, sintáctica o científica (nuevos barrios del lenguaje-ciudad) siempre es posterior y viene a la zaga del desarrollo mediante diversos *materiales* y en vista de variadas funciones del núcleo histórico (centro fundacional del lenguaje-ciudad) y ambos sectores no cesan de modificarse e incrementarse.

El lenguaje consiste en un conjunto de prácticas sociales complejas que se definen de acuerdo con las reglas que gobiernan los diferentes usos de las palabras y que poseen distintos fines; el fundamento del sentido del lenguaje está en la praxis humana: el uso del lenguaje está en concordancia con las demás prácticas que los usuarios realizan; las palabras no pueden ser entendidas fuera de la utilización que hacen de ellas los hablantes y hacia ésta hay que orientar las investigaciones. El significado de los signos lingüísticos no se obtiene por definiciones ostensivas, ni puede encontrarse tratando de imponer modelos ideales a lo que sucede en la realidad, sino que debe buscarse en la vida cotidiana el sentido que adquieren las palabras de acuerdo con el uso (*Gebrauch*) que se hace de ellas, como lo expresa Wittgenstein en su célebre apotegma:

para una *gran* clase de casos de la palabra “significado” –aunque no para *todos* los casos de su utilización– puede explicarse esta palabra así: el significado de una palabra es su uso en el lenguaje. (*Investigaciones* §43)<sup>3</sup>

Así, Wittgenstein resalta la naturaleza social y el carácter institucional del lenguaje, todo lenguaje consiste en actividades regladas, las cuales se estipulan en base a usos, hábitos, costumbres, que suponen necesariamente la regularidad propia de las instituciones; el lenguaje está constituido por variadas prácticas que requieren la adquisición de habilidades en base al dominio de técnicas a lo largo del tiempo, por ello, es tan importante la instrucción para ser capaces de seguir reglas *adecuadamente* y participar de las acciones lingüísticas colectivas de modo satisfactorio:

¿es lo que llamamos “seguir una regla” algo que pudiera hacer sólo *un* hombre sólo *una vez* en la vida? (...) No puede haber sólo una única vez en que un hombre siga una regla. (...) Seguir una regla, hacer un informe, dar una orden, jugar una partida de ajedrez son *costumbres* (usos, instituciones). Entender una oración significa entender un lenguaje. Entender un lenguaje significa dominar una técnica. (*Investigaciones* §199)

Si bien la utilización del lenguaje está en dependencia del seguimiento de reglas, estas reglas no determinan unívocamente el uso del

<sup>3</sup> Esta concepción ya había sido presentada expresamente por Wittgenstein en la etapa del *Cuaderno azul* (99 108) y también recibe nuevas formulaciones en *Sobre la certeza* (§§ 61-2 65).



lenguaje, quedan abiertas diferentes posibilidades de continuación; las posibles conexiones entre lenguaje y realidad son tan amplias como la multiplicidad de *juegos de lenguaje*, pues de acuerdo con las tareas y fines de cada *juego* se instituye la manera propia en que las expresiones adquieren sentido. Por ello, los *juegos de lenguaje* son variados, no pueden ser reducidos a unidad alguna, son actividades regladas y autosuficientes dadas dentro de una comunidad determinada, que conforman un todo irreductible a un meta-juego de lenguaje y que cristalizan en costumbres e instituciones. El concepto de *juego de lenguaje*, además, no puede separarse de la idea que el lenguaje es un instrumento que cumple una gran diversidad de funciones sin que exista una función arquetípica u originaria, la agrupación de tan disímiles roles y entidades bajo el término de lenguaje se da gracias a que comparten rasgos análogos, *parecidos de familia* (*Familienähnlichkeiten*). Los *juegos de lenguaje* están en intrínseca relación con la acción, su alta capacidad realizativa es lo que los vuelve tan importantes; no son sus capacidades representativas, expresivas o lógicas donde reside el principal interés, sino en que producen e institucionalizan los aspectos sobresalientes que sirven de patrones para nuestra vida<sup>4</sup>. La predominancia de los *juegos de lenguaje* sobre las instancias individuales y subjetivas hace que aquel o aquello que es tomado como criterio o patrón para establecer algo no lo sea por sí mismo, sino porque en el *juego* en que está inmerso se le ha atribuido esa función; de aquí que no *domine* o pueda arrogarse la potestad sobre ese *juego de lenguaje*, por el contrario, es éste el que le otorga esa autoridad<sup>5</sup>.

En resumen, en la concepción del lenguaje wittgensteiniana sobresale la importancia y la atención que deben prestarse al carácter social e institucional del lenguaje, a las reglas y los factores constituyentes en

<sup>4</sup> La centralidad otorgada a la acción como elemento cardinal de las prácticas lingüísticas queda inmejorable y sintéticamente expresada en la aplicación que hace Wittgenstein al lenguaje de la célebre frase del *Fausto* de Goethe: “el lenguaje –quiero decir– es un refinamiento, *im Anfang war die Tat* (“en el comienzo fue la acción”)” (*Cause and Effect* 420).

<sup>5</sup> El *alcance* de los conceptos de la teoría de los *juegos de lenguaje* de Wittgenstein es fuente de fervientes discusiones entre dos grandes líneas. Según una, aceptar la tesis de que cada juego no puede ser conocido, ni juzgado desde los fundamentos de otras *formas de vida* implica que son inconmensurables –no hay jerarquía entre ellos, pues no existe un juego ideal que sirva de patrón de comparación– y, entonces, la mayoría de las afirmaciones hechas en las *Investigaciones* carecen de sentido universal: el singularismo de los *juegos de lenguaje* vuelve ilegítima la pretensión de brindar cualquier teorización general sobre ellos. En esta postura se enrolan Winch, Lyotard y los posmodernistas. En tanto que la otra línea supone que la teoría de los *juegos de lenguaje* puede funcionar como base para un meta-juego de lenguaje que tenga por misión aclarar y fijar ciertos parámetros para los demás juegos, por lo que sería capaz de alcanzar una extensión superior a la de la propia *forma de vida* en que fue generado. A esta segunda posición son afines Habermas, Apel y los partidarios del modernismo crítico.

toda actividad lingüística –que van mucho más allá de los procesos racionales–, a la diversidad de *juegos de lenguaje*, a las múltiples formas y funciones que presentan los lenguajes de acuerdo con los intereses y propósitos que persiguen y a la prioridad que posee lo que llevamos a cabo con el lenguaje antes que cualquier otra característica. Sin abordar el lenguaje –en tanto acción lingüística–, todo acercamiento a las demás acciones humanas es parcial y erróneo; y si *desprendemos* del lenguaje las múltiples acciones entrelazadas y realizadas con él, éste se vuelve una entidad fantasmagórica o bien comprensible sólo de modo muy defectuoso. La profundización del carácter institucional de las acciones hechas *en* y *con* el lenguaje es la vía para desentrañar aquello que incluyen y posibilitan los procesos lingüísticos, en esto el pensamiento de Wittgenstein ha abierto múltiples senderos que no cesan de ser transitados en diversas direcciones y que tal vez en parte por el tono dialógico y aforístico de sus exposiciones, ha sido fuente de muy divergentes continuaciones, las cuales, sin embargo, coinciden en señalar la originalidad y ascendencia de sus reflexiones para el establecimiento y desarrollo de la perspectiva pragmática dentro de la filosofía del lenguaje contemporánea.

### III

#### **Habermas: acción comunicativa e institucionalización moderna**

En las elaboraciones teóricas de Jürgen Habermas el bagaje conceptual proveniente de las reflexiones iniciadas por el *giro pragmático* está puesto en relación con un amplísimo abanico de corrientes intelectuales que ligán con un contexto de discusión filosófica y sociológica mucho más extendido –conceptual y temporalmente– que el del pragmatismo lingüístico. La propuesta habermasiana se caracteriza por ser una crítica de la racionalización mediante el despliegue de una teoría de la racionalidad en términos reconstructivos o dicho de manera más grandilocuente: una crítica de la razón para defender y salvar a la razón –de otras críticas más devastadoras–. En este marco, la significación de los actos de habla depende de su orientación a lograr el entendimiento y acuerdo entre los participantes de la comunicación; el lenguaje es considerado como el medio donde se puede lograr el consenso entre los integrantes de una comunidad, comunidad que en última instancia tiene alcance universal, ya que el hecho lingüístico de la argumentación es irrefutable –no puede ponerse en duda sin presuponerlo a la vez, sin caer en *autocontradicción performativa*–.

Habermas deja en claro que los problemas de racionalización y cosificación que son de su interés central pertenecen a la línea alemana del pensamiento que sigue el derrotero Kant-Hegel-Marx-Weber-Lukács y desemboca en la *Escuela de Frankfurt*; por ello, lo que Habermas intenta es renovar y poner en nuevos senderos ciertos aspectos de esta amplia línea teórica a partir de los aportes del pragmatismo, por lo que la *teoría de la acción comunicativa* -TAC- implica una ampliación de la perspectiva pragmática inicial hacia nuevos ámbitos y líneas teóricas. A través de su reapropiación del pragmatismo lingüístico, Habermas busca enmendar los *excesos* de la crítica hiperbólica de la teoría frankfurtiana y superar las consecuencias negativas atribuidas a la modernización por Weber, para lo cual recurre a una lectura que hace retroceder las múltiples capacidades de las fuerzas ilocucionarias del lenguaje, señalada señeramente por Austin, en favor del modelo de la producción de entendimiento y retoma al segundo Wittgenstein desde una óptica netamente consensualista. Habermas recurre a la *teoría de los actos de habla* de Austin y Searle como fundamento para sus elaboraciones sobre el lenguaje, se basa en el *principio de expresabilidad* de Searle y la reducción del análisis a la forma estándar de los actos de habla y le resultan de mayor interés los actos de habla que no están ligados a instituciones particulares, sino que pueden encajar con múltiples instituciones, por lo que está más cercano al modo de análisis del Searle de *Speech Acts* que al de Austin.

La influencia wittgensteniana se manifiesta especialmente en que Habermas, también vincula la racionalidad al ámbito de la lingüística y la intersubjetividad, remite el sentido compartido a la capacidad basal de seguir reglas comunitariamente -no a la esfera de la intencionalidad, ni de la conciencia privada- y coincide en que es totalmente erróneo postular la existencia de una razón trascendente y de procesos de argumentación o crítica con pretensiones de fundamentación última. Además, también para Habermas el lenguaje ordinario es el ámbito de intersubjetividad que hace posible la interacción, el entendimiento y la comprensión mutua y debido a su reflexividad constituye su propio meta-lenguaje, es el marco de referencia para la interpretación del mundo histórico. En todos estos puntos es notable la ascendencia de Wittgenstein ya desde los escritos tempranos habermasianos; "lenguaje y acción se interpretan recíprocamente; es lo que desarrolla Wittgenstein en su concepto de juego de lenguaje" (Habermas, *Conocimiento e interés* 175). Asimismo, Habermas encuentra en Wittgenstein las elaboraciones pioneras en dos puntos de suma importancia para el armado de su

propio sistema teórico: en primer lugar, la inclusión de la dimensión pragmática del empleo de las oraciones en la teoría del significado como uso, y, en segundo lugar, la sustitución del paradigma de la conciencia por el del lenguaje y el tratamiento reductivo que hace de la intencionalidad frente a las propiedades del lenguaje:

Wittgenstein efectuó sin vacilación alguna el paso desde la filosofía de la conciencia a la teoría del lenguaje. Trata de entrada a los contenidos intencionales con independencia de las vivencias intencionales. (Habermas, *Teoría... Complemento 4*)<sup>6</sup>

En cuanto a la vinculación entre lenguaje y procesos de institucionalización, Habermas propone una doble imbricación: el lenguaje contribuye a través de sus usos a los procesos de institucionalización moderna y la institucionalización de la racionalidad moderna especifica las funciones del lenguaje y *libera* sus potencialidades comunicativas. El interés de Habermas está puesto en la relación específica entre el lenguaje y las instituciones legadas por la modernidad, es decir, en la interrelación entre la función del lenguaje como reaseguro de la racionalidad ya destrascendentalizada y los procesos de institucionalización que son producto de la modernización occidental, relación en la que encarnan las posibilidades de avanzar en los logros positivos de dicha modernidad según Habermas.

Habermas defiende el carácter universal de la racionalidad contra las posturas particularistas y relativistas y, además, separa a la racionalidad del modo parcial en que terminó desarrollándose en la modernidad. Frente a las posturas *detractoras* de la razón, Habermas toma un concepto de razón al que considera *no reducido*, sustentado en la racionalidad comunicativa y en las notas del consenso intersubjetivo no coactivo presupuesto en el habla argumentativa. La racionalidad comunicativa se refiere:

---

<sup>6</sup> A modo de registro radiográfico de los posteriores cambios teóricos de Habermas, es interesante rescatar la afirmación que hace en 1965 sobre los escritos de período intermedio de Wittgenstein: "sin embargo, Wittgenstein no emprende una reflexión sobre la conexión de las formas de vida con los juegos del lenguaje. Sigue siendo lo bastante positivista como para no pisar la dimensión de la historia, que es en la que se establece esa conexión. Renuncia así al lado crítico que sus análisis comparativos siguen, empero, presuponiendo implícitamente" (*Perfiles filosóficos* 198). Lo particular es que luego, con las variaciones que va introduciendo a su teoría a lo largo de las décadas del 70' y 80', Habermas termina optando por tampoco *pisar la dimensión de la historia* en la construcción de su TAC y realiza un análisis del lenguaje en términos exclusivamente reconstructivos ahistóricos.

a la experiencia central de la capacidad de aunar sin coacciones y de generar consenso que tiene un habla argumentativa en que diversos participantes superan la subjetividad inicial de sus respectivos puntos de vista y merced a una comunidad de convicciones racionalmente motivada se aseguran a la vez de la unidad del mundo objetivo y de la intersubjetividad del contexto en que desarrollan sus vidas. (Habermas, *Teoría... I 27*)

Es una racionalidad basada en las condiciones de simetría –exenta de violencia y coacción– para todos los participantes propias del diálogo argumentativo, está encaramada en las pretensiones de validez susceptibles de crítica (inteligibilidad, verdad, rectitud y veracidad) y consiste en una racionalidad de tipo formal-procedimental; formal, pues no determina, ni fija sus contenidos y procedimental, ya que consiste en establecer un modo específico de proceder, su principal tarea es ofrecer una base normativa de carácter procedimental que sirva como marco para la búsqueda de consensos basados en razones argumentativas.

Para Habermas las potencialidades universales de la racionalidad comunicativa recién se revelan y quedan a disponibilidad con la modernidad, lo que a su vez hace que por primera vez ellas se tornen abordables y realizables, en cuanto a que sólo pueden ser tematizadas y concretadas a partir de la modernidad –con las condiciones e instrumentos provistos por ésta– y convierte en un problema central a resolver el modo de asegurar socialmente la reproducción de los consensos alcanzados a través de las prácticas comunicativas, así como la obtención de nuevos consensos, lo que implica la institucionalización de la racionalidad comunicativa. Para asegurar niveles crecientes de racionalización, Habermas deja en claro la necesidad de que se formen y consoliden nuevas instituciones que se estructuren a partir de las esferas comunicativas del *mundo de la vida* y considera separables la modernización ocurrida con el capitalismo y el sistema político-económico conformado por éste. La posibilidad de erigir nuevas instituciones que contrarresten la *colonización* producida por los *sistemas de acción* está en total dependencia de la capacidad del lenguaje de coordinar las acciones a través del entendimiento y el acuerdo, es decir, de la institucionalización de ámbitos en que dicha capacidad pueda plasmarse, o lo que es lo mismo, de la institucionalización de la racionalidad comunicativa.

Para Habermas la tradición cultural debe proveer de consensos de modo tal, que en la acción orientada al éxito no tenga que renovarse constantemente el acuerdo ya establecido entre los integrantes de esa tradición; es decir, a través de la institucionalización de la acción racional con arreglo a fines en determinadas esferas de acción (básicamente los sistemas económico y de administración) se logra que las funciones de la acción orientada al entendimiento no se vean sobrecargadas, cuidando, a su vez, que estas acciones no queden *aplastadas* por las primeras. Es por ello que busca medios de estabilización que logren la institucionalización de los acuerdos, al mismo tiempo que se aparta de la visión weberiana acerca de que la institucionalización de las esferas de acción guiadas por la racionalidad con arreglo a fines sea lo que define la racionalización moderna. El acceso a nuevos niveles evolutivos está dado por “una materialización institucional de estructuras de racionalidad que están ya acuñadas en el plano de la cultura” (Habermas, *Teoría... II* 445); por lo que el potencial de racionalidad depende de las esferas y prácticas de la cultura estructuradas comunicativamente y debe quedar inserto en las estructuras de las instituciones, las que actúan como reaseguro de los procesos racionalmente dirigidos. Así, más allá del desarrollo y autonomización de los *sistemas de acción*, son las instituciones del *mundo de la vida* las que actúan como soporte necesario de los procesos de institucionalización de la racionalidad. Así mismo, la acción comunicativa desempeña una función trascendental para Habermas, no en el sentido que sus presuposiciones y condiciones sean inviolables o no puedan ser quebradas, sino que de no cumplirlas se incurre inevitablemente en formas patológicas o distorsionadas de comunicación.

Por tal motivo, para Habermas la acción comunicativa se desempeña como medio de institucionalización de la racionalidad y como mecanismo de integración social por excelencia, no existen otros usos del lenguaje o elementos que puedan cumplir estas funciones de modo similar, ya que carecen de las condiciones y capacidades estructurales dirigidas al entendimiento, cooperación y solidaridad, si bien debe *competir* con ellos para llevar a cabo la coordinación e integración;

mientras sólo se utilice el lenguaje como medio para la transmisión de informaciones y de «redundancias», la coordinación de la acción discurre a través del influenciamiento mutuo de actores que operan los unos sobre los otros con el fin de conseguir cada uno sus propios



finés. Pero en cuanto las fuerzas ilocucionarias de los actos de habla asumen un papel coordinador de la acción, es el lenguaje mismo el que aparece como fuente primaria de integración social. Sólo en este caso ha de hablarse de «acción comunicativa». (Habermas, *Facticidad y validez* 79)

Sin embargo, las acciones comunicativas orientadas al entendimiento poseen un carácter esporádico y las condiciones fácticas que posibilitan que se lleven a cabo deben avenirse a las condiciones instauradas por las acciones estratégicas que buscan suprimir el entendimiento. La coordinación lingüística de la acción es el procedimiento racional por excelencia para concretar el consenso y la normatividad sociales, pero es poco seguro, pues depende de la aceptación continua de las pretensiones de validez expuestas en los actos de habla, con lo que incluye grandes posibilidades de disentimiento y, consecuentemente, de desorden, por lo cual termina siendo un medio de reproducción de la vida social y de integración social de alto grado de inestabilidad. De aquí que la propuesta de solución a la *colonización* pasa en la TAC por descolonizar el *mundo de vida* no en el sentido de aislarlo de los procesos de modernización, sino de desarrollar instituciones que limiten la dinámica de los *sistemas de acción*; la única redefinición posible de las relaciones entre *sistemas* y *mundo de la vida* para salir de la crisis contemporánea está dada por la potencialidad comunicativa de los espacios públicos autoorganizados y no por la performatividad funcional dada por la eficacia de los *sistemas de acción*.

### Reflexiones finales

En relación con el carácter institucional y meta-institucional del lenguaje, las ventajas y mayores alcances teóricos de la perspectiva pragmática están en dependencia de la gran potencia teórica inicial del *giro pragmático* para dar cuenta de las múltiples capacidades realizativas del lenguaje y su vinculación con los factores extra-lingüísticos y las fuerzas sociales e institucionales intervinientes en las acciones lingüísticas. En este sentido, las reflexiones de Wittgenstein representan verdaderas contribuciones originales que aportan elementos de importancia, ineludibles a la hora de encarar un estudio pragmático del lenguaje centrado en su carácter institucional. Los distintos aspectos resaltados de la teoría de Wittgenstein detentan la fuerza heurística propia de las invenciones intelectuales que poseen la capacidad de abrir todo un ámbito de reflexión y marcar el sendero de las investigaciones



posteriores, por lo que constituyen el suelo común al que remiten los abordajes pragmático-institucionalistas del lenguaje.

En cuanto a Habermas, en primer orden, es necesario resaltar que su lectura de los procesos contemporáneos y de la participación del lenguaje en ellos, no sólo es realmente abarcadora y de gran alcance, sino también de una complejidad e interés de verdadero relieve. Es claro que el análisis de las relaciones entre lenguaje e instituciones adquiere en Habermas un nivel de integración muy superior a los presentes en el grueso de los autores de la tradición analítica-pragmática; la introducción en forma decisiva de la intersubjetividad como base y ámbito de la actitud realizativa en el lenguaje y la incorporación de elementos y procesos macro-sociales en el análisis sitúan su abordaje teórico en un nivel de complejidad de mayor amplitud. Ésta es una característica que los autores del pragmatismo anglosajón tienden a no incorporar, por lo que el doble registro en que se mueve Habermas (lenguaje como componente meta-institucional determinante en la institucionalización y procesos de institucionalización que especifican y sirven de reaseguro a las funciones del lenguaje comunicativo) sin dudas otorga nuevo relieve y textura al análisis institucionalista del lenguaje.

Ahora bien, esto convive con distintos puntos discutibles dentro de la TAC en lo que refiere específicamente a su reapropiación de la teoría wittgensteiniana y a la relación entre lenguaje y procesos de institucionalización. Por ello, a continuación realizamos señalamientos particulares acerca de la forma en que Habermas retoma la visión del lenguaje de Wittgenstein y luego concluimos con una evaluación de la propuesta habermasiana sobre nuestra temática.

Habermas considera que el fin inmanente del lenguaje es el entendimiento y la producción de acuerdos discursivos e, independientemente de si es acertado tomar esa función como la original y fundamental, se basa para ello en las reflexiones de Wittgenstein sobre el lenguaje: “con Wittgenstein estoy convencido que «lenguaje» y «entendimiento» son conceptos cooriginarios, conceptos que se explican mutuamente” (*Teoría... Complemento* 417). También sirve de claro ejemplo de esta interpretación la siguiente cita:

Wittgenstein se ha percatado de que el concepto de acuerdo reside en el concepto de lenguaje. Sólo en un sentido auto-explicativo podemos decir que la comunicación lingüística

«sirve» al acuerdo. Todo acuerdo se acredita, tal y como nosotros decimos, en un consenso racional; de lo contrario no es ningún acuerdo «real». Los hablantes competentes saben que todo consenso obtenido fácticamente puede ser engañoso; pero en la raíz del concepto de consenso engañoso (o meramente forzado) deben haber puesto ya el concepto de consenso racional. (Habermas, *Teoría y praxis* 29)

En esta reapropiación de la teoría wittgensteiniana consideramos que Habermas incurre en dos errores fundamentales. El primero, referido a la interpretación de Wittgenstein en forma específica, para Wittgenstein –especialmente en *Investigaciones filosóficas* y *Sobre certeza*– el concepto de acuerdo no reside en el lenguaje, sino, por el contrario, el funcionamiento del lenguaje únicamente puede ser entendido en base a los acuerdos ya dados en las *formas de vida* compartidas, son éstas las que sirven de fundamento a los acuerdos reflejados en el lenguaje o bien que el lenguaje ayuda a conformar. Para Wittgenstein, en todo caso, los que son cooriginarios son acuerdo y *formas de vida*, justamente la posibilidad de entendimiento y consenso dentro del lenguaje está basado en la coincidencia de *formas de vida*. Habermas reduce todo consenso válido al cumplimiento de las cuatro pretensiones universales dadas en la base de validez del habla (inteligibilidad, verdad, rectitud y veracidad), pero con esto produce una reducción lingüística de los consensos a todas vistas indebida si partimos de los principios wittgensteinianos acerca de que el consenso puede expresarse en el lenguaje, pero no necesariamente parte del lenguaje y de sus pretensiones de validez.

Y el segundo error, devenido del primero, según las premisas pragmáticas en la raíz de todo consenso (sea engañoso o no) no se encuentra el consenso racional, sino el consenso de *formas de vida*; el consenso racional únicamente puede desempeñar una función dentro de las comunidades que lo plantean como un objetivo regulativo (y subrayamos aquí lo de regulativo, no lo de constitutivo), de lo contrario no puede ser entendido como siempre ya funcionando, tal como pretende Habermas de la *situación ideal de habla* como elemento de tensión ya operante en la facticidad de la realidad. Esta interpretación forzada de Wittgenstein no es casual dentro de las reflexiones habermasianas, sino que está emparentada con su afán de establecer la prioridad de la función del entendimiento por sobre cualquier otra a través de la hipostatización que realiza de dicha función dentro del lenguaje y la

posterior absolutización de la acción comunicativa como único modelo válido de acción lingüística.

En lo que respecta a la relación entre lenguaje e institucionalización, en Habermas existe una hipostatización de la función comunicativa ya que únicamente los procesos de institucionalización basados en el modelo del habla comunicativa son entendidos como válidos, hipostatización que se revela como indebida tanto porque el lenguaje en su totalidad –no sólo el uso comunicativo– contribuye a la institucionalización de los distintos ámbitos de racionalidad y esferas de acción (cognitivo-instrumental, estético-expresivo y jurídico-moral), como porque es posible institucionalizar los diversos usos del lenguaje, no únicamente el comunicativo. Habermas desconoce la posible autonomía y capacidad normativa incardinada en aquellas acciones que no sean comunicativas; al ser éstas el único criterio para el análisis y comprensión de las demás, los otros tipos de acción se ven reducidos a modelos defectivos de interacción y carecen de valor positivo en la constitución e institucionalización de prácticas racionales en la esfera social. Al tomar la parte por el todo en sus análisis (*una* función del lenguaje como *la* función), Habermas comete con el lenguaje, aunque en dirección conceptual inversa, una sinécdoque idéntica a la que habrían realizado los primeros frankfurtianos con la razón según la acusación del mismo Habermas (tomar *un* tipo de razón, la instrumental, como *la* forma de racionalidad moderna).

La acción comunicativa como mecanismo de coordinación de la acción y el orden social racionalmente estructurado necesita del sometimiento de los participantes a la fuerza del mejor argumento en pos de lograr acuerdos objetivamente válidos; sin embargo, la instauración de las condiciones que los hacen posibles depende de la coacciones y luchas mediante relaciones institucionalizadas de poder –en las que también participa el lenguaje– antes que de la interacción comunicativa libre. El consenso discursivo debe sobreponerse a las prácticas y procedimientos jerárquicos que dentro de marcos institucionalizados buscan su perpetuación a través de las acciones lingüísticas; en relación con esto, la propuesta de Habermas adolece de partir de acciones comunicativas pensadas en base a condiciones fácticas y normativas de igualdad, por lo que no tiene debidamente en cuenta los problemas generados por un entramado socio-político con abismales desigualdades donde el debate y las decisiones públicas lejos están de poder comprender a todos los implicados. Al estar institucionalizadas las coerciones y las inequidades

es imposible un ámbito institucional que esté exento de coacciones estructurales; del mismo modo, que al portar el lenguaje dentro de sí componentes y prácticas que son consecuencia de dominaciones e imposiciones sociales, es imposible un uso del lenguaje que esté libre por completo de coacciones.

En su intento de asegurar los requisitos incluidos en este concepto formal de racionalidad y aplicar las condiciones de la racionalidad comunicativa a los demás tipos de interacción, Habermas somete su programa teórico a una idealización creciente y necesita basarse en un análisis reconstructivo ahistórico, por lo que quita entidad al dominio fáctico –vía los productos de la racionalidad técnica– engarzado con la liberación de las fuerzas racionales, subordina las consecuencias patológicas de la modernidad a meras desviaciones de la dinámica evolutiva y se desentiende mayoritariamente de las retroalimentaciones no unidireccionales entre la esfera comunicativa-deliberativa y la esfera político-estratégica. La TAC, en cuanto teoría reconstructiva formal, toma como único modelo a los actos de habla en que se cumplen las pretensiones universales de validez y, por tanto, prescinde de desarrollar elementos teóricos y metodológicos que le permitan abordar las condiciones fácticas que hacen posible el surgimiento y realización de los procesos discursivos, por lo que las conexiones entre su amplio registro conceptual y las situaciones y usos reales se ven notoriamente debilitadas y la pragmática lingüística queda *encorsetada* en líneas formales reconstructivas que la orientan con exclusividad hacia un uso deliberativo del lenguaje sólo dable en esferas limitadas y, además, dependiente de factores extra-discursivos no regulables por las prácticas comunicativas. Por ello, si bien la TAC presenta desafíos y diagnósticos que no carecen de interés, las sucesivas abstracciones a que es sometido el funcionamiento lingüístico por parte de Habermas conducen a que desaproveche buena parte de los hallazgos teóricos y la potencia heurística de los abordajes institucionalistas del lenguaje propios de la perspectiva pragmática.

REFERENCIAS BIBLIOGRÁFICAS

Cabanchik, S., Penelas, F. y Tozzi, V. (comp.). *El giro pragmático en filosofía*. Barcelona: Gedisa, 2003. Impreso.

Habermas, J. *Teoría y praxis. Estudios de filosofía social*. Madrid: Tecnos, 1987. Impreso.

---. *Conocimiento e interés*. Madrid: Taurus, 1990. Impreso.

---. *Teoría de la acción comunicativa: complemento y estudios previos*. Madrid: Cátedra, 1994. Impreso.

---. *Facticidad y validez. Sobre el derecho y el Estado democrático de derecho en términos de teoría del discurso*. Madrid: Trotta, 1998. Impreso.

---. *Perfiles filosófico-políticos*. Madrid: Taurus, 2000. Impreso.

---. *Teoría de la acción comunicativa I. Racionalidad de la acción y racionalización social*. Madrid: Taurus, 2003. Impreso.

---. *Teoría de la acción comunicativa II. Crítica de la razón funcionalista*. Madrid: Taurus, 2003. Impreso.

Naishtat, F. *Problemas filosóficos en la acción individual y colectiva: una perspectiva pragmática*. Buenos Aires: Prometeo, 2004. Impreso.

Rorty, R. (ed.). *The linguistic Turn: Recent essays in philosophical method*. Chicago: University of Chicago Press, 1967. Print.

Wittgenstein, L. "Cause and effect: Intuitive awareness". *Philosophia*. Sep. 1976: 409-425. Print.

---. *Sobre la certeza*. Barcelona: Gedisa, 1991. Impreso.

---. *Los cuadernos azul y marrón*. Madrid: Tecnos, 1998. Impreso.

---. *Tractatus lógico-philosophicus*. Madrid: Alianza, 1999. Impreso.

---. *Investigaciones filosóficas*. Barcelona: UNAM-Crítica, 2004. Impreso.

# EL CAMPO DE LA ARGUMENTACIÓN ESENCIALMENTE POLÉMICA\*

*THE FIELD OF ESSENTIALLY CONTESTED ARGUMENT*

JULDER A. GÓMEZ

Universidad Eafit, Colombia. jgomezp5@eafit.edu.co

RECIBIDO EL 11 DE JULIO DE 2012 Y APROBADO EL 5 DE OCTUBRE DE 2012

## RESUMEN ABSTRACT

Este artículo hace parte de un proyecto de investigación que busca determinar una clase de desacuerdos comprensible como un campo de argumentación que no se produce porque haya malentendidos, sino por la naturaleza de algunos fenómenos sociales. Este artículo se aproxima a ese campo mediante la noción de los conceptos esencialmente polémicos (Gallie) y arguye que se puede comprender satisfactoriamente, siguiendo la retórica clásica, como un campo constituido por una clase de cuestiones de definición. Para ello se disponen dos partes, en la primera, se exponen la noción de concepto esencialmente polémico de Gallie y algunas notas de la recepción de su propuesta; en la segunda parte, se expone la noción de campo de argumentación, el debate en torno al modo en que los campos de argumentación pueden ser comprendidos y la cuestión de cómo entender el campo de argumentación relativo a los conceptos esencialmente polémicos.

This article is part of a research project that seeks to determine a class of disagreements understandable as a field of arguments that do not obtain due to misunderstandings but because of the nature of some social phenomena. This paper approaches this field through the notion of essentially contested concepts (Gallie), and argues that one can understand this kind of disagreement successfully, following classical rhetoric, as a field constituted by a class of questions of definition. This article has two parts: the first explains Gallie's notion of essentially contested concept and some notes on its reception. The second part clarifies the notion of field of argumentation, the debate on how to understand this notion, and the question of how to understand the field of argumentation related to essentially contested concepts.

## PALABRAS CLAVE KEY WORDS

campo de argumentación, conceptos esencialmente polémicos, definición, Gallie, retórica.

field of argumentation, essentially contested concepts, definition, Gallie, rhetorics.

---

\* Este artículo es producto de la investigación: "Sobre la posibilidad objetiva de la aplicación de las reglas retóricas a los desacuerdos relativos a la acción social", realizada en el grupo de investigación "Estudios sobre Política y Lenguaje", de la Escuela de Ciencias y Humanidades de la Universidad Eafit. Este artículo ha sido realizado durante el período sabático que la Universidad Eafit me ha concedido en el segundo semestre del año 2011.

## Gallie, conceptos esencialmente polémicos

A pesar de que puede ser adoptada como un marco teórico para el estudio de importantes problemas relativos a conceptos sociales (Collier, Hidalgo, and Maciuceanu 215), la propuesta de Gallie tiene la vocación de una hipótesis explicativa de algunos desacuerdos inevitables. Gallie, en contra de una práctica habitual desde la modernidad, pretende que hay casos de desacuerdo para los cuales las explicaciones psicológicas no son correctas (Gallie). El filósofo británico ofrece una hipótesis explicativa que tiene que ver con los conceptos. Él describe estos desacuerdos como casos en los cuales hay grupos de personas que discuten entre sí acerca del uso apropiado de conceptos carentes de una regla de uso aceptable para todos; los describe como casos en los que la conciencia de la multiplicidad de usos y funciones del término en cuestión, conciencia agenciada en los querellantes, no sólo no motiva la disolución de la discusión, sino que además no obsta para que las partes involucradas en la disputa sigan reclamando que su uso del término es el apropiado.

Estos desacuerdos no son, ni enteramente sustantivos o sobre las cosas, ni enteramente lingüísticos o sobre el significado de las palabras (Miller 40-2 48-9). Si por sustantivos se entienden los desacuerdos en los cuales las partes involucradas comparten enteramente la definición de los términos que aplican y discrepan acerca de si algo reúne o no las características que la definición exige, o si algo es o no tan valioso como algo más, entre otros, entonces, los desacuerdos conceptuales que Gallie pretende explicar no son sustantivos, pues en ellos las partes pueden no estar de acuerdo en alguna definición del término que aplican. Del mismo modo, si por lingüísticos se entienden los desacuerdos en los cuales las partes involucradas asienten a una misma descripción de todos los elementos pertinentes para la discusión, pero discrepan acerca del significado de una palabra, entonces, los desacuerdos conceptuales que Gallie pretende explicar no son lingüísticos porque en ellos los querellantes no aceptan la misma descripción de todos los elementos relevantes para la disputa.

Se ofrece como ejemplo de desacuerdo sustantivo el caso en que dos personas discuten sobre si es bueno abortar alegando, la una, que no lo es, porque el feto tiene derecho a vivir y alegando, la otra, que sí lo es, porque la madre tiene derecho a controlar su cuerpo; él señala, que el anterior no es un ejemplo de desacuerdo sobre el concepto de bueno



porque en él, quienes discuten pueden aceptar por igual una definición tal como la de:

llamar a algo bueno es prescribir que sea hecho por cualquiera que se encuentre en una situación determinada e implicar que esta prescripción puede ser justificada con referencia a los intereses humanos. (Miller 40-2)

Como ejemplo de desacuerdo lingüístico ofrece el caso de una discusión sobre si los términos “voluntario” e “involuntario” cubren todo el campo de la acción, de suerte que toda acción tenga que ser o bien voluntaria o bien involuntaria o si el significado de estos términos es tal que a algunas acciones no se les puede aplicar ninguno de los dos; una situación como esta se distinguiría de un desacuerdo conceptual en que ella requiere un acuerdo en la descripción de las acciones que no se constata en las discusiones conceptuales que Gallie intenta explicar.

Gallie define los conceptos que generan esta clase de desacuerdos, a los cuales llama conceptos esencialmente impugnados o polémicos, mediante las siguientes condiciones: (i) deben ser conceptos evaluadores en el sentido que significan o acreditan algún tipo de logro valorado. (ii) Este logro debe tener un carácter internamente complejo, su valor le debe ser atribuido como un todo. (iii) En consecuencia, cualquier explicación de este valor debe incluir referencias a las respectivas contribuciones de sus diversos aspectos o partes. (iv) El logro acreditado debe ser de un tipo que admita una modificación considerable a la luz de las circunstancias cambiantes. Y (v) para que el concepto cuente como esencialmente polémico cada uno de los usuarios del concepto debe reconocer que su propio uso es impugnado por otros usuarios y debe tener por lo menos alguna apreciación de los diferentes criterios a la luz de los cuales los otros grupos afirman que están aplicando el concepto en cuestión (Gallie 10-2).

A estas cinco condiciones hay que añadir otras dos para justificar la continuación del uso de un concepto esencialmente polémico, es decir, para establecer que los querellantes discuten acerca del mismo concepto, para explicar que los usuarios del concepto insistan en continuar utilizándolo, en lugar de distinguir un concepto distinto para cada uno de los usos rivales. Esas dos condiciones son: (vi) los usos rivales del concepto deben haberse derivado de un proceso de imitación y adaptación de un mismo modelo, abierto, complejo y diversamente

descriptible, que tanto puede tener la forma de un prototipo como de una tradición, cuya autoridad reconocen por igual todas las partes comprometidas en la disputa. Y (vii) debe ser verosímil que el logro valorado por el modelo representado no llegaría hasta el óptimo permitido por las circunstancias reales de no ser por la competencia continua entre quienes pretenden que su propia interpretación del modelo es la mejor (Ibíd. 16 20).

Cada una de estas siete condiciones ha sido valorada e interpretada de diferentes modos a lo largo de la discusión de la tesis de Gallie (Collier, Hidalgo, and Maciuceanu 216 222) y, con frecuencia, la tesis ha sido concebida haciendo énfasis en algunas condiciones más que en otras. Gray asume como tenor principal de la propuesta de Gallie una idea consistente en que la mayor fuente de controversias conceptuales reside en que hay conceptos cuyos criterios de uso incorporan estándares normativos, es decir, en la interpretación de Gray, un concepto es esencialmente polémico si los usos rivales del concepto invocan un trasfondo normativo o una serie de juicios de valor (Gray 392). Grafstein le concede mayor importancia al carácter internamente complejo y múltiplemente descriptible de los conceptos esencialmente polémicos; a su juicio, la polemicidad conceptual se deriva de que los logros acreditados por los conceptos esencialmente polémicos se pueden describir de distintos modos (Grafstein 24). MacIntyre encuentra en el carácter abierto, histórico, de los conceptos esencialmente polémicos la razón de ser de la imposibilidad de clausurar las disputas en torno a ellos (MacIntyre 35). En cambio, ninguno de los estudiosos de la tesis de Gallie hace énfasis especial, ni en el carácter reconocido de la disputa, ni en la derivación de los usos conceptuales a partir de un modelo, ni en la especial relación entre la continuidad de la discusión y la optimización del logro acreditado por los conceptos que lo refieren.

En cuanto al dominio y extensión de esta clase de desacuerdos, Gallie los inscribe en los ámbitos de la estética, la filosofía política, la filosofía de la historia y la filosofía de la religión (Gallie 6); Connolly muestra su presencia en el ámbito de las ciencias políticas (Connolly 45 179); Lakoff afirma que cubren todo el conjunto de las ideas procedentes de la ilustración (Lakoff); y un amplio número de estudiosos ha tratado como esencialmente polémicos una serie importante de conceptos, entre ellos se pueden nombrar los conceptos de liberalismo (Abbey), desarrollo sostenible (Connolly), interés, poder, libertad (Connolly; Lukes), amor (Hamilton), autoridad, responsabilidad, equidad, igualdad (Lakoff), bien

(Mason), terrorismo (Moten), acceso carnal violento (Reitan) e imperio de la ley (Waldron), entre otros.

### La cuestión del campo de la argumentación esencialmente polémica

Si los conceptos esencialmente polémicos permiten delimitar un ámbito o campo de argumentación y de controversias debe ser posible describirlo más allá de su carácter esencialmente polémico, debe ser posible caracterizar los criterios de la elección y evaluación de argumentos que rigen en él, debe ser posible caracterizar el tipo de racionalidad y validez al que en él tiene sentido aspirar. Pero antes de formular preguntas en este sentido es necesario considerar una dificultad previa a la que en gran medida se destina este artículo: ¿hay realmente una clase única de desacuerdos para la cual es adecuada como explicación la propuesta referida? Dicho de otro modo, ¿pueden pertenecer a un mismo tipo de desacuerdos los relativos a cosas tan disímiles como lo son la política, el arte, la religión, entre otros?

Esta pregunta presenta una objeción. Respecto de ella, hay que empezar por reconocer que desde la publicación de *Los usos de la argumentación* (Toulmin), por lo menos, la noción de campo de argumentación es muy importante y puede ser interpretada de una manera tal que resulte difícil comprender en qué sentido puede constituirse un campo de argumentación a partir de los desacuerdos de ámbitos tan aparentemente distantes entre sí.

La noción de campo de argumentación (*field of arguments*) se introduce en *Los usos de la argumentación* para plantear un problema que surge cuando se intentan contestar las preguntas ¿cómo se elabora un argumento justificatorio de una pretensión? Y ¿cómo se evalúan esos argumentos? El problema que surge cuando se procede al desarrollo de estos interrogantes es que estas preguntas suponen que, a pesar de la multiplicidad de campos en que se ofrecen argumentos, hay elementos generales, invariables respecto al campo, que pueden ser tenidos en cuenta para la resolución de los interrogantes, el problema consiste en saber,

¿hasta qué punto los argumentos justificatorios adoptan una sola y la misma forma o apelan a una misma serie de estándares en todos los diversos tipos de casos que

tendremos ocasión de considerar? Y en consecuencia, ¿hasta qué punto, cuando evaluamos los méritos de estos argumentos diferentes, podemos basarnos en el mismo tipo de cánones o estándares de argumentación para realizar su crítica? ¿Poseen el mismo tipo de méritos o son diferentes? ¿Y en qué aspectos podemos buscar uno solo y el mismo tipo de valor en un argumento de todas esas clases? (Toulmin 32-3)

Toulmin afirma que “dos argumentos pertenecen al mismo campo cuando los datos y las conclusiones en ambos argumentos son, respectivamente, del mismo tipo lógico” (Ibíd. 33).

El filósofo británico ilustra la noción de campos de argumentación con los siguientes ejemplos: un problema puede ser adoptar una opinión frente a la cuestión de quién será seleccionado en el equipo americano para jugar la copa Davis contra Australia, si frente a él se adopta una posición, favoreciendo a un jugador en lugar de a otros, como a Budge Patty, el tipo de hechos y de argumentos con los que esta posición se puede justificar tendrá en cuenta cosas como el estado físico de los jugadores. Pero si el problema es sobre la cuestión de saber si fue justa la condena de Crippen por la muerte de su esposa y si la opinión adoptada frente a esto es que no lo fue, el tipo de hechos y de argumentos a elaborar tendrá en cuenta cosas como las pruebas aducidas durante el proceso jurídico. Mas, si la cuestión es la de si el pintor Piero della Francesca merece los elogios de sir Kenneth Clark y la opinión a justificar es escéptica, el tipo de hechos y de argumentos a elaborar se referirán a los rasgos característicos de la pintura de Piero y a la importancia que sir Kenneth Clark les otorga. Y si la cuestión es si la teoría de la super-conductividad del profesor Frölich es satisfactoria y la opinión es positiva, los argumentos versarán sobre los experimentos de super-conductividad y sus relaciones con las predicciones de la teoría de Frölich. Ahora que, si la cuestión es la de cuándo tendrá lugar el próximo eclipse de Luna, para defender una opinión definida se tendrán en cuenta las posiciones pasadas y presentes de la misma. O si, por último, la cuestión es cuál sea la naturaleza exacta de la relación entre los ángulos rectos de un triángulo equilátero, los argumentos justificatorios de una respuesta tendrán en cuenta cosas como los axiomas de Euclides (Toulmin 31-2).

Dada la multiplicidad de campos en los que los argumentos tienen lugar, el problema al que se enfrentan las primeras dos preguntas de Toulmin

(¿cómo se elaboran y cómo se evalúan, en general, los argumentos?), consiste en determinar hasta qué punto hay justificación para suponer que es posible apelar a un mismo procedimiento y a un mismo conjunto de estándares en la elaboración y evaluación de argumentos.

La noción de campo de argumentación adquiere una mayor determinación en la respuesta que Toulmin ofrece a este problema. A pesar de todas las diferencias que hay entre los campos, en todos los campos de argumentación se justifican las opiniones o pretensiones mediante enunciados particulares, hechos o datos que resultan pertinentes como elementos justificatorios en virtud de la aceptación de enunciados universales o garantías que se expresan en la forma de reglas o leyes cuya aceptación se deriva del reconocimiento de la autoridad de una fuente ya de conocimiento o ya de prescripción. Del mismo modo, a pesar de las diferencias entre las garantías de los distintos campos de argumentación, en todos ellos éstas le confieren una cierta modalidad de aceptación a la conclusión, una modalidad que se expresa mediante términos como necesariamente, posiblemente, presumiblemente y demás, en todos ellos las garantías son aplicables a los datos bajo ciertas condiciones y no son aplicables bajo otras que por ello se conciben como excepciones. Así, pues, a pesar de todas las diferencias, en todos los campos hay datos, garantías, respaldos, modalizadores y condiciones de excepción que se articulan entre sí para la justificación de conclusiones. Según esto, un campo difiere de otro en cuáles son los elementos que integran cada una de estas categorías y en cómo varían las relaciones que estos elementos observan entre sí (Toulmin 132 143).

Pero ¿cómo se definen los campos de argumentación? ¿Cuál es el enunciado de las características constitutivas y diferenciadoras de los campos de argumentación? ¿Cuáles son los elementos que hay que tener en cuenta para realizar una descripción, un análisis y una evaluación de una argumentación determinada, parte de una controversia concreta, en un ámbito dado? La cuestión de la definición de los campos de argumentación ha conocido diversas respuestas.

Un campo de argumentación puede ser concebido como un dominio temático. Lo que es pertinente ofrecer como argumento depende del tema de la discusión. Lo que es pertinente y aceptable como argumento depende de las garantías relativas a un tema que gozan de aceptación en una comunidad específica. Las garantías relativas al tema se erigen así en el criterio mediante el cual se selecciona la información que puede ser

aportada como justificación y también como criterio de la evaluación de la pertinencia y aceptabilidad de esa información. Por tanto, el campo de argumentación puede concebirse como delimitado por la noción de tema y esta noción puede concebirse como el punto de partida de su definición (Klumpp 49-0). Sin embargo, esta concepción es insuficiente, porque no consigue dar cuenta de las discusiones en las que una pluralidad de temas se encuentran inextricablemente imbricados, como lo es la discusión acerca del aborto; en ella, en efecto, son pertinentes las garantías procedentes de dominios temáticos tan diversos como la ley, la ética, la política y la medicina (Rowland, *Argument fields* 470-71).

Dificultades semejantes se le plantean a la concepción de los campos de argumentación como dominios disciplinares. Un campo de argumentación puede concebirse como un dominio disciplinar porque las comunidades de estudiosos organizados de manera disciplinar son las encargadas de establecer lo que respecto de cada asunto cuenta como válido y racional (McKerrow; Willard). Puede concebirse cada campo como lo que sus estudiosos hacen. Pero esta forma de concebir los campos de argumentación es insuficiente porque hay discusiones en las que se consideran problemas desde el punto de vista de diversas disciplinas (Rowland, *Argument fields* 471-72).

Una tercera opción consiste en definir los campos de argumentación como perspectivas. En el ámbito académico, las perspectivas de las escuelas de pensamiento definen lo que es y lo que no es relevante en un campo. Es evidente que tienen muchas más cosas en común dos miembros de una misma escuela que dos miembros de una misma disciplina. Así, por ejemplo, dos conductistas, uno psicólogo y el otro sociólogo, tienen muchos más puntos de vista en común que dos psicólogos de distintas escuelas. Esto, también explicaría cómo se pueden integrar en un mismo campo consideraciones procedentes de diversas disciplinas. En el ámbito del lenguaje ordinario, las perspectivas ligadas a la manera en que los agentes definen sus acciones determinan los campos de las discusiones. Así, una persona que entra en una discusión determinada la delimita de manera distinta según adopte la posición de cónyuge, vecino, religioso, entre otros, bien entendido que cada cual se verá abocado a adoptar tantas posiciones cuantos tipos de actividades y de relaciones le correspondan (Willard, *Field theory* 22-26). Pero tampoco esta manera de concebir los campos de argumentación parece satisfactoria. En primer lugar, no es evidente que haya equivalencia entre la perspectiva entendida como manera en que los intereses y actividades



cotidianas de un agente delimitan un asunto y la perspectiva entendida como la manera en que una escuela de pensamiento académico delimita un asunto. En segundo lugar, las escuelas de pensamiento no siempre tienen tanto influjo como parece suponerse. Tienen muchas más cosas en común un estudioso de la crítica literaria marxista y cualquier otro estudioso de la crítica literaria que un estudioso de la crítica literaria marxista y un economista marxista. En tercer lugar, existe la posibilidad de que en algunos casos la adopción de una escuela de pensamiento sea producto de, y no condición para, la participación en una discusión (Rowland, *Argument fields* 473-74).

Otra manera sociológica de definir los campos de argumentación consiste en concebirlos como dominios contextuales. Toda argumentación se realiza en un contexto social y para la audiencia de ese contexto. Es posible definir lo que es relevante, lo que es válido, lo que es racional, lo que puede suponerse y, en fin, el campo, a partir de la audiencia a la cual se dirige la argumentación. Mas, para ser una definición de la noción de campo de argumentación que permita una clasificación de los campos esta concepción requiere una clasificación de los contextos o de las audiencias (McKerrow, *Argument communities*). Con este fin se puede distinguir entre contextos primarios, secundarios y terciarios. Los contextos primarios son personales, sociales, técnicos o filosóficos. Ejemplos de contextos secundarios son ley, negocios, medicina, religión, ciencia, entre otros. Son terciarios la nación, el Estado, la clase social, entre otros. Los primarios son tales porque cuando una discusión tiene lugar en alguno de los contextos llamados secundarios siempre lo hace con algunas dimensiones del contexto primario en mente. Los contextos terciarios son características que el agente del discurso, voluntaria o involuntariamente, lleva al discurso argumentativo (Ibíd. 28-30). La objeción que se le plantea a esta definición de los campos consiste en que hay argumentaciones que no caen bajo ninguno de estos contextos y hay otras que pueden caer bajo varios contextos a la vez. La argumentación científica, por ejemplo, no es social, porque no está dirigida al auditorio para su aprobación, sino que está reglamentada por criterios que no obligan al auditorio menos que al orador. Tampoco es filosófica la argumentación científica, porque no es abstracta y conceptual, sino empírica y constatable. Por supuesto tampoco es personal. Y del otro lado, una discusión política particular, por ejemplo, puede desarrollarse simultáneamente en todos los contextos señalados (Rowland, *Argument fields* 474-75).



Una quinta forma de definir un campo de argumentación consiste en delimitarlo como el dominio de una forma de argumentación. Esta manera de concebir los campos puede, por ejemplo, proceder a través de una distinción entre campos de argumentación inductiva, deductiva, analógica, entre otros. Y, a partir de ello, una especificación de campos menores en virtud de las características que adopta la argumentación en cada uno de los campos mayores. Pero esta concepción se enfrenta inmediatamente con que, de un lado, una misma clase de argumentos puede ser empleada a propósito de muy distintos asuntos y, de otro lado, a propósito de un mismo asunto en una discusión determinada se pueden ofrecer argumentos de muy distintas clases (Ibíd. 471).

Una sexta manera de definir los argumentos consiste en concebirlos como el producto de propósitos compartidos. Parece claro que los temas, las formas de argumentación, las perspectivas adecuadas y demás factores integrantes de un campo de argumentación surgen, no por azar, sino como consecuencia de que una comunidad determinada tiene una serie de propósitos compartidos. Por tanto, parece que un campo de argumentación puede definirse a partir de los propósitos que en él se persiguen (Ibíd. 476-479). Tampoco esta concepción parece completamente satisfactoria porque sólo puede dar cuenta de los campos de argumentación cooperativa. En las discusiones acerca del aborto, por ejemplo, no todos los participantes tienen el mismo propósito. Unos quieren defender a la mujer, otros quieren defender la vida de los fetos y otros más quieren regular con alguna claridad una práctica que se percibe como problemática. La concepción de los campos a partir de los propósitos compartidos se ve obligada en este caso a hablar de tres campos: el campo de las feministas, el campo de los fundamentalistas y el de los jueces (Rowland, *Argument fields* 498). Pero es evidente que estos no son tres campos de una argumentación, sino tres posiciones no cooperativas a propósito de una discusión a la que le corresponde un sólo campo.

Hay, sin duda, otras maneras de definir los campos de argumentación. Pero el debate entre las diversas concepciones de cuáles son los criterios adecuados para la definición y clasificación general de los campos de argumentación parece ofrecer tres moralejas disuasivas: primero, cada una de las concepciones es más o menos adecuada para algunos casos, pero ninguna consigue una definición adecuada a todos los casos ni, por tanto, una clasificación exhaustiva. Segundo, las diversas concepciones de los campos de argumentación hacen énfasis en diversos aspectos de

algunos campos de argumentación, pero no son incompatibles entre sí. Por tanto, tercero, aunque no parezcan asequibles una definición y una clasificación de los campos sí es posible determinar cada vez el campo de cada caso, de cada una de las controversias particulares, mediante una integración de los elementos más relevantes para el caso de las teorías propuestas acerca de los campos de argumentación (Rowland, *Purpose argument* 238).

Así, pues, hay que reconocer la importancia de la noción de los campos de argumentación. Esta noción sirve para indicar que en distintas argumentaciones puede haber distintos criterios para producir y para evaluar argumentos. La noción sirve para advertir de los peligros de una generalización apresurada de las características de algunas argumentaciones. Dada la enorme variedad de discusiones posibles, dada la multiplicidad de aspectos en los que las argumentaciones pueden variar entre sí, no parece asequible y, en todo caso, nunca se ha aclarado cuál es la necesidad de, una definición y una clasificación, adecuada la primera y exhaustiva la segunda para todos los casos.

De ello se desprende, a su vez, que la pretendida confutación contenida en la pregunta de si pueden pertenecer a un mismo campo de argumentación los desacuerdos relativos a cosas tan disímiles entre sí como política, arte, religión, entre otros, esa pretensión, supone una distinción de los campos de argumentación, probablemente fundada en las nociones de tema y disciplina, que no es apremiante. Es posible que para ciertos propósitos la distinción subyacente a la pregunta sea adecuada, pero a buen seguro no lo es para todos los propósitos. Así, pues, se plantea la pregunta por los criterios adecuados para una comprensión de los campos de argumentación que permita entender las discusiones argumentativas aparentemente interminables acerca de los asuntos aquí llamados esencialmente polémicos.

Pero antes de avanzar en la dirección de esta pregunta hay que insistir en el intento de señalar los límites de la tarea: la pregunta supone que hay un número de discusiones argumentativas interminables por razones esenciales. Se da por supuesto que bajo ciertos criterios esas discusiones hacen parte de distintos campos, por ejemplo, bajo criterios temáticos y disciplinares. Ser esencialmente interminables, no es la única característica de esas discusiones y no es necesario suponer que sea la más importante; pero es una característica importante para la comprensión del desarrollo de estas discusiones y para establecer qué tiene sentido

y qué no tiene sentido pretender racionalmente en estas discusiones. Con este fin parece conveniente preguntar por ellas como campo de argumentación. El preguntar por ellas como campo de argumentación puede ayudar a comprender su carácter esencialmente polémico, porque la teoría de un campo de argumentación puede contribuir al esclarecimiento del por qué se producen acuerdos y desacuerdos en él. En efecto, el propósito de las teorías de los campos es ayudar a entender la manera en que se producen y se evalúan argumentos en ellos; hay acuerdos cuando hay criterios para evaluar cuál es el mejor argumento en una situación discursiva determinada y no los hay en caso contrario; así que el acuerdo y el desacuerdo son, por lo menos, en parte, productos de los modos en que se evalúan argumentos en los campos; por tanto, en la medida en que el propósito de las teorías de los campos sea contribuir a la comprensión de la evaluación de argumentos, en esa medida, las teorías de los campos contribuyen a la explicación del acuerdo y del desacuerdo. Por eso, para comprender mejor el desacuerdo persistente que caracteriza algunas discusiones, parece procedente preguntar por ellas como campo de argumentación. No hay razones para suponer que la respuesta a esa pregunta haya de tener la forma de una teoría general a partir de la cual sea posible describir y evaluar de manera exhaustiva las discusiones en cuestión. No obstante, parece lícito suponer que la respuesta a esa pregunta puede aportar criterios para la descripción y evaluación de algunos rasgos importantes de esas discusiones.

Así, pues, ¿qué criterios son adecuados para una comprensión de los campos de argumentación que permita entender las discusiones argumentativas aparentemente interminables acerca de los asuntos aquí llamados esencialmente polémicos? Como respuesta a esta pregunta, no se puede ofrecer ninguno de los considerados en las concepciones antes referidas de los campos de argumentación porque esta clase de polémicas, así lo muestran los ejemplos desde Gallie hasta Lakoff, tienen diversos temas, ocupan a diversas disciplinas tanto se pueden producir al interior de una escuela de pensamiento, como entre varias, tanto pueden tener lugar en contextos personales, como filosóficos, no se desarrollan mediante una sola clase de argumentos y habitualmente no son cooperativas. Un antecedente más adecuado para el desarrollo de esta pregunta puede ofrecerlo la retórica clásica.

En la retórica clásica, la pregunta por los campos de argumentación se inscribe en el estudio de la materia del discurso. En esta parte del estudio de la elaboración de la obra retórica se consideran los posibles objetos

del discurso. La clase y el número de los discursos retóricos posibles se determinan a partir de la clase y el número de los objetos accesibles a las técnicas del discurso retórico (Lausberg 100). El problema es aquí el de la división de la materia: ¿cuáles y cuántos son los “géneros” de los objetos y discursos retóricos? Para contestar esta pregunta se tienen en cuenta, como elementos fundamentales de la situación retórica, el orador, el discurso, el oyente y sus relaciones (Ibíd. 104). En esta problemática la retórica clásica se orienta, pues, por la situación. La pregunta obtiene a partir de los elementos de la situación, diversas y complementarias respuestas. De la consideración de la relación entre el orador y su objeto se deriva la división entre cuestiones civiles y cuestiones propias de un arte. De las relaciones entre el objeto del discurso y el auditorio o entre el orador y el auditorio se desprenden las divisiones aristotélicas de los géneros del discurso retórico. Asimismo, de la relación entre los discursos contrapuestos de las partes se obtiene la clasificación de los objetos y discursos según el punto sintáctico discutido, la pregunta o problema que debe hacerse quien puede decidir en la situación (Ibíd. 104-05).

La primera división, en cuestiones civiles y propias de un arte, es producto de la consideración de cuán fácil o accesible es el objeto del discurso para el orador. Aunque también se conoce una respuesta minimalista, esta división es una respuesta maximalista al interrogante de si la clase y el número de los discursos posibles es limitada o ilimitada (Ibíd. 100 103). La respuesta minimalista es afín a la concepción de la retórica como arte liberal, como *ars bene dicendi* y delimita los objetos del discurso retórico a los asuntos civiles, éticos y políticos para los cuales deben ser competentes todos los ciudadanos, a los asuntos cuyo conocimiento está al alcance de todos (Lausberg 83-4 101). La respuesta maximalista extiende los objetos posibles del discurso retórico hasta abarcar el ámbito de las artes específicas, ciencias, profesiones especializadas y cuestiones filosóficas, para las cuales se requieren conocimientos que ya no están al alcance de todos (Lausberg 101).

La segunda división, la aristotélica, entre discursos epidícticos, deliberativos y judiciales, es producto de la consideración de la relación entre el objeto del discurso y el oyente, o entre el hablante y el oyente, en la medida en que esta última está atravesada por el discurso. Aristóteles parte de la constatación de que el oyente puede estar llamado a contemplar el discurso o a tomar una decisión acerca del asunto discutido a partir de su captación del discurso. En el último

caso, la decisión puede versar sobre cosas futuras o sobre cosas pasadas. Cuando el oyente está llamado a contemplar el discurso, y la capacidad del orador, el discurso se llama epidíctico, su tiempo es el presente y su función el elogio y la censura. Cuando el oyente está llamado a decidir sobre el futuro, el discurso se llama deliberativo y su función es proporcionar consejo. Cuando, en fin, el oyente está llamado a decidir sobre el pasado, el discurso se llama judicial y su función es acusar o defender (Aristóteles 1358a 36 1358b 21).

La tercera división toma el género judicial como modelo y se concentra en la cuestión, el problema o la pregunta que debe contestar quien debe adoptar la decisión. En esta división se distinguen las clases de discursos según las clases de cuestiones. Y éstas, por su parte, se discriminan haciendo uso de tres criterios: primero, el grado de concreción de la cuestión; segundo, el grado de complejidad de la cuestión; y, tercero, el tipo de asunto discutido en la cuestión.

Haciendo uso del criterio de la concreción, se diferencian las cuestiones finitas de las cuestiones infinitas (Lausberg 118 121). Se entienden por finitas las cuestiones relativas a agentes y circunstancias espacial y temporalmente determinadas. Se entienden por infinitas las cuestiones que, como las filosóficas, trascienden las circunstancias espaciales y temporales. Esta distinción entre cuestiones finitas e infinitas se cruza, si bien sólo parcialmente, con la distinción entre cuestiones civiles y propias de artes especiales.

Mediante el criterio de la complejidad se clasifican las cuestiones en simples, conjuntas y comparativas (Ibíd. 118). Se les llama simples a las cuestiones en las que hay una sola cosa controvertida, conjuntas a las cuestiones en las que hay varias simples y, por último, se les llama comparativas a las cuestiones que consisten en una alternativa de dos o más miembros.

Por último, según el criterio del asunto discutido, del *status* o cuestión capital, se disciernen las cuestiones conjeturales, de definición, cualitativas y de translación (Ibíd. 122 148). Son conjeturales las cuestiones en las que se interroga si un hecho ha tenido lugar o no ha tenido lugar, de definición las que requieren la determinación de si el hecho que ha tenido lugar cae o no bajo un concepto o ley, cualitativas las que inquieren si el hecho ha tenido lugar de cierto modo y, para finalizar, de translación las que piden que se establezca si el llamado a decidir es competente para hacerlo.

Ahora bien, si frente a esta clasificación de los discursos y sus objetos se pregunta ¿para qué clase de desacuerdos son adecuadas las explicaciones referidas? ¿Cuál es el campo de la argumentación interminable, de la esencialmente polémica? Se llega a la consideración de que las explicaciones referidas son apropiadas para las cuestiones cuyo *status* es el de definición.

Aquí es donde se articulan con mayor claridad las explicaciones que se orientan por el texto de la narración y las que se orientan por el análisis de la definición. Esto resulta más comprensible si, a la manera usual en la retórica clásica, se adopta como modelo la argumentación jurídica y conforme al criterio de la división de las cuestiones, se adopta como punto de referencia la pregunta que debe plantearse el juez en la resolución de las controversias: las cuestiones de definición surgen cuando el acusado rechaza la denominación, "peligrosa" para él, con la que el acusador se refiere a un acto suyo. Estas cuestiones son posibles, porque no siempre hay una relación obvia y armónica, e incluso puede haber una relación conflictiva, entre la normalización del lenguaje contenida en la ley y la multiplicidad de los hechos en el mundo. El problema consiste, por tanto, en coordinar las palabras y las cosas. De allí que las cuestiones de definición abarquen dos dominios, el de las palabras y el de las cosas. De allí también, que el punto de partida de las argumentaciones contrapuestas en las controversias cuyo estatus es el de definición puedan ser tanto las cosas, como las palabras. Así se explica el que característicamente estas argumentaciones puedan tener la forma de un análisis de las palabras cuya aplicación es problemática, es decir, una definición, o de un análisis de los hechos a los que se pretende aplicar la palabra, es decir, una narración. De sobra se advierte que ambos procedimientos son parciales y encarecedores. Sin embargo, para el juez, la cuestión que se plantea no es tanto la del conjunto de posibilidades lingüísticas para la designación cuanto la de si el contenido del hecho justifica o no justifica la denominación con la que el acusador se refiere a él. La cuestión tiene, pues, la forma de sí o no este hecho particular y determinado cae bajo la esfera del concepto X. Se trata, en principio, de una cuestión finita, relativa a un particular; en efecto, la cuestión más general de la definición del término problemático surge sólo como condición previa para la determinación de sí o no este particular es un caso suyo y, en menor medida, como condición previa para la regulación de los posibles casos futuros (Lausberg 133 138).



Ahora bien, ni todas las cuestiones de definición son esencialmente polémicas, ni se caracterizan suficientemente por el uso argumentativo de las definiciones. De un lado, sólo son esencialmente polémicas las cuestiones que en el plano de la definición tienen las características explicadas por Gallie, Perelman y Lakoff, así como las cosas que, en el plano de su descripción y narración, tienen las características textuales e interpretativas explicadas por Ricoeur, a las que se puede aplicar lo que Gadamer entiende como literatura: “textos que no desaparecen, sino que se ofrecen a la comprensión con una pretensión normativa y preceden a toda posible lectura nueva del texto” (Gadamer 338). De otro lado, aún si estas cuestiones se nombran como de definición, es evidente la importancia en ellas de la descripción y de la narración de los hechos a los cuales se pretende aplicar la denominación problemática. Por esta última razón es importante distinguir con claridad entre el estudio del uso argumentativo de las definiciones y el estudio de las cuestiones de definición, en muchos sentidos este último puede aprender del primero, pero no debe agotarse en él.

Así las cosas, para la comprensión del campo de las cuestiones de definición esencialmente polémicas, vale la pena preguntarse ¿cómo se establecen en él los elementos impletivos de las categorías de la forma estándar de los argumentos determinada por Toulmin? ¿A cuáles de estos elementos se debe el señalado carácter de esencial polémicidad? Suponiendo que a las garantías, ¿qué particularidades se pueden observar en ellas si se las compara con las pertenecientes a otros campos de cuestiones de definición? Suponiendo que a los datos, ¿qué los distingue de los de otros campos de argumentación relativa a cuestiones de definición?

Es claro que para la resolución de estas preguntas resulta indispensable trazar una subdivisión del campo de las cuestiones de definición en términos de la polémicidad conceptual. Para ello puede ser muy oportuno proceder a partir de la división de los conceptos en conceptos incontestables o que no pueden dar lugar a discusiones (*uncontestable concepts*), conceptos concebiblemente contestables o que pueden dar lugar a discusiones (*conceivably contestable concepts*) y conceptos esencialmente contestables o esencialmente polémicos (*essentially contested concepts*) (Hurley 43 50). Los primeros son conceptos como el de rojo o el de sumar uno más cuatro. A la pregunta: ¿por qué se dice que esto es rojo? O ¿por qué uno más cuatro es cinco? No se contesta ofreciendo argumentos, sino reglas que constituyen el significado de esos



conceptos; si alguien agrupara un objeto rojo junto a otros verdes como reacción a la orden de mostrar cosas rojas, no se diría que discrepa de los demás, sino que no entiende el significado del concepto en cuestión, lo mismo si dice que uno más cuatro es algo distinto de cinco; en estos casos el descuerdo es imposible, la diferencia es aquí indicio de que se poseen distintos conceptos. Los conceptos concebiblemente contestables son conceptos categoriales que subsumen conceptos incontestables, se trata de conceptos como el de color o el de suma, son conceptos que normalmente no dan lugar a controversias porque sus criterios no se relacionan conflictivamente entre sí. Los conceptos esencialmente polémicos son los caracterizados por Gallie. A partir de esta clasificación de los conceptos en términos de su polemicidad, la tarea de la definición del campo de la polemicidad esencial consiste en establecer mediante comparación respuestas a las preguntas formuladas.

#### REFERENCIAS BIBLIOGRÁFICAS

Abbey, Ruth. "Is liberalism now an essentially contested concept?" *New Political Science*. Dec. 2005: 461-480. Print.

Aristóteles. *Retórica*. Madrid: Gredos, 1990. Impreso.

Collier, David, Fernando Daniel Hidalgo, and Andra Olivia Maciuceanu. "Essentially contested concepts: Debates and applications". *Journal of Political Ideologies*. Oct. 2006: 211-246. Print.

Connelly, Steve. "Mapping sustainable development as a contested concept". *Local Environment*. Jun. 2007: 259-278. Print.

Connolly, William. *The terms of political discourse*. 3rd ed. London: Wiley-Blackwell, 1993. Print.

Gadamer, Hans-Georg. *Verdad y método II*. Salamanca: Ediciones Sígueme, 2004. Impreso.

Gallie, Walter Bryce. *Conceptos esencialmente impugnados*. Vol. 49. México: Instituto de Investigaciones Filosóficas de la UNAM, 1998. Impreso.

Grafstein, Robert. "A realist foundation for essentially contested political concepts". *The Western Political Quarterly*. Mar. 1988: 9-28. Print.

Gray, John. "On liberty, liberalism and essential contestability". *British Journal of Political Science*. Oct. 1978: 385-402. Print.

Hamilton, Richard Paul. "Love as a contested concept." *Journal for the Theory of Social Behaviour*. Sep. 2006: 239-254. Print.

Hurley, S. L. *Natural reasons: Personality and polity*. Oxford: Oxford University Press, 1992. Print.

Klumpp, James. "A dramatistic approach to fields". Ziegelmüller, G. and J. Rhodes (eds.). *Dimensions of argument: Proceedings of the Summer Conference on Argumentation*. Utah: Speech Communication Association, 1981. Print.

Lakoff, George. *The political mind: Why you can't understand 21st-Century American politics with an 18th-Century brain*. New York: Viking Adult, 2008. Print.

Lausberg, Heinrich. *Manual de retórica literaria Tomo I*. Madrid: Gredos, 1976. Impreso.

Lukes, Steven. *El poder: un enfoque radical*. Madrid: Siglo XXI de España, 1985. Impreso.

MacIntyre, Alasdair. "The essential contestability of some social concepts". *Ethics*. Oct. 1973: 1-9. Print.

Mason, Andrew. "Justice, contestability, and conceptions of the Good". *Utilitas*. Nov. 1996: 295-305. Print.

McKerrow, R. E. "Argument communities". Trapp, R. and J. Schuetz (ed.). *Perspectives on argumentation: Essays in honor of Wayne Brockriede*. Illinois: IDEBATE Press, 2006. Print.

---. "On fields and rational enterprises: A Reply to Willard". Ziegelmüller, G. and J. Rhodes (eds.). *Dimensions of argument: Proceedings of the Summer Conference on Argumentation*. Utah: Speech Communication Association, 1981. Print.

Miller, D. "Linguistic philosophy and political theory". Miller, D. and L. Siedentop (ed.). *The nature of political theory*. Oxford: Clarendon Press, 1983. Impreso.

Moten, Abdul. "Understanding terrorism: Contested concept, conflicting perspectives and shattering consequences". *Intellectual Discourse*. Jan. 2010: 35-63. Online.

Reitan, Eric. "Rape as an essentially contested concept". *Hypatia*. May. 2001: 43-66. Print.

Rowland, Robert. "Argument fields". Benoit, William L., Hample, Dale

and Pamela J. Benoit. *Readings in argumentation*. New York: Foris Publ., 1992. Print.

---. "Purpose, argument fields, and theoretical justification". *Argumentation*. May. 2008: 235-250. Print.

Toulmin, Stephen. *Los usos de la argumentación*. Barcelona: Península, 2007. Impreso.

Waldron, Jeremy. "Is the rule of law an essentially contested concept (in Florida)?" *Law and Philosophy*. Mar. 2002: 137-164. Print.

Willard, Arthur. "Field theory: A cartesian meditation". Ziegelmüller, G. and J. Rhodes (eds.). *Dimensions of argument: Proceedings of the Summer Conference on Argumentation*. Utah: Speech Communication Association, 1981. Print.

---. "Some questions about Toulmi's view of argument fields". Ziegelmüller, G. and J. Rhodes (eds.). *Dimensions of argument: Proceedings of the Summer Conference on Argumentation*. Utah: Speech Communication Association, 1981. Print.



# OBJETO, TIEMPO Y COLECTIVIDAD EN *EL OTOÑO DEL PATRIARCA*

OBJECT, TIME AND COLECTIVITY IN *THE AUTUMN OF THE PATRIARCH*

ESTEBAN QUESADA

Universidad Central, Colombia. equesadas@ucentral.edu.co

RECIBIDO EL 5 DE JULIO DE 2012 Y APROBADO EL 5 DE OCTUBRE DE 2012

## RESUMEN      ABSTRACT

Este artículo describe la construcción de la dimensión temporal en *El otoño del patriarca* a través de la aplicación de una "arqueología objetual" que retoma elementos del análisis noético-noemático de la fenomenología eidética y que se aplica a la lectura de la novela. El análisis revela que en *El otoño del patriarca* hay un complejo de relaciones entre los conceptos de eternidad, historia y muerte que son resultado de la manera como la colectividad se enfrenta a la reconstrucción de los espacios y de los objetos que permanecieron a través de la dictadura, particularmente, del palacio presidencial y del cuerpo muerto del patriarca. Se mostrará finalmente, cómo esta "reconstrucción de la realidad" plantea un problema fenomenológico y una posibilidad política para los pueblos de América Latina.

This paper describes the construction of *The autumn of the patriarch's* temporal dimension, through the application of an "objectual archeology" that incorporates elements of the noetic-noematic analysis of the eidetic phenomenology, which are applied to the reading of the novel. The analysis reveals that in *The autumn of the patriarch* there is a complex of relations among the concepts of eternity, history and death that result from the way collectivity is faced with the reconstruction of spaces and objects that remained through dictatorship, particularly the presidential palace and the patriarch's dead-body. The paper shows how this "reconstruction of reality" raises a phenomenological question and a political possibility for Latin America peoples.

## PALABRAS CLAVE      KEY WORDS

arqueología, eternidad, historia, muerte, naturaleza, patriarca, palacio presidencial, rumor.

archeology, eternity, history, death, nature, patriarch, presidential palace, rumor.

“El otoño del patriarca es un universo caótico que inicia cerca del final y retorna a él al final del libro sin que todo ello tenga el sentido de una recurrencia cíclica. Todo sucede como si cada cosa ocurriera simultáneamente al final, sin orden”  
Roberto González.

## I

### Propuesta metodológica para una “arqueología objetual” de *El otoño del patriarca*

Partimos de un modelo de interpretación que hemos denominado “arqueología objetual” y que diferenciamos respecto de formas científicas de *explicación* arqueológica y de algunas otras formas *explicativas* propias del análisis genealógico. Lo que hemos denominado en este ensayo “arqueología” se concentra (i) en la *descripción* fenomenológica del acontecer del “tiempo” en tanto que residuo de la constitución de objetos en *El otoño del patriarca* y (ii) en la *mostración* de una posibilidad de “constitución histórico-objetual” propia de un “discurso circulante” (de un discurrir, a decir de Heidegger), mejor aún, de una voz anónima constituyente que es el verdadero sujeto de la novela. Y semejante tipo de descripción se desarrolla desde un doble recurso teórico:

En primer lugar, se trataría de mostrar que, aunque el recurso mnemotécnico usado por García Márquez en *El otoño del patriarca* guarde una simetría formal respecto del usado por M. Proust en su novela *En busca del tiempo perdido*<sup>1</sup>, en García Márquez la tendencia sería no sólo hacia la “recuperación” del tiempo, sino hacia su “invención” a partir de una configuración doble del olvido que fungiría, además, como tópico ideológico permanente de la novela: (a) el olvido como acontecimiento *necesario* para la instauración y para la permanencia en el poder en una dictadura y (b) el olvido como una *ausencia estructural* de “memoria histórica”, que sería, además, bastante típica de los pueblos latinoamericanos.

<sup>1</sup> Nos referimos a la “función” del té de Magdalena que, al tocar la lengua de Marcel, automatiza todo un complejo sensualista y lingüístico que “recupera” para la memoria lo que había devorado el olvido. Si este recurso mnemotécnico guarda una simetría formal con el de García Márquez, es porque “la observación” de las ruinas del palacio presidencial cumple la misma función formal de recuperar la dimensión temporal de la novela, que en *El otoño del patriarca*, no es, sin embargo, la dimensión de un “tiempo recuperado”. Hay importantes diferencias que hacen muy específica la situación de ambas novelas: así por ejemplo, el sentido que en cada una adquiere “recuperar” algo del olvido o la determinación del “quien” en la recuperación de ese algo.

En segundo lugar, la descripción arqueológica que desarrollaremos aquí como metodología para la lectura y el análisis de textos de ficción se diferencia de la descripción fenomenológica propuesta por Günther Müller y cuestionada por Ricoeur (78ss). De acuerdo con Müller, habría una diferencia entre “tiempo de la narración” y “tiempo narrado”, entre un tiempo propio del *objeto* de la narración, que siempre es narración-de-algo (“tiempo narrado”) y el tiempo de la *vivencia del narrar*, que en este caso sería el tiempo en que se coloca el narrador (“tiempo de la narración”), una forma de presentar, para el caso de la narración de ficción, la diferencia entre el análisis noemático y el noético en fenomenología. Siguiendo a Ricoeur y aplicando el análisis al caso de *El otoño del patriarca*, semejante descripción tendría fundamento si se considera que “lo narrado” hace parte del tiempo-vivido del mundo-de-vida abierto por la novela, esto es, si lo narrado no es, a su tiempo, una narración, como por el contrario ocurre de hecho. Una novela es un hecho ficcional, esto es, es la iteración (en la imaginación) de otra iteración (el recuerdo) de la percepción que utiliza el recurso al lenguaje, todo lo cual altera completamente la estructura del tiempo que podría asignársele, por ejemplo, a la percepción misma.

Cuando “lo narrado” es, además, en sí mismo una “narración”, por ejemplo, la narración de lo que otro me ha narrado, surgiría una paradoja sobre el tiempo que haría eclosionar la distinción propuesta por Müller. Es este el caso de *El otoño del patriarca*. Aquí no hay seguridad de la diferencia entre lo narrado y el tiempo de la narración, no sólo porque difícilmente se puedan separar los tiempos propios de la narración de los tiempos de vida de los personajes, sino además, porque parece ser que el único personaje que hay es una pluralidad de voces, el rumor del pueblo, de la colectividad. La perspectiva, entonces, no sólo es la de la reconstrucción de personajes sin identidad que representan, más bien, ciertas características de una heterogeneidad que los supera, sino también la de la narración de algo que ha sido narrado en más de una versión y que no tiene (ni necesita tener) una versión original.

La arqueología objetual que proponemos no intenta recuperar la “objetividad del mundo” de *El otoño del patriarca*, sino la particular manera en que el rumor de la colectividad genera sentido (es decir,



*realiza el mundo*)<sup>2</sup>. “Realizar el mundo” sucede en la novela a partir de la observación sobre los espacios y sobre las cosas, de un *hacerlas* visibles a través de un rumorear incesante. La colectividad, además, no funciona en la novela simplemente como el “polo tiempo” o el “polo-subjetivo” de la narración, pues no tiene la forma de la unidad de una *conciencia colectiva*, ni de un *imaginario colectivo* narrador que permita identificar un punto de dónde surja el flujo del tiempo o de la constitución. Es más bien un caso de fragmentación esencial de la conciencia (Laverde) y, por tanto, también del tiempo subjetivo, una fragmentación que, finalmente, es también *objeto* del rumorear, esto es, que también es constituida.

Nuestra arqueología toma por principio metodológico la *descripción* del inicio de los seis capítulos de la obra, no sólo porque en cada caso el inicio de la narración toma como referencia el palacio presidencial (*lugar* por excelencia de desarrollo de la narración), sino porque es a través de los objetos que emergen en él, y entre ellos el cuerpo muerto del general (*objeto* por excelencia de la narración) que se reconstituye la dimensión de “realidad” del mundo abierto por *El otoño del patriarca*, que es siempre el mundo del rumor de una colectividad fragmentada (*sujeto* de la narración).

Si la realidad es una *realización* de la dimensión y del horizonte del sentido a la que llamamos *mundo*, esta realización no puede, sino referir a una actividad, que en nuestro análisis muestra una doble operación referida a la estructura del tiempo: primero, el rumor que *realiza* el mundo sucede siempre en la triple forma temporal iteración-recuerdo-suposición y segundo, los objetos sobre los que opera la realización tienen, ellos mismos, la triple dimensión eternidad-historia-muerte. En los dos apartados siguientes veremos cómo funciona esta *doble operación realizadora* en referencia al “lugar”, el “objeto” y el “sujeto” identificados como fundamentales en la narración (palacio presidencial, cuerpo muerto y colectividad). En el segundo apartado, se procederá al análisis de la relación constituyente desde el punto de vista noemático; se describirá así, el modo de aparecer de los objetos en *El otoño del*

<sup>2</sup> Recuperar la objetividad del mundo sí es, en cambio, la pretensión de Celeita & Pardo (1991), quienes diseñan un modelo lingüístico para el análisis de *El otoño del patriarca*, que pretende develar una estructura lineal en la historia bajo la suposición de que hay en *El otoño del patriarca* una estructura susceptible de un análisis motivado por la necesidad de coherencia (Celeita y Pardo 25ss). Tres son las posiciones que plantea mi interpretación frente a este modelo: (i) no es posible establecer una georreferenciación “correcta” de los espacios físicos en que suceden las acciones; (ii) ellas, los personajes y los objetos no mantienen su sentido, identidad y fisionomía a través de la narración; y (iii) no es correcto afirmar que “los tiempos” de la novela sean claramente diferenciables entre sí.

*patriarca*, en tanto que objetos visuales primero y posteriormente, en tanto que objetos incluidos en el horizonte del mundo (dotados de sentido por el rumor); y luego, de ese peculiar objeto que es el cuerpo muerto. En el último apartado, se realizará el análisis de la “relación paradójal” de la constitución noemática de lo noético en la novela o la paradoja de un “constituyente constituido”; se describirá la modalidad realizante del rumor y, a modo de cierre, la subsiguiente conceptualización de la colectividad rumorosa como “objeto realizado” por la narración como posibilidad política.

## II

### Religión, mito y finitud en los objetos de *El otoño del patriarca*

“...volando entre el rumor oscuro de las últimas hojas heladas de su otoño hacia la patria de tinieblas de la verdad del olvido”. García Márquez (*El otoño del patriarca*).

En una entrevista realizada en 1987 por R. Williams a Gabriel García Márquez en su casa en México, el entrevistador intenta determinar la influencia de “lo visual”, más específicamente, de las “artes visuales” y de “la imagen” en la construcción de los espacios, de los lugares y de los objetos en la obra de García Márquez (Williams 134-137). En la conversación, en la que García Márquez señala, entre otras cosas, que la Cartagena del siglo XX es el prototipo de ciudad descrita en *El otoño del patriarca*, éste informa a Williams que, una vez iniciada la escritura de la novela, aún no tenía claro cómo describir la arquitectura, ni la paisajística del palacio presidencial, pensando además, en que semejante descripción apareciera paradójal frente a las descripciones de la ciudad en la novela; cuenta García Márquez a Williams, que se le aclaró la situación cuando, revisando un libro antiguo de paisajes asiáticos, vino a dar con la foto de un templo aparentemente hindú, habitado por animales de diversas especies, que servía perfectamente a sus fines. Esta era, dice allí García Márquez, “la imagen que yo estaba necesitando” (Williams 134).



Fotografía (Williams 133).

El “efecto visual” de la lectura de las primeras páginas de *El otoño del patriarca* depende, por supuesto, de la maestría narrativa de García Márquez, aun cuando también del hecho de que lo que se describe sea una fotografía, mejor, la percepción de una fotografía aunada a la imaginación de una serie simultánea de sucesos (sobre la simultaneidad volveremos unos párrafos más abajo). Como señalamos con anterioridad, cada una de las seis partes de la novela es una descripción sobre la percepción y las opiniones que despierta a una colectividad determinada por el “nosotros”, la observación de los “lugares” y de los “objetos”, y entre ellos del “cuerpo muerto” ubicados en esos espacios. La descripción tiene, además, dos momentos, el primero, eminentemente material y el segundo, donde a través de la materialidad de los objetos emerge la especulación y se van entremezclando las descripciones de las tres formas del tiempo señaladas con anterioridad: el histórico, la eternidad y la muerte. Y todo sucede cuando nos familiarizamos poco

a poco con una fotografía a la que primero observamos y a la que a través, por ejemplo, del encuadre, de la calidad de la imagen y del papel, tratamos de introducir una fecha, encuadrar en un tiempo histórico del que sabemos, además, que puede tener un importante margen de error y sobre la que finalmente sospechamos un suceso, un tema e incluso una posible narración, una posible historia y una posible forma de narrarla que quedan ambas suspendidas en el plano de lo posible, a no ser que tengamos a nuestro lado al camarógrafo o al fotografiado. En el caso de *El otoño del patriarca*, no tenemos posibilidad de acceder, ni al camarógrafo, ni al fotografiado, es decir, no contamos con una narración de primera mano, las versiones originales o privilegiadas de los sucesos, de modo que, ciertamente como en el caso de la fotografía, es el lector quien debe completarlos. Como quiera que sea, a medida que la colectividad se acerca al palacio, la novela se va permitiendo la introducción cada vez más evidente de elementos propios de la “opinión”, mejor aún, del “mero rumor”, sobre la historia de la nación, la vida del patriarca y, desde luego, sobre todo sobre su muerte.

A continuación, se describen la estructura y los detalles de la estrategia. Cada capítulo inicia con la descripción de los lugares aledaños al palacio y luego del palacio mismo y se enfoca en los detalles de aquellos objetos o de aquellos animales que, en lo seguido, volverán a aparecer para dar contexto a algún suceso, o que serán tema de una reflexión especial<sup>3</sup>. El primer momento es la “descripción material” de todo el sector del palacio, que está compuesto de dos patios exteriores, además de los interiores de la casa presidencial. El colectivo ingresa al palacio por el primer patio, a través una puerta herrumbrosa, el galpón y la gallera, la alberca bautismal y la caballeriza; pasa a un segundo patio, que dirige a la casa a través del establo, la galería de arcadas, la cocina, las albercas de lavar y el “cagadero común”. Ingresan a la casa presidencial por las áreas comunes, las salas de oficiales, la sala de audiencias y la “oficina lateral”, donde se halla el cuerpo fenecido de quien no se sabe si es el patriarca; finalmente, intentan reconstruir los detalles de su muerte o confirmar la identidad de su cuerpo buscando pistas en las áreas

<sup>3</sup> Es forma de “iteración” es una estrategia típica de la forma de escritura del “realismo mágico” de García Márquez: se describe con lujo de detalles un lugar o un objeto para, luego, hacerlo emerger en un contexto completamente diferente y con un sentido *sui generis*, de tal forma que el lector tiene una sensación de continuidad dominada por un constante extrañamiento por la metamorfosis y mutagenidad de las cosas. Esto sucede con casi todos los objetos emergentes de los inicios de capítulo, pero es en particular repetitivo con los animales, particularmente con las vacas.

privadas, el dormitorio de Bendición Alvarado, en el de Leticia Nazareno y finalmente en el dormitorio presidencial<sup>4</sup>.

Es característica de todos estos espacios y objetos que parecen en el primer momento de la descripción el hecho de ser eminentemente visuales<sup>5</sup> y de permanecer siendo tratados como tales cada vez que reaparecen; no hay referencias táctiles o sonoras sobre los objetos, aun cuando algunas propiedades olfativas puedan “derivarse” de la descripción de los mismos (“la ropa podrida al Sol de las albercas de lavar, la sentina abierta del cagadero común de concubinas y soldados” (García Márquez 4), pero como un efecto de la lectura (esto es, como algo completado por el lector) y no como un contenido propiamente tal de la narración. Por otro lado, aún cuando estos objetos estén en un espacio determinado (por lo menos, más que el espacio amorfo e infinito de “la ciudad” o de “el país”), todos tienden a desdibujar sus mutuas relaciones referenciales en la medida en que su sentido empieza a emerger gracias a un rumor imparable que los “pone en su lugar”, de unas situaciones que no sólo, no se relacionan cronológicamente (pues las situaciones y los objetos corresponden a época divergentes), sino que, además, tampoco tienen todos una única y exclusiva dimensión temporal. Hay por el contrario un constante juego de superposiciones de los tiempos en el rumorear sobre los objetos que, como resultado del modelo, hemos colegido, se resume a una triple contraposición que es fundamental para la construcción de los escenarios, los objetos y los sujetos a lo largo de toda la novela: la contraposición entre un “tiempo estancado” del espacio y de los objetos en el palacio presidencial, un “tiempo histórico” propio de la ciudad y del país, y un “tiempo de la muerte”, mejor, un tiempo “puesto entre paréntesis” por el cuerpo de la muerte indemostrable e imposible del patriarca.

Es aquí donde entramos en propiedad al segundo momento de la descripción: el de la determinación de las formas de *realización* del

---

<sup>4</sup> Esta lista, que pretende ser exhaustiva (si es que la exhaustividad tiene, aquí, algún sentido), no reproduce necesariamente el orden cronológico de los lugares por lo que atraviesa la colectividad en busca del cuerpo muerto y posteriormente de algo que confirme la identidad del cuerpo con el patriarca. Es una reconstrucción realizada a través de pasajes que narran situaciones sin ninguna conexión aparente y que en muchos casos introducen una variedad de personas, animales, artefactos y meras cosas en esos mismos espacios, por lo que uno termina por dudar de sus dimensiones.

<sup>5</sup> “Que García Márquez utilice imágenes visuales para organizar su escritura se clarifica cuando él afirma que los detalles son ‘siempre’ algo que él ve” (Hinds 897). De la entrevista en mención, donde dice García Márquez: “es siempre algo que yo veo. Así es siempre, siempre una imagen, sin excepciones” (Williams 132).

tiempo de la “eternidad” y “del tiempo de la historia” (al tiempo de la muerte y a su relación con estos otros será dedicado un aparte al final de este apartado, que enfocará explícitamente el objeto “cuerpo muerto”). El “*tiempo estancado*” (García Márquez 4) es una metáfora para exponer una doble forma de la eternidad: una eternidad natural, que se superpone a la descripción de los lugares y objetos del palacio, y la eternidad concedida a todo cuanto tiene que ver con el discurso sobre el poder o para todo lo que está implicado por ello. El “tiempo histórico”, por su parte, es un tiempo fluido donde nada permanece en su sitio, excepto las figuras del patriarca y del palacio determinadas por la segunda forma de eternidad; es el tiempo “que habita” el rumor, en el doble sentido de una voz reprimida por el poder y de una voz que despierta a darse su propia historia. El *espacio* descrito por el “tiempo histórico”, esto es, el espacio propiamente cultural, es tanto el espacio del carnaval y de la imaginación como el espacio de la opresión y de la violencia instituidas por el Estado.

A continuación, expondremos los hallazgos de la aplicación del modelo planteado en I sobre los espacios y objetos del palacio presidencial y a la interpretación de las determinaciones “eternidad” e “historia” y de su relación en *El otoño del patriarca*.

### **La eternidad natural**

En la fotografía que incluimos al inicio de este apartado puede apreciarse la condición de ciertas edificaciones aparentemente abandonadas por los hombres que parecieran retornar a la naturaleza o a su “condición de naturaleza” y todo sucede como si esta fuera una condición propia de su arquitectura. Desde luego que esto es posible en regiones como los trópicos o en grandes extensiones de selva en que las ciudades hayan emergido no “a pesar”, sino “justo en medio” de la naturaleza. Pero el proceso de naturalización del palacio presidencial de *El otoño del patriarca* implica la emergencia de formas eminentemente naturales (maleza, flores, animales), no tan sólo en ese “tiempo histórico” en el que paulatinamente el palacio se iría llenando de vacas, gallinas y perros, sino en un tiempo que se ha estancado, justamente, por el fenecimiento del patriarca, suceso que, de acuerdo con la descripción, implicaría una clara superposición de un tiempo cosmológico-natural, amorfo, creador de vida y transformador sobre el tiempo cultural de la edificación o de la monumentación tal cual puede apreciarse y que también sucede por el abandono en ciertos palacios en la India, Vietnam o Camboya y en



las “ciudades perdidas” de las culturas prehispánicas en las tupidas selvas de América Latina. Como en *Cien años de soledad* “al final”, en el momento del deceso del patriarca, que es lo mismo que el momento del decaimiento de la historia, que era la suya, por cuanto él constituía su origen, es decir, el origen de toda posible interpretación (como veremos en seguida), todo sucede, decíamos, como si en ese momento la naturaleza terminase literalmente por *devorar* los monumentos de los hombres, de su humanidad e incluso, de sus cuerpos.

En este segundo momento de la descripción de *El otoño del patriarca*, se muestran los espacios y los objetos intervenidos por una actividad natural imparable, una topografía sinfín de cuerpos de animales vivos y muertos que parece apropiarse de los ambientes del palacio presidencial que otrora hubieran determinado el “curso de la historia” y del cuerpo de quien fura el contra maestre de la orquestación de esta historia: gallinazos destruyendo las mallas de alambre de las rejas y ventanales del palacio (García Márquez 4-8) y devorando el cuerpo del patriarca (Ibíd. 21 38 90), que por otro lado estaba ya cubierto de plantas y animales de mar (Ibíd. 5-6); maleza crecida destruyendo las losas de los patios, enredaderas y todo tipo de plantas traspasando los muros y los retratos, polillas destrozando los fastuosos ropajes (Ibíd. 21); vacas destrozando la porcelana, las alfombras con sus pezuñas, comiendo todo cuanto hallan a su paso (los jardines presidenciales, las cortinas, los cueros de los sillones, los lienzos (Ibíd. 4 5 21 38 71 104) y bebiendo de la pila bautismal y de los inodoros (Ibíd. 11); “gallinas poniendo en las gavetas de los escritorios” (Ibíd. 6), restos de comida de animal, deposiciones y micciones humanas y animales (Ibíd. 4 7 11 71 81); micos colgando de los tejados de los teatros, “aves del paraíso, leopardos dormidos” (Ibíd. 9 91); “animales raros de colores intensos” (Ibíd. 26).

### **Eternidad y sacralidad del poder**

La “casa mostrenca del poder” (García Márquez 104), esto es el palacio presidencial, es monstruosa por lo menos en dos sentidos: (a) por cuanto el poder mismo que detenta el patriarca es absoluto y porque se trata, en todo caso, de determinar el poder dictatorial como un “desorden de fábula” (Ibíd. 71) en el que es “imposible establecer dónde estaba el gobierno” (Ibíd. 6) y porque al fin y al cabo el gobierno no necesita ningún tipo de orden, pues “los asuntos del Estado se arreglaban solos, la patria andaba” (Ibíd. 17); y (b) por cuanto es monstruoso lo que supera lo finito y lo terrenal, porque el poder del Patriarca, su vida y su



personalidad misma tienen ese tinte *sacro* de todo aquello que revuelve lo pantagruélico con lo que carente, con lo excesivamente simple; y porque el palacio es el monumento-representación del poder dictatorial y de su centramiento en la vida y en la personalidad del patriarca.

En consideración a esta particular forma de sacralidad, pantagruélica e irrisoria, que recae sobre el poder, la vida y la personalidad del dictador, el palacio presidencial es habitualmente descrito como un templo, como un santuario: “encontramos en el santuario desierto los escombros de la grandeza” (Ibíd. 6). Y el mero ingreso a los patios exteriores del palacio deja en la colectividad una sensación de penetrar “en el ámbito de otra época” uno en el que “el aire era más tenue” y “el silencio era más antiguo” (Ibíd. 4). La imagen de unas ruinas de un antiguo templo en la India, rodeado de animales, hace juego con esta concepción del palacio presidencial como templo; y la descripción de las imágenes mismas del palacio recuerdan siempre la descripción bíblica del templo de Jerusalén en la conocida parábola donde Jesús expulsa a los mercaderes y en la que exclama “Dadle al César lo que es del César, y dadle a Dios lo que es de Dios”, parábola que, por demás, tendría también su correlato en la expulsión de los religiosos (Ibíd. 66-7) de la nación del patriarca. Al fin y al cabo, con ambas expulsiones se trata del mismo acto auto-afirmatorio del poder en el poder, de una única verdad, de una única religión.

Es que, “siendo percibido desde el relato colectivo, desde la fábula popular, su figura [la del patriarca] adquiere una dimensión mitologizante” (Ortega 426): la de un origen, una dimensión fundacional del tiempo que, por tanto, está “fuera del tiempo”, que no está, ni antes, ni después de la historia, sino como “por detrás”, constituyendo su sentido. La historia es, aquí, la realización de lo sagrado, la venida del héroe sagrado (momento pantagruélico), de un héroe que, por otro lado, en su vida concreta es nadie, un héroe sin nombre cuyos atributos personales dejan más bien mucho que desear, un anti-héroe prototipo de su propia in-necesidad, la voz del aislamiento, del miedo y de la locura, de la soledad; precisamente, un despojo que sólo cobra sentido a través de la muerte.

### **El sentido de la historia**

Lo anterior, desde la perspectiva no del “objeto sacro”, sino de la historia, puede describirse con la afirmación de Ortega de que “la dimensión arquetípica del patriarca ocupa la historia misma como

distorsión" (Ibíd. 426); generando un bucle, una abertura en el tiempo que reconduce todo pasado y todo futuro a sí mismo, tanto desde el punto de vista del fluir del tiempo (pues el patriarca es inicio y es final de la historia) como desde el punto de vista teleológico (pues el patriarca es orientación y explicación última del sentido de lo real): "el origen no es sino la proyección del presente: o sea, el vaciado de sentido histórico, su ocupación por el poder" (Ibíd. 435) y con la muerte del patriarca "el silencio y las tinieblas se volverían a establecer en el universo porque aquél había de ser el término de la creación" (García Márquez 54). Sólo la figura del patriarca como sentido y como destino explica la ausencia de una cronología específica de la historia en *El otoño del patriarca*: la instauración del patriarca como origen histórico único hace del tiempo una mera simultaneidad de eventos, y de la historia misma, un presente que no es presente, puro "anacronismo y anticipación" (Ortega 428).

Antes y después de la historia, antes del nacimiento y después de la muerte del patriarca, antes y después del emerger del sentido no queda, sino *posibilidad*: la posibilidad de un relance, de un nuevo comienzo, de la creación de una o de múltiples historias no para recuperar, sino para re-inventar el destino.

### **Historia, poder y tradición oral**

Quiero llamar la atención sobre el hecho de que el patriarca nunca gobernó por ley escrita, es decir, nunca transmitió ninguna orden por escrito, sino que gobernó vía oral e incluso vía oral, también eran transmitidas órdenes que él nunca emitió, pero que, sin embargo, eran "suyas sin la menor duda" (García Márquez 74). Ésta no es tan sólo una metáfora sobre el analfabetismo de nuestros pueblos, es, ante todo, la determinación del poder dictatorial como iteración histórica. "El dictador analfabeto... puede ser re-producido, y es esta reproducción la que gobierna" (González 214). El dictador gobierna de voz y sus mandatos los repiten oralmente múltiples voces; es, pues, realmente esta oralidad no localizada la que gobierna, justo porque esta infinita iteración de la voz es la auténtica voz y el auténtico cuerpo del dictador sin nombre que ya no necesita existir. El dilema que plantea la muerte del dictador es, justamente, el dilema de su existencia y la ausencia constante de una referencia real al nombre propio es también la posibilidad de tener todos

los nombres y de ser todos los hombres<sup>6</sup>. Cuando el rumor da paso a ser la voz del patriarca, es decir, cuando es él quien habla a través de la voz de la colectividad, es ésta una voz de aislamiento, de miedo, de la soledad: es la voz de un despojo. El dictador, surgido del colectivo como exceso, es en su vida cotidiana un despojo; es este un “dictador cuya dimensión mitológica es totalizante pero cuya individualidad es reductora” (Ortega 437). Su muerte plantea, así, no una cuestión individual, sino una cuestión universal: la de la coherencia de la historia, la de la exactitud de la representación.

### Un cuerpo ajeno: la iteración de la muerte

Agostin & Pierce resumen en una hermosa frase lo que hemos determinado como una “simetría formal” en el recurso mnemotécnico de *En busca del tiempo perdido* de Proust con *El otoño del patriarca* de García Márquez. A decir de suyo, la narración no sería la búsqueda por “recobrar” el tiempo perdido, sino la “búsqueda de un cuerpo perdido” (Agostin & Pierce 329): el cuerpo de la evidencia para una muerte lógicamente imposible.

Lo mismo que el palacio, el “cuerpo muerto” del patriarca es un objeto cuya descripción se itera en cada inicio de capítulo y esta descripción se hace también en dos órdenes descriptivos complementarios: el primero, relacionado con la mera “visión” del cuerpo muerto y el otro, referido a la realización de la imposibilidad de su finitud a través del rumor. Al igual que los objetos y los lugares del palacio, las descripciones del cuerpo muerto están plagadas de elementos naturalistas, de una imposición devorante de la naturaleza sobre el cuerpo. El cuerpo del patriarca, “más viejo que todos los hombres y todos los animales viejos de la tierra y del agua” (García Márquez 5), no sólo está siendo devorado por los gallinazos, sino que, además, se describe siempre como “retoñado de líquenes minúsculos y animales parasitarios de fondo de mar” (Ibíd. 6). Si el patriarca, como arquetipo mítico, es el origen y el sentido del tiempo histórico, su muerte será entonces una imposibilidad de principio y su *cuerpo* muerto, parecerá siempre como una suplantación. No hay nada que confirme la muerte de patriarca, no sólo porque es inconfirmable, sino porque no hay nada que permita identificar su despojo con la

<sup>6</sup> “Aquí importa esa tautología sin misterio del yo: y soy el que soy yo; esto es, la pérdida del nombre es la ganancia del sujeto, su individuación no se da en un nombre sino en la representación de su figura arquetípica que el poder totaliza. Yo no soy tu: o sea, tú tienes un nombre, y ese nombre, como tu persona, pertenecen al poder”. (Ortega 426)

imagen que la colectividad se había formado de él. Ni en los elementos de su ropa, ni en la posición de su cuerpo, ni en las múltiples salas y cuartos del palacio halló la colectividad ninguna posible confirmación de su identidad: “tampoco el escrutinio meticuloso de la casa aportó ningún elemento válido” (García Márquez 21) y no “habíamos encontrado rastro alguno de su vida que pudiera conducirnos al establecimiento inequívoco de su identidad” (Ibíd. 38).

Este cuerpo muerto ajeno a su imagen, lo es porque no hay una experiencia original de esta imagen, porque la cara del poder es dada por experiencias de oídas, porque las monedas y los retratos “eran copias de copias de retratos que ya se consideraban infieles... cuando nuestros propios padres sabían quién era él porque se lo habían oído contar a los suyos, como éstos a los suyos” (García Márquez 5). Pero no sólo su imagen es una iteración de otras iteraciones, una repetición esencialmente inexacta, sino que es inexacta también su muerte; su cuerpo, que ya ha sido enterrado, no será si no la iteración de otros cuerpos que ya han sido muertos por él sin que fueran suyos, como el de Patricio Aragonés y ese incesante retorno suyo del reino de los muertos a la confirmación de su eternidad y santidad.

La repetición de la imagen del patriarca y de su muerte dota a ambas de una necesaria inestabilidad, necesaria, desde luego, para el poder; es la misma inestabilidad de la muerte que se manifiesta con la santificación de Bendición Alvarado luego de su muerte, con la posterior muerte de la monja, Leticia Nazareno, y de su hijo, devorados por una naturaleza representada por sesenta perros (Ibíd. 83), la desaparición de Manuela Sánchez durante un eclipse (Ibíd. 37) y la absurdidad propia de la violencia política masiva.

### III

## La constitución del mundo y la ideología: el dilema del constituyente-constituido

“...ajeno para siempre jamás a las músicas de liberación y los cohetes de gozo y las campanas de gloria que anunciaron al mundo la buena nueva de que el tiempo incontable de la eternidad había por fin terminado” García Márquez  
(*El otoño del patriarca*).

Eternidad, historia y finitud se juegan en ese momento en que la “visión del cuerpo muerto” es, a un tiempo, una forma de recordar “el tiempo incontable” de la enajenación y un “despertar” de la colectividad al porvenir de una historia que carecerá, desde ahora, de sentido, de una historia que habría que “inventar”. Desde esta perspectiva *El otoño del patriarca* sería, además de una metáfora sobre las dictaduras latinoamericanas, una apuesta ideológica por la liberación, por el “despertar del pueblo”. Desarrollar el concepto de “colectividad” es, pues, urgente, en la medida que es la colectividad la que a través de un incesante rumorear, recuerda e reinventa su destino. Este análisis valdría, así pues, tanto para el análisis sobre el “tiempo de la narración” de *El otoño del patriarca*, cosa que todavía debemos resolver, como para indicar el sentido total que adquiere, en su marco, el concepto de “tiempo histórico” de la novela.

### La constitución de lo real

La “indeterminación de lugar, tiempo y consciencia” (Córdova 70) es una consecuencia directa de la estrategia de la narración de *El otoño del patriarca*, estrategia que apunta a la constitución del sentido a través de una voz plural y fragmentada, de un nosotros que no es unitario y del cual puede un personaje “apropiarse” sólo de manera provisional. En esta narración, “el ‘Yo’ es un pronombre flotante que cualquiera puede asumir en cualquier momento, pero que nunca permite concretizar una identidad” (Ibíd. 71), ni siquiera en el sentido de un posible señalamiento del allí de una acción en la medida en que la colectividad es una multiplicidad de voces sin cuerpo. Hemos determinado esta voz con el concepto de rumor; el sentido constituido por el rumor, es decir, la realización de lo real por parte de esta polifonía, acontece en la descripción de la experiencia de la colectividad sobre el poder; es

una experiencia que “se cuenta” de oídas, una serie de versiones y de versiones sobre versiones en la que no hay hechos propiamente dichos, ni una “experiencia originaria” que permita reconstituirlos. Hemos determinado, finalmente, como fundamentalmente visual el modo en que el rumor opera el ingreso a la experiencia a través de los objetos, y entre ellos, a través del “cuerpo muerto” como objeto paradigmático del poder dictatorial. *El otoño del patriarca* es, pues, un “relato de lo visto que desencadena la suma de lo oído para recomponer el escenario de lo vivido” (Ortega 439), una narración constituyente de tipo testigo-ocular, propia además de la tradición oral en que ancla la comprensión de la historia de los pueblos latinoamericanos: “Un pueblo sin historia escrita se reconstruye... en la historia oral: es en su propio relato donde conoce, reconoce y discierne” (Ibíd. 439).

### La constitución de la colectividad

Que el sujeto de la constitución sea la “colectividad” es una asunción meramente formal. A esta constitución, que tiene la misma estructura que la constitución subjetiva de la fenomenología, algunos autores, pero particularmente Heidegger, Sartre y Merleau-Ponty han planteado un dilema que aquí queremos reproducir: el hecho de que las condiciones trascendentales para la constitución de lo real son, ellas mismas, constituidas en el marco de la experiencia; en una palabra, que el sujeto constitutivo es habitualmente sobrepasado por el sentido de eso que constituye, mejor aún, que termina siendo constituido por el sentido del mundo<sup>□</sup>.

Queremos ver cómo funciona esta constatación teórica en *El otoño del patriarca*. Decíamos anteriormente que, antes del nacimiento y después de la muerte del patriarca, o lo que es lo mismo, antes y después del sentido, no queda sino la *posibilidad* de la narración misma, una narración cuya función sería la de re-inventar el destino. Todo lo demás es eternidad natural, renovación y amorfidad cosmológica. Hay una “*inocencia animal*” antes del nacimiento y después de la muerte del patriarca, donde la ciudad es un “*extenso animal dormido*” que, sin saberlo, ingresa de pronto a un “*lunes histórico*” (García Márquez 5); después de la muerte del patriarca no habría, así pues, posibilidad de un retorno a lo natural, siempre y cuando asumamos que la colectividad no muere con el patriarca. De allí que el emerger de lo natural, el devorar cosmológico del monumento, sólo pueda suceder en el palacio del patriarca, esto es, en un lugar ausente de cualquier forma de humanidad.

En *El otoño del patriarca*, ciudad (historia) y palacio (eternidad) constituyen una dialéctica que sólo puede solucionar “la visión del cuerpo muerto”, justo en la medida en que “la multitudinaria confrontación actual con el cadáver del dictador frustra su deseo de ver o conocer su muerte” (López 304), esto es, justo en la medida en que acontece la inexplicabilidad de un hecho, además, imposible, que obliga el despertar de la colectividad a la necesidad de la formación de un discurso propio, de una narración donde los hechos, en ausencia, vuelvan a ser explicables. Esto significa, en primer lugar, siguiendo a Laverde, que la colectividad ha pasado de un estado pasivo a uno donde es “agente y se convierte en el protagonista de un drama en el que se ha olvidado o nunca se conoció la totalidad de los hechos” (95). Y significa también, en segundo lugar, que colectividad ha visto ante todo la necesidad de darse a sí misma un sentido, pues ha reconocido su capacidad de pensarse sin la figura constituyente de la eternidad como sentido de la historia.

No compartimos la conclusión de Córdova de acuerdo con la cual,

si hemos de guiarnos por *El otoño del patriarca*, la novela hispanoamericana no sería otra cosa que una parodia de la historia como un proceso sin dialéctica. Los latinos están condenados a vivir su historia como un círculo vicioso infinito, e incluso el dictador debe soportar la carga de la impotencia. (74)

Nuestro análisis ha mostrado, por el contrario, que *El otoño del patriarca* no sólo cuestiona, en el sentido que aduce Córdova, la realidad latinoamericana y en ella el papel que ha tenido nuestra colectividad, sino que, además, ofrece una alternativa, ofrece una posibilidad de liberación al pueblo, en la medida en que muestra cómo el pueblo “ha visto”, cabe el cuerpo muerto, la necesidad de narrarse a sí mismo.

#### REFERENCIAS BIBLIOGRÁFICAS

Agostin, Marjorie. “Inhabitants of decayed places: The dictator in the Latin American novel”. *Human Rights Quarterly*. May. 1990: 328-335. Print.

Celeita, Lola and Neyla Pardo. *Un modelo lingüístico para el análisis integral del discurso: propuesta metodológica aplicada a “El otoño del patriarca”*.



Bogotá: Instituto Caro y Cuervo, 1991. Impreso.

Córdova, José. "Looking for the Indian in Columbus journal and Garcia Marquez The Autumn of the Patriarch". *Journal of Comparative Poetics*. Jan. 1984: 63-76. Print.

García Márquez, Gabriel. *El otoño del patriarca*. Buenos Aires: Editorial Sudamericana, 1975. Impreso.

González, Roberto. "The dictatorship of rhetoric/the rhetoric of dictatorship: Carpentier, García Márquez, and Roa Bastos". *Latin American Research Review*. Sep. 1980: 205-228. Print.

Hinds, Elizabeth. "'Interview with Gabriel Garcia Marquez". *PMLA*. Oct. 1989: 897. Print.

Laverde, Alfredo. "El narrador incompetente, clave para una lectura de El otoño del patriarca". *Hojas Universitarias*. Jul. 2001: 90-97. Online.

López, Adelaida. "Burying the dead: Repetition in El otoño del patriarca". *MLN*. Mar. 1992: 298-320. Print.

Ortega, Julio. "El otoño del patriarca: texto y cultura". *Hispanic Review*. Dec. 1978: 421-446. Online.

Ricoeur, Paul. *Time and narrative. Vol. II: The configuration of time in fictional narrative*. Chicago: The University of Chicago Press, 1987. Print.

Williams, Raymond. "The visual Arts, the Poetization of Space and Writing: An Interview with Gabriel García Márquez". *PMLA*. Mar. 1989: 131-140. Print.



# ph

## COLABORADORES

COLLABORATORS

### STEPHEN BARKER

Filósofo, doctor por la Universidad de Melbourne y postdoctor en la UNAM y Monash University. Actualmente es profesor asociado y profesor adjunto de filosofía de la Facultad de Artes de la Universidad de Nottingham. Sus líneas de investigación son: filosofía del lenguaje, la metafísica y la metaética. Entre sus publicaciones más importantes se encuentran: *The emperor's new metaphysics of powers mind* y *Forget semantics: How speech-act theory and expressivism illuminate names with non-existing referent's monist*, entre otros.

### BRADLEY ARMOUR-GARB

Filósofo, doctor de la CUNY y miembro del Wolfson College de Oxford. Actualmente es profesor asociado en el Departamento de Filosofía de la Universidad de Albany. Sus líneas de investigación son: filosofía del lenguaje, la filosofía de la lógica, la metafísica y la epistemología. Entre sus publicaciones más importantes se encuentran: *Deflationism and paradox* y *Philosophical pathology*, entre otros.

### JAMES A. WOODBRIDGE

Físico y filósofo, doctor por la Universidad de Michigan. Actualmente es profesor asociado en el Departamento de Filosofía de la Universidad de Las Vegas. Sus líneas de investigación son: filosofía de la ciencia, la metafísica, filosofía del Lenguaje y lógica simbólica. Entre sus publicaciones más importantes se encuentran: *Truth as a pretense* y *The story about propositions*, entre otros.

### WŁODEK RABINOWICZ

Filósofo, doctor por la Universidad de Uppsala. Actualmente es catedrático del Departamento de Filosofía de la Universidad de Lund. Sus líneas de investigación son: filosofía moral, teorías del valor, teorías de la decisión y lógica. Entre sus publicaciones más importantes se encuentran: *Value and choice: Some common themes in decision theory and moral philosophy* y *Patterns of value. Essays on formal axiology and value analysis*, entre otros.

**GUSTAF ARRHENIUS**

Filósofo, doctor por la Universidad de Toronto. Actualmente es profesor de la Universidad de Estocolmo y Fellow del Colegio Sueco de Estudios Avanzados. Sus líneas de investigación son: filosofía moral, filosofía política y ciencias sociales. Entre sus publicaciones más importantes se encuentran: *The impossibility of a satisfactory population ethics* y *Population change and inequality*, entre otros.

**MICHAEL ISTVAN**

Filósofo con estudios de postgrado. Actualmente es profesor del Departamento de Filosofía de la Universidad de Texas A&M. Sus líneas de investigación son: epistemología, metafísica, historia de la filosofía, filosofía de la acción. Entre sus publicaciones más importantes se encuentran: *Concerning the resilience of Galen Strawson's basic argument*, entre otros.

**ISHA GAMLATH**

Filósofa con estudios de postgrado. Actualmente es Directora del Departamento de Cultura Cristiana y Cultura Clásica Occidental de la Universidad de Kelaniya. Sus líneas de investigación son: presocráticos, filosofía antigua, filosofía medieval, filosofía de la religión. Entre sus publicaciones más importantes se encuentran: *Buddhism as an environmental religion* y *Perversions and promiscuities: Manifestations of divine madness*, entre otros.

**LAUREN DE LACERDA NUNES**

Filósofa con estudios doctorales de la Universidad Federal de Santa María (UFSM-RS). Actualmente es profesora asistente del área de humanidades de la Universidad Federal do Pampa (UNIPAMPA-RS). Sus líneas de investigación son: ética, ética aplicada, metaética. Entre sus publicaciones más importantes se encuentran: *Lying* y *O dilema moral segundo Bernard Williams*, entre otros.

**GABRIEL GARMENDIA DA TRINDADE**

Filósofo con estudios de postgrado. Actualmente pertenece al Programa de Postgrado en Filosofía de la Universidad Federal de Santa María (UFSM-RS). Sus líneas de investigación son: ética, ética aplicada, metaética. Entre sus publicaciones más importantes se encuentran: *Lying* y *Conflitos morais insolúveis e sistemas racionalistas: uma abordagem sobre consistência moral*, entre otros.

**TÁRIK DE ATHAYDE PRATA**

Filósofo, doctor por la Ruprecht-Karl Universität Heidelberg. Actualmente es profesor adjunto del Departamento de Filosofía de la Universidad Federal de Pernambuco (UFPE). Sus líneas de investigación son: filosofía de la mente, filosofía de la acción y filosofía de la psicología. Entre sus publicaciones más importantes se encuentran: *Irreducibilidade ontológica versus identidade: John Searle entre o dualismo e o materialismo* y *É o naturalismo biológico uma concepção fisicalista?*, entre otros.

**JOHN R. SEARLE**

Filósofo, doctor por la Universidad de Oxford. Actualmente es profesor Slusser de filosofía en la Universidad de Berkeley. Sus líneas de investigación son: filosofía del lenguaje, filosofía de la mente, filosofía de las ciencias sociales. Entre sus publicaciones más importantes se encuentran: *The construction of social reality*, *The rediscovery of the mind* y *Making the social world: The structure of human civilization*, entre otros.

**LAURA GIRALDO**

Filósofa con estudios de postgrado. Actualmente es profesora del Departamento de Filosofía de la Corporación Universitaria Minuto de Dios. Sus líneas de investigación son: filosofía del lenguaje, filosofía de la mente, filosofía de la acción. Además, posee diversas publicaciones especializadas.

**ALEXANDER BIRD**

Filósofo, doctor por la Universidad de Cambridge. Actualmente es profesor asociado de filosofía de la Facultad de Artes de la Universidad de Bristol. Sus líneas de investigación son: epistemología, metafísica, filosofía de la ciencia, historia de la ciencia, filosofía de la medicina. Entre sus publicaciones más importantes se encuentran: *Nature's metaphysics* y *Kuhn, naturalism, and the social study of science*, entre otros.

**JUAN CARLOS AGUIRRE**

Filósofo con estudios de postgrado. Actualmente es profesor del Departamento de Filosofía de la Universidad del Cauca. Sus líneas de investigación son: filosofía de la ciencia, metafísica, fenomenología. Entre sus publicaciones más importantes se encuentran: *Sufrimiento, verdad y justicia* y *Husserl y la justificación indirecta*, entre otros.

**JUAN MANUEL JARAMILLO URIBE**

Filósofo con estudios de postgrado. Actualmente es profesor jubilado del Departamento de Filosofía de la Universidad del Valle. Sus líneas de investigación son: filosofía de la ciencia, estructuralismo, epistemología. Entre sus publicaciones más importantes se encuentran: *Filosofía de las ciencias sociales* y *Filosofía de la tecnología*, entre otros.

**JAVIER ALEGRE**

Filósofo con estudios doctorales. Actualmente es profesor de filosofía de la Universidad Nacional del Nordeste. Sus líneas de investigación son: filosofía contemporánea, filosofía del lenguaje, filosofía de las ciencias sociales. Entre sus publicaciones más importantes se encuentran: *Trabajo y procesos de subjetivación e institucionalización en la crítica contemporánea* y *Hechos y des(h)echos en el trabajo*, entre otros.

**JULDER A. GÓMEZ**

Filósofo con estudios de postgrado. Actualmente es profesor del Departamento de Humanidades de la Universidad EAFIT. Sus líneas de investigación son: teoría de la argumentación, filosofía del lenguaje, hermenéutica y retórica. Entre sus publicaciones más importantes se encuentran: *Del comprender en la elocución retórica* y *Las "expresiones" en el conocimiento lógico y argumentativo*, entre otros.

**ESTEBAN QUESADA**

Filósofo de la Universidad Nacional de Colombia. Actualmente es profesor de la Universidad Central. Sus líneas de investigación son: ética, metafísica, fenomenología, historia de la filosofía. Entre sus publicaciones más importantes se encuentran: *Hölderlin y su Empédocles: el poetizar, la muerte y la conciencia de los tiempos* y *Un acercamiento al análisis onírico desde Friedrich Nietzsche: aportaciones y diferencias frente al psicoanálisis freudiano*, entre otros.

## ORIENTACIONES PARA POTENCIALES AUTORES DE *DISCUSIONES FILOSÓFICAS*

La Revista *Discusiones Filosóficas* recibe colaboraciones originales en forma de artículos que reportan resultados de investigación, reseñas, traducciones y ponencias sobre sus temas de interés (filosofía y filosofía de la literatura). Estas colaboraciones se publican en español, inglés o portugués, a menos que los autores soliciten explícitamente la traducción al español. La revista también invita a los lectores a enviar textos cortos de réplica a los artículos publicados, así como reseñas de libros especializados publicados recientemente. Las colaboraciones deben ser enviadas a su Consejo Editorial.

Los trabajos deben ceñirse a las normas conocidas como MLA (ver ejemplo en la página web de la revista) y estar acompañados de un resumen (abstract) no mayor de 150 palabras, preferiblemente incluyendo también su versión inglesa, y no menos de cuatro palabras clave o descriptores. Las propuestas serán enviadas en copia de papel y en disquete (programas Word). No se considerarán propuestas cuya extensión exceda las 7000 palabras (incluyendo el resumen). Todos los artículos serán sometidos a evaluación anónima por parte de lectores expertos, en el mismo orden de llegada. La versión electrónica también se puede remitir como archivo adjunto al E-mail: [discufilo@ucaldas.edu.co](mailto:discufilo@ucaldas.edu.co) En cada artículo se debe identificar la institución a la que pertenece su autor o una mínima información profesional e investigativa que, a su criterio, pueda ser mencionada. Los potenciales autores deben certificar que no han enviado su propuesta de publicación a ninguna otra revista y que se abstendrán de someterla a consideración de otra revista mientras dura el proceso de evaluación en la nuestra.

Las referencias bibliográficas deben aparecer integradas en el texto, entre paréntesis, así: (apellido del autor número de página). Las referencias completas deben aparecer al final del trabajo.

Para el fascículo correspondiente al semestre enero-junio, la revista recibe propuestas de publicación hasta el 31 de marzo. Para el fascículo correspondiente al semestre julio-diciembre, las propuestas se reciben hasta el 30 de septiembre. Los artículos recibidos después de esas fechas serán considerados para la publicación del semestre siguiente. Todos los artículos son sometidos a una evaluación doblemente anónima. Los autores serán notificados de la decisión de los árbitros en los dos meses siguientes a la recepción de sus propuestas. Sólo se publicarán los artículos que superen satisfactoriamente el proceso de evaluación.



## GUIDELINES FOR POTENTIAL CONTRIBUTORS

The journal *Discusiones Filosóficas* accepts original contributions such as research papers, reviews, translations and articles on its topics of interest (philosophy on all fields and philosophy of literature). These contributions are published in Spanish, English or Portuguese, unless the author explicitly requests a Spanish Translation. The journal also invites the readers to send short texts in response to previously published articles, as well as reviews of specialized books published recently. The collaborations should be sent to the Editorial Board.

The texts should follow the MLA norms (see an example in the web page of the journal), and be accompanied by an abstract no longer than 150 words, preferably including its Spanish translation, and at least four key words. A printed (Word program) and a CD copy of the proposals should be sent to the Editorial Board. Proposals exceeding 7000 words (including the abstract) will not be considered. All of the articles will be subjected to anonymous evaluation by expert readers, in the order of arrival. Electronic versions can also be mailed to the following e-mail address: [discufilo@ucaldas.edu.co](mailto:discufilo@ucaldas.edu.co). In each article the authors must include their institutional affiliation as well as some basic professional and investigative information. Prospective authors should certify that they have not sent their proposal for publication to any other journal and that they will refrain from submitting it for consideration of another journal for as long as our evaluation process takes.

The bibliographical references should appear integrated in the text, in parenthesis, in the following way: (author's last name page number). The complete references should appear at the end of the text.

For the issue corresponding to the semester January-June, the journal receives potential articles until March the 31st. For the issue corresponding to the semester July-December, the journal receives potential articles until September the 30th. The contributions received after these dates will be considered for publication in the following semester. All of the articles are subjected to a double blind evaluation. The authors will be notified of the referees' decision within the two months following the reception of the proposals. Only articles that satisfactorily pass the evaluation process will be published.

# Entregando lo mejor de los **colombianos**



Línea de atención al Cliente Nacional: **01 8000 111 210**

Línea de atención al Cliente Bogotá: **(57-1) 4199299**

► [www.4-72.com.co](http://www.4-72.com.co)

4°N  
72°W

# Revista Discusiones Filosóficas

## FORMATO DE SUSCRIPCIÓN

Nombre / Name	
Cédula / Identification number	
Dirección / Address	
Ciudad / City	
Departamento / State	Código Postal / Zip Code
País / Country	
Teléfono / Phone Number	
Profesión / Profession	
Institución / Employer	
Correo Electrónico / E-mail	
Dirección de envío / Mailing Address	

### Suscriptores Nacionales por un año. (2) Ejemplares

Se debe consignar en Bancafé, cuenta de ahorros No. 255050114 código 00HD005  
Promoción e indexación de publicaciones científicas.

### Mayores informes:

Vicerrectoría de Investigaciones y Postgrados  
Universidad de Caldas. Calle 65 N° 26 - 10  
A.A. 275 Manizales - Colombia  
Tel: 8781500 ext. 11222  
Fax: 8781500 ext. 11622  
E-mail: [revistascientificas@ucaldas.edu.co](mailto:revistascientificas@ucaldas.edu.co)

Último ejemplar recibido / Last issue mailed:

Año/Year      Volumen/Volume      Número/Number      Fecha / Date



**Ventas, suscripciones y canjes**  
Vicerrectoría de Investigaciones y  
Postgrados  
Universidad de Caldas  
Sede Central  
Calle 65 No. 26 - 10  
A.A. 275  
Teléfonos: (+6) 8781500  
ext. 11222  
e-mail:  
revistascientificas@ucaldas.edu.co  
Manizales - Colombia



Revista  
Agronomía



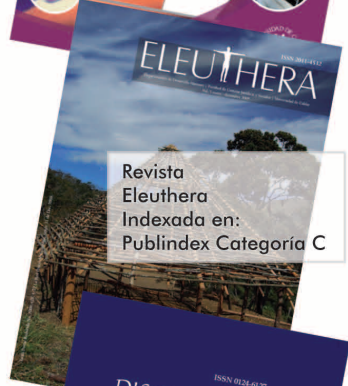
Revista  
Cultura y Droga



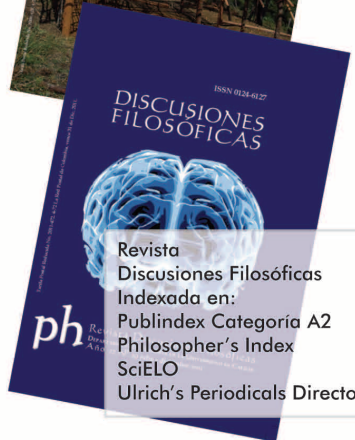
Revista  
Luna Azul (On Line)  
<http://lunazul.ucaldas.edu.co>  
Indexada en:  
Publindex Categoría B  
Index Copernicus, DOAJ



Revista  
Biosalud  
Indexada en:  
Publindex Categoría B  
Lilacs



Revista  
Eleuthera  
Indexada en:  
Publindex Categoría C

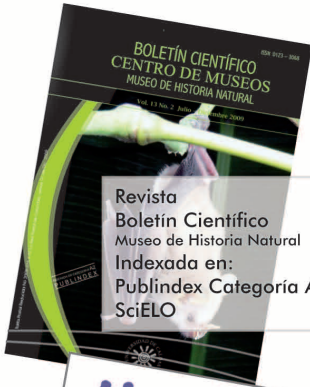


Revista  
Discusiones Filosóficas  
Indexada en:  
Publindex Categoría A2  
Philosopher's Index  
SciELO  
Ulrich's Periodicals Directory

# Revistas







Revista  
Boletín Científico  
Museo de Historia Natural  
Indexada en:  
Publindex Categoría A2  
SciELO



Revista Colombiana de  
las Artes Escénicas



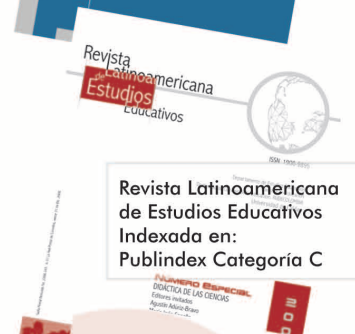
Revista  
Veterinaria y Zootecnia  
Indexada en:  
Publindex Categoría C



Revista  
Hacia la Promoción  
de la Salud  
Indexada en:  
Publindex Categoría A2  
Lilacs  
SciELO



Revista  
Jurídicas  
Indexada en:  
Publindex Categoría C  
DialNet



Revista Latinoamericana  
de Estudios Educativos  
Indexada en:  
Publindex Categoría C



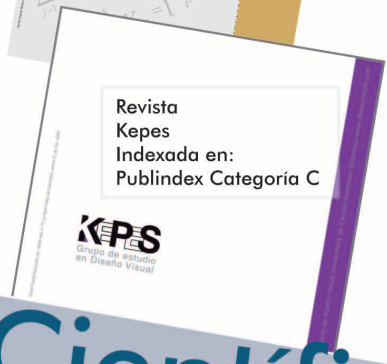
Revista  
Vector  
Indexada en:  
Publindex Categoría C



Revista de Antropología  
y Sociología (Virajes)  
Indexada en:  
Publindex Categoría C



Revista  
Universidad de Caldas



Revista  
Kepes  
Indexada en:  
Publindex Categoría C



Revista Latinoamericana  
de Estudios de Familia

# Científicas



ph

**Esta Revista se terminó de imprimir  
en diciembre de 2012 en los talleres de  
Capital Graphic  
Manizales - Colombia**