

cuadernos salmantinos de filosofía

SECCIÓN MONOGRÁFICA:
FILOSOFÍA E INTELIGENCIA
ARTIFICIAL

CUADERNOS SALMANTINOS DE FILOSOFÍA
Revista anual de la Facultad de Filosofía
Universidad Pontificia de Salamanca - Salamanca / España
Vol. 48, 2021 ISSN: 0210-4857
E-ISSN: 2660-9509

UNIVERSIDAD PONTIFICIA DE SALAMANCA

2021



48

CUADERNOS SALMANTINOS DE FILOSOFÍA

cuadernos salmantinos de filosofía

**SECCIÓN MONOGRÁFICA:
FILOSOFÍA E INTELIGENCIA
ARTIFICIAL**

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca – Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857

E-ISSN: 2660-9509

UNIVERSIDAD PONTIFICIA DE SALAMANCA

2021



48

CUADERNOS SALMANTINOS DE FILOSOFÍA

Vol. 48, 2021 / ISSN: 0210-4857 / E-ISSN: 2660-9509

Revista anual de investigación filosófica,
fundada en 1974,
a cargo de la Facultad de Filosofía
de la Universidad Pontificia de Salamanca

DIRECTORA

Ana María ANDALUZ ROMANILLOS

DIRECCIÓN

SECRETARÍA Y ADMINISTRACIÓN

Compañía, 5. Teléf. 923 277108
37002 SALAMANCA (España)
cuadernossalmantinos@upsa.es

WEB

<https://revistas.upsa.es>

ARCHIVO DIGITAL

<http://summa.upsa.es/about.vm?lang=es>

SUSCRIPCIÓN

		Núm. suelto
España:	43 €	45 €
Europa:	48 €	49 €
Demás naciones:	57 €	59 €

Para comprar ejemplares o suscribirse: publicaciones@upsa.es

PERIODICIDAD Y TIRADA

Volumen anual con una tirada de unos 220 ejemplares

Depósito Legal: S. 643-1974

I.S.S.N.: 0210-4857

Imprenta «KADMOS»

Teléf. 923 28 12 39

Salamanca, 2021



El contenido de esta revista, excepto que se establezca de otra forma, cuenta con una **Licencia Creative Commons Reconocimiento-No comercial-Sin obras derivadas 3.0 España**, que puede consultarse en <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

© **1974 Universidad Pontificia de Salamanca**. Para su uso impreso o reproducción del material publicado en esta revista se deberá solicitar autorización a la Dirección de la revista.

Las opiniones expuestas en los trabajos publicados por la revista son de la exclusiva responsabilidad de sus autores.

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca – Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857 E-ISSN: 2660-9509

CONSEJO DE REDACCIÓN EDITORIAL BOARD

DIRECTORA

EDITOR

Ana María Andaluz Romanillos (Universidad Pontificia de Salamanca, España)

SECRETARIO DE REDACCIÓN

EDITORIAL SECRETARY

Jesús Manuel Conderana Cerrillo (Universidad Pontificia de Salamanca, España)

VOCALES

MEMBERS OF THE BOARD

Sixto J. Castro (Universidad de Valladolid, España)

Juan José García Norro (Universidad Complutense de Madrid, España)

Modesto M. Gómez Alonso (Universidad de La Laguna, Tenerife, Canarias, España)

Rosa María Herrera García (Universidad Pontificia de Salamanca, España)

María del Carmen Paredes Martín (Universidad de Salamanca, España)

† Jacinto Rivera de Rosales (Universidad Nacional de Educación a Distancia, Madrid, España)

Pedro Jesús Teruel (Universidad de Valencia, España)

COMITÉ CIENTÍFICO ADVISORY BOARD

Gabriel Amengual Coll (Universidad de las Islas Baleares, España)

Mauricio Beuchot Puente (UNAM, México)

Fernando Broncano Rodríguez (Universidad Carlos III de Madrid, España)

Jesús Conill Sancho (Universidad de Valencia, España)

Adela Cortina Orts (Universidad de Valencia, España)

John Cottingham (University of Reading / University of London / Oxford University, Reino Unido)

Dulce María Granja Castro (UNAM, México)

Diego Gracia Guillén (Universidad Complutense de Madrid, España)

Danièle Moyal-Sharrock (University of Hertfordshire, Reino Unido)

Jesús Padilla Gálvez (Universidad de Castilla La Mancha, España)

Chon Tejedor (University of Hertfordshire, Reino Unido)

Nuria Sánchez Madrid (Universidad Complutense de Madrid, España)

Jesús Vega Encabo (Universidad Autónoma de Madrid, España)

Nuno Venturinha (Universidade Nova de Lisboa, Portugal)

MIEMBROS HONORARIOS HONORARY MEMBERS

† Saturnino Álvarez Turienzo (Universidad Pontificia de Salamanca, España)

Antonio Pintor Ramos (Universidad Pontificia de Salamanca, España)

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca – Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857 E-ISSN: 2660-9509

INDIZACIÓN

INDEXING INFORMATION

Cuadernos salmantinos de filosofía sigue el procedimiento de revisión externa y anónima por pares. La revista aparece en los siguientes índices, repertorios, directorios, rankings y bases de datos:

CIRC Clasificación Integrada de Revistas Científicas

DICE Difusión y Calidad Editorial de las Revistas Españolas de Humanidades y Ciencias Sociales y Jurídicas

Dialnet

EBSCO

ERIHPLUS European Reference Index for the Humanities and Social Sciences

ISOC-CSIC

LATINDEX 33/33

MIAR Matriu d'Informació per a l'Avaluació de Revistes

PIO Periodical Index Online

RBPH Répertoire Bibliographique de la Philosophie

REBID Red Iberoamericana de Innovación y Conocimiento Científico

REGESTA IMPERII

RESH Revistas Españolas de Ciencias Sociales y Humanidades. Cat. B

SCOPUS SJR

Ulrich's Periodical Directory

IBZ online

Políticas OA: Dulcinea color verde

REBIUN Red de Bibliotecas Universitarias

Está además en:

ATLA American Theological Library Association

Base. Biefeld Academic Search Engine

Google Scholar

International Directory of Philosophy

PhilPapers

Hispana

OpenAire

OpenDoar

Recolecta

Worldcat

También se ha solicitado su inclusión en **ESCI**

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca – Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857 E-ISSN: 2660-9509

LISTA DE REVISORES

REFEREES

Vol. 48, 2021

El Consejo de Redacción de *Cuadernos salmantinos de filosofía* agradece profundamente a los árbitros el esfuerzo dedicado a la evaluación de los trabajos. Su experiencia y juicio académico contribuyen decisivamente en la calidad de nuestra publicación. Durante los años 2020-2021 actuaron como evaluadores los siguientes expertos:

Miguel Acosta López (Universidad CEU San Pablo, Madrid, España)
Esperanza Aguilar de la Morena (Universidad de Málaga, Málaga, España)
Daniel Álvaro (CONICET, Argentina)
José Antúnez Cid (Universidad Eclesiástica San Dámaso, Madrid, España)
Fernando Bárcena Orbe (Universidad Complutense de Madrid, Madrid, España)
Cristina Basili (Universidad Complutense de Madrid, Madrid, España)
María José Binetti (CONICET, Argentina)
Marcelo D. Boeri (Pontificia Universidad Católica de Chile, Santiago, Chile)
Alcira Beatriz Bonilla Llanes (Universidad de Buenos Aires / CONICET, Argentina)
Carles Brasó Broggi (Universitat Oberta de Catalunya, Barcelona, España)
Francisco Manuel Carballo Rodríguez (Universidad de Salamanca, Salamanca, España).
David Casacuberta (Universidad Autónoma de Barcelona, Barcelona, España)
Ricard Casadesús Castro (Institut de Teologia Fonamental, Barcelona, España)
Salvador Cayuela Sánchez (Universidad de Murcia, Murcia, España)
Cecilia Celeste Danesi (Universidad de Buenos Aires / Universidad de Perugia, Argentina / Italia)
Stelio Cro (MacMaster University, Hamilton, Canadá)
Lorenzo Cotino (Universitat de València, València, España)
Graciela Marta Chichi (Universidad Nacional de la Plata, Buenos Aires, Argentina)
José Manuel Chillón Lorenzo (Universidad de Valladolid, Valladolid, España)
Antonio Diéguez Lucena (Universidad de Málaga, Málaga, España)
Luciano Espinosa Rubio (Universidad de Salamanca, Salamanca, España)
Ricardo Espinoza Lolas (Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile)
José Luis Fernández Fernández (Universidad Pontificia de Comillas, Madrid, España)
Diego Fonti (CONICET, Argentina)
Ana Belén Gil González (Universidad de Salamanca, España)
Joan Gimeno Simó (Universidad de Valencia, Valencia, España)
Francisco Guevara (Universidad Nacional de Villa María, Córdoba, Argentina)
Rodolfo Gutiérrez Simón (Universidad Complutense de Madrid, Madrid, España)
Juan Jesús Gutierrez Carrasco (Universidad Católica de Ávila, Ávila, España)
Carmen Herrando Cugota (Universidad San Jorge, Zaragoza, España)
Fernando Infante del Rosal (Universidad de Sevilla, Sevilla, España)
Manuel Lázaro Pulido (Universidad Nacional de Educación a Distancia, Madrid, España)
Salvador López Arnal (CEMS Universidad Pompeu Fabra, Barcelona, España)
Martín López Corredoira (Instituto de astrofísica de Canarias, Tenerife, Islas Canarias, España)
Pablo Martínez Becerra (Universidad de Playa Ancha, Valparaíso, Chile)
Ricard Martínez Martínez (Universidad de València, València, España)
Óscar Parcero Oubiña (Universidad de Santiago de Compostela, La Coruña, España)
Mora Perpere Viñuales (Pontificia Universidad Católica de Argentina, Buenos Aires, Argentina)
María Luisa Pro Velasco (Universidad Católica de Ávila, Ávila, España)
Óscar Quejido Alonso (Universidad Complutense de Madrid, Madrid, España)
Carles Ramió Matas (Universidad Pompeu Fabra, Barcelona, España)
Sergio Rodero Cilleros (Universidad de Salamanca, Salamanca, España)
Carlos Rodríguez-Gordo (Salamanca, España)
Francisco Rodríguez Valls (Universidad de Sevilla, Sevilla, España)

José María Santamaría García (Universidad de Alcalá de Henares, Alcalá de Henares, Madrid, España)
Cedric Guillermo Steinlen Cuevas (Universidad Santo Tomás, Santiago, Chile)
Jorge Teixeira da Cunha (Universidade Católica Portuguesa, Portugal)
Gabriel Terol (Universidad de Valencia, Valencia, España)
Javier Toscano Guerrero (APRA Foundation, Berlín, Alemania)
Luca Valera (Universidad Católica de Chile, Santiago, Chile)
Álvaro Vallejo Campos (Universidad de Granada, Granada, España)
José Domingo Villaplana Guerrero (Universidad de Sevilla, España)
Alicia Villar Ezcurra (Universidad Pontificia de Comillas, Madrid, España)
Raúl Villarroel Soto (Universidad de Chile, Santiago, Chile)
Carolina Villegas-Galaviz (Universidad Pontificia de Comillas, Madrid, España)
Dominique Vinck (Université de Lausanne, Lausanne, Suiza)
Jairo Andrés Villalba Gómez (Universidad Militar Nueva Granada, Bogotá, Colombia)
María Villaro Mañes (Pamplona, España)
Ángel Viñas Vera (Universidad Loyola Andalucía, Sevilla, España)

EDITORIAL

La repercusión de la inteligencia artificial en nuestras vidas, tanto en nuestra vida cotidiana como en los ámbitos empresarial, financiero, comunicativo, médico, educativo, jurídico, entre otros, es un hecho. En todos esos campos su desarrollo puede convertirse en una fuente de grandes oportunidades que mejoren nuestra existencia individual y colectiva. Pero, por otro lado, sus efectos podrían tener también implicaciones negativas, si llegara un momento en el que por este medio quedaran amenazados nuestros derechos fundamentales y resultara mermada la capacidad de los seres humanos para dirigir su destino. De acuerdo con Habermas, del conocimiento científico “se sigue un saber técnicamente utilizable, pero no un saber normativo, orientador de la acción”, afirmación que hace juego con esta otra de Nietzsche en sus fragmentos póstumos, en *Das philosophen-Buch*: “La ciencia investiga el curso de la naturaleza, pero nunca puede dar órdenes al hombre” (*cit. apud J. Habermas, Conocimiento e interés*, Madrid: Taurus, 1987, p. 288). Y, sin embargo, ¿son aplicables, sin más, estas afirmaciones al actual desarrollo y al desarrollo previsible de la inteligencia artificial? Entre sus potencialidades, se habla ya de su capacidad para tomar decisiones. Como todo desarrollo científico y técnico, el de la inteligencia artificial necesita, sin duda, de una orientación de su uso, de manera que se garantice que el progreso técnico va alineado con un progreso del género humano hacia mejor, por utilizar la expresión kantiana, es decir, con un progreso en el sentido de un desarrollo mayor y mejor de nuestras capacidades, nuestra autonomía, en suma, con un progreso ético de la humanidad. Ante el hecho del impacto de la inteligencia artificial en nuestra existencia, debemos preguntarnos, también al estilo kantiano, por sus posibilidades y por sus límites; esta vez, no por las posibilidades y los límites del uso metafísico de la razón, sino por las posibilidades y los límites del uso de la inteligencia artificial.

¿Qué tiene que decir la filosofía ante el nuevo hecho de la IA? ¿Cuáles son sus grandes potencialidades? ¿Qué problemas filosóficos plantea el desarrollo de la inteligencia artificial? ¿Qué efectos o repercusiones tiene sobre nuestras vidas, sobre los derechos humanos, sobre las dimensiones social y política de las

personas? ¿Qué exigencias se plantean en cuanto a su regulación? La cuestión no es trivial, si se tiene en cuenta, además, que el desarrollo de la IA podría comportar incluso un cambio cualitativo en el proceso de la evolución humana. *Cuadernos salmantinos de filosofía* dedica la sección monográfica de su volumen 48 a pensar esas cuestiones.

El tema es abordado desde diferentes perspectivas por los 10 artículos que componen dicha sección monográfica, “Filosofía e inteligencia artificial”.

Empezando por la filosofía misma, las potencialidades de la Humanidades Digitales para la investigación en ese campo y, en general, para la investigación en el campo de las humanidades, son inmensas, tal como se pone de relieve en el artículo con el que se abre la sección, dedicado a la revolución de las Humanidades Digitales, su concepto, su génesis y su uso en la investigación. El segundo artículo, que se centra en las aportaciones del matemático Alan Turing al desarrollo de la inteligencia artificial, pone al descubierto cómo las potencialidades de esta plantean problemas filosóficos relacionados con la comparación entre los humanos y las máquinas. En la díada humano-máquina profundiza otro artículo, que aborda el concepto de “emocionalidad artificial”. Por otro lado, según se pone de relieve en otro trabajo, algoritmos, *big data* e inteligencia artificial, pueden incidir en nuestras vidas en el sentido de una planificación y control, con el consiguiente riesgo de socavar los fundamentos de la libertad humana y conducirnos a un nihilismo. En esta línea del riesgo de un control sobre la existencia del ser humano se desarrolla otro artículo, que se apoya en el argumento del *Deus ex machina*, al mismo tiempo que plantea los nuevos fenómenos del transhumanismo y el posthumanismo. Desde una perspectiva también crítica, otro trabajo destaca la polarización social y el control político, como algunas consecuencias de la inteligencia artificial y de las redes sociales para la razón práctica, llevando el análisis, además, a una reflexión sobre la racionalidad y la tecnología, en diálogo con algunos grandes exponentes de la filosofía contemporánea y actual. El incremento del desarrollo y de la aplicación de la inteligencia artificial en todos los ámbitos plantea, sin duda, nuevos retos y desafíos y, con ello, tal como pone de manifiesto otro artículo, la necesidad de una gobernanza global de la IA que tenga un carácter ético, de modo que se garantice que el desarrollo de la inteligencia artificial es beneficioso para la humanidad. El tema de la gobernanza ética de la inteligencia artificial es abordado también por otro artículo, el cual examina el papel de la Unión Europea en ese sentido; ante los riesgos que comporta el uso de la IA, especialmente en el ámbito de los derechos fundamentales y los derechos humanos de los ciudadanos, se hace necesaria una gobernanza ética de la IA y, en ese sentido, la diplomacia de la UE puede convertirse en un potente mediador entre las grandes potencias de China y Estados Unidos. Los dos últimos artículos de esta sección monográfica se centran justamente, desde otras perspectivas,

en las contribuciones de la Unión Europea y China. Una de esas perspectivas hace referencia a la smartificación de la Administración pública o aplicación de la IA en las administraciones públicas, en los estados europeos. El último artículo centra su reflexión en la filosofía de la tecnología; en esta perspectiva analiza la incorporación a este ámbito del filósofo chino Yuk Hui y reflexiona críticamente sobre sus aportaciones a la “tecnodiversidad” en el paradigma de una “filosofía posteuropea”.

En sintonía con la sección monográfica, la sección miscelánea, compuesta de 12 artículos, comienza con un trabajo en el que se analizan los conceptos de transhumanismo y posthumanismo; con ellos queda también planteado el tema de las posibilidades y los límites de la acción humana a través de la técnica. Sobre la problemática de los seres humanos, versan otros dos trabajos más: Uno de ellos se ocupa de la teoría del reflejo y su importancia para la construcción de la identidad; el otro se pregunta por el papel del ser humano en el cosmos y por la posibilidad de una filosofía de la naturaleza. Otro artículo estudia el impacto del clima en los sistemas de gobierno. Siguen otros ocho trabajos, que desarrollan aspectos concretos de autores relevantes de la historia de la filosofía: Hume y la experiencia estética; la experiencia religiosa en Kierkegaard; Kierkegaard y su libro *Forord*; los vínculos de arte y poder en Nietzsche; el tema de la atención en Simone Weil; el concepto de felicidad en Ortega y Gasset; la teoría del lenguaje en *El ser y la nada*, de Sartre; la concepción del diálogo en Emmanuel Levinas.

A la sección miscelánea sigue una Nota crítica sobre los escritos de Zubiri *Ciencia y realidad*, y *Filosofía primera*, con motivo de los cuales el autor informa sobre el estado actual del *Corpus zubiriano*. A esta Nota crítica sigue un conjunto de cinco reseñas sobre publicaciones recientes, referentes a Manuel Sacristán sobre Jean Paul Sartre; Peter Michael Hacker; María Martín Gómez, Rico Sneller y Byung Hang. Cierra el volumen una traducción de *Mi propósito principal*, de Gabriel Marcel.

A mi juicio, que una Universidad publique revistas científicas y, en este caso, una revista de impacto, es muy relevante desde el punto de vista de la investigación. Como mínimo, es señal de su compromiso con la misma; no solo con la que se hace dentro, sino también con la que se hace en otras universidades; con ello, se potencia la necesaria comunicación del conocimiento. En este volumen colaboran investigadores procedentes de unas 28 instituciones de España, Europa y América.

El Consejo de redacción de la revista y la Dirección de la misma desean expresar su agradecimiento a todos ellos. Igualmente, nuestro reconocimiento hacia quienes han colaborado como expertos en la evaluación de los trabajos: un total de 59, de más de 40 lugares diferentes.

Apenas había terminado de escribir estas líneas cuando recibimos la noticia del inesperado fallecimiento del Profesor Jacinto Rivera de Rosales, miembro del Consejo de redacción de la revista. ¡Qué gran crueldad y qué inmensa tristeza! El seis de septiembre había asistido a nuestra reunión del Consejo de redacción, donde participó activamente en el diseño del próximo volumen de la revista. He compartido con él simposios, grupo de investigación, tribunales, congresos de la SEKLE. Sirvan estas breves palabras para expresar nuestro gran reconocimiento a su intensa y creativa actividad académica e investigadora y para darle las gracias por su generosidad y su amistad.

Unos meses antes, en agosto de 2021, fallecía el Profesor Saturnino Álvarez Turienzo, fundador de *Cuadernos salmantinos de filosofía*, primer Director y Miembro Honorario de la misma. Colaboró con la revista hasta donde le llegaron las fuerzas y no pasaba un año en el que no nos felicitara por la edición de un nuevo volumen. Como muestra de nuestro reconocimiento y profunda gratitud, en la última reunión del Consejo, acordamos dedicarle una sección en el volumen 49 de *Cuadernos*, correspondiente a 2022.

ANA MARÍA ANDALUZ ROMANILLOS

Directora de *Cuadernos salmantinos de filosofía*

Salamanca, octubre de 2021

ORCID: 0000-0002-9033-0710

ESTUDIOS

**SECCIÓN MONOGRÁFICA:
FILOSOFÍA E INTELIGENCIA ARTIFICIAL**

LAS HUMANIDADES DIGITALES EN Y PARA LA HISTORIA DE LA FILOSOFÍA. UNA INTRODUCCIÓN

*DIGITAL HUMANITIES IN AND FOR THE HISTORY
OF PHILOSOPHY. AN INTRODUCTION*

MANUEL LÁZARO PULIDO

Doctor en Filosofía (UPSA)

Catedrático. Facultad de Derecho.

UNIR-Universidad Internacional de La Rioja

Investigador asociado de la Universidad Bernardo O'Higgins, Santiago, Chile

Madrid/España

mlazarop@fsf.uned.es

ORCID: 0000-0002-0064-5293

Recibido: 16/06/2020

Revisado: 17/07/2021

Aceptado: 6/09/2021

Resumen: El presente artículo pretende ser una introducción a las Humanidades digitales (HD), intentando presentar qué son desde su evolución. El objetivo no es solo presentarlas, sino indicar la potencialidad de su uso en y para la historia de la Filosofía. Intentaremos señalar en su recorrido los diferentes usos que se han ido dando, señalando proyectos que sirven de referente e incluiremos algunas reflexiones de corte filosófico. Partiremos de las Humanidades Computacionales para derivar a las Humanidades Digitales, en sintonía con el desarrollo de la Web, de la 1.0 a la 3.0.

Palabras Clave: Humanidades Computacionales, Humanidades Digitales, Web, Historia de la filosofía, Procesamiento del Lenguaje Natural.

Abstract: This paper is aimed to be an introduction to the digital Humanities (HD), trying to present what they are since their evolution. The aim is not only to present them, but to try to indicate the potential of their use in and for the History of Philosophy. We will try to point out the different uses that have been made of them, pointing out projects that serve as a reference and including some philosophical reflections. We will start from the Computational Humanities to derive the Digital Humanities, in line with the development of the Web, from 1.0 to 3.0.

Keywords: Humanities Computing, Digital Humanities, Web, Natural Language Processing.

1. LA INTRODUCCIÓN DE LA TECNOLOGÍA Y EL LENGUAJE COMPUTACIONAL EN LA HISTORIA DE LA FILOSOFÍA. UN NUEVO LENGUAJE: *HUMANITIES COMPUTING (HC)*

Uno de los nuevos lugares (*tópicos*) que iluminan el quehacer humanista del siglo XXI es la llamada revolución de las Humanidades Digitales (HD). No hay universidad que se precie y quiera presentarse como novedosa o, al menos, en sintonía con los desafíos que se plantean en el siglo XXI que no contemple de alguna u otra forma en el área de las humanidades algún desarrollo curricular o de investigación, o ambas a la vez, de HD. Es lo que hay, y es por lo que ahora mismo se financia en la investigación en humanidades tanto pública como privadamente, en Europa y en España¹.

Si bien todos hablan de las HD, sin embargo, no resulta sencillo definir “qué sean las HD”. Existen diversas formas de definir las, conceptualizarlas o etiquetarlas. Se trata de una historia no muy larga, pero que ya cuenta con una nutrida literatura y que va desde la pregunta que se hiciera John Unsworth –“What Is Humanities Computing and What Is Not?” (2010)–, pasando por las reflexiones de Willard McCarty (2005), a las más recientes, que tienen en cuenta el cambio constante que experimenta este campo. Un área que ha pasado –veremos brevemente– de las humanidades computacionales o computación de las humanidades (HC) (Burnard, 2002) a las HD (Svensson, 2010). En 2013 contábamos ya con más de veinte definiciones (Terras, Nyhan y Vanhoutte, 2013) y hoy en día podríamos multiplicar el número por diez, solo en el área anglosajona (Callaway *et al.*, 2020).

No obstante, más allá de las divergencias, podemos convenir que las HD hacen referencia tanto a una tecnología aplicada a un objeto o campo de estudio,

1 Ejemplos lo tenemos en EADH (European Association For Digital Humanities) que “reúne y representa a las Humanidades Digitales en Europa en todo el espectro de disciplinas que investigan, desarrollan y aplican los métodos y la tecnología de las humanidades digitales. Estas incluyen historia del arte, estudios culturales, historia, procesamiento de imágenes, estudios de lenguaje y literatura, estudios de manuscritos y musicología, entre otros”. URL: <https://eadh.org/>. Otro ejemplo lo tenemos en los proyectos en Humanidades financiados por el BBVA. Según la última convocatoria, en este caso 2019 (URL: <https://www.fbbva.es/ayudas/ayudas-equipos-investigacion-cientifica-humanidades-digitales-2019/>):

“En el área de Humanidades Digitales, los proyectos de investigación deben incorporar de manera central:

- La aplicación de las tecnologías de la información y de técnicas estadísticas avanzadas para el tratamiento de objetos propios de las humanidades.
- El abordaje desde la perspectiva de las humanidades de los nuevos objetos digitales, y el análisis de los efectos de Internet y el espacio digital en el plano de la cultura.
- La difusión innovadora, a través de la web, de los objetos y la producción propia de las humanidades.”

en nuestro caso objetos que pueden digitalizarse, y que forman la fuente del campo de las humanidades –objetos que van desde el soporte físico (libros, edificios, localizaciones, objetos artísticos, restos arqueológicos...), hasta contenidos con capacidad de ser digitalizados (etiquetados...). Se trata de una actividad digital que tiene vocación social, es decir, se entiende como un trabajo o empresa social: una red de personas compartiendo conocimientos, campos de estudio y tecnologías de estudio clásicas y digitales... Hacer, conectar, interpretar y colaborar son elementos unificadores de las HD. Todo ello en un entorno móvil y de crecimiento en el que se combinan diferentes enfoques en distintos marcos, integrando el desarrollo de nuevos campos de la cultura ingenieril, tales como el aprendizaje por máquina, la ciencia de los datos y la inteligencia artificial.

Nos encontramos ante una nueva, y a la vez tradicional, encrucijada en las humanidades, y en el caso que nos conviene, en la filosofía. Y es que la elaboración filosófica ha tenido siempre una estrecha relación con el soporte en el que se ha realizado. La filosofía surge en una cultura oral. Es verdad que, en sí, la filosofía es una cosa difícil de explicitar, sin duda; pero, por otra parte, sabemos que es una reflexión que se hace “en lo que es” en el lenguaje. La filosofía surge en el “decir”. Hablamos de “decir” como lenguaje vivido y hablado –claro está que siempre viviéndose en cuanto hablándose– y lo hacemos, aunque sepamos que el decir se iba reflejando en el “escribir” como lenguaje vivido y escrito (Lázaro, 2011, pp. 25-26). En este sentido, la realidad filosófica escapa a la dialéctica latente en la pugna entre *logos* y *gramática*. Nos recuerda Derrida (1987) que para Platón “el arte o la técnica (*tecne*) de la escritura era un *fármakon* (droga o tintura, saludable o maléfica)” (pp. 367-368). Por otra parte, la idea de la importancia de la escritura fue señalada por Walter J. Ong (2002) para quien la filosofía debería ser consciente de que es el producto de la tecnología –como señala Jack Goody, una razón gráfica (1979)– o mejor de un producto muy humano, al que se somete la propia lógica: la tecnología de la palabra. La disputa filosófica daría para un ensayo nuevo. No es nuestro propósito aquí, pero quizás conviene recordar la advertencia impulsada por Derrida respecto a la filosofía, “en cuanto a no confundir el conocimiento de la verdad (*episteme tes alezeias*) con la técnica de la gramática (*téchne grammatiké*)” (Lázaro, 2011, p. 36). Sin duda, la reflexión filosófica no solo se ve mediada por la escritura, sino por la extensión de la técnica libresca, y especialmente con la introducción de la imprenta. En todo caso, señala Klock-Fontanille (2014): “Todos los investigadores están de acuerdo en la importancia del soporte y en que es un elemento en la construcción del significado. En otras palabras, el soporte y el texto se entienden como objetos semióticos de escritura” (p. 30).

Si la oralidad, la escritura... el soporte, han planteado una reflexión filosófica, el hecho de que este soporte sea acrecentado por la computación y la digitalización

—que suponen no solo una suma de los soportes, sino una interrelación y su sobre-dimensión—, nos lleva a plantearnos cuestiones en sí filosóficas. Más aún cuando la introducción de la tecnología digital lleva tras de sí cuestionamientos epistemológicos, ontológicos y/o semánticos en tanto que no solo es intrínsecamente un soporte, sino un medio con herramientas de investigación en sí misma.

Más allá de las cuestiones filosóficas que entraña y que precisan de ser pensadas, la tecnología computacional y digital, aplicada a las humanidades, también introduce una reflexión filosófica (racional, axiomática, ética...) relativa a la propia cultura digital en el entorno vital y existencial humano. Efectivamente, uno de los retos del siglo actual se plantea en los términos acerca de la influencia que los nuevos soportes y herramientas de producción racional tienen en la cultura, en la sociedad, en el acceso cultural y en la propia definición de realidad humana, en definitiva, en las áreas de conocimiento que tradicionalmente estudian las humanidades. Más aún, en cuanto cuestionamiento en sí, en la filosofía, lo que ha provocado un pensamiento sobre la “interacción renovada entre ciencias, tecnologías y culturas” (Dominique, 2013, p. 64). Detrás de las tecnologías computacionales y digitales existe todo un contexto de racionalidad, proyectos humanos y culturales, representación y recreación del mundo... que desafía la propia lógica digital, toda vez que esta no deja de ser desarrollada por seres humanos (Verry y Fagerjord, 2017).

A pesar de todo lo que hemos comentado, la realidad es que aquí nos referimos a la historia de la filosofía y nuestra cuestión descansa en una breve presentación del contexto de las HD en la historia de la filosofía y cómo sirven para su desarrollo, a través de sus herramientas básicas.

En principio, podríamos decir que la historia de la filosofía tiene como objetivo “la lectura, comprensión e interpretación de las obras que los filósofos nos han transmitido, y el de su ordenación en una secuencia que forme una «historia»” (Sánchez, 2003, p. 9). Para ello es necesario considerar que la relación de la filosofía con su historia es diferente a la de otras ciencias, puesto que la propia filosofía va evolucionando en cuanto tal en su desarrollo histórico, de modo que la historia no constituye en sí un mero apéndice, como señala Martial Gueroult (1979): “Quien recurre a la filosofía existente encuentra, no un cuerpo de verdades, sino su historia” (p. 26). Más aún, incluso la propia concepción histórica de la narración del tiempo pasado está imbuida de una construcción teórica y una visión filosófica —por otra parte, la propia historia como tal en cuanto que precisa de la interpretación de los datos también tiene una relación filosófica en la que la filosofía de la historia se hace presente al menos implícitamente—.

Las relaciones de la filosofía con la historia están transidas de criterios que atienden, pues, a la problematicidad inherente a la interpretación. Como señala Antoni Bordoy, algo tan especulativo como “la metafísica debe ir necesariamente

de la mano de una historiografía que se rija por los mismos métodos y patrones que en el resto de las ciencias” (Bordoy, 2020, p. 306). Eso lleva siempre un componente creativo, pero también una sensación de provisionalidad, sino de cierta arbitrariedad o, mejor, inseguridad. La filosofía se debate –señala Jacques Bouveresse (2017)– “entre la amenaza del escepticismo radical y la ilusión del conocimiento”.

Ante tal incertidumbre puede resultar útil la introducción de mecanismos metodológicos que asistan al investigador a ofrecer elementos de objetividad y que sirvan, a la vez, como instrumentos que apoyen la labor hermenéutica a la que está abocada dicha investigación. Creo que no debemos negar que los métodos que proporcionan las tecnologías de la computación y el entorno digital, tales como los bancos de datos (los *big data*), ... pueden ser de utilidad tanto para la identificación de los textos filosóficos, como para su comprensión y análisis en la formación de las ideas, e implementación de una nueva hermenéutica digital (Tatarinov y Fickers 2021). Los métodos computacionales y los métodos digitales nos ofrecen enfoques de aplicación a la historia de la filosofía que no tienen por qué competir con los tradicionales métodos interpretativos, sino que pueden y deberían ser una alternativa complementaria e, incluso, enriquecedora, en la medida que se introduce una nueva lógica (y, por lo tanto, una racionalidad) y, a la vez, un enfoque interdisciplinario. Ahora bien, ello supone, también, un renovado esfuerzo de pensamiento “meta-” en el quehacer filosófico. Si como afirman Anheim, Lilti y Van Damme (2009), “la historia de la filosofía, pensada como una exigencia intelectual y como una práctica pedagógica, ha sido durante mucho tiempo, a los ojos de los filósofos, un modo de escritura privilegiado”, la implementación de herramientas de análisis de texto, como puede ser la estilometría, pueden ser, sin duda, de gran utilidad para el historiador de la filosofía.

La percepción de la utilidad de las nuevas tecnologías en el contexto de la investigación de la historia de la filosofía, en particular, y de las humanidades, en general, han ido paralelas al propio desarrollo tecnológico. Podemos decir que, como veremos, su implementación fue bastante intuitiva y que, en ese proceso –pues no podemos pensar que este uso es algo exclusivo del siglo XXI–, la historia de la filosofía, especialmente la medieval y la tradición escolástica, ha tenido y tienen un papel protagonista. Si hoy en día hablamos de HD en el entorno de la web 4.0, hay que pensar que es así porque han existido cuatro versiones anteriores y que ese subsistema se sustenta en un sistema que tiene como esqueleto el entorno de la ingeniería computacional, algo que hunde sus raíces en el siglo XX. Las HD son un último (en tanto que presente) estadio de una historia que descansa, en relación con las humanidades, en las HC.

1.1. HUMANITIES COMPUTING (HC)

El origen de las HC va a la par con el desarrollo de la propia tecnología informática, desde la automatización a la computación. El surgimiento de nuevas herramientas de tratamiento de información fue visto como una oportunidad, entonces visionaria, del estudio de las humanidades. En este sentido, podríamos decir que existen varias tareas aplicadas desde la informática a los objetos humanísticos, especialmente los textos, que ya se pudieron realizar desde el inicio, piénsese en el trabajo estructural y de metadatos con los textos mediante la codificación o los sistemas de etiquetados.

En este sentido, el desarrollo de la aplicación de las tecnologías de la computación a las humanidades ha tenido una doble vertiente. Por una parte, han ido acompañando los avances de dichas tecnologías. Por otra parte, simultáneamente, han aprendido a usar, cada vez mejor y de forma más eficiente, dicha tecnología en sus actualizaciones. Pero, a la par que iban acompañando el avance informático e ir profundizando en su aplicación como herramienta al estudio de las humanidades –en nuestro caso la historia de la filosofía, especialmente medieval– su recorrido ha servido también de impulso a la implementación de otros recursos y herramientas en el campo de la tecnología.

Si hoy hablamos de HD y no de HC, no es porque la computación aplicada a las humanidades haya desaparecido, sino porque el desarrollo tecnológico se ha ampliado en sus aspectos epistemológicos, en las herramientas ingenieriles, en los lenguajes formales, en la recreación de nuevos espacios –como el caso de Internet y sus subsistemas–, en la entrada de realidades complejas como la I.A..., que afectan también a la propia cultura de la investigación, tal como hemos señalado en el apartado anterior.

En este sentido, en algunos aspectos, la HC fue sustituida por la HD, por ejemplo, en las transformaciones significativas de la web, que no solo evolucionó en el desarrollo de sus herramientas, sino en los conceptos y en su enfoque multidisciplinar ingenieril y social. Pero en muchos otros aspectos, el hecho de que hablemos del paso de las HC al HD no significa la anulación de la primera frente a la segunda, en el sentido en que, efectivamente, la informática (computación), sigue siendo, lógicamente, básica en el desarrollo de las HD (Svensson, 2009).

1.1.1. *Las humanidades computarizadas*

Ciertamente el término inglés *Humanities Computing* (HC), admitiría la traducción al español tanto de “computación en humanidades”, como de

“humanidades computarizadas”. En este sentido, me parece pertinente recordar la reflexión de quien era entonces profesor de “*Humanities Computing*” en el King’s College de Londres, Willard McCarty. En su libro *Humanities Computing* realizaba un análisis sobre el modo en que la informática contribuye a responder a la cuestión ineludible en las ciencias humanas de plantear cada vez mejores y más desafiantes preguntas. En el desarrollo del trabajo plantea una triple caracterización del desarrollo que han experimentado las relaciones entre la informática (computación) y las humanidades. Según el autor estas han pasado del concepto de “*computers and the humanities*” (computadoras y humanidades) –es decir, del deseo por parte de algunos humanistas de poder aplicar la herramienta informática a las humanidades–, a la impresión de “*computing in the humanities*” (computación en las humanidades) –que describe el hecho de la entrada de los humanistas en el mundo informático y su aplicación–, para al fin situarse en las “*humanities computing*” (humanidades computarizadas) –concepto que describe la situación en la que el investigador en humanidades confía en la informática, si bien sigue siendo algo enigmático– (McCarty, 2005, p. 3). El investigador en humanidades se pregunta en este contexto cómo se desarrolla el encuentro entre máquina y humanidades, cuestiones en torno al entendimiento de la computación como una máquina, la metodología, la naturaleza del trabajo de la computadora, la epistemología... (Orlandi, 2002).

En este devenir de la relación entre informática y humanidades, destaca un proceso comunicativo que va del objeto a la herramienta informática, de la herramienta al investigador (que ha diseñado la relación primaria y analiza los datos obtenidos de dicha relación) y del investigador a la comunidad científica, mediante un soporte unidireccional que, aunque diferente, cumple básicamente el mismo cometido comunicativo: el papel, el soporte informático (disco...), por último, la Web en Internet. De este modo, esta etapa viene caracterizada por el desarrollo de herramientas informáticas y la creación de archivos y bases de datos para textos, obras de arte y otros materiales. En base a esta direccionalidad, incluimos en este apartado la primera versión de la Web (1.0).

Si miramos con la perspectiva clásica podríamos decir que el inicio de la HC es reciente. Mirada desde la perspectiva del tiempo actual, especialmente en la línea de historia del mundo computacional, podríamos decir que se sitúa ya en el origen de los tiempos (de la era informática). Y en ese inicio destaca, precisamente, los estudios en historia de la filosofía, más aún de historia de la filosofía medieval, y concretando si cabe: los estudios de historia de la filosofía (y la teología) medieval cristiana.

El protagonista es un referente por todo el mundo conocido, el jesuita Roberto A. Busa y el objeto de estudio, la obra de santo Tomás de Aquino. Roberto A. Busa empezó a barruntar el potencial de análisis computacional de los textos durante

la Segunda Guerra Mundial, recién terminada su tesis doctoral sobre la *Summa Theologica*, al observar las máquinas de automatización de análisis lingüístico de textos escritos (Busa, 2008, p. xvi). El jesuita quería trabajar sobre el vocabulario de la presencia de la encarnación en las obras del doctor Angélico. Para ello debía estudiar las concordancias en los sustantivos, pero también en las partículas, lo que supone el rastreo de millones de términos. El proyecto de Roberto Busa, consistía, en una primera instancia, en la adaptación de la localización textual computarizada, en el conteo y la comparación de textos. Esa tarea inhumana podía encontrar solución en la utilización de las nuevas técnicas de la ingeniería de la computación a los campos de la concordancia y la lingüística del corpus (Berra, 2012). Con un equipo de cinco personas, desde la alianza con IBM, empezó a reescribir la obra de Tomás de Aquino en tarjetas perforadas, de modo que llegó a generar automáticamente un índice de cada palabra del corpus. La historia exitosa de la persistencia y competencia del padre Busa –que ha dado lugar a más de un artículo (Mounier, 2018)– dio como resultado el *Index Thomisticus*² (11 millones de palabras) publicado en 1972, del que disponemos una edición Web alojada en el *Corpus Thomisticum*, editada por Eduardo Bernot y Enrique Alarcón y sostenida por la Fundación Tomás de Aquino, la Universidad de Navarra e IBM.

Desde este momento pionero se va fraguando un trabajo centrado en la recogida de datos que componen el texto. El trabajo del equipo de Busa fue clave para toda la labor de codificación, marcando el inicio de la computación lingüística. Siendo precursora la reflexión literaria de los textos de la historia de la filosofía (y la teología) medieval, se sucederán una serie de proyectos que irán implementando la idea original, especialmente aplicados al campo de la filología. Así, en 1963, Roy Wisbey fundó, en Cambridge, el *Centro de Computación Literaria y Lingüística (Centre for Literary and Linguistic Computing-LCC)* para apoyar su trabajo con textos complejos escritos en lengua alemana. Se inicia una época en la que se desarrolló un trabajo intenso en los centros de computación que revertirá en el estudio de las humanidades. Esta labor crea conciencia entre los humanistas fundándose diferentes sociedades académicas. Entre estas podemos citar, en la década de los setenta del siglo pasado, la *Fundación de la Asociación para la Computación Literaria y Lingüística (The Association for Literary and Linguistic Computing-ALLC)*, en 1973, con el fin de apoyar la aplicación de la computación en el estudio de la lengua y la literatura, siendo el origen de la actual *Asociación Europea de Humanidades Digitales (European Association For Digital Humanities-EADH)*. Cinco años más tarde, en 1978, surge en Boston la *Asociación para la Informática y las Humanidades (Association*

2 Mantenemos hiperenlaces para la versión online de la revista soportada en OJS.

for *Computers and The Humanities-ACH*), en 1978, que actualmente abre su campo a las HD³. Este es un muestreo somero de la actividad societaria que fue animando e impulsando el trabajo de textos usando la herramienta informática.

Para este trabajo de la historia de la filosofía medieval también han resultado básicos los intentos de trabajos sobre bases de datos computarizados, especialmente, en lo que se refiere a la labor lexicográfica del latín medieval, cuyas fuentes son, a menudo, los teólogos de la época. Benoit-Michel Tock subrayaba a principios del presente siglo, en la transición de la Web 1.0 a la Web 2.0, los límites de este trabajo informático, señalando la importancia de la herramienta a la hora de poder agilizar su trabajo. Entonces aún se trataba de explotar las bases de datos en la búsqueda de citas o en el aprovechamiento de repositorios (en la época aún limitados), pero el autor ya barruntaba algunas perspectivas de lo que entonces era aún la orientación de las HC: “nos permiten dar un nuevo enfoque a los textos en los que estamos trabajando. Este nuevo enfoque no sustituye, por supuesto, a ninguno de los métodos tradicionales, sino que se suma a ellos” (Tock, 2001). Esta mentalidad traslucía lo que ha sido uno de los escollos que acontecen en HC y HD, que es de índole mental y de cultura de investigación y que atañe al hecho de que la forma de pensar e investigar con textos electrónicos está influida por los hábitos de la cultura libresca (Johansson, 2004).

El desafío de esta época era dar un salto más allá de la digitalización o la datación parcial de los documentos. Desafortunadamente, hay muy pocas bases de datos de textos mediolatinos lematizados: ni la *Patrología Latina10*, ni la *Biblioteca Cetedoc de Textos Latinos Cristianos*, ni el *cd-rom MGH*, ni el *The-saurus Diplomaticus*, presentan esta herramienta y, en este sentido, no permiten implementar plenamente los requisitos de extracción de conocimientos profundos en los datos de los documentos, de proporcionar un análisis visual y de apoyar eficazmente la investigación de los estudiosos de las humanidades de la historia de la filosofía. Esto no supone el fin, ni mucho menos, sino que abra camino a los nuevos retos que vendrán de la mano de las HD y cuyos primeros pasos se están realizando, como veremos más tarde al hablar del Procesamiento de Lenguaje Natural (PLN).

3 Esta inclusión de las HC en las HD se ve en su presentación: “ACH aboga y apoya a todos nuestros miembros en su trabajo de humanidades digitales. Humanidades digitales es un término amplio que abarca una amplia gama de dominios temáticos, métodos y comunidades de práctica, incluyendo (pero no limitado a) la investigación asistida por ordenador, la pedagogía y el software; la creación, la conservación y el compromiso de recursos; la informática física; el uso de las tecnologías digitales para escribir, publicar y revisar la erudición; y la investigación humanística en y sobre los objetos y la cultura digitales. ACH reconoce que esta labor es intrínseca e inextricablemente sociopolítica, por lo que aboga por el cambio social mediante el uso de computadoras y tecnologías conexas en el estudio de temas humanísticos”. <http://ach.org/about-ach/> [Consulta: 3 de abril de 2020].

Si nos centramos en la historia de la filosofía medieval, nos encontramos con diversas herramientas computarizadas que responden a esta fase de desarrollo. Jean Luc Solère ha recopilado las diferentes ediciones electrónicas (portales, basas de datos de autores múltiples y de autores individuales), libros impresos digitalizados y manuscritos en portales y Web sites (formato Web 2.0), disponible desde la SIEPM – Société Internationale pour l'Étude de la Philosophie Médiévale: en las pestañas *Editions* y *Manuscripts* del site *Medieval Philosophy Digital Resources*.

A pesar de los límites de las HC en el campo del estudio de las humanidades, estas tuvieron y siguen teniendo un papel relevante. Como hemos dicho, las HC no han desaparecido, ahora conviven con otros elementos que han ampliado la informatización, básica en el entorno de la biblioteca digital, mediante el desarrollo de la rama algorítmica, metatextual y representacional (Willard McCarty, 2002), a la digitalización. Pero la computarización como herramienta y como concepto sigue siendo imprescindible. Por otra parte, las HC no solo fueron pioneras, sino que como tales implementaron herramientas que alentaron el trabajo y la simbiosis entre la máquina y el investigador. Su éxito estriba en su evolución intrínseca –en tanto que HC en constante desarrollo– y en su evolución conceptual hacia nuevas formas como las HD. Además, su uso –su potencialidad al principio y su realidad más tarde– animó al mundo académico a su incorporación curricular. Si estamos escribiendo este trabajo ahora, si se realizan máster en HD, si se aumenta el número de cursos y programas de licenciatura en informática para humanidades..., es por el éxito de las HC.

1.1.2. *La Web 1.0*

Como hemos señalado, el mundo de la computación, las tecnologías digitales y el entorno de herramientas Web (World Wide Web, o www) se define por su dinamismo. Mientras que Internet es una enorme red pública de redes, un laberinto de conexiones cableadas e inalámbricas entre unidades dispares en incontables dispositivos dispersos por todo el mundo, la Web es un sistema estandarizado para acceder y navegar por Internet. Es, pues, una plataforma sobre la que se construyen tecnologías innovadoras y un espacio en el que participan los usuarios. No es el único (el correo electrónico y aplicaciones para móviles, por ejemplo, no utilizan la Web para conectarse a la Internet), pero es, con mucho, el más común. La Web es, pues, un subsistema –el más común– de Internet, junto a otros como el correo electrónico, las aplicaciones, o las redes sociales, blogs y wikis. Se distingue por el uso de HTTP (Hypertext Transfer Protocol) que, a su vez, es sólo un sistema para estandarizar el uso de HTML (HyperText Markup Language). Este

no es un lenguaje de programación, sino un lenguaje descriptivo, utiliza etiquetas que el navegador interpretará de una u otra forma para mostrar distintos contenidos por pantalla. Esencialmente HTML es el lenguaje de la Web, y HTTP son las reglas gramaticales para su uso.

Tanto las tecnologías como el modo de participación van cambiando y han creado con el tiempo distintas versiones de la Web, así hablamos de la Web 1.0, 2.0, 3.0, 4.0 y pensamos las siguientes generaciones. En cierta manera, la Web 1.0 es la introducción, la primera fase de la Web 2.0. Esta primera fase de la Web estaba liderada, principalmente, por compañías que buscaban una presencia en la misma. El contenido debía ser publicado principalmente por las empresas, si bien también existieron algunos individuos que lo hicieron a través de páginas web personales. Este contenido era típicamente accesible en modo de solo lectura. Esos sitios web solían presentar contenido estático o contenido de bases de datos de catálogos que se presentaban al usuario, pero con poco contenido posible proporcionado por el usuario. La Web 1.0 queda definida por su unidireccionalidad, el webmaster ejercía el control absoluto del contenido en un entorno definido por las páginas estáticas. El énfasis de la Web 1.0 era “publicar” el contenido en la Web *para* los usuarios, no el contenido proporcionado *por* el usuario. Otros aspectos de la Web 1.0 son los navegadores poco sofisticados y con un acceso limitado en velocidad a Internet, a menudo a través de velocidades analógicas de 56 K o menos (Anandajaran, et al., 2010, p. 505).

La Web 1.0 evolucionó antes de ser caracterizada, como veremos, en Web 2.0 a través de la Web 1.5. Esta evolución supuso la introducción de web dinámica frente a la web estática, pero sin llegar a la Web colaborativa de 2.0. El dinamismo se iba implementando debido al uso del correo electrónico, que era un medio de introducción de información, de los grupos de noticias, de las páginas webs (documentos con hipertexto), que, aún estáticas, iban permitiendo enlaces y textos hipervinculados, aplicaciones interactivas como los chats o foros de discusión... Aún así, “el perfil del usuario tipo en la Web 1.0 y la Web 1.5 es principalmente, pasivo y consumidor de información. El usuario final es el receptor de la información que transmite el sitio web, limitándose a la toma de información desde la red, asemejándose esta a una gran biblioteca a la que puede acceder fácilmente” (Reuelta y Pérez, 2009, p. 63).

A pesar de las limitaciones que podamos ver de la Web 1.0 (Web 1.5.), si lo observamos retrospectivamente, observamos que la llegada de la World Wide Web (www) implicó un avance en la concepción de las HD. Entendidas, en este entorno inicial, como una especie de oportunidad de ampliar el espacio de la biblioteca y el archivo y, por lo tanto, de una inicial base de datos, la Web 1.0 supuso el inicio del registro electrónico de los datos al trabajo digital usando información tecnológica.

La Web 1.0 inició, además, la digitalización, y, por lo tanto, la accesibilidad, mediante procesos computarizados de información valiosa para la investigación en humanidades, y en lo que nos concierne para la historia de la filosofía. Múltiples fuentes a las que no se tenían acceso empezaron a ser digitalizadas, catalogadas y disponibles en repositorios. En este momento, los proyectos de digitalización envolvían procesos de conversión de materiales a un formato digital y se iniciaba, tímidamente, la introducción de mejora y calidad de imágenes y el reconocimiento óptico de caracteres (OCR). Los resultados se presentaban en una página web, pero sin incluir aún la interoperabilidad para que la colección digital fuera incorporada fácilmente a otras colecciones, algo que pertenece ya a una fase posterior.

Proliferan las bases de datos de información bibliográfica, como la pionera *Philosopher's Index*, que es un ejemplo del salto del papel al soporte informático primero y a la red, después, y está en la raíz de las bases de datos. Estas bases terminarán, no pocas veces, fagocitando la investigación en la medida en que, con el avance de la digitalización y la optimización de sus recursos métricos para evaluar, mediante algoritmos, el espíritu de los artículos de la filosofía, sustituyan, así, la evaluación del lector. Esto hace pensar en que no se puede caer en la tiranía del algoritmo, o mejor en la reducción de la masa crítica administrativa y universitaria a los datos cuantitativos y las agencias, las nuevas academias y sociedades científicas que constituyen el criterio de demarcación de una investigación (*Scopus*, *WoS*..., por citar los más representativos).

En cierta forma la Web 1.0 es la pieza que posibilita el salto a las HD. Algo así como la primera fase, unas Humanidades 1.0, a las que le seguirán las Humanidades 2.0 asimiladas a la Web 2.0. Así, mientras, que la primera se basa en datos de primera generación, la segunda supone interactividad y la apertura sobre la participación basada en un conjunto diferente de premisas teóricas que descentran el conocimiento y la autoridad (Davidson, 2008, pp. 711-712).

Efectivamente, las HC, en tanto que un enfoque basado en la herramienta, se centran en la construcción de herramientas, infraestructuras, normas y colecciones, pero la Web 1.0 ya introduce la capacidad de presentar, de forma limitada y muy unidireccional –pero no exclusivamente–, medios de comunicación en red e intercambio de información. De este modo, presentan y abren el camino a herramientas académicas, bases de datos, escritura en red..., en los que se empieza a aprovechar el potencial de los medios visuales y auditivos que forman parte de la vida contemporánea, es lo que McPherson denomina humanidades multimodales (2009).

La publicación en el año 2004 de la obra *A Companion to Digital Humanities* (Schreibman, Siemens, y Unsworth, 2004) sugería que, efectivamente, el trabajo previo de las HC, junto al primer desarrollo de la Web (1.0 y 1.5) ya habían formado un corpus suficiente que permitiera hablar de nueva concepción de la interacción entre tecnología y humanidades: las HD.

2. LA INTRODUCCIÓN DE LA TECNOLOGÍA Y EL LENGUAJE DIGITAL (HD): REFLEJO EN LA HISTORIA DE LA FILOSOFÍA

La evolución de la incorporación de nuevas herramientas y los soportes en la sociedad en general y, también, en la sociedad científica, entendida está también las humanidades, nos ha llevado ahora a la situación de la presencia de las HD. Surge la pregunta inicial, ¿qué es la tecnología digital, y en qué se diferencia de la computación?

La computación o informática y la introducción de lo digital son el escenario visible, no incompatible, al contrario, de diferentes etapas u olas tecnológicas. Se trata de un flujo que se inicia con la informática o computación, se desarrolla a través de la tecnología y sistemas de información, y que ha desembocado actualmente en la digitalización. Y todas estas fases se van retroalimentando.

La informática, en cuanto que desarrolla tecnología y diseño de artefactos aplicando lenguajes algorítmicos, no ha desaparecido y sigue constituyendo la base del edificio tecnológico, en cuanto que es la herramienta fundamental. Con ella podemos implementar tanto la tecnología de la información como sus sistemas, que constituyen la apertura mental de la interoperabilidad que será aplicada en el diseño del desarrollo de la Web del 1.0 al 2.0.

Por su parte, lo digital explicita la ingeniería informática y de información entendida en múltiples facetas y diseños ingenieriles y socio-comunicativos que interactúan entre sí, no solo en redes de actuación conjunta, sino, también, mediante sistemas de análisis de las partes (los datos) que conforman la red y de la propia red en tanto que sistema. Lo que supone este desarrollo es la integración tecnológica de comunicación y conocimiento (TIC's), que no solo es el almacenaje del conocimiento y una primera visualización, sino que permite su universalización y análisis mediante herramientas de ingeniería e inteligencia artificial (y ello supone, lógicamente el desarrollo de la computarización o informática). Se realiza, así, lo que se conoce como convergencia tecnológica.

Por su parte, la convergencia tecnológica es el “entrelazamiento de especies o tecnologías”, que “se refiere a una tendencia en la que un solo producto como el teléfono móvil, utilizado en el pasado únicamente para la comunicación, evoluciona hasta convertirse en un producto que funciona no solo como un dispositivo de comunicación, sino que incorpora las funciones específicas de una serie de otras tecnologías, lo que permite a los usuarios tomar fotografías, escuchar música, acceder a la Web, enviar y recibir mensajes de correo electrónico, encontrar una dirección, etc., con el mismo éxito” (Kurbanoglu et al, 2010, p. 1).

Este desarrollo tecnológico afecta el estudio, la investigación y la transferencia de conocimiento –incluida la docencia– de las humanidades. Efectivamente,

la forma de realizar la investigación en los centros de investigación y en las universidades, tanto en su elaboración como en la transferencia de conocimiento ya ha cambiado, es un hecho que no podemos escamotear. El momento actual, tal como hemos señalado, está dominado por la tecnología digital, que es, a su vez, dinámica, por definición está en continuo movimiento tanto cuantitativa como cualitativamente. El encuentro puede y debe ser feliz, sumativo e incluso exponencial, a la par que no debe ser reductivo.

Mirando nuestro objeto de reflexión en humanidades –la historia de la filosofía–, no se trata de sustituir su esencia, sino de procurar su encuentro fecundo con las tecnologías digitales, comprendidas las ciencias computacionales. Que la historia de la filosofía se abra a la *trans(multi-)disciplinarietà*, supone, a su vez, la capacidad de integrar equipos y comunidades de investigación donde existen miembros que podemos denominar genéricos (aunque vayan especializándose en aplicaciones tecnológicas específicas), llámense ingenieros, documentalistas, especialistas de lingüística computacional, informáticos... e investigadores específicos del área que, veremos, son claves para proporcionar el detalle preciso de la investigación tanto a nivel de fines y objetivos, como incluso, en el desarrollo de una ontología y epistemología precisa donde los primeros desarrollen su colaboración. En nuestro caso: el historiador de la filosofía o el equipo básico en investigación en historia de la filosofía (junto a los filósofos, también, filólogos, historiadores..., piénsese en un trabajo sobre fuentes manuscritas y ediciones de la Escuela de Salamanca) que constituye el núcleo duro del proyecto. A su vez, la orientación de base del trabajo del núcleo duro temático ayudará al desarrollo tanto de las herramientas específicas para dicho estudio, como de las propias herramientas en sí, en la medida que tienen que hacer frente a nuevos desafíos. Imaginemos, por ejemplo, en el desarrollo de ontologías específicas para webs semánticas o en la estilometría en el campo de lenguas clásicas como el latín o el griego.

Por su parte, los catálogos de las bibliotecas son ahora, probablemente, la forma mínima en el que un académico puede acceder a libros y artículos de investigación sin el uso de un ordenador, pero, con los índices de tarjetas sufriendo una muerte lenta y segura (Baker, 1996, p. 2001), quedan pocos resultados para que el erudito no digital emprenda la investigación en la universidad moderna. Correo electrónico, búsquedas en Google y bibliografía, las bases de datos..., son cada vez más cruciales. No entendemos una biblioteca actual en la que parte de su contenido no esté escaneado, y su catálogo puesto en línea. Frente a esto, algunos denuncian la pérdida de las habilidades y técnicas de las tradiciones de investigación más antiguas y otros han abrazado calurosamente la accesibilidad de las HD (Schreibman Siemans y Unsworth, 2008; Schnapp y Presner, 2009; Presner, 2010; Hayles, 2011). Como señala Moretti (2007, pp. 3-4) la metodología tradicional de las humanidades se centra en una porción limitada del contenido

global. Nos tomaría una vida leer y evaluar una mínima porción de las fuentes documentales de la historia de la filosofía. Pero no solo ello, sino que el propio análisis se termina individualizando, es difícil realizar una visión estructural que entienda el pensamiento como un todo colectivo. En este sentido, las HD (junto con las HC) pueden ayudar a afrontar retos que afectan no solo al contenido, sino a la metodología de aproximación del corpus filosófico. Las HD nos ayudarán a entenderlo como un sistema colectivo, como un todo.

Lo que señala Moretti no creo que deba ser incompatible con el trabajo minucioso del investigador, ni tampoco creo que debamos tomarlo como la panacea de la investigación. No se debe caer en un reduccionismo de uno u otro lado. No podemos negarnos a la evidencia de la utilidad de las nuevas herramientas, como tampoco podemos “morir de éxito”, es decir, dejarnos llevar por el exceso de datos, por la lluvia del análisis cuantitativo. La labor del historiador de la filosofía es contextual, y el contexto y el mapeo de ideas siempre implica una situación hermenéutica que los *big data* no lo pueden solucionar. Pero, a la vez, estos datos y sus herramientas de análisis, nos proporcionan útiles valiosos para desarrollar teorías hermenéuticas, y, también, para realizar comprobaciones de conjeturas, a nivel discursivo, o de la presencia de conceptos, de semántica y sintácticas presentes en las fuentes que estudiamos y que nosotros no podríamos hacer sin estas herramientas o que nos costaría un esfuerzo enorme. Supongamos el caso que vamos a presentar brevemente del análisis del discurso filosófico desde el Procesamiento de Lenguaje Natural, por ejemplo.

No podemos, pues, negar el uso de las HD. Estas se caracterizan por aprovechar la plasticidad de las formas digitales, lo que viene muy bien para trabajar sobre el patrimonio cultural, o filosófico, desplegando una nueva forma de trabajar con la representación y la mediación. Permite un “plegado” digital de la memoria y de los archivos que implica una nueva forma de estudiar la cultura, en nuestro caso, la filosofía. Pero para poder realizar esta transformación en el estudio y acercamiento a la filosofía (a la cultura), precisamos de objetos con los que trabajar y convertirlos a códigos digitales. Esto ni es bueno ni es malo, es lo que es. Quiero decir, trabajar la historia de la filosofía desde sus objetos, implica, en primer lugar, identificar sus objetos y, después transformarlos. En nuestro caso los objetos más evidentes son las fuentes (derivadamente podríamos operar con las fuentes secundarias). Como he dicho ni es bueno ni es malo, con esto quiero decir que, si reparamos en ello, esta operación es una herramienta que no sustituye, en ningún caso, la capacidad de análisis del historiador de la filosofía, que partirá de unos principios hermenéuticos, una cosmovisión filosófica y un posicionamiento historiográfico determinado, al menos *a priori*.

Pero la intervención del historiador de la filosofía, como hemos señalado antes, no se reduce a un análisis de los datos vertidos por las HD, sino que es

preciso que intervenga, a su vez, en el propio proceso de transformación del objeto cultural (en nuestro caso de la historia de la filosofía, normalmente las fuentes). Efectivamente, en la transformación del objeto en un código digital que pueda ser entendido, se efectúa mediante un mecanismo de entrada de un dispositivo socio-técnico, con el fin de que se establezca un modelo o imagen. Esto supone que internamente será transformado. Dicha transformación precisa de una serie de intervenciones, procesos o filtros, que tendrá como resultado un cálculo final que se presenta, normalmente, en forma visual. El trabajo se realiza mediante la representación numérico-algorítmica (el algoritmo trabaja sobre dicha numeración) con la que trabaja la herramienta digital, normalmente la computadora. Para desarrollar dicha representación computacional y/o digital se precisa tanto un técnico que sepa realizar dicha representación numérica (semántica), como la creación y aplicación de una regla algorítmica (sintáctica). Ello precisa de una ontología lingüística que refleje lo que el objeto cultural precisa y en ese diseño está implicado el investigador en humanidades que conoce el contexto cultural en el que se inserta dicho objeto que debe ser transformado.

El historiador de la filosofía no tiene por qué realizar dicha transformación, pero debe conocer las posibilidades que se le ofrecen para poder orientar dicho trabajo, así como manejar herramientas que se derivan del trabajo ingenieril y computacional o informático, con el fin de sacar provecho de dicho avance. Efectivamente, todos usamos ordenadores. No sabemos como se crea el hardware, no conocemos normalmente el lenguaje de programación que subyace al sistema operativo ni a los diferentes programas. Del mismo modo, si usamos dispositivos móviles, posiblemente, no sepamos desarrollar una aplicación (app), pero para sacar provecho a estas herramientas necesitamos saber que existen y cómo manejarlas. Todos usamos programas de texto, como Word, pero hay muchas funciones que no utilizamos y ello hace que no nos aprovechemos de todas sus funcionalidades, porque las desconocemos. Pero Word no condiciona lo que vamos a escribir, aunque sí que provoca que escribamos diferente a como lo hacíamos antes de su llegada, cuando los más mayores usábamos máquinas de escribir. Con la máquina de escribir, tenías que pensar bien antes de darle a la tecla, puesto que esta te permitía hacer alguna corrección (se toleraba algún tono blanco del Tipp-Ex). Esto suponía que la precipitación no estaba contemplada, bajo la responsabilidad de tener que repetir la hoja y si la corrección era estructural, todo el trabajo. Ese hándicap ha pasado a los anales de la historia, la herramienta informática nos permite revisar más, pero, también, algunas veces precipitarnos, puesto que podemos borrar y rescribir, otras veces. Con las hojas de la máquina de escribir, debíamos tener cuidado con el café, sus manchas podían ser peligrasas, ahora hemos de tener cuidado con guardar bien el archivo, el programa puede desestabilizarse y quedarse “colgado”, se nos puede romper el dispositivo

de memoria... En la época anterior teníamos que poner la taza a distancia prudencial, ahora la prudencia es ir guardando el contenido frecuentemente y en un dispositivo que no sea perecedero. Antes el manuscrito estaba a buen recaudo en el escritorio del despacho, ahora el contenido seguro en la *icloud* (en la nube) tiene más posibilidades de ser disponible para otros...

La llegada de lo digital ha supuesto también un desarrollo temporal, un salto de la Web 1.0 a las sucesivas versiones y en cada ola tecnológica surgen nuevas implementaciones, operatividades, funcionalidades y, en fin, posibilidades y retos.

2.1. HUMANIDADES DIGITALES. HERRAMIENTAS DEL 2.0

En el apartado anterior hicimos referencia a la Web 1.0. Una de las diferencias esenciales entre la Web 1.0 y la Web 2.0 es que los creadores de contenido eran pocos en la Web 1.0 y la gran mayoría de los usuarios actuaban simplemente como consumidores de contenido, mientras que en la Web 2.0 cualquier participante puede ser un creador de contenido, compartiendo numerosas ayudas tecnológicas para maximizar el potencial de creación de contenido. El enfoque de la Web 1.0 era el comercio electrónico, mientras que la arquitectura de la Web 2.0 está orientada a la interconectividad y la interacción. El término Web 2.0 fue acuñado en el año 2004 in el encuentro organizado por O'Reilly Media para describir el desarrollo de aplicaciones online (Tuten, 2010, p. 1),

O'Reilly (2005) señalaba algunos aspectos visibles de este cambio operado, en una especie de lluvia de ideas:

Web 1.0	Web 2.0
DoubleClick	Google AdSense
Ofoto	Flickr
Akamai	BitTorrent
mp3.com	Napster
Britannica Online	Wikipedia
personal websites	blogging
evite	upcoming.org and EVDB
domain name speculation	search engine optimization
page views	cost per click
screen scraping	web services
publishing	participation
content management systems	wikis
directories (taxonomy)	tagging ("folksonomy")
stickiness	syndication

Entender la Web 2.0 es sencillo, al menos en lo que implica para la vida diaria, puesto que, en cierto sentido, la mayoría de nuestras interacciones con la web se encuentran en esta fase de desarrollo. Nuestra generación se siente informatizada, cuando ha llegado a esta fase. Es decir, es la común. No es la actual, no es la que se va haciendo presente en las diversas facetas de la vida, pero es en la que la mayoría de las personas estamos, la que manejamos de forma ya interiorizada.

En este momento de desarrollo de las HD no podemos pensar en una simple bidireccionalidad comunicativa ni transferencial, sino en una comunidad de trabajo donde los actores de la investigación ya son múltiples y multidisciplinares en cada área de conocimiento, lógicamente en el campo de la historia de la filosofía también: investigadores en humanidades, informáticos, lingüistas, ingenieros, docentes, estudiantes, especialistas en información y documentación, como ya señalamos...

La Web 1.0 y 1.5 era la primera etapa de la ejecución de las bibliotecas y repositorios: era el momento de la digitalización. El desarrollo de Google como herramienta, ya en una web con interfaz comunicativa amplió las perspectivas de las HC y de las librerías digitales proyectándolas a la nueva fase de HD (Susan Hockey, 2005). Se trata de la explosión de las bases de datos hacia exhibiciones digitales: nos llegan las ediciones príncipes y fuentes primarias, los manuscritos se ponen a disposición, se multiplica el proceso de OCR de textos, podemos mapear conceptos y términos. La Web 1.0 estaba orientada a reunir, compilar, preservar, presentar y comunicar datos. La Web 2.0 ya puede manipular, organizar, combinar, moderar, revisar y editar los datos. Al incluir la participación, permite al usuario operar en la edición de contenidos en línea (de forma *on-line*), así como trabajar con datos múltiples y mapearlos. Esto supone que no solo somos capaces de generar un *Corpus* en la red y ponerlo al alcance de todos, sino que podemos trabajar con él: podemos empezar a analizar el corpus, en esta fase, al menos, la sintaxis.

Las HD eclosionan a partir de la Web 2.0. La entrada de lo digital supone una propia forma de ser, lo que se llama una propia ontología que se manifiesta en las tecnologías y relaciones humanas que definen y rodean lo digital, y que se define por la abstracción. Los tipos de estructuras sociales agrupadas bajo la etiqueta Web 2.0 ejemplifican la abstracción materializada de lo digital (Evens, 2012).

Las estrategias de investigación en HD desde el entorno de la Web 2.0 son variados. En el sentido en el que venimos diciendo está la actitud de la interconectividad y la participación que ha permitido diseños webs más dinámicos. A modo de ejemplo podemos utilizar, entre los muchos que existen, los *Archives Henri-Poincaré. Philosophie et Recherches sur les Sciences et les Technologies* de la Université de Lorraine en Francia. Podemos ver en esa página web de inicio, una página de acogida en la que se presentan la agenda y notas de anuncios. Por otra parte, tenemos una presentación, una página sobre las actividades, otra sobre

fondos y archivos, otra llamada *philosophia scientiae* –una publicación en formato digital que da acceso a trabajos en línea–, otra llamada gran público, y formación. Esta página es una muestra de los logros y los límites que la presentación HD en versión 2.0 se manifiestan en la academia. Efectivamente, esta página es más dinámica que una página web que pudiera haberse diseñado antes de 2004. Tiene hiperenlaces, por ejemplo, a la *Stanford Encyclopedia* y, en la página “Fondos y archivos” nos lleva a la obra de Henri Poincaré, también mediante hiperenlace al *Henri Poincaré Papers* que, a su vez nos proporciona datos e imágenes jpg de los documentos. Pero carece de la naturaleza colaborativa e interconectada, quizás porque en la academia se mira mucho la calidad en la creación de contenidos. Esta preservación de la calidad, sin embargo, tiene como hándicap el hecho de que el público puede ir al contenido digital sin entrar en la propia digitalización. La única interactividad queda relegada a la página de contactos, donde se añade, por lo menos un *widget de google maps* para la geolocalización de los *Archives*.

Pongamos ahora el ejemplo de *Fragmentarium*. Laboratory for Medieval Manuscript Fragments, cuyo objetivo, tal como señala en la propia página web es el de permitir a las bibliotecas, coleccionistas, investigadores y estudiantes publicar imágenes de fragmentos de manuscritos medievales, permitiéndoles catalogar, describir, transcribir, ensamblar y reutilizarlos. Este concepto mantiene la calidad, pero también fomenta la participación. El concepto es el del *Laboratorio digital*, como señala en la web (traducimos):

“El objetivo principal del *Fragmentarium* es desarrollar un laboratorio digital especializado en la investigación de fragmentos de manuscritos medievales. Aunque se basa en los muchos años de experiencia de e-codices-Biblioteca Virtual de Manuscritos de Suiza, el Laboratorio Digital del *Fragmentarium* tiene una orientación internacional. En primer lugar, se concibe como una plataforma para que las bibliotecas, los estudiosos y los estudiantes realicen trabajos académicos sobre los fragmentos. Se ajusta a las últimas normas establecidas por las bibliotecas digitales y establecerá nuevas normas, especialmente en el ámbito de la interoperabilidad.

La aplicación web contiene una serie de herramientas:

- Una herramienta de catalogación que permite a las bibliotecas, coleccionistas, investigadores y estudiantes reunir y describir fragmentos a través de un CMS.
- Una herramienta para etiquetas, facetas y palabras clave, que permite una investigación eficiente a través de la comparación y la comprobación cruzada.
- Una herramienta para vincular y ensamblar fragmentos, que ofrece la posibilidad de organizar los recortes, los fragmentos de hojas y las hojas individuales en cualquier orden.

En la fase inicial del proyecto (2015-2018), la aplicación web *Fragmentarium* funcionó como un laboratorio “cerrado” con el fin de evaluar el potencial de la fragmentación digital mediante una serie de estudios de casos, que se han desarrollado en cooperación con nuestras 16 instituciones asociadas y con los becarios designados.

En la segunda fase del proyecto (2019-2022), el laboratorio se está abriendo gradualmente y se está haciendo accesible a un público más amplio, lo que permite a las bibliotecas y colecciones, así como a los investigadores y profesores, utilizar esta plataforma para catalogar, transcribir, ensamblar e investigar fragmentos. Se invita a los interesados a que se pongan en contacto directamente con nosotros”.

Esta iniciativa incluye todas las virtualidades de la Web 2.0 y se abre ya en la perspectiva de sus posteriores desarrollos, pero con estos dos ejemplos podemos ir acercándonos a lo que este mundo significa. Muchas más cosas podríamos ver tal como señala, por ejemplo, en el campo de la historia contemporánea Frédéric Clavert al indicar los problemas asociados a la verificación de datos, la cobertura legal especialmente de la propiedad intelectual, la dispersión de los datos que en esa época se vislumbraban... (2013, p. 347-350).

2.2. HERRAMIENTAS DE HD EN UN ENTORNO DE LA WEB 3.0 (LA WEB SEMÁNTICA) EN CAMINO A LA 4.0

La situación actual se define por la Web 3.0 orientada a refundir los fundamentos epistemológicos de la cultura de Internet, colocando la computadora o la red en la posición dejada por el individuo creativo. No se trata de colgar y mostrar contenido (1.0), o de interactuar en dicha presentación y generación, mediante uso de contenidos y comentarios, o procesamiento sintáctico (2.0), se trata de utilizar las máquinas inteligentes para procesar el contenido incluyendo elementos semánticos en su extensión. Nos referimos a la Web Semántica, una Web “más inteligente”, basada en el trabajo con metadatos interpretables provistos a los programas de software mediante máquina de información y datos publicados. Esta Web se basa en buscadores que se encuentran más cerca de entender el lenguaje natural humano (PLN). Esto supone añadir más descriptores de datos a los contenidos y datos existentes en la web. Como resultado, las computadoras son capaces de hacer interpretaciones significativas similares a la forma en que los humanos procesan la información para lograr sus objetivos. Es el resultado de una idea que tuvo su origen en 2006, por parte de Berners-Lee, James Hendler y Ora Lassila, quienes hablaban de la posibilidad “de que los ordenadores no sólo son capaces de presentar la información contenida en las páginas web, como hacen ahora, sino que además pueden ‘entender’ dicha información” (Sánchez, 2008,

p. 66). Otra característica es la movilidad en su uso, es, pues una “web móvil” o “web flexible”, en tanto que se desplaza con nuestros dispositivos móviles almacenando los datos fuera de dispositivos rígidos, alojándose en “la nube”.

Una versión de esta revolución digital lo podemos ver en *Open Journal System*. Creado por *Public Knowledge Project*, liberado bajo licencia *GNU General Public License*. OJS respondía (versión 2.0) a la necesidad de un entorno de conocimiento en sintonía con la Web 2.0, es decir, que respondiera a la apertura y a la participación, a la Web Social que la caracterizaba, por lo que se pudiera realizar tanto el acceso libre de la información académica de publicaciones periódicas abiertas. Los módulos (*plugins*) instalados respondían a ello. Así, el modelo de presentación de los trabajos era el pdf creado a partir de un procesador de textos, normalmente Word. La evolución hacia la necesidad de poder trabajar más y mejor con los metadatos, hizo que la comunidad pkp tuviera que emplearse en ir creando constantes módulos para ir ganando en interoperabilidad y en el aprovechamiento del etiquetado, así como en interfaces más flexibles y manejables. Llega un momento en el que es preciso dar un salto a la versión 3.0 en un entorno de digitalización y de Web 3.0 que garantiza la explotación de las posibilidades de la revista en un etiquetado, proyección de los elementos del contenido, como las referencias bibliográficas o la ejecución de versiones html, así como diversas funcionalidades del XML. Se mejora la operatividad social (discusiones editoriales, flujo de trabajo y roles flexibles), el entorno visual (mejores temas y un diseño web más adaptable), o se actualiza la estructura mediante la implementación de módulos básicos diseñados para la exportación a bases de datos y para el manejo de los propios metadatos y datos, por ejemplo, las estadísticas de artículos.

La Web Semántica se define por usar interfaces móviles, por el uso de la nube (*Cloud Computing*), así como la minería de datos (*Data mining*) del cuerpo lingüístico (*Corpus Linguistics*). Esto implica, para las HD, la introducción de un etiquetado semántico, a partir de una ontología diseñada, con el fin de introducir una compleja regla algorítmica en la que participan máquinas inteligentes (en camino hacia la inteligencia artificial). Esta nueva perspectiva, abre caminos, también, para la historia de la filosofía.

2.2.1. Análisis lingüístico y PNL

Una de las herramientas que pueden servir al análisis de las fuentes desde las HD tienen que ver con herramientas que tienen relación con la Lingüística Computacional, muy presente en los planes de estudios filológicos y, especialmente, de lingüística, como es lógico. Esta perspectiva hace que comprendamos una realidad de las HD (y antes de las HC): el hecho de que estas no son una nueva disciplina

de las humanidades, sino que proporcionan un nuevo enfoque en el estudio de las humanidades en las que se plantea una aproximación multidisciplinar, como ya hemos señalado en repetidas ocasiones. Es lógico que, en un entorno de Web Semántica basada en una inteligencia más cercana a la comprensión del lenguaje natural humano, este acercamiento sea el que tenga una importancia esencial.

Ahora bien, la versatilidad de las HD actuales es otra seña de identidad. Un ejemplo es el desarrollo de modelos que trascienden el propio estudio de las fuentes y los textos (que es a lo que nos vamos a referir a continuación) para, por ejemplo, realizar la interacción de modelos de reconstrucción en 3D para recreación de espacios históricos y culturales. Una muestra de ello, y que aquí no podemos desarrollar por el límite de espacio, es el caso –que ha de citarse– de *Time Machine Europe*. Time Machine es un proyecto, que recuerda no solo a la “máquina del tiempo” de la ciencia ficción sino, también, a la función de copia de seguridad integrada en el Mac que sirve para realizar automáticamente una copia de seguridad de los archivos y que permite restaurar archivos desde la copia de seguridad si los originales se borran en el Mac, o si el disco duro (o SSD) del Mac se borra o se reemplaza. En su presentación señalan que: “La red *Time Machine*, está integrada por más de 14.000 instituciones que representan a más de 100.000 profesionales que van desde expertos en ciencia y tecnología, investigadores académicos, historiadores, científicos culturales y sociales, humanistas, expertos en museos y bibliotecas, archiveros, científicos ciudadanos, genealogistas, hasta investigadores aficionados y entusiastas de *Time Machine*. Con sede en Viena (AT), *Time Machine* tiene oficinas adicionales en Lausana, Ámsterdam y Budapest. Con su red de más de 50 embajadores de la máquina del tiempo en toda Europa, la iniciativa de la máquina del tiempo está firmemente asentada en el continente europeo... permite viajar no sólo a través del espacio, sino también a través del tiempo” (Time Machine, 2020).

Se trata, pues, de la recreación de una máquina del tiempo que sea capaz de restaurar nuestro pasado, mediante la construcción de un “simulador a gran escala que traza un mapa de 2000 años de historia europea, transformando kilómetros de archivos y grandes colecciones de museos en un sistema de información digital”, usando una compilación de “Grandes Datos del pasado” (*Big Data of the Past*) con el fin de cartografiar la evolución social, cultural y geográfica de Europa a través de los tiempos. “Esta infraestructura de digitalización e informática a gran escala permitirá a Europa convertir su larga historia, así como su multilingüismo e interculturalidad, en un recurso social y económico vivo”. Este es un ejemplo muy acabado de lo que supone la potencialización de las HD en el entorno de la Web 3.0.

Volviendo al entorno de la Web Semántica y el lenguaje computacional, destaca el uso en el estudio de las fuentes en historia de la filosofía a partir del ámbito

del procesamiento computacional del lenguaje natural. Se trata del uso de las herramientas y recursos lingüísticos en la investigación de HD en los que se ponen en juego las herramientas de procesamiento de idiomas existentes para poder abordar las cuestiones de investigación en las HD y que lleva implícito en sí, tanto la implementación de herramientas existentes, como el desarrollo de nuevas herramientas de procesamiento del lenguaje.

Una visualización común de estos procesos de Web Semántica y uso de PLN en el desarrollo del lenguaje aplicado a generación de idiomas lo tenemos, por ejemplo, en el traductor de google (*Google Translate*) que funciona a partir del sistema de traducción automática llamado *Google Neural Machine Translation* (GNMTS) que está basado, a su vez, en las llamadas Redes Neuronales Recurrentes (RNN), “en el que no solo se toman en cuenta las palabras o las frases, sino también el contexto en el cual aparecen esas palabras y frases y qué otras aparecen cercanas a ellas” (Pastor, 2019); y en el traductor *DeepL* que “hace uso también de sistemas de inteligencia artificial, y en concreto redes neuronales de convolución, algo distintas a las recurrentes”. Basado en la base de datos *Linguee*, “*DeepL* ofrece una traducción ‘más natural’” (Pastor, 2019).

Empezando por lo complicado, y para ver casos que se están desarrollando, podemos fijarnos, en el proyecto de desarrollo de métodos de lenguaje computacional en latín. Algo esencial para estudios de las fuentes de filosofía. Nos fijamos en el trabajo de Marco C Pasarotti y el proyecto europeo *LiLa: Linking Latin. Building a Knowledge Base of Linguistic Resources for Latin*, desarrollado dentro del *Centro Interdisciplinare di Ricerche per la Computerizzazione dei Segni dell'Espressione* (CIRCSE). Como se expone en su página web y aquí resumimos “El proyecto LiLa construye una base de conocimientos de recursos lingüísticos y herramientas de procesamiento del lenguaje natural (PLN) para el latín. La Base de Conocimientos consiste en diferentes tipos de objetos conectados a través de un vocabulario explícitamente declarado para la descripción del conocimiento”. Se trata de la construcción de una ontología a partir de lemas. El modelo se basa en el léxico, a partir de los lemas, que son “son el tipo de nodo clave en la Base de Conocimientos” (Lila), de modo que recoge y conecta metadatos existentes (corpus, léxico, ontologías, diccionarios, tesauros) y herramientas de PNL (fichas, lematizadores, etiquetadores PoS, analizadores morfológicos y analizadores de dependencia). Así, en relación con el trabajo iniciado con Busa, amplía la cobertura textual del banco de árboles del *Index Thomisticus*.

Este proyecto europeo es un producto acabado de lo que se puede conseguir con las HD aplicadas a las fuentes, no solo como un producto final, sino como una herramienta a utilizar en tanto que proporcione metadatos lingüísticos (lemas, etiquetas, bancos...) y metadatos relacionales que ayuden a los investigadores. El tema es amplio, vamos a exponer ahora, brevemente, dos herramientas para el

trabajo de PLN (*Natural Language Toolkit*) para los investigadores de humanidades con el fin de realizar investigaciones procesamiento de lenguaje natural (PLN), de minería de textos (*Text Mining*), problemas de atribución de autoría, lingüística forense⁴, modelado de tópicos (*Topic Modelling*) o análisis de sentimientos, como son el entorno *Python* y lenguaje de programación *R*. Antes, no obstante, precisamos definir, al menos, estos procesos a los que aplicamos *Python* y *R*.

La minería de textos (*Text Mining*) –también denominada análisis de textos (*Text Analytics*)– es una tecnología de inteligencia artificial (IA) cuyo objetivo es la búsqueda de conocimiento en grandes colecciones de documentos. Utiliza el procesamiento del lenguaje natural (PNL) para transformar el texto libre (no estructurado) de los documentos y bases de datos en datos normalizados y estructurados adecuados para el análisis o para impulsar algoritmos de aprendizaje automático (ML). Para ello se apoya en la minería de datos (*Data Mining*), que se “refiere al conjunto de métodos estadísticos que proporcionan información (correlaciones o patrones) cuando se dispone de muchos datos”. Mientras en la primera “obtenemos información nueva a partir de grandes cantidades de texto, en la que la información suele estar no estructurada”, en la segunda “el conocimiento se obtiene de bases de datos, en la que la información está estructurada” (Rochina, 2017). Para poder realizar la minería de textos, proporcionaremos a la herramienta de minería de datos, tanto una información recuperada, es decir, una selección de textos, en nuestro caso las fuentes que queremos analizar, como extraer la información de dichos textos mediante PLN.

El Procesamiento del Lenguaje Natural (PLN) simula la capacidad humana de entender un lenguaje natural como el inglés, el español o el latín, como hemos visto. Comprende tanto la comprensión como la generación del Lenguaje Natural, ello no sirve, también para extraer y resumir información, utilizando herramientas analíticas que permiten extraer datos precisos. Para ello, y aquí se reúnen las diferentes habilidades (las del filósofo, las del lingüista, las del informático, las del ingeniero...), se han de crear ontologías, vocabularios y diccionarios personalizados que permitan realizar búsquedas, extracción e integración de datos eficaces. Entre los desarrolladores de PNL tenemos OpenNLP (Morton et al., 2005), CoreNLP (Manning et al., 2014), spaCy (Honnibal y Johnson, 2015), y NLTK (Bird y Loper, 2004; Bird, 2006).

Como hemos señalado, *Python* y *R* son dos lenguajes de programación accesibles y eficaces. *Python* fue desarrollado por Guido van Rossum al final de la década de los 80 del siglo pasado y se ha ido enriqueciendo mediante una

4 La lingüística forense es una rama de la lingüística aplicada que consiste en el examen de las pruebas lingüísticas en un asunto penal o civil (Fadden y Disner, 2014).

programación de fuente abierta (*open source programming language*) a través de discusión pública, lo que le ha hecho muy popular. Se puede ejecutar en casi todas las arquitecturas de sistemas y está detrás de muchas de las plataformas que usamos desde Mozilla a Google o empresas como *CapitalOne* o *Blommborg*..., y, lógicamente, es útil para las HD. De entre las muchas aplicaciones destaca para nuestro interés el análisis textual (encontrar frecuencias de palabras y “palabras de parada” en textos de dominio público), como la estilometría (comparar las características medidas de las imágenes gráficas). El análisis estilométrico o estilometría nos permite determinar, por ejemplo, la atribución de un autor en un texto dudoso, o la utilización de elementos en un autor para saber cómo estructura su discurso (Kokensparger, 2018, p. 70). Nos permiten gráficas de análisis de textos y discursos.

R, por su parte, es una potente herramienta en la medida en que se pueden descargar paquetes gratuitos para tantas herramientas y tipos de análisis como se vayan desarrollando, permitiendo ir estructurando los datos, especialmente los del texto, desde uso semántico y sintáctico hasta el carácter emocional para ello usamos paquetes (*packages*) y se va generando un *corpus*, que es un formato para almacenar datos textuales que se utiliza en toda la lingüística y el análisis de textos, que normalmente “contiene cada documento o conjunto de textos, junto con algunos meta-atributos que ayudan a describir ese documento” (Ward, 2019). A partir de ahí podemos ir operando, transformando y explorando el texto, por ejemplo, comparando y analizando el vocabulario mediante el proceso de colapsar las palabras (*Stemming*), usando una matriz de términos de documentos (DTM) lo que nos permite comparar todos los términos o palabras de cada documento, o creando una nube de palabras de los términos más frecuentes, lo que permite elegir el número de palabras más frecuentes de un texto. E incluso, podemos, realizar, minería de sentimientos (*Opinion Mining*), un análisis de polaridad sentimental, usando el paquete *Sentiment Analysis* mediante el diccionario Harvard-IV (*General Inquirer*) que es un diccionario de palabras asociadas con el sentimiento positivo (1.915 palabras) o negativo (2.291 palabras).

Ante esta avalancha de potencialidad de análisis, el historiador de la filosofía no puede no sensibilizarse. Este análisis no debe chocar con el acercamiento hermenéutico. El uso de estas herramientas, como otras similares, imaginemos la prospectiva bayesiana, no deben leerse desde el desencuentro. Quizás podemos realizar estas operaciones con las gafas de la hermenéutica. “Y no podemos obviar que la hermenéutica filosófica contemporánea se ha esforzado por hacerlo. Más que ninguna otra corriente de pensamiento, ha intentado, sin negar nada de sus propios principios, entablar un diálogo con otras tradiciones, pues precisamente lo ajeno, según Hölderlin, es indispensable para toda apropiación de lo propio” (Lázaro, 2019, pp. 523-524). Así pues, la contrario de lo que pueda

pensarse, y evitando reduccionismos, es positivo acercarse a las HD con la seguridad de que este enfoque complementa y contextualiza los métodos con los que están familiarizados los filósofos.

Un ejemplo reciente de este trabajo lo tenemos en el trabajo de Mark Alfano (2019). Su trabajo arroja resultados a partir de la utilización de la fuente, en este caso la obra de Friedrich Nietzsche. A partir de aquí, el autor ha seleccionado los conceptos básicos. Para ello ha etiquetado cada pasaje de los manuscritos publicados y autorizados de Nietzsche por la presencia o ausencia de varias docenas de conceptos psicológicos morales, como impulso, instinto, virtud, valor, vida, salud, integridad y una gama de virtudes y emociones discretas. Ha partir de la creación del etiquetado y el diccionario ha operado los conceptos para buscar su fuente. Esto le permite tanto realizar búsquedas, como limpiar los datos, analizarlos y visualizarlos. La utilización de este conjunto de datos le permite realizar análisis estadísticos de la prevalencia y co-ocurrencia de los conceptos, así como la visualización de la red semántica, de modo que pueda cerrar la lectura de los pasajes relevantes. Es un ejemplo, actual de las potencialidades de las HD para la historia de la filosofía.

3. REFLEXIÓN CONCLUSIVA

La historia de la filosofía siempre ha utilizado aquellas herramientas conceptuales y metodológicas de su época para acometer su tarea que es a la vez filosófica como histórica o, incluso, filológica, toda vez que los registros que tenemos de las doctrinas y pensamiento filosóficos anteriores tienen como lugar privilegiado de expresión el texto escrito. La elaboración de una investigación en historia de la filosofía actual no puede ni debe prescindir de las herramientas propias del siglo XXI, ni de las nuevas implicaciones de un trabajo en red y en comunidad, como no se prescindió de la oportunidad que supuso la imprenta en un primer momento, o la industria de la edición en su día.

Las nuevas posibilidades de investigación que suponen las HD no deben sustituir a las formas estandarizadas y clásicas tales como la recopilación de datos, la investigación de los textos, la elaboración de equipos de investigación, el análisis de datos, la compilación bibliográfica, la transferencia de conocimiento mediante publicaciones científicas, la comunicación en encuentros científicos, la docencia... Pero las posibilidades que nos ofrecen las HD pueden ayudarnos, por una parte, a multiplicar las posibilidades y los resultados que estas actividades clásicas nos ofrecen, al tiempo que, en otras ocasiones las pueden posibilitar cuando por circunstancias excepcionales no podrían hacerse sin la mediación digital. Ya en el 2010, Teresa Numerico, Domenico Fiormonte y Francesca Tomasi en su clásica

introducción compilatoria *L'humanista digitale*, señalaban –desde el contexto de la web 2.0– en el intento de aplicar una cultura interdisciplinar, encontrar sinergias y subrayar la entonces dimensión emergente de la cultura Digital. Así afirmaban que “Desde los años ochenta hasta hoy, el trabajo de HD ha sido mostrar la naturaleza epistemológica de los cambios en la metodología. No se trata realmente de nuevos instrumentos, sino de una forma diferente de representar (y acceder) al conocimiento y a la cultura a través de los instrumentos digitales” (2015, p. 208)⁵.

Precisamente el confinamiento provocado por la crisis de la Pandemia de Covid-19 puso en evidencia este segundo aspecto, si bien ya nos habíamos beneficiado de ello con más regularidad de lo que a simple vista parece. ¿Quién no ha acudido a las bases de datos a disfrutar de manuscritos y ediciones antiguas que antes eran casi inaccesibles? ¿Quién no ha puesto a su disposición bases de datos para descubrir fuentes y bibliografía secundaria que antes hubiera sido imposible disponer? ¿Quién no ha leído en acceso abierto o disfrutado de conferencias y clases grabadas en plataformas como *youtube* o *vimeo*, o en multi plataformas como *zoom* o *teams*? ¿Quién no ha disfrutado de su participación en Encuentros organizados a distancia bien en *streaming* bien mediante grabaciones visualizadas? Si uno ha estudiado a Tomás de Aquino ¿es que no ha utilizado el motor de búsqueda de términos del *Corpus Thomisticum*?

Esta misma Pandemia ha revelado también problemas que no podemos desarrollar sino enumerar, como el de la realidad de la brecha digital existente en los distintos espacios geográficos y sociales del planeta o el de la apropiación de lo digital lo que lleva a realizar una lectura crítica capaz de transformar los modelos.

Sea como fuere, el tiempo transcurre y afecta a la investigación. Este recorrido no se trata tan solo de un itinerario sobre el papel. Muchos de los investigadores en historia de la filosofía hemos ido acompañando este itinerario. No solo lo hemos visto en la forma de escribir, pues los que pintamos canas venimos de experimentar del paso de la máquina de escribir, a la máquina de escribir electrónica, a los primeros Amstrad y Spectrum, basados en MS-DOS, hasta llegar a los PC (con su sistema Word 5.0) y Apple..., sino en el propio acompañamiento computarizado y digital de nuestro estudio de las humanidades. La juventud (1994-1995) la recuerdo en las bibliotecas y en la Colaboración en el *Centre de Recherches sur la Pensée Médiévale-Bureau de recherches des études médiévales et Glossaire de lexicographie latine et médiévale del C.N.R.S. (Centre Nationale de Recherches Scientifiques)*, un trabajo pionero lexicográfico en la década de los noventa del siglo XX. El trabajo de investigación lo recuerdo asociado a la búsqueda en las bases de datos y la participación como colaborador en MEL (Medioevo

5 Usamos la traducción inglesa del libro 2010, publicada en 2015.

latino) *Bolettino bibliografico della cultura europea da Boezio a Erasmo (secoli VI-XV)* della “Società Internazionale per lo Studio del Medioevo Latino” della “Fondazione Ezio Franceschini” del “Dipartimento di studi sul Medioevo e Rinascimento” dell’Università di Firenze, que publica SISMEL – Edizioni del Galluzzo. Firenze (Italia), en la primera década del siglo XXI. Sin duda la digitalización masiva de las fuentes y el acceso a los archivos ha facilitado el trabajo empírico del historiador en general (Dufournaud, p. 405) y en él, el de la filosofía. Pero no solo ha facilitado el acceso, sino el trabajo en y con los mismos datos.

La actualidad afecta también a la formación en la segunda década del siglo XXI. Este es un recorrido que se va insertando, y que seguirá con nosotros. Ya no vale con saber lenguas muertas (latín, griego, hebreo...), con los que acercarnos a las fuentes, y los idiomas modernos (inglés, francés...) que acompañan las lecturas de las fuentes secundarias, es preciso aprender otras lenguas de las que hemos hablado (*R, Python...*) para desarrollar la interacción que suponen las HD.

Esta realidad implica un desarrollo en los planes de estudio. Las distintas instituciones abordan esta cuestión de maneras muy diferentes. No resulta fácil insertar de forma ordenada los campos del mundo informático, digital y de las humanidades. Su inclusión está destinada a cambiar tanto lo que aprenden los estudiantes de humanidades como la forma en que lo hacen, su integración en las humanidades es un proceso que exige reflexión, formación de profesores y diferentes acciones colaterales. Afecta también a los equipos de investigación que han de ser, por definición, interdisciplinares. Ya en 2005 Jennifer Edmons (2005) sugería la necesidad de un “Intermediario de Humanidades Digitales (DHI) dedicado” capaz de proporcionar las habilidades y el apoyo muy necesarios de naturaleza no específicamente técnica en los momentos críticos del desarrollo de los proyectos. Según la autora, este tipo de intermediario podría ser decisivo para aumentar el uso de la tecnología entre el personal de investigación en humanidades, proporcionando una amplia gama de conocimientos técnicos y otras funciones de apoyo. Para ello se precisa de una estructura sólida de investigación interdisciplinar universitarias. Se trata de contar en los equipos de humanidades, en nuestro caso de historia de la filosofía, con filósofos, historiadores, filólogos, técnicos en informática, ingenieros de I.A., especialistas en documentación, y humanistas que se manejen en HD como intermediarios y orientadores en las reuniones de estrategia de investigación. Quizás las HD no impliquen una revolución epistemológica, pero sí que han supuesto una actualización de las técnicas

Hemos visto que las HC y HD han estado presentes en y para la historia de la filosofía, tan presentes que fueron básicas, especialmente desde la figura de Tomás de Aquino, en la aventura de las HD, y siguen siendo la inspiración de los programas más avanzados como los de PLN en el texto latino.

No hemos visto problemas epistemológicos y filosóficos relacionados con las HD en historia de la filosofía, algo que gusta mucho al filósofo. Nuestra pretensión ha sido presentar de forma introductoria esta realidad para animarnos a conocerla de forma amable como historiadores de la filosofía que las vamos a tener que usar. Esto supone huir del *a priori* crítico (que será importante en un segundo momento y en otra esfera) y del complejo informático, pues no tenemos porqué ser expertos informáticos, sino historiadores de la filosofía.

REFERENCIAS BIBLIOGRÁFICAS

- ALAFANO, M., "Methodology: Introducing Digital Humanities to the History of Philosophy". *Nietzsche's Moral Psychology*. Cambridge: Cambridge University Press, pp. 12-46, 2019.
- ANANDAJARAN, M, et al., "Emerging Web Techonologie", en Narayanan, V. K. y O'Connor, G.C., *Encyclopedia of Technology and Innovation Management*, Chichester: John Wiley & Sons, pp. 505-512, 2010.
- ANHEIM, É., LILTI, A. y VAN DAMME, St., "Quelle histoire de la philosophie ?", *Annales. Histoire, Sciences Sociales*, 64 (1), pp. 5-11, 2009.
- BERRA, A., "Faire des humanités numériques", en Mounier, P. (dir.), *Read/Write Book 2. Une introduction aux humanités numériques*. Marseille: OpenEdition Press. , 2012. doi: <https://doi.org/10.4000/books.oep.238> (Consulta: 3 de abril de 2020).
- BERRY, D.M. y FAGERJORD, A., *Digital Humanities: Knowledge and Critique in a Digital Age*. Cambridge: Polity Press, 2017.
- BIRD, S., "NLTK: The natural language toolkit". *Proceedings of the COLING/ACL on interactive presentation sessions*. Sydney, Association for Computational Linguistics, pp. 69-72, 2006. Disponible en <https://www.aclweb.org/anthology/P06-4018> (Consulta: 16 de junio de 2020).
- BIRD, S. y LOPER, E., "NLTK: The Natural Language Toolkit". *Proceedings of the ACL Interactive Poster and Demonstration Sessions*. Barcelona, Association for Computational Linguistics, pp. 214-217, 2004. Disponible en <https://www.aclweb.org/anthology/P04-3031> (Consulta: 16 de junio de 2020).
- BORDOY, A., "Metafísica e historia de la metafísica: consecuencias de un olvido". *Cauriensia*15, pp. 303-329, 2020.
- BOUVERESSE, J., "L'histoire de la philosophie, l'histoire des sciences et la philosophie de l'histoire de la philosophie", en *L'histoire de la philosophie, l'histoire des sciences et la philosophie de l'histoire de la philosophie*. Paris: Collège de France, 2017. doi: <https://doi.org/10.4000/books.cdf.4936> (Consulta: 1 de abril de 2020).
- BUSA, R.A., "Foreword: Perspectives on the Digital Humanities", en Schreibman, S., Siemens, R. y Unsworth, J. (eds.), *A Companion to Digital Humanities*. Oxford: Balckwell Publishing, pp. xvi-xxi, 2004.
- CALLAWAY, E., Turner, J., Stone, H., Halstrom, A., "The Push and Pull of Digital Humanities: Topic Modeling the "What is digital humanities?" Genre", *Digital*

- Humanities Quarterly*. 14 (1), 2020. Disponible en <http://www.digitalhumanities.org/dhq/vol/14/1/000450/000450.html> (Consulta: 1 de abril de 2020).
- CLAVERT, F., NOIRET, S. (eds.), *L'histoire contemporaine à l'ère numérique*. Bruxelles: Peter Lan, 2013. Disponible en <http://hdl.handle.net/10993/34981> (Consulta el 20 de junio de 2021).
- DAVIDSON, C.N., "Humanities 2.0: Promises, Perils, Predictions". *Publications of the Modern Language Association of America*, 123 (3), pp. 707-717, 2008.
- DERRIDA, J., *De la Gramatología*. Trad. Oscar del Barco y Conrado Ceretti, Rev. Ricardo Potschart. 4 ed. México: Siglo XXI, 1987.
- DOMINIQUE, V., "Las culturas y humanidades digitales como nuevo desafío para el desarrollo de la ciencia y la tecnología en América Latina", *Revista Universitas Humanistica* 76, pp. 51-72, 2013.
- DUFOURNAUD, N., "La recherche empirique en histoire à l'ère numérique". *La Gazette des archives*, 240, 397-407, 2015. Doi: <https://doi.org/10.3406/gazar.2015.5321>
- EDMONS, J., "The Role of the Professional Intermediary in Expanding the Humanities Computing Base", *Literary and Linguistic Computing*, 20 (3), pp. 367-380, 2005. doi: <https://doi.org/10.1093/lc/fqi036>
- EVENS, A., "Web 2.0 and the Ontology of the Digital", *Digital Humanities Quarterly*, 6 (2), 2012. Disponible en <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html> (Consultado: 11 de junio de 2020).
- FADDEN L. y DISNER S. F., "Forensic Linguistics". Bruinsma G.y Weisburd D. (eds) *Encyclopedia of Criminology and Criminal Justice*. New York, NY: Springer, 2014. doi: <https://doi.org/10.1007/978-1-4614-5690-2> (Consultado: 16 de junio de 2020).
- GUEROULT, M., *Philosophie de l'histoire de la philosophie*. Paris: Aubier Montaigne, 1979.
- GOOFY, J., *La raison graphique*. Paris: Minuit, 1979.
- HOCKEY, S., "Living with Google: Perspectives on Humanities Computing and Digital Libraries: Busa Award Lecture, June 2004", *Literary and Linguistic Computing*, 20 (1), pp. 7-24, 2005. doi: <https://doi.org/10.1093/lc/fqh040>
- HONNIBAL, M., y Johnson, M., "An improved non-monotonic transition system for dependency parsing". *Proceedings of the 2015 conference on empirical methods in natural language processing*. Lisboa, Association for Computational Linguistics, pp. 1373-1378, 2015. Disponible en <https://www.aclweb.org/anthology/D15-1162> (Consulta: 16 de junio de 2020).
- JOHANSSON, K. G., "Computing Medieval Primary Sources from the Vadstena Monastery: Arguments for the Primary Source Text", *Literary and Linguistic Computing*, 19 (1), pp. 93-105, 2004. doi: <https://doi.org/10.1093/lc/19.1.93>
- KLOCK-FONTANILLE, I., "Penser l'écriture : corps, supports et pratiques", *Communication & langages*, 182 (4), pp. 29-43, 2014.
- KOKENSPARGER, B., *Guide to Programming for the Digital Humanities. Lessons for Introductory Python*. Berlin, Springer, 2018.
- KURBANOGLU et al. (eds.), *Technological Convergence and Social Networks in Information Management. Second International Symposium on Information*

Management in a Changing World, IMCW 2010, Ankara, Turkey. Proceedings. Berlin, Heidelberg: Springer, 2010.

- LÁZARO, M., “Filosofía Antigua: La filosofía como Decir”, en Natário, M.C., Epifânio, R. (coord.), *Entre filosofía e literatura. Ciclo de Conferências*. Sintra: Zêfiro, pp. 25-36, 2011.
- LÁZARO, M., “La hermenéutica jurídica desde la perspectiva filosófica”. *Anuario de la Facultad de Derecho. Universidad de Extremadura*, 25, pp. 481-532, 2019.
- MANNING, C.D., SURDEANU, M., BAUER, J., FINKEL, J.R., BETHARD, S., y McCLOSKEY, D., “The Stanford CoreNLP Natural Language Processing Toolkit”. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Baltimore, Ma.: Association for Computational Linguistics, pp. 55-60, 2014. Disponible en <https://www.aclweb.org/anthology/P14-5010> (Consulta: 16 de junio de 2020).
- MCCARTY, W., “Humanities Computing: Essential Problems, Experimental Practice”, *Literary and Linguistic Computing*, 17 (1), pp. 103-125, 2002. doi: <https://doi.org/10.1093/lc/17.1.103>
- , (2005) *Humanities Computing*. New York/ Basingstoke: Palgrave Macmillan.
- MCPHERSON, T., “Introduction: Media Studies and the Digital Humanities”. *Cinema Journal* 48 (2), pp. 119-123, 2009.
- MORETTI, F., *Graphs, Maps, Trees: Abstract Models for a Literary History*. London: Verso, 2007.
- MORTON, T., KOTTMANN, J., BALDRIDGE, J., y BIERNER, G. (2005) “Opennlp: A java-based NLP toolkit”. EACL. Disponible en <https://opennlp.apache.org>.
- MOUNIER, P., “IBM ou International Busa Machines ? De l’informatique aux humanités”, en *Les humanités numériques: Une histoire critique*. Paris: Éditions de la Maison des sciences de l’homme, 2018. doi: <https://doi.org/10.4000/books.edition-smsh.12033> (Consulta: 3 de abril de 2020).
- NUMERICO, T., FIORMONTE, D., TOMASI, F., *The Digital Humanist: A Critical Inquiry*. New York, NY: Punctum books (Edición original: (2010). *L’umanista digitale*. Milano: Le edizioni del Mulino), 2015.
- O’REILLY, T., “What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software”, en *O’Reilly*, 2005. Disponible en <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> (Consulta: 5 de junio de 2020).
- ONG, W. J., *Orality and Literacy: The Technologizing of the Word*. 2 ed. New York: Routledge, 2012.
- ORLANDI, T., “Is Humanities computing a discipline?”, *Jahrbuch für Computerphilologie*, 4 pp. 51-58, 2002. Disponible en <http://computerphilologie.uni-muenchen.de/jg02/orlandi.html> (Consulta: 1 de abril de 2020).
- PASTOR, J., “DeepL vs Google Translate: ¿quién gana la batalla de los traductores online?”, 2019. Disponible en <https://www.xataka.com/servicios/deep-l-vs-google-translate-quien-gana-batalla-traductores-online> (Consulta: 15 de junio de 2020).
- PRESNER, T., “Digital Humanities 2.0: A Report on Knowledge”, 2010. Disponible en <http://cnx.org/content/m34246/1.6/?format=pdf> (Consulta: 11 de junio de 2020).

- REVUELTA, F.I. y PÉREZ, L., *Interactividad de los entornos en la formación on-line*. Barcelona: Editorial UOC, 2009.
- ROCHINA, P., “¿Qué es y cuáles son las aplicaciones del Text Mining?”. *Revista digital INESEM* 24/04/2017, 2017. Disponible en <https://revistadigital.inesem.es/informatica-y-tics/text-mining/> (Consulta: 15 de junio de 2020).
- SÁNCHEZ, D., *Teoría y metodología de la Historia de la Filosofía*. Madrid: Universidad Nacional de Educación a Distancia, 2003.
- SÁNCHEZ, J. R., “Perspectivas de la información en Internet: ciberdemocracia, redes sociales y web semántica”. *Zer* 13 (25), pp. 61-81, 2008.
- SCHNAPP, J. y PRESNER, P., “Digital Humanities Manifesto 2.0”, 2009. Disponible en http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf (Consulta: 11 de junio de 2020).
- SCHREIBMAN, S., SIEMENS, R. y UNSWORTH, J. (eds.), *A Companion to Digital Humanities*. Oxford: Balckwell Publishing, 2004.
- , *A Companion to Digital Humanities*. London: Wiley-Blackwell, 2008..
- SOLÈRE, J.-L., Medieval Philosophy Digital Resources. Disponible en <http://capricorn.bc.edu/siepm/index.html> (Consulta: 7 de junio de 2020).
- SVENSSON, P., “Humanities Computing as Digital Humanities”, *Digital Humanities Quarterly* 3 (3), 2009. Disponible en <http://www.digitalhumanities.org/dhq/vol/3/3/000065/000065.html> (Consulta: 1 de abril de 2020).
- TATARINOV, J., FICKERS, A., *Digital History and Hermeneutics. Between Theory and Practice*. Berlin: DeGruyter, 2021.
- TERRAS, M., NYHAN, J., y VANHOUTTE, E. (“Defining Digital Humanities: A Reader. Terras, Melissa, Julia Nyhan, and Edward Vanhoutte. n.d. ‘Further Reading’”, en *UCL Defining Digital Humanities (blog)*. Disponible en <https://blogs.ucl.ac.uk/definingdh/further-reading/> (Consulta: 5 de abril de 2020).
- TIME MACHINE (2020). Disponible <https://www.timemachine.eu>. (Consulta: 15 de junio de 2020).
- TOCK, B.-M., “Le latin médiéval et l’ordinateur”, en: *Les historiens et le latin médiéval*. Paris: Éditions de la Sorbonne, 2001. doi: <https://doi.org/10.4000/books.pSORBONNE.21061> (Consulta: 4 de abril de 2020).
- TUTEN, T. L., “The strategir Enterprise”, en Tuten, T.L. (ed.), *Enterprise 2.0: How Technology, Ecommerce, and Web 2.0 are Transforming Business Virtually. Volume 1: The Strategic entrprise*, Santa Bárbara, Ca.: Praeger, ABC-CLIO, 2010, 1-14, 2010.
- WARD, B., “A Light Introduction to Text Analysis in R. Working with Corpora, Document-Term Matrices, Sentiment Analysis, etc...”. *Towards Data science*, 2019. Disponible en <https://towardsdatascience.com/a-light-introduction-to-text-analysis-in-r-ea291a9865a8> (Consulta: 16 de junio de 2020).

ALAN TURING: BASES, FORMA Y CRÍTICAS A LA INTELIGENCIA ARTIFICIAL

ALAN TURING: BASE, FORM AND CRITICISM
TO ARTIFICIAL INTELLIGENCE

GUSTAVO ESPARZA

Doctorado en Ciencias Sociales y Humanidades
Profesor Investigador
Universidad Panamericana
Aguascalientes/México
gaesparza@up.edu.mx
ORCID: 0000-0002-9470-6519

Recibido: 9/04/2021
Revisado: 6/08/2021
Aceptado: 6/09/2021

Resumen: El objetivo del presente artículo es definir los elementos que aportó Alan Turing al desarrollo del concepto de *Inteligencia Artificial (AI)*. Como problema se propone reflexionar las aportaciones del matemático a la génesis del término como lo conocemos hoy en día. Metodológicamente se estudian las obras centrales del autor para delimitar los fundamentos lógicos y epistemológicos a través de los cuales opera una *Computadora Inteligente*, para con ello valorar las aportaciones a la disciplina computacional y la eventual formulación del término. La presente investigación ofrece tres resultados: (1) Turing concibe a las computadoras como auténticos objetos materiales pensantes; (2) el modelo de diseño de una computadora con estas características, de acuerdo al autor, debe considerar el patrón de desarrollo cerebral (y aprendizaje) infantil; (3) el debate póstumo sobre el pensamiento de Turing fue decisivo para la construcción de la agenda de problemas que dio origen al término *AI*.

Palabras Clave: Computación, Matemática, Filosofía, Inteligencia Artificial, Turing.

Abstract: The objective of this article is to define the elements that Alan Turing contributed to the development of the concept of Artificial Intelligence (AI). As a problem, it is proposed to study the contributions of the mathematician to the genesis of the term as we know it today. Methodologically, the central works of the author are studied to define the logical and epistemological foundations through which an Intelligent Computer operates, in order to assess the contributions to the computational discipline and the eventual formulation of the term. The present investigation offers three results: (1) Turing conceives computers

as authentic thinking material objects; (2) the design model of a computer with these characteristics, according to the author, must consider the pattern of brain development (and learning) in children; (3) the posthumous debate on Turing's thought was decisive for the construction of the agenda of problems that gave rise to the term *AI*.

Keywords: Computation, Mathematics, Philosophy, Artificial Intelligence, Turing.

1. INTRODUCCIÓN. INTELIGENCIA NATURAL Y ARTIFICIAL¹

El objetivo del siguiente artículo es estudiar los fundamentos y concepción de una Computadora Inteligente propuesta por Alan Turing y las aportaciones que se derivaron de esta propuesta al desarrollo de una *Artificial Intelligence (AI)*; el interés es mostrar tres fases cruciales en la configuración de una máquina pensante: primera, la creación de un mecanismo que fuese capaz de resolver si una fórmula de primer orden es lógicamente válida; segunda, la apología que realizó el matemático inglés sobre la posibilidad de fundamentar que las computadoras fuesen objetos inteligentes; tercera, el debate que se generó en torno estas formulaciones previas y que contribuyeron a la construcción de una agenda de problemas filosóficos y computacionales relacionados con el surgimiento de una *AI*².

Es importante mencionar que Alan Turing no ofreció una definición de *AI*, pero sus contribuciones fueron decisivas para el posterior surgimiento del término. Como tal, fue John McCarthy y un grupo de investigadores quienes durante un encuentro en Darmouth en 1956, propusieron la siguiente agenda de objetivos como parte de las tareas a desarrollar:

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, from abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves³.

1 Agradezco la lectura y comentarios de Daniel Martínez-Barba, Belen Prado y de los revisores anónimos por su lectura y comentarios a versiones previas de este manuscrito. Sus aportaciones fueron valiosas para la precisión de la tesis central de esta investigación. El presente artículo se encuentra bajo el auspicio del Fondo Fomento a la Investigación 2020, con código: UP-CI-2020-MEX-13-FIL: Lógica, Epistemología y Ética de la Inteligencia Artificial.

2 TURING, Alan. *The essential Turing. Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus The Secret of Enigma*, (Edición de Jack Copeland. Oxford: Oxford University Press, 2004). Esta edición de Copeland reúne las obras más importantes del autor y, además, ofrece estudios críticos que aclaran la importancia de los trabajos del matemático inglés; en adelante: *The essential Turing*.

3 MCCARTHY, John; MINSKY, Marvin; ROCHESTER, Nathaniel; SHANNON, Claude. *A proposal for the Darmouth Summer Research Project on Artificial Intelligence*, 1955, p. 1.

De este planteamiento se propusieron siete tareas específicas, tales eran: (1) *Computadoras automáticas*: la velocidad y la memoria de las computadoras de ese tiempo era insuficiente para simular las funciones superiores del cerebro humano, sin embargo, consideraban que el problema central estaba en la (in)capacidad del programador para escribir los programas. (2) *¿Cómo programar una computadora para que use lenguaje?*: al asumir que el cerebro humano utiliza y codifica signos lingüísticos, estudiar el modo en que la generalización de nuevas palabras implica el seguimiento de reglas previas y nuevas. (3) *Redes neuronales*: buscar el modo en que un grupo de neuronas se organizan para formar un concepto. (4) *Tamaño del cálculo*: desarrollar criterios eficientes para medir modos eficientes para resolver un problema. (5) *Mejora autónoma*: encontrar los medios a través de los cuales una computadora inteligente busca cambios automáticamente de modo independiente. (6) *Abstracciones*: describir y precisar los procesos mediante los cuales se clarifique el proceso de abstracción. (7) *Aleatoriedad y creatividad*: estudiar los mecanismos mediante los cuales la aleatoriedad se vincula a un pensamiento intuitivo que establece un principio de orden.

John McCarthy, en una serie de ensayos recolectados posteriormente, reconocería que ese encuentro de Dartmouth en 1956 sería considerado decisivo para la génesis del concepto, pero sobre todo para el desarrollo de una agenda de investigación cuyo marco de objetivos partiera de estas siete tareas. En la retrospectiva sobre los logros, el autor menciona que la actividad central del encuentro fue el estudio de “las relaciones entre el cerebro y la computadora”⁴, cuestión que tuvo como referencia los trabajos de Turing: “[he] proposed that mental processes be studied by programming a computer to carry them out rather than by building machines that imitate the brain at the physiological level”⁵. De esta manera, para los iniciadores del término ‘AI’ la visión del matemático inglés fue determinante para la construcción de la agenda de problemas ya mencionados.

El presente artículo estudia el desarrollo del pensamiento de Alan Turing, así como sus aportaciones al campo de la computación. En el progreso de las investigaciones del matemático inglés, se aprecian referencias a conceptos y fundamentaciones sobre “motores de cálculo automático”, “máquinas pensantes”, “computadoras inteligentes” y “computadoras digitales pensantes”, pero no propiamente el término *inteligencia artificial*⁶, aún con ello, la reconstrucción de los

Disponible en línea: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> (última consulta 2 de agosto de 2021).

4 McCarthy, John, *Defending AI Research. A collection of Essays and Reviews* (Stanford: CSLI Publications, 1996), p. 45.

5 McCarthy, John, *Defending AI Research*, p. 46.

6 Los detalles y fundamentos de cada uno de los términos, se pueden consultar en: *The essential Turing*, pp. 362-506, más adelante revisaremos las definiciones y argumentos aducidos para

resultados que llevaron a preguntar *Can Machines think?* fueron decisivos para que McCarthy y su grupo lograran sus objetivos en el Dartmouth College⁷.

En el resto del artículo, se expone el pensamiento computacional de Turing desde el contexto histórico que animó a imaginar una solución al problema matemático de la completitud, en donde su primera propuesta concebía un procesamiento de la información del tipo “sí... entonces”, hasta sus reflexiones sobre las posibilidades de construir una computadora inteligente, cuyas características presentan una primera aproximación a lo que hoy se conoce como “aprendizaje por refuerzo”⁸; el conjunto permitirá delimitar los aportes que el grupo de Dartmouth tuvo en cuenta para el desarrollo del término *AI*.

2. MATEMÁTICA Y COMPUTACIÓN. LA VALIDACIÓN DEL TEOREMA DE LA COMPLETITUD

Con respecto a los antecedentes del concepto *AI* es necesario remitirse a una conferencia dictada en 1900 en París por David Hilbert, en donde delimitaba una serie de problemas matemáticos que debían ser resueltos por esta ciencia; más tarde, él mismo, en 1927 durante el Seminario de Matemáticas de Hamburg, pronunciaría una conferencia titulada *The Foundations of Mathematics* en la cual sostendría lo siguiente:

The source of pure existence theorems is the logical c-axiom, upon which in turn the constructions of all ideal propositions depends... This formula game enables us to express the entire thought-content of the science of mathematics in a uniform manner and develop it in such a way that, in a same time, the interconnections between the individual propositions and facts become clear⁹.

La aspiración de Hilbert, si bien se colocaba dentro de la esfera puramente matemática, requería de la lógica para estructurar los fundamentos. De acuerdo a sus postulados había que probar que los axiomas no eran contradictorios entre sí, para lo cual había que demostrar que partiendo de éstos no se llegaría nunca

cada uno de estos términos.

7 Aunque en la convocatoria del encuentro no se menciona como antecedente la obra de Turing, años después el mismo McCarthy reconocería que el trabajo “Computing Machinery and Intelligence” (*The essential Turing*, pp. 441-464) fue una referencia central para situar la génesis de problemas de una *AI*.

8 Agradezco a Belen Prado por hacerme notar esta distinción en el pensamiento de Turing.

9 HILBERT, David. “On the Foundations of mathematics”, (en: VAN HEIJENOORT, Jean, *From Frege to Gödel. A Source Book in Mathematical Logic, 1879-1931*. Cambridge: Harvad University Press, pp. 464-479, 1967), p. 495.

a resultados contrarios siguiendo un número finito de deducciones lógicas¹⁰. Para probar la veracidad de esta cuestión, había que solventar, primero, lo que el matemático alemán llamaría *Entscheidungsproblem* (problema de la decisión)¹¹. Otro matemático inglés, Alonzo Church, definiría este dilema del modo siguiente:

By the *Entscheidungsproblem* of a system of symbolic logic is here understood the problem to find an effective method by which, given any expression Q in the notation of the system, it can be determined whether or not Q is provable in the system¹².

De acuerdo a esta definición, era comprensible que los estudiosos de la época buscaran demostrar el postulado de Hilbert; tres serían las más importantes refutaciones a este proyecto: la de Gödel, la de Church y la de Turing. Primero, en 1931 Kurt Gödel afirmaría que no era posible determinar un axioma que estableciera la veracidad o falsedad de una proposición aritmética, pues existían limitaciones para tales efectos: “In any of the formal systems... there are undecidable problems of the restricted functional calculus (that is, formulas of the restricted functional calculus for which neither validity nor the existence of a counterexample is probable)”¹³. Por su parte, Church, desarrollando el cálculo lambda, demostraría que era imposible determinar si una fórmula bien formada era o no un teorema: “no effective decision procedure exists which suffices to determine of an arbitrary [well formed formula] whether or not it is a theorem”¹⁴. En general, tanto Gödel como Church explicaban que dada la afirmación primera de un teorema **L** como falso, podía admitirse, en una segunda instancia, una sustitución de sus postulados iniciales por otros que fueran verdaderos, pero dado que la proposición inicial había sido demostrada como falsa o irresoluble, al sustituirse por **L'** (cuya redefinición permitía resolver el teorema cumplidas las nuevas condiciones) entonces se implicaría que **L=L'**; en este nuevo contexto, siendo **L** verdadera o siendo plausible su solución, se derivaba un conflicto con la primera afirmación en

10 HILBERT, David, *Mathematical Problems* (ICM París, 1990): “But above all I wish to designate the following as the most important among the numerous questions which can be asked with regard to the axioms: *To prove that they are not contradictory, that is, that a definite number of logical steps based upon them can never lead to contradictory results*”, p. 8. Énfasis en el original.

11 Recomiendo ampliamente el trabajo de Belen Prado (“Ambivalence in Machine Intelligence: The epistemological Roots of the Turing Machine”, *Signos Filosóficos*, Vol. 23, No. 45 (2021), pp. 57-61) para valorar las implicaciones epistemológicas de este problema.

12 CHURCH, Alonzo. “A Note on *Entscheidungsproblem*”, (*The Journal of Symbolic Logic*, Vol. 1, No.1 (1936), pp. 40-41), p. 41, nota 6.

13 GÖDEL, Kurt. “On formally undecidable propositions of *Principia Mathematica* and related systems I”, (en: Kurt Gödel, *Collected Works, Vol. I (Publications 1929-1936)*. Edited by Solomon Feferman. Oxford: Oxford University Press, pp. 144-195, 1986), p. 187. En adelante *Kurt Gödel, I*.

14 CHURCH, Alonzo. *Introduction to Mathematical Logic. Vol. I*, (New Jersey: Princeton University Press, 1956), p. 246.

la que **L** era falsa. Finalmente Turing¹⁵, propondría una demostración del teorema de Hilbert recurriendo a un experimento computacional, concluyendo que, en aquellos casos en las que una proposición **L** falsa, al sustituirse por **L'** verdadera, dicha proposición se traduciría en la circularidad de la máquina generando que esta trabajase indefinidamente. Ello se interpretaría como la imposibilidad para demostrar el problema de la decisión: “the Hilbertian *Entscheidungsproblem* [have] no solution”¹⁶. De las tres aportaciones, en lo que resta de la sección, únicamente se retomará la ruta computacional considerada.

En el artículo de Turing, “On Computable Numbers, with an Application to the *Entscheidungsproblem*”¹⁷, la demostración comenzaba asumiendo la existencia de una máquina que tenía la siguiente fórmula de programación:

$$\delta: Q \times \Gamma \rightarrow \Gamma \times \{L, R\} \times Q^{18}$$

En donde:

δ : es la función de transición.

Q : es el estado en el que se encuentra la máquina ($q_1, q_2, q_3 \dots q_n$)

Γ : es el alfabeto leído por la máquina, el cual puede ser ($\square, 0, 1$)

L, R : la dirección que puede tomar la máquina en la cinta o memoria (izquierda o derecha).

Un ejemplo sería:

$$\delta(q_0, 0) = (0, R, q_1)$$

[ubicada la máquina en un estado inicial q_0 , lee 0, deja 0, y pasa a la derecha de la cinta para ubicarse en el estado q_1].

La explicación de Turing¹⁹ sobre el funcionamiento de la máquina (inicialmente presentada como parte de un experimento mental), concebía algunas alternativas de funcionamiento en las que se posibilitaba la impresión de números o “Digital

15 TURING, Alan, *The essential Turing*, pp. 58-90.

16 TURING, Alan, *The essential Turing*, p. 59.

17 TURING, Alan, *The essential Turing*, pp. 58-90.

18 Para la elaboración de estos ejemplos me he guiado de la exposición de Javier García (*La máquina de Turing (explicada)*). Conferencia 2016, en: <https://www.youtube.com/watch?v=NS-NQ5mCSs8>), pues aprovecha una simbología hoy en día mayoritariamente aceptada.

19 Turing (*The essential Turing*, pp. 59-68) ofrece una explicación más amplia de estas ideas, pero usando otros símbolos y apoyándose de los números binarios. Un ejemplo que él propone es el siguiente:

“In this way we reduce each line of the table to a line of one of the forms (N_1), (N_2), (N_3).

From each line of form (N_1) let us form an expression $q_i S_j S_k L q_m$; from each line of form (N_2) we form an expression $q_i S_j S_k R q_m$; and from each line of form (N_3) we form an expression $q_i S_j S_k N q_m$ ” (p. 67).

Numbers” (1,2,3,4,5,6 y 7) y letras o “Standar Description” (“A”, “C”, “D”, “L”, “R”, “N”, así como el caracter “;”)²⁰. Con esta máquina se determinaría si los números computables eran enumerables. Eventualmente, debido al problema de la detención (*Halting problem*), Turing concluiría que la hipótesis de la completitud de Hilbert era falsa, lo que se traducía en la imposibilidad de demostrar si el total de las proposiciones eran falsas o verdaderas universalmente²¹.

De acuerdo a esto, la completitud aducida por Hilbert encontraría tres respuestas definitivas que demostraban su falsedad; si bien cada una de ellas establecía que no era posible construir un algoritmo general que valorara si una fórmula del cálculo del primer orden era un teorema, se confirmaría que la solución de Turing y su máquina eran un recurso de demostración válido²²; al respecto Marvin Minsky otorgaría el reconocimiento este procedimiento del siguiente modo:

We define the computable real numbers... as sequences of digits interpreted as decimal fractions. But we add one key restriction. *The digits must be generated sequentially by a Turing machine.* That is, we require that in order that the real number $.a_0a_1a_2\dots$ be a *computable real number*, there must be a Turing machine which starts with a blank tape and prints out a tape of the form:

$$\dots 000Xa_0Xa_1Xa_2X\dots 000^{23}.$$

Como se aprecia, la relevancia de este procedimiento no descansa en la novedad de la conclusión de Turing, sino en los recursos implementados para alcanzar este resultado. La demostración aritmética de Gödel o el cálculo de Church, a pesar de haber argumentado que el teorema de la completitud era falso, se obtenía como logro importante el diseño de una máquina con la cual se demostraba un problema matemático.

A pesar de este avance, el dilema no se hizo esperar: ¿eran las computadoras un recurso que podía sustituir el trabajo de los matemáticos? ¿y a los humanos en

20 TURING, Alan, *The essential Turing*, p. 66.

21 Belen Prado (“Ambivalence in Machine Intelligence”, pp. 66-69) considera que no sólo se tratan de consecuencias para la computación, sino que pueden ampliarse sus resultados a la demostración de que no es posible establecer una determinación para ningún sistema inteligente. Las implicaciones antropológicas, como se ve, son bastante interesantes.

22 Son necesarias varias precisiones sobre esta afirmación, por ello se remite al detalle histórico elaborado por Claudio Gutiérrez (“Kurt Gödel y Alan Turing: una nueva mirada a los límites de lo humano”, *Revista BITS de Ciencia, DCC*, No. 9, 2013, pp. 19-27). La resistencia de los matemáticos de la época, y del propio Gödel, a ciertas generalizaciones son fundamentales en lo que respecta a los fundamentos de las matemáticas y la computación misma. En el siguiente apartado se retoman algunas de estas ideas.

23 MINSKY, Marvin. *Computation. Finite and Infinite machines*, (Nueva Jersey: Prentice Hall, 1967), p. 158, énfasis en el original.

lo general? La cuestión central, al paso de los años, se tradujo en debate importante ya no sólo dentro de las definiciones matemáticas de “función recursiva”²⁴ o dentro del funcionamiento correcto de la propia máquina de Turing y su programación, sino que alcanzó la esfera de la filosofía (antropología y de la mente)²⁵. La cuestión básica apuntaba a si la mecanización de un proceso podía traducirse como la esencia misma del acto intelectual²⁶; si fuese el caso, entonces, la posibilidad para sustituir la inteligencia natural humana por una *AI* no sólo era evidente sino deseable –dados los beneficios que traía consigo resolver los problemas a partir de su mecanización²⁷. Pero la complejidad de estos temas se tratará más adelante, en una nueva sección.

El siguiente apartado presenta los argumentos que ofreció el propio Turing con respecto a la plausibilidad de una máquina capaz de pensar; gran parte de las ideas que propuso estaban focalizadas a satisfacer ciertas posturas sociales que se resistían a aceptar la idea de crear una inteligencia no natural, lo que sirvió para que Turing profundizará en el concepto, proceso metodológico y límites epistemológicos del concepto de *AI*.

3. TURING Y LAS MÁQUINAS INTELIGENTES: ARGUMENTOS DE PLAUSIBILIDAD

A pesar de los avances logrados y de los varios resultados alcanzados por las máquinas que Turing ayudó a diseñar para resolver algunas de las operaciones de comunicación durante la Segunda Guerra Mundial²⁸, la resistencia a una cierta inteligencia cuya naturaleza fuese mecanizada mantuvo por igual el escepticismo como el rechazo entre el mundo matemático. En contramedida, a partir de 1947

24 GUTIÉRREZ, Claudio. “Kurt Gödel y Alan Turing”, pp. 24-27.

25 ROBINSON, William S. “Philosophical Challenges”, (*The Cambridge Companion of Artificial Intelligence*, Edited by Keith Frankish; William M. Ramsey, Cambridge: Cambridge University Press, 2014), pp. 69-72; MINSKY, Marvin. *The Society of Minds*, (New York: Simon & Schuster, 1986), pp. 38-70; BOŁTUĆ, Piotr. “Church-Turing Lovers”, (en: Lin, Patrick; Jenkins, Ryan; Abney, Keith, eds. *Robot Ethics 2.0. From autonomous Cars to Artificial Intelligence*. Oxford: Oxford University Press, 2017), pp. 214-229.

26 ARKOUDAS, Konstantine; BRINGSJORD, Selmer. “Philosophical foundations”, (en: *The Cambridge Companion of Artificial Intelligence*, Edited by Keith Frankish; William M. Ramsey, Cambridge: Cambridge University Press, pp. 36, 63, 2014), pp. 36-39.

27 LIN, Patrick; JENKINS, Ryan; ABNEY, Keith, eds. *Robot Ethics 2.0. From autonomous Cars to Artificial Intelligence*. (Oxford: Oxford University Press, 2017).

28 COPELAND, Jack, *The essential Turing*, pp. 313-335.

Turing comenzó a justificar la posibilidad de las computadoras como objetos físicos capaces de pensar artificialmente²⁹.

Una de las cuestiones centrales –constantemente referidas por el autor– era si la intelección tenía que situarse, necesariamente, al lado de la infalibilidad de las computadoras. El tema no resultaba menor pues si la creación de un sistema inteligente ofrecía mejores recursos de solución a los problemas que por siglos habían ocupado la mente de los investigadores, entonces se erigía la necesidad de reconsiderar las capacidades humanas y, por ende, el nuevo rumbo que podría asumir algunas ciencias como la Matemática. La complejidad estaba en la posibilidad de una computadora para elegir. De acuerdo a su opinión, mientras que una máquina programada no podría frenar su búsqueda ante una solución imposible, un investigador lograría, en cambio, detenerse y elegir una nueva ruta de indagación, lo que reflejaba que la mejor vía era la que ofrecía una intelección natural.

Esta comparación de habilidades entre humanos y computadoras, para Turing, era errónea pues implicaba una inequidad en los términos de valoración de lo que podía entenderse por el concepto de ‘inteligencia’:

It might be argued that there is a fundamental contradiction in the idea of a machine with intelligence. It is certainly true that ‘acting like a machine’ has become synonymous with lack of adaptability. But the reason for this is obvious. Machines in the past have had very little storage, and there has been no question of the machine having any discretion. The argument might however be put into a more aggressive form. It has for instance been shown that with certain logical systems there can be no machine which will distinguish provable formulae of the system from unprovable... if a machine is expected to be infallible, it cannot also be intelligent³⁰.

Como se aprecia, él consideraba que el centro de las críticas se colocaba en la, por entonces, imposibilidad de las máquinas para encontrar vías de solución alternativas dado que la programación inicial impedía una elección más flexible; él mismo advierte que en aquellos casos en los que la incorporación de los recursos permitiese a la computadora considerar algunas estrategias para evitar la circularidad, la designación de un objetivo de solución concentraba las operaciones en un marco operativo delimitado: “The possibility of letting the machine alter its own instructions provides the mechanism for this, but this of course does not get us very far.”³¹.

29 Se recomienda el detalle histórico y sistemático elaborado por Jack Copeland (*The essential Turing*, pp. 217-264) en su edición crítica a las obras de Turing.

30 TURING, Alan, *The essential Turing*, pp. 393-394.

31 TURING, Alan, *The essential Turing*, p. 393.

A pesar de ello, el matemático inglés consideraría que estas restricciones sobre las posibilidades electivas de las computadoras podían encontrar bordes de restricción más amplios en la medida en que el *Hardware* mejorara. La reflexión central se dirigía a encontrar un recurso material que tuviese una memoria lo suficientemente grande como para garantizar que trabajara indefinidamente hasta alcanzar una solución al problema: “it [is] essential in this theoretical arguments that the memory should be infinite”³². Como alternativas para lograr esta idealidad, propone tres rutas de investigación de materiales que hicieran efectiva dicha condición: almacenamiento magnético, patrones de carga en la pantalla de un tubo de rayos catódicos y líneas de retardo acústico como unidades de almacenamiento³³. El esfuerzo de mostrar estas nuevas alternativas para mejorar la memoria, si bien no garantizaban que de ello surgiera una especie de voluntad artificial, se delineaba la idea de que era posible mejorar el marco de recursos electivos derivado de situaciones previas enfrentadas por la máquina y por el programador.

Con la búsqueda de recursos para mejorar la memoria se ampliaba el marco de investigación sobre patrones de aprendizaje, así como el valor del diseño y desarrollo de mecanismos de solución de problemas computacionales. El propio Turing, en 1947, atenuaba este avance para el campo de la computación, pues mesuraba los resultados a los que se podía aspirar con una máquina con memoria casi infinita, ya que consideraba que el papel del analista sería siempre indispensable para mejorar los procesos de la máquina. En contraparte, el beneficio de las computadoras digitales sería la supresión del “sentido común” en ciertas esferas del trabajo humano, pues con la programación de una computada se garantizaba un recurso de operación sistemático: “With a digital computer we can no longer rely on common sense, and the bounds of error must be based on some proved inequalities”³⁴.

Estas y otras indagaciones en la misma línea de reflexión tenían como interés mostrar la plausibilidad del uso dentro de las comunidades matemáticas, por lo que se puede decir que las reflexiones hasta antes de 1940 eran únicamente experimentales. Hacia 1948, en un artículo titulado *Intelligent Machinery*³⁵, propondrá una investigación apegada a las posibilidades reales hasta entonces logradas e iniciará una defensa del desarrollo de una inteligencia no natural. Para tales efectos, comienza por coordinar las objeciones estableciendo que son cinco restricciones contra el diseño de una máquina pensante. Tales son:

32 TURING, Alan, *The essential Turing*, p. 394.

33 TURING, Alan, *The essential Turing*, pp. 380-384.

34 TURING, Alan, *The essential Turing*, p. 391.

35 TURING, Alan, *The essential Turing*, pp. 410-432.

- a) Una renuencia dogmática para aceptar la existencia de un comportamiento inteligente distinto al del ser humano.
- b) La concepción de que religiosamente la creación de una máquina inteligente contravenía a los fundamentos en las que se sostenían tales creencias.
- c) El estado de las máquinas y sus posibilidades de operación alcanzados hasta 1940, lo que, al parecer, demostraba sus límites naturales y necesarios de desarrollo. (Impulsados por una serie de argumentos teológicos, había que garantizar dicho estado a riesgo de intervenir en la forma de la creación deseada por Dios).
- d) La posibilidad limitada por las máquinas de demostrar ciertos teoremas que, por el contrario, los matemáticos podrían resolver ya que podían seguir rutas alternativas de solución reflejándose la capacidad flexible de la mente humana y acentuando la rigidez de las soluciones mecanizadas.
- e) Cualquier tipo de inteligencia 'artificial' no era más que el reflejo de su creador, por ende, la 'inteligencia' correspondía al programador y no podía atribuirse tal cualidad a un objeto material³⁶.

Turing considera que las objeciones (a) y (b), al ser posturas emotivas y carecer de una racionalidad lógica, no requerían respuesta pues carecían de un núcleo argumental; la objeción (c), tampoco requería respuesta pues ya existían máquinas que podían realizar una cantidad de $10^{60.000}$ operaciones sin repetirse, lo que implicaba que, en todo caso, se trataba de una ignorancia sobre el estado alcanzado por la ingeniería computacional hacia 1948. La objeción (d) implicaba un contrasentido, pues de fondo se proponía que una máquina estuviera exenta de errores para demostrarse así su inteligencia siendo que el ser humano había alcanzado un estado de conocimientos importante sólo después de ensayar y corregir sus equivocaciones continuamente a lo largo de la historia. Finalmente, la objeción (e) es refutada diciendo que un alumno no puede ser considerado obra de su maestro una vez el primero logra actuar por cuenta propia o superar al segundo, en el caso de las máquinas, existía evidencia de ambos hechos³⁷.

Una vez Turing resuelve estas objeciones, reflexiona sobre las posibilidades y formas de una *AI*. Él comienza por mostrar sus tipos, los cuales son: Discrete and continuous, Controlling and Active, Logical Computing, Practical Computing, Universal Practical Computing, Paper y, finalmente, Partially random and apparently random machines; la finalidad es mostrar la gran variedad de alternativas de recursos de solución de problemas. Aunado a ello, muestra un cierto tipo

36 TURING, Alan, *The essential Turing*, pp. 410-411; 449-458.

37 TURING, Alan, *The essential Turing*, pp. 412-413.

de máquinas cuya operación dependía de las posibilidades de interrelación de los datos procesados en lugar de la operación como tal. Puesto que el funcionamiento implicaba la vinculación de los datos introducidos (programados), parecía, por ende, que los resultados serían limitados, sin embargo, al centrarse en la interconexión, ello representaría un avance en la metodología de procesamiento de la información. En palabras del propio Turing: “The behavior of a machine with so few units is naturally very trivial. However, machines of this character can behave in a very complicated manner when the number of units is large”³⁸.

Gracias a las posibilidades de interconectar la información, se propondría que una máquina podía alcanzar ciertos estados de procesamiento que el ser humano no era capaz, por lo cual, una computadora inteligente aventajaba en algunas cuestiones a la inteligencia humana; sin embargo, para Turing el objetivo no era la recreación de una máquina humanoide, sino resaltar algunas actividades que él consideraba centrales para el diseño y desarrollo de una *AI*: “...by applying appropriate interference, mimicking education, we should hope to modify the machine until it could be relied on to produce definite reactions to certain commands”³⁹. Para lograr ésto, sitúa las hipótesis de una arquitectura cuyo procesamiento de la información estuviese constituida a partir de un modelo neuronal humano; específicamente consideraba el patrón de flujo cerebral acaecido cuando el individuo experimenta situaciones de dolor y placenteras, pues consideraba que tales experiencias remitían a situaciones modelo para comprender el procesamiento de la información. Específicamente, proponía a la infancia para estudiar el proceso de aprendizaje durante el crecimiento⁴⁰.

El autor propuso desarrollar una línea de investigación para construir una máquina capaz de operar con patrones aleatorios, basándose en el modo en que el cerebro humano procesa cierta información; con ello propuso diseñar un sistema de alimentación *input-output* y un sistema de procesamiento semejante a la estructura neuronal del modo en que operan las redes neuronales. Lo relevante filosóficamente de esta hipótesis era la idea de la existencia de una simetría conceptual entre el diseño de una máquina (programación) y el crecimiento (aprendizaje) de un infante, pues de ese modo se proponía la unidad entre el mundo teórico, el de la percepción y el natural. Turing, en una delimitación a sus alcances, resume así sus reflexiones:

The analogy with the human brain is used as a guiding principle. It is pointed out that the potentialities of the human intelligence can only be realized if suitable

38 TURING, Alan, *The essential Turing*, p. 417.

39 TURING, Alan, *The essential Turing*, p. 422.

40 TURING, Alan, *The essential Turing*, pp. 422-428.

education is provided. The investigation mainly centers round an analogous teaching process applied to machines. The idea of an unorganized machine is defined, and it is suggested that the infant human cortex is of this nature... In one case the education process is carried through until the organization is similar to that of an [Automatic Computing Engine]⁴¹.

En general, era claro que para el matemático la analogía humano-computadora debía aparejarse tanto como hipótesis de experimentación e investigación, como modelo de desarrollo para el desarrollo de una cierta forma artificial de inteligencia deseable para la creación de una nueva generación de computadoras.

Estas reflexiones se verán profundizadas en otro celebre texto de 1950 titulado *Computing Machinery and Intelligence* en donde sugiere el famoso "Imitation game"⁴²; la tesis central del artículo plantea el siguiente experimento mental: supongamos que tres personas, (A) una mujer, (B) un hombre y un (C) participante (sexo indistinto) que no sabe si A o B es la mujer. Como parte de las condiciones del juego se encuentran aislados unos de los otros y se encuentran comunicados por tarjetas mecanografiadas que pueden enviarse entre sí. El objetivo del juego es determinar, por parte de C, cuál de los participantes es la mujer (A o B). La cuestión se vuelve más compleja cuando Turing propone la siguiente variación:

'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'⁴³.

Si bien la recepción de la experimentación cuestionaba la identidad entre la definición de inteligencia, el procesamiento del lenguaje y la representación conceptual lograda de dicha vinculación⁴⁴, la propuesta de Turing se seguía manteniendo en la exploración de los procesos de aprendizaje infantil y las posibilidades de modelar procesos cognitivos a partir de dicho patrón de desarrollo evolutivo. De acuerdo con él, la evidencia psicopedagógica demostraba que la inteligencia infantil dependía de los estímulos que el medio circundante le ofrecían (input) para producir estados conductuales específicos (output) de tal manera que el procesamiento de la información (aquello que se socioculturalmente se reconoce como 'Inteligencia'), estaba directamente vinculado con la capacidad del agente

41 TURING, Alan, *The essential Turing*, pp. 431-432.

42 TURING, Alan, *The essential Turing*, pp. 441-464.

43 TURING, Alan, *The essential Turing*, p. 441.

44 DREYFUS, Hubert, *What Computers can't do. A critique of Artificial Reason*, (New York: Harper & Row, 1973), en adelante: *Computers*; SEARLE, John. "Minds, Brains, and Programs", (*The Behavioral and Brain Sciences*, Vol. 3 (1980), pp. 417-457), en adelante *Minds*; ARKOUDAS, Konstantine; BRINGSJORD, Selmer. "Philosophical foundations", pp. 46-52.

para procesar una cierta información recibida y, eventualmente, responder en consistencia con dicha proposición. En ese contexto, para Turing⁴⁵, la conclusión era evidente: si un ser humano determina que las respuestas dadas por una computadora digital provienen de un ‘ser humano’, entonces había que asumir que la máquina era inteligente.

La conclusión, como se aprecia, asumía al menos dos referentes epistémicos centrales para construirse, los cuales podrían formularse de la forma siguiente:

- a) La inteligencia es el producto de una formulación lingüística consistente con ciertos parámetros culturalmente constituidos.
- b) El desarrollo cerebral, sobre todo durante el proceso de maduración infantil, muestra una estructura base del intelecto humano, el cual, por ende, sirve como base para la arquitectura del software.

Para el autor era fundamental que el participante humano determinara que una computadora digital era específicamente un ‘ser humano’ (asumiendo tal respuesta a partir de las respuestas dadas por la máquina), pues con ello se implicaba que un agente epistémico reconocía un conjunto de patrones conceptuales aprendidos culturalmente (y atribuidos convencionalmente a la inteligencia humana) en los textos intercambiados. Para alcanzar esta imitación, Turing proponía asumir a las redes neuronales como un criterio de desarrollo, pues nuclearmente el proceso de aprendizaje moldeaba el cerebro del niño de acuerdo a ciertos modelos culturales para que actuara bajo patrones conductuales reconocidos como inteligentes. Por ello, él mismo concluiría lo siguiente: “Intelligent behaviour presumably consists in a departure from the completely disciplined behaviour involved in computation, but a rather slight one, which does not give rise to random behaviour, or to pointless repetitive loops”⁴⁶; a partir de esto, será fundamental explorar los procesos y patrones de aprendizaje para mejorar las respuestas inteligentes esperables de una computadora, pues de los procesos de programación se lograría una mejor respuesta por parte de la máquina.

Como se aprecia, para el autor la idea de una computadora inteligente no constituía una simple analogía semántica o algún tipo de metáfora para referirse a una nueva área de la computación, en cambio implicaba el esfuerzo real para examinar los modelos de investigación relacionados con esta área de conocimiento: “My contention is that machines can be constructed which will simulate the behaviour of the human mind very closely”⁴⁷. Es importante remarcar que la finalidad última de Turing no era otra que demostrar lo siguiente: “machines will eventually

45 TURING, Alan, *The essential Turing*, pp. 448-449.

46 TURING, Alan, *The essential Turing*, p. 463.

47 TURING, Alan, “Intelligent Machinery. A Heretical Theory”, *The essential Turing*, p. 472.

compete with men in all purely intellectual fields”⁴⁸. En ese sentido, no es extraño que hubiese impulsado una reflexión sobre las condiciones de posibilidad a través de las cuales las máquinas fuesen capaces de pensamiento y cognición.

Hasta aquí se han mostrado los argumentos aducidos por Turing con respecto a la posibilidad de crear una computadora inteligente. Se ha visto que parte de la apología se proponía justificar los fundamentos y estructura de una máquina pensante procurando mostrar la factibilidad antropológica y ética de crear un objeto material capaz de manifestar tanto cognición como actividad mental. En el siguiente apartado se muestran algunos de los debates y lecturas desprendidos de las propuestas de Turing. La finalidad de esta selección es mostrar una doble cuestión: primero, los cuestionamientos por parte de matemáticos (inicialmente) y filósofos (posteriormente) a la factibilidad de que una máquina resolviese problemas o que fuese capaz de pensamiento. La segunda cuestión apuntará a mostrar cómo el proyecto de construir una computadora inteligente gradualmente se orientó a la configuración del término *AI*.

4. LA INTELIGENCIA ARTIFICIAL SEGÚN TURING: LIMITACIONES Y RELECTURAS

La presente sección se subdividirá en dos sub-apartados para mostrar algunas limitaciones a las posibilidades de una computadora inteligente: revisiones de matemáticos y relecturas filosóficas. A través de la presentación de estas reflexiones que se harían sobre las propuestas de Turing, se valoraran los aportes definitivos, las limitaciones y los fundamentos que servirían para una posterior definición del concepto de *AI*.

4.1. EL TRABAJO Y LA FLEXIBILIDAD DEL PENSAMIENTO MATEMÁTICO

La propuesta de una máquina que determinara si la hipótesis de Hilbert era posible, aunque encontró el respaldo y validación de Alonzo Church en su momento, ello no implicaba que la comunidad estuviese de acuerdo en la totalidad de los argumentos y, sobre todo, en los procesos que Turing proponía para demostrar la falsedad de la proposición⁴⁹. Casi de inmediato Emil Post publicaría una crítica al estudio sobre los números computables; algunas de las restricciones pro-

48 TURING, Alan, *The essential Turing*, p. 463.

49 Cfr. COPELAND, Jack, en: *The essential Turing*, pp. 52-53; 91-93.

puestas por el autor serían: (1) algunos de los comandos propuestos no presentan los resultados determinados por Turing, ya sea porque sus cálculos no fueron precisos o porque habían algunos errores menores de notación; (2) se establecían algunos procedimientos iniciales para el movimiento (la máquina únicamente se puede mover a la izquierda al iniciar, nunca borra el símbolo dado inicialmente, ni imprime alguno cuya definición estuviese dada; Post se pregunta si esta decisión arbitraria incidía en la demostración de la insolubilidad de la hipótesis de Hilbert. (3) La tercera crítica, remite, nuevamente, a la selección sin criterios definidos de los símbolos cuyo cálculo implicaba que la máquina de Turing operara de un distinto al definido inicialmente⁵⁰.

En la misma línea, Donald Davies, presentaría un informe detallando dichos errores en la redacción original de Turing, al tiempo que insistía que éstas provocarían confusiones entre los lectores: “I realized that, even though the feasibility of the universal computing machine was not in doubt, the mistakes in Turing’s exposition could puzzle future readers and plague anyone who tried to verify Turing’s design by implementing his universal machine in practice”⁵¹. La respuesta de Turing, en una carta a Alonzo Church, explicaría que estaba de acuerdo en la existencia de dichos errores, sin embargo, reafirmaba que las fallas señaladas a su trabajo no alteraban el funcionamiento general propuesto.

La relevancia de estas precisiones remitía a la viabilidad de confirmar el teorema de la incompletitud de Gödel a través de medios computacionales. De acuerdo con Jack Copeland⁵², editor de las obras de Turing, algunos matemáticos se opusieron al surgimiento de una máquina como vía de demostración de teoremas, el argumento central era que se trataba de una tarea humana que no podría asumir a una computadora⁵³. De acuerdo con Claudio Gutiérrez⁵⁴, la tensión central estaba tanto en las aplicaciones que proponía, como en los alcances que la máquina de Turing ofrecía para la demostración de teoremas de primer orden. Si bien la aportación de una máquina que operara de acuerdo a las especificaciones del matemático inglés no aportaban un conocimiento distinto al desarrollado por Gödel en 1931, sí, al menos, ofrecía un criterio nuevo de interpretación de los mismos resultados presentados por el matemático alemán: la demostración de que la elección de los axiomas puede seguir un patrón intuitivo y, por tanto, su justificación proviene de la efectividad de la operación de una

50 POST, Emil, en: *The essential Turing*, pp. 98-99.

51 DAVIES, Donald, *The essential Turing*, p. 103.

52 COPELAND, Jack, *The essential Turing*, p. 53.

53 PRADO, Belen, “Ambivalence in Machine Intelligence”, pp. 61-64.

54 GUTIÉRREZ, Claudio, “Kurt Gödel y Alan Turing”, pp. 24-27.

máquina programada bajo un esquema también intuitivo. Precisamente esta es una de las objeciones de Gödel:

The [lack of statements] necessarily presuppose a kind of Platonism, which cannot satisfy any critical mind and which does not even produce the conviction that they are consistent⁵⁵.

Al hablar de “platonismo” Gödel apuntaba a que un sistema de la forma S justificaba un conjunto de teoremas $s_1, s_2, s_3 \dots s_n$ a partir los argumentos derivados de S , lo que haría que el sistema operara bajo un “idealismo formal” indemostrable⁵⁶. De acuerdo con Turing, las computadoras inteligentes, si bien proponían un sistema de programación para operar de acuerdo a criterios propuestos por el propio programador (sin necesariamente ofrecer una justificación inicial de dicha elección), ello no implicaba que los resultados carecieran de fundamentos demostrables, pero en lugar de atribuir argumentos en favor de su plausibilidad, el propio funcionamiento mostraría si el código de programación era correcto.

En resumen, se aprecia que las objeciones matemáticas se dirigían a cuestionar las posibilidades de una máquina que fuese capaz de inteligencia, pues se aducía que un proceso mecánico se ajustaba la rigidez de la programación predefinida y que, además, parecía no justificar los principios de los que partía. Turing, al respecto, objetará dos cuestiones: primera, que las posibilidades de una computadora inteligente no deben valorarse a partir de la ‘infalibilidad’, pues esta no es una característica de la inteligencia natural⁵⁷; en ese sentido, las posibilidades correctivas (ensayo y error) permitían la reestructuración de patrones de modelación cada vez más precisos. Segunda, aun cuando la teoría de la incompletitud apuntaba la necesidad de incorporar sistemas o funciones alternativos (S') para evaluar la factibilidad de un sistema dado (S), para el caso de la computación había que, primero, delimitar el campo semántico de posibles problemas que podían ser resueltos por una computadora; estratégicamente Turing proponía que la delimitación de los problemas cuya solución aceptaba métodos computacionales, no sólo serviría para definir el propio campo epistemológico de esta área

55 GÖDEL, Kurt, “On undecidable sentences” (en: Solomon Feferman, John Dawson, Warren Goldfarb, Charles Parsons, Robert Solovay, eds., *Kurt Gödel. Collected Works*, Vol. III, Oxford: Oxford University press), p. 35. En adelante: *Kurt Gödel*, III.

56 *Kurt Gödel*, III, “The present situation in the foundations of mathematics”, p. 50.

57 El propio Turing explica que sólo se puede saber lo que una computadora realizará hasta que finalice su procedimiento, pues si el programador fuese capaz de conocer el resultado, el acto mismo de programar carecería de sentido: “Sometimes a computing machine does do something rather weird that we hadn’t expected. In principle one could have predicted it, but in practice it’s usually too much trouble. Obviously if one were to predict everything a computer was going to do one might just as well do without it”, *The essential Turing*, p. 500.

de conocimiento, sino para garantizar la universalidad del método empleado así como el resultado alcanzado.

Con esto Turing ofrecería una justificación al término maquina o computadora “inteligente”, pues, así como la práctica de la matemática no estaba exenta de errores y, sin embargo, a través del gradual avance de las investigaciones se habían encontrado soluciones a los problemas que por años pudieron parecer irresolubles, de modo análogo, no se podía esperar que un campo emergente como el de la computación ofreciera respuestas definitivas a todas las cuestiones sin ofrecerle un periodo de tiempo suficiente para que los resultados se clarificaran y, con ello, pudiera definirse con precisión la agenda de problemas propios de una actividad vinculada al uso e investigación de las computadoras inteligentes. Al respecto, el propio Turing en 1951 plantearía la siguiente idea:

It might be argued that there is a fundamental contradiction in the idea of a machine with intelligence... It is certainly true that ‘acting like a machine’ has become synonymous with lack of adaptability... Against it I would say that fair play must be given to the machine. Instead of it sometimes giving no answer we could arrange that it gives occasional wrong answers... in other words then, if a machine is expected to be infallible it cannot also be intelligent... No man adds very much to the body of knowledge, why should we expect more of a machine?⁵⁸

Precisiones como esta, a pesar de su simpleza, aclaraban tanto el rumbo hacia el que debía orientarse una agenda de problemas relacionados con el estudio de la “inteligencia” de las computadoras, como el modelo del que debía partir para alcanzar resultados sólidos y definitivos. Esta apología, no podría resultar más profética, pues como se mencionó en la introducción apenas dos años después de su muerte en 1954, en Darmouth, un grupo de investigadores encabezados por McCarthy definirían el campo de trabajo de una *AI* reconociendo el funcionamiento del cerebro humano como modelo de procesamiento de información como referencia, pero también explicando que se trataba de objetivos propios e independientes de otras ciencias, pues a pesar de que se nutría de, por ejemplo, la Matemática o la Neurociencias, tenía una ruta de investigación propia y con tareas a desarrollar; dicha independencia, debe mucho a las defensas y proposiciones que realizaría Turing a lo largo de su actividad profesional. Pasamos ahora a revisar las objeciones filosóficas a las que fueron sometidas las propuestas del autor.

58 TURING, Alan, *The essential Turing*, p. 394.

4.2. SENTIDO Y FORMA DE UNA INTELIGENCIA ARTIFICIAL ¿PUEDEN LAS MÁQUINAS ENTENDER?

Si Turing había resuelto que la circunscripción de los temas y problemas posibles de una disciplina que investigara la “inteligencia” de las computadoras, debía ajustarse a una agenda propia de problemas cuyo lenguaje, metodología y objetivos se entendieran dentro de sus propios términos, para los filósofos el problema no se zanjaría únicamente por establecer límites epistemológicos y metodológicos entre los investigadores en el área computacional y el resto de saberes. Aunque son varios los dilemas que se pueden establecer⁵⁹, uno de los más importantes debates remitió al sentido y forma del significado de una forma de inteligencia no natural. Es claro que Turing, entre 1948 y 1950 defendió la idea de una cierta capacidad intelectual que podía ser reproducida en objetos materiales una vez se asumiera el modelo neuronal de desarrollo en la etapa infantil. De esta comparación análoga, a los filósofos les interesaría principalmente las implicaciones cognitivas: ¿puede una estructura diseñada a partir de un modelo neuronal humano entender?

Hubert Dreyfus y John Searle abrirían el debate al reconocer que en un cierto nivel la comprensión blanda y fuerte de la *AI* (*weak and strong AI*)⁶⁰ ayudarían a clasificar el grado de racionalidad viable de modelación neuronal para la organización de un programa computacional. La oposición central era suponer que, necesariamente, el cerebro humano operaba como una computadora y, por ende, había que asumir a la inteligencia como un cierto tipo de programa avanzado que habilitaba el entendimiento humano. Hay que recordar que Turing consideraba lo siguiente: “our digital computer, suitably programmed, will behave like a brain”⁶¹. A partir de ello, Dreyfus, años más tarde, sostendría:

a neural net can be simulated using a program, but such a program is in no sense a heuristic program. Thus the mere fact that the brain might be a digital computer is in no way ground for optimism as to the success of artificial intelligence as defined by Simon or Minsky⁶².

59 Para una síntesis de estos problemas filosóficos se recomiendan los siguientes trabajos: ARKOUDAS, Konstantine, pp. 34-63; BRINGSJORD, Selmer. “Philosophical foundations”; ROBINSON, William S. “Philosophical Challenges”, pp. 64-88; Prado, Belén, “Ambivalence in Machine Intelligence”, pp. 57-61.

60 El término es propuesto por Searle (*Minds*, p. 417), pero se aprecia que Dreyfus (*Computers*, pp. 71-72) acepta a un cierto nivel la modelación neuronal como referente para el desarrollo de ciertos sistemas de procesamiento de información.

61 TURING, Alan, *The Essential Turing*, p. 483.

62 DREYFUS, Hubert, *Computers*, p. 71.

Por su parte, para Searle:

according to strong *AI*, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to *understand* and have other cognitive states⁶³.

Como se aprecia, la crítica central no se remitía a la posibilidad de diseñar y desarrollar a las computadoras a partir de un modelo neuronal, sino, justamente, invertir los términos: que el cerebro fuese entendido a partir de un descubrimiento reciente generaba en ambos autores una primera oposición; ¿Qué criterios justificaban el definir la intelección natural a partir de una cierta intelección artificial? Pero la cuestión, más allá del anacronismo, recaía directamente en asumir que las computadoras eran objetos materiales que, debido al procesamiento de la información, debían reconocerse como (i) agentes epistémicos, (ii) capaces de entendimiento. Específicamente Dreyfus se opondrá a esta cuestión al señalar enfáticamente que “the evidence is against the possibility of using digital computers to produce intelligence”⁶⁴. Searle hará eco de la objeción al sostener que “the claim that the appropriately programmed computer literally has cognitive states and that the programs thereby explain human cognition”⁶⁵. La resistencia filosófica se oponía a la analogía entre un cierto modo de procesar la información cuyo referente de diseño consideraba la organización neuronal.

Estas críticas de Dreyfus y Searle tienen como base las posturas de Turing, para quien la estructuración de un programa que reproducía el modelo neuronal para procesar la información debía, por tanto, esperar resultados similares a los alcanzados por el cerebro humano. Es cierto que él aceptaba que el estado de la *AI* se encontraba en un momento incipiente, pero no por ello se debía desacreditar el rápido avance hacia el que se dirigiría la evolución de esta área de estudios. De acuerdo a ello, entonces, existía un contexto de racionalidad a través del cual se podía atribuir que las computadoras, independientemente de lo que culturalmente se definiera como procesamiento de información, no contravenía a que una *AI* ofrecía resultados similares a los esperados para un agente epistémico: “The idea... is that the machine has to try and pretend to be a man, by answering questions put to it, and it will only pass if the pretence is reasonably convincing”⁶⁶.

63 SEARLE, John, *Minds*, p. 417.

64 DREYFUS, Hubert, *Computers*, p. 74.

65 SEARLE, John, *Minds*, p. 417.

66 TURING, Alan, “Can Automatic Calculating Machines Be Said to Think?”, *The essential Turing*, p. 495. El objetivo del diálogo en el que se enmarca el presente texto citado es, precisamente, definir el grado de comparación entre la forma de la inteligencia humana y la arquitectura de programación inteligente de una computadora. Además, es aquí donde Turing proyecta que para que

Una de las cuestiones troncales de esta equivocidad provenía de las dificultades para definir el concepto de “aprendizaje”. La introducción de este nuevo término se justificaba porque si el acto intelectual se configuraba a través de un proceso de aprendizaje a partir de la relación “*Inteligencia si sólo si Aprendizaje*” ($I \leftrightarrow A$), entonces, para Turing, era posible revertir el estado bicondicional de tal modo que la relación se mantenía al sostener que $A \leftrightarrow I$. Al invertir los términos, era aceptable aducir que se habilitaba una ruta de investigación a partir de las condiciones generales que permitían al ser humano aprender. Considerando esta ruta, Turing proponía que dada una memoria lo suficientemente grande como para almacenar una buena cantidad de datos y un diseño *Hardware* que simulara una red neuronal, restaba por desarrollar la arquitectura de programación que reprodujera operaciones consideradas como inteligentes por un ser humano.

Dreyfus⁶⁷ se opondrá a estas ideas al sostener que se suele asumir que el acto reflexivo, así como el propio acto cognitivo de ordenamiento de la información es considerado evidente por parte de buena parte de los investigadores de la AI, por lo que parecería que no hay razones para dudar de una conducta ‘inteligente’. Para el autor, en cambio, dichas actividades, filosóficamente hablando, requieren una amplia justificación para garantizar tanto origen, desarrollo como interrelación de la información con la cual expresamos buena parte de nuestras emociones estados cognitivos básicos (miedo, creencias, seguridad, certeza, etc.). De acuerdo a ello, lo inteligente no es evidente y, por ende, la relación analógica entre el actuar de una computadora y el del ser humano no implicaba que una acción fuese ‘inteligente’.

La respuesta a estas objeciones, provendrían de, precisamente, los impulsores del término AI en Darmouth en 1956. De acuerdo con Marvin Minsky –un “entusiasta” de la AI⁶⁸, si bien existen algunos estatutos básicos atribuibles a esta facultad humana (resolución de problemas complejos, capacidad de adaptación, autoconciencia, entre otros), con el progresivo descubrimiento de resultados desde

las computadoras puedan superar el denominado *Test de Turing* deberán pasar, al menos, 100 años (para 2050 aproximadamente).

67 DREYFUS, Hubert, *Computers*, pp. 75-81.

68 Dreyfus (*Computers*) califica así, entre otros a Minsky, de tal modo que las objeciones están dirigidas tanto a la AI y a los investigadores entusiastas de esta área de estudio: “In 1957 Simon predicted that within ten years psychological theories would take the form of computer programs, and he set out to fulfill this prediction by writing a series of programs which were meant to simulate human cognition by simulating the conscious and unconscious steps a person goes through to arrive at a specific cognitive performance. And we have seen that despite the general inadequacy of such programs, admitted even by enthusiasts such as Minsky, all those involved in the general area of artificial intelligence (Minsky included) share the assumption that human beings, when behaving intelligently, are following heuristic rules similar to those which would be necessary to enable a digital computer to produce the same behavior”, p. 76.

áreas interdisciplinarias demostraba que las concepciones de lo ‘inteligente’ podía reajustarse a nuevos marcos de comprensión, lo que implicaría que no todo acto intelectual era reducible a un estatuto natural de nuestra capacidad. En todo caso, se trata tanto de una actividad natural propia del ser humano y de una concepción que reconoce la interconectividad de funciones como ‘acción inteligente’⁶⁹.

Además, a diferencia de Dreyfus, Minsky sostiene que un sistema inteligente no sólo remite a un sistema cuya operación es capaz de resolver problemas que ninguna “mente humana” puede, eso sólo implicaría una región desconocida por explorar a fin de encontrar una respuesta. En contramedida, para Minsky un sistema “experto” es cualquier actividad cuya programación implica la interrelación de información que habilita acciones consideradas “de sentido común (de acción regular)”; sin embargo, la facilidad de una acción, únicamente demuestra la gran capacidad neuronal del ser humano para resolver situaciones complejas, a pesar calificarse como una resolución simple. En ese sentido, el que se pueda programar una computadora para realizar acciones “de sentido común” refleja un gran avance en el campo de la *AI*:

To be considered an ‘expert,’ one needs a large amount of knowledge of only a relatively few varieties. In contrast, an ordinary person’s ‘common sense’ involves a much larger variety of different types of knowledge-and this requires more complicated management systems⁷⁰.

Otros argumentos en favor de situar una *AI* como un campo que había logrado una independencia luego de las propuestas de Turing, fue McCarthy quien –en su respuesta a las críticas de Dreyfus⁷¹– consideraba que una aportación que se podía esperar de las computadoras era una nueva clase de “intuicionismo” para la fundamentación del conocimiento, particularmente de la lógica. En lugar de entender este proceso cognitivo como un acto creativo sobre el cual se pretendía sostener la programación de las computadoras se debía asumir como un proceso sistemático que implicaba un modo de proceder organizado, pero del cual únicamente se establecían la lógica de procedimiento general, para luego esperar el resultado de una computadora con la cual se simplificaban las operaciones, pero no así los principios que debían procesarse por parte del ordenador: “Common sense knowledge and reasoning is at the core of *AI*, because a human or an intelligent machine always starts from a situation in which the information available to it has a common sense character”⁷².

69 MINSKY, Marvin, *The Society of Mind*, (Londres Pan Books, 1988), p. 71.

70 MINSKY, Marvin, *The Society*, p. 72. Énfasis en el original.

71 Se refiere particularmente a los planteamientos realizados en *Computers*.

72 McCARTHY, John, *Defending AI Research*, p. 89.

Con todo esto, se aprecia una cuestión fundamental que había sido defendida por Turing y que luego sería objeto de debate en el seno filosófico y por parte de quienes propusieron formalmente el término *AI*; el dilema central indagaba si el concepto de Inteligencia, de acuerdo a la concepción de los filósofos, necesariamente se ligaba a las acciones humanas vinculadas con el ejercicio del conocimiento, o si, en cambio, remitía a la capacidad para procesar información asumiendo como modelo de procesamiento a la actividad cerebral. Siendo esta una cuestión que se ha reflexionado aún después de la muerte de Turing, se puede pensar por qué el matemático inglés encontró resistencias cuando propuso la idea de una máquina capaz de entendimiento. Pero aun cuando éste último no hubiese encontrado una respuesta definitiva, sin duda sus aportaciones resultaron cruciales para el desarrollo del término, así como para la delimitación de las tareas propias de este campo.

En resumen, el concepto de *inteligencia* y *entendimiento* que Turing consideraba podían desarrollar una máquina, remitían a las posibilidades de interrelación de datos e información que se podía habilitar para que una computadora operara. Si bien la principal crítica de Dreyfus y Searle se enfocan a cuestionar que no todo sistema cuya arquitectura tome como modelo el cerebro humano operará inteligentemente, ni, por ende, entenderá la información que procesa, para Turing (y posteriormente McCarthy y Minsky) la conducta inteligente remite a la posibilidad de interrelacionar la información a través de un modelo neuronal de interconexión. En ese sentido, la cuestión artificial de este modelo de comportamiento, no se propone alcanzar conductas humanas por parte de una computadora, sino programar acciones cuyo procesamiento informacional (por simple que este pueda ser) sea análogo al modo en que es procesado por un ser humano.

CONCLUSIONES

Aunque no existe un concepto de *Artificial Intelligence* (AI) propuesto por Turing, se aproxima a él a través de su propuesta de una *máquina pensante*, y la cual, a su vez, fue determinante para el surgimiento del término en 1956, en Darmouth en el encuentro entre McCarthy, Minsky y otros. De acuerdo con lo desarrollado hasta aquí, tres cuestiones fueron las aportaciones de Turing para la eventual definición del concepto: (1) el concepto de *Inteligencia* no remite, necesariamente, a la esfera humana pues esta capacidad refiere a la posibilidad de resolver problemas a partir de un sistema de principios definidos inicialmente y a partir de los cuales se operarán y definirán acciones posteriores. (2) Una *máquina* es capaz de procesar información a partir de la delimitación de funciones y

operaciones generales, lo cual cumple con la definición de inteligencia. (3) El progreso de las investigaciones eventualmente cristalizaría en un campo más definido para el desarrollo de los problemas relacionados con la intelección (resolución de problemas) y la computación (máquinas que procesarían y calcularían información) para con ello simplificar procesos cuyas aplicaciones podrían favorecer a distintos campos de saberes.

El proyecto de Turing, a pesar de ofrecer soluciones a diversos campos de problemas en la matemática y la criptografía, por citar ejemplos, encontró resistencias a su concepción de un campo vinculado con las máquinas inteligentes. Dos, fueron las principales resistencias: por parte de los *matemáticos*, las objeciones cuestionaban los procedimientos lógicos y metodológicos de entre las cuales una computadora elegía para proceder a completar una demostración; la resistencia central estaba, por tanto, en la falta de flexibilidad inherente a la línea de programación. Los *filósofos*, por su parte, se oponían al concepto “inteligencia” pues implicaba la capacidad volitiva y cognitiva atribuible a un objeto material que, por definición, carecía de capacidad para actuar bajo parámetros humanos. Sin embargo, la aportación central de Turing remite al desarrollo de un sistema de procesamiento de la información a través de la interconectividad de datos con los cuales se habilite una toma de decisiones cuyo resultado favorezca la resolución de un problema que requiere la mecanización de los procesos. Sin embargo, la paradoja central de este modo de proceder es que a pesar de que existe una programación predefinida, aún con ello no se garantiza que se puedan predecir los resultados, lo cual implica que el procesamiento de la información con *AI* cumple con las condiciones básicas esperadas para una inteligencia humana: la toma de decisiones sigue siendo indeterminable e imposible de definir a pesar de que opera con datos e instrucciones conocidos; una cuestión que sin justificar que las computadoras son capaces de pensar, sí, al menos, habilita la posibilidad de responder afirmativamente a la pregunta de Turing: “Can Machines think?” y que años más tarde permitiría el surgimiento de un campo denominado *Artificial Intelligence*.

REFERENCIAS BIBLIOGRÁFICAS

- ARKOUDAS, Konstantine y BRINGSJORD, Selmer, “Philosophical foundations”. En: FRANKISH, Keith; RAMSEY, William M., eds., *The Cambridge Companion of Artificial Intelligence*, Cambridge University Press, 2014, 34-63.
- BOŁTUĆ, Piotr, “Church-Turing Lovers”. En: LIN, Patrick; JENKINS, Ryan; ABNEY, Keith, eds. *Robot Ethics 2.0. From autonomous Cars to Artificial Intelligence*. Oxford: Oxford University Press, 2017, 214-229.
- CHURCH, Alonzo, “A Note on *Entscheidungsproblem*”, *The Journal of Symbolic Logic*, Vol. 1, No.1, 1936, 40-41.

- , *Introduction to Mathematical Logic. Vol. I.* New Jersey: Princeton University Press, 1956.
- DAVIES, Donald W., “Corrections to Turing’s Universal Computing Machine”. En: TURING, Alan. *The essential Turing. Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus The Secret of Enigma.* Edited by Jack Copeland. Oxford: Clanderon Press/Oxford University Press, 2004, 103-124.
- DREYFUS, Hubert, *What Computers can’t do. A critique of Artificial Reason*, (New York: Harper & Row, 1973).
- GARCÍA, Javier, *La máquina de Turing (explicada)*. Conferencia 2 de diciembre de 2016, en: <https://www.youtube.com/watch?v=NS-NQ5mCSs8>
- GÖDEL, Kurt, “On undecidable sentences”. En: FEFERMAN, Solomon; DAWSON, John; GOLDFARB, Warren; PARSONS, Charles, SOLOVAY, Robert, eds., *Kurt Gödel. Collected Works*, Vol. III, Oxford: Oxford University press, 1995, 30-35.
- , “On formally undecidable propositions of *Principia mathematica* and related systems I”. En: GÖDEL, Kurt, *Collected Works, Vol. I (Publications 1929-1936)*. Edited by Solomon Feferman. Oxford: Oxford University Press, 1986, 144-195.
- GUTIÉRREZ, Claudio, “Kurt Gödel y Alan Turing: una nueva mirada a los límites de lo humano”. *Revista BITS de Ciencia, DCC*, No. 9, 2013, 19-27. Disponible en: <https://users.dcc.uchile.cl/~cgutierr/otros/godel-turing-bits.pdf>
- HILBERT, David, *Mathematical Problems*. ICM París, 1990.
- , “On the Foundations of mathematics”. En: HEIJENOORT, Jean van, ed. *From Frege to Gödel. A Source Book in Mathematical Logic, 1879-1931*. Cambridge: Harvard University Press, 1967, 464-479.
- LIN, Patrick; JENKINS, Ryan y ABNEY, Keith (eds.), *Robot Ethics 2.0. From autonomous Cars to Artificial Intelligence*. Oxford: Oxford University Press, 2017.
- McCARTHY, John; MINSKY, Marvin; ROCHESTER, Nathaniel y SHANNON, Claude, *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 1955, p. 1. Disponible en línea: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> (última consulta 2 de agosto de 2021).
- McCARTHY, John, *Defending AI Research. A collection of Essays and Reviews*. Stanford: CSLI Publications, 1996.
- MINSKY, Marvin, *Computation. Finite and Infinite machines*. Nueva Jersey: Prentice Hall, 1967.
- , *The Society of Minds*. New York: Simon & Schuster, 1986.
- POST, Emil, “On computable numbers, with an Application to the *Entscheidungsproblem*. A critique”. En: TURING, Alan. *The essential Turing*, 97-101.
- PRADO, Belen, “Ambivalence in Machine Intelligence: The epistemological Roots of the Turing Machine”, *Signos Filosóficos*, Vol. 23, No. 45, 2021, 54-73.
- ROBINSON, William S., “Philosophical Challenges”, *The Cambridge Companion of Artificial Intelligence*, Edited by Keith Frankish; William M. Ramsey, Cambridge: Cambridge University Press, 2014, 64-88.
- SEARLE, John, “Minds, Brains, and Programs”, *The Behavioral and Brain Sciences*, Vol. 3, 1980, 417-457.

TURING, Alan, *The essential Turing. Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus The Secret of Enigma*. Edited by Jack Copeland. Oxford: Clarendon Press/Oxford University Press, 2004.

ALGORITMOS, BIG DATA E INTELIGENCIA ARTIFICIAL: ¿UN NIHILISMO ANUNCIADO?

*ALGORITHMS, BIG DATA AND ARTIFICIAL
INTELLIGENCE: A FORETOLD NIHILISM?*

DUNE VALLE JIMÉNEZ

Doctor en Filosofía
Docente Tiempo Completo
Universidad Sergio Arboleda
Bogotá/Colombia
dunevalle@hotmail.com
ORCID: 0000-0001-7653-9182

Recibido: 1/06/2021
Revisado: 27/08/2021
Aceptado: 6/09/2021

DIEGO GARCÍA RAMÍREZ

Doctor en Comunicación
Docente Tiempo Completo
Universidad del Rosario
Bogotá/Colombia
garcia.ramirez.diego@gmail.com
ORCID: 0000-0002-6173-8047

Resumen: Este trabajo analiza los vínculos entre la técnica de la era digital y el nihilismo. Ante el anuncio de un camino certero hacia la perfección y la plenitud, a través del desarrollo exponencial de las tecnologías de la información y la comunicación, contraponemos las filosofías de Nietzsche, Jünger y Heidegger. El advenimiento de una sociedad de la certeza, en la que el cálculo y la planificación controla las esferas existenciales individuales y colectivas, en diálogo con estos autores resuena a nihilismo tecnológico. La inteligencia artificial y sus aplicaciones, entendida como el gran relato del Siglo XXI, genera un nihilismo pasivo e incompleto, que reemplaza la teleología divina con una profecía tecnológica, en la que la inteligencia artificial, comprendida como valor absoluto, nos conduce hacia el mejor de los mundos posibles.

Palabras clave: Nietzsche, Heidegger, Jünger, nihilismo, tecnología, inteligencia artificial, algoritmos, big data.

Abstract: This paper analyzes the links between the technique in the digital era and nihilism. We challenge the announcement of a road to perfection and fulfilment through the exponential development of information and communication technologies, from the philosophies of Nietzsche, Jünger and Heidegger. The future of a society of certainty, in which computation and planning shape the individual and collective existential spheres, resonates, in dialogue with these authors, as a technological nihilism. Artificial intelligence and its applications, understood as the great story of the 21st century generates a passive and incomplete nihilism, which replaces divine teleology with a technological prophecy, in which artificial intelligence, understood as an absolute value, leads us to the best of all possible worlds.

Keywords: Nietzsche, Heidegger, Jünger, nihilism, technology, artificial intelligence, algorithms, big data.

INTRODUCCIÓN

Sin lugar a dudas, el siglo XXI puede ser caracterizado por la omnipresencia y omnipotencia que las tecnologías de la información y la comunicación han adquirido en la sociedad, lo que se manifiesta de forma indiscutible y sin contrapeso en todas las esferas y quehaceres de la vida humana en un presente en el cual toman mayor protagonismo, los algoritmos, el *big data* y la inteligencia artificial, dando lugar así a una especie de círculo de poder u oráculo, que gracias a la correlación activa de sus componentes tiene la llave de acceso para descifrar el futuro, decirnos realmente quiénes somos, qué debemos hacer y qué debemos esperar. Adicionalmente, llevarnos hacia al reino de la certeza, la perfección; incluso hacia la inmortalidad. Todo esto, gracias al desarrollo de las tecnologías bajo el dominio de la tecnociencia.

Precisamente, dentro de este escenario de promesas, plenitud y avance tecnológico, ciertos pensadores del presente siglo han planteado una postura crítica respecto a estos desarrollos. Especialmente, en lo relacionado con las consecuencias humanas y sociales que enfrentamos o podríamos enfrentar si no ponemos nuestra atención en lo que ocultan o disfrazan dichas promesas o discursos. Llama la atención, sin embargo, la presencia de cierta coincidencia en la caracterización de la contemporaneidad, comprendida como nihilista. Algunos pensadores se *aventuran* a enjuiciar, o, mejor dicho, diagnosticar como nihilista el mundo actual dominado por la técnica. Algunos de ellos, relacionan directamente el nihilismo con la importancia, la presencia y la influencia que han adquirido los algoritmos, el *big data* y la inteligencia artificial en nuestra vida cotidiana. Influencia que para gran parte del planeta es todavía apenas perceptible.

Teniendo en consideración estos planteamientos, es que analizaremos qué tienen que ver esta serie de diagnósticos que proponen una lectura nihilista de la era digital, con aquella otra visión nacida a finales del siglo XIX presentada

por Friedrich Nietzsche. Entendida esta última como una crítica a la cultura que pone en el centro de la mirada al nihilismo. Propuesta recogida por Ernst Jünger y profundizada por Martin Heidegger, en consonancia y, al mismo tiempo, disonancia con el pensamiento del filósofo del eterno retorno. Aquellos precisaron y focalizaron parte de su reflexión en el carácter nihilista del fenómeno técnico. Preguntaremos, asimismo, en qué sentido dichas lecturas respecto a la era digital pueden ser denominadas como nihilistas y de qué hablamos cuando mentamos la palabra nihilismo. Algunos autores señalan que Nietzsche no fue capaz de captar con total claridad –al menos en sus textos–, la relación indisoluble entre tecnología y nihilismo; pero a la vez caracterizó como nunca antes el fenómeno del nihilismo y su manifestación en el mundo contemporáneo. Esta lectura nos permitirá entrelazar las interpretaciones y usos de las tecnologías en la actualidad con el fenómeno del nihilismo.

1. NIHILISMO EN LA ERA DIGITAL

No es un misterio que en las últimas décadas la humanidad ha vivido una serie de transformaciones radicales con carácter planetario, esto merced al desarrollo exponencial de las tecnologías de la información y la comunicación en el contexto de la globalización. Dichas tecnologías han configurado un escenario marcado a fuego por los beneficios, ventajas, comodidades y soluciones que nos brinda el aparato instrumental propio de la tecnociencia, entre los cuales encontramos *smartphones*, internet, redes sociales, aplicaciones de la más diversa índole, con la inteligencia artificial, como fundamento ordenador que posibilita estos dispositivos. Así, un discurso tecnológico, con claros tintes ideológicos, se impone sobre toda otra consideración, en lo que parece contravenir una de las tesis fundamentales de la Posmodernidad, entendida como el fin de los metarrelatos o el fin de las ideologías, pues esta ideología técnica tiene todas las características propias de ese tipo de discursos¹.

Al mismo tiempo, esta interpretación ha venido a llenar ese vacío que en la historia del pensamiento ha dejado la muerte de Dios, como nueva interpretación impuesta por la voluntad de poder técnico científica que promete toda suerte de respuestas y soluciones salvíficas, como una especie de “mesianismo técnico”²;

1 Cfr., GARCÍA, D & VALLE, D. (2020). Los impactos de la ideología técnica y la cultura algorítmica en la sociedad: una aproximación crítica. *Revista de Estudios Sociales*, 2020 (71), pp. 18-19.

2 SADIN, E., *La siliconización del mundo*. Margarita Martínez (trad.). Buenos Aires: Caja Negra, 2018, p.100.

pero no es necesario esperar la vida eterna, la segunda venida del Mesías o el más allá para alcanzar su cumplimiento. Todo se puede conseguir aquí y ahora, incluso la ansiada inmortalidad³. En este escenario, los algoritmos, el *big data* y la inteligencia artificial, son los poseedores de los códigos que permitirán alcanzar una sociedad de la certeza, perfección, moldeada por la planificación y cálculo algorítmico de la realidad y el mundo.

Precisamente, la tarea propuesta es pensar en estos alardes de la tecnociencia y sus consecuencias no solamente desde una óptica sociológica, política o económica, sino desde una perspectiva filosófica. Este escenario había sido avizorado hace ya tiempo por los heraldos del nihilismo: Nietzsche, Jünger y Heidegger. Por ende, esta interpretación del fenómeno técnico tiene que ir más allá de una postura meramente ideológica, como acostumbramos a escuchar en la mayoría de discursos políticos, académicos y de los medios de comunicación. Parece ser que esos espacios han sido completamente colonizados por una visión de la tecnología que impide ver otras cuestiones profundamente problemáticas insertas en su creación, modelo y aplicación.

Como hemos señalado anteriormente, algunos pensadores del presente siglo han destacado el carácter nihilista de la era digital, en lo que, a buenas y a primeras, podría ser solo una proyección o continuidad del pensamiento de los anunciadores del nihilismo. Dice Volpi,

entre los principales factores del avance del nihilismo, o bien entre las causas que han contribuido mayormente a la consunción de los valores y los ordenamientos tradicionales, no pocos analistas ponen hoy a la técnica. Juntamente con la difusa conciencia de que ella se ha convertido en uno de los elementos dominantes de nuestra época en el factor principal de globalización, ha emergido también la preocupación en torno a su verdadera naturaleza, su desarrollo y su posibilidad de controlar su dinámica⁴.

En este mismo sentido se expresa Han cuando resalta el valor adquirido por el *big data* y los algoritmos en el mundo contemporáneo. Bajo este prisma la existencia se transforma en *homo digitalis*, quien a merced de la numeración y calculabilidad de sus conductas completamente registradas y rastreables en la red termina sumergido en una visión o interpretación del mundo completamente sustentada en los datos; a saber, una barbarie de los datos con su propia filosofía emergente :el dataísmo. Barbarie que deviene en nihilismo, pues el dataísmo implica necesariamente renunciar al sentido y al significado; vacío que se pretende

3 Cfr., DIÉGUEZ, A., *Transmumanismo*. Barcelona: Herder. 2017, p. 15.

4 VOLPI, F., *El nihilismo*. Buenos Aires: Biblos, 2005, p. 145.

llenar con números y cálculos⁵. En otras palabras, en esta forma de nihilismo, la ausencia de sentido se disfraza y oculta a través de la cortina del cálculo maquinal.

Otro pensador contemporáneo que se ha hecho reconocido por su crítica a ciertos desarrollos surgidos en la era digital es Eric Sadin, quien también hace hincapié en el carácter nihilista del presente, en la denominada Cuarta Revolución Industrial, cuya característica distintiva es la inteligencia artificial. Aquello que está en juego en estas interpretaciones, según él, es:

una *visión del mundo* basada en el postulado tecno-ideológico de que hay una deficiencia humana fundamental que va a ser salvada por los poderes afectados a la inteligencia artificial, que aumentan y varían sin descanso. Y esto la inteligencia artificial representa la mayor potencia política de la historia...dotada en todo momento de la intuición de la verdad y que orienta el curso de la nuestras acciones individuales y colectivas hacia el mejor de los mundos posibles. En los hechos se trata de un nihilismo tecnológico o de un antihumanismo radical⁶.

Lo que quiere decir el pensador francés, es que la inteligencia artificial como postulado tecnoideológico, socavaría los fundamentos del humanismo occidental, cuyos principios básicos son la soberanía, la autodeterminación, el libre albedrío; es decir, el poder de decidir cada persona y sociedad su propio destino, dado que cada vez más nuestras decisiones serían determinadas por los designios de las nuevas formas de poder expresados en las tecnologías y sus aplicaciones.

Respecto al carácter ideológico de estas representaciones y a su carácter nihilista, Severino afirma que: “la técnica europea es nihilismo no en cuanto es usada “mal” sino en cuanto es técnica, y las ideologías que llevan a usarla “bien” o “mal” tienen en común con la llamada naturaleza neutral de la técnica el rasgo decisivo del nihilismo...el nihilismo es la morada de Occidente: su estructura”⁷.

Ahora bien, esta idea respecto a la relación entre nihilismo y tecnología no es una interpretación novedosa u original, sino que más bien está enquistada en la filosofía contemporánea con Jünger y Heidegger en primera línea. No obstante, el desarrollo tecnológico alcanzado en el nuevo siglo da nuevos bríos a la interpretación y nos permite ver de forma más clara todavía cómo se ajustan y relacionan las intuiciones adelantadas por estos filósofos con la situación del ser humano en el presente. Mostrando así la conveniencia y necesidad de retomar dichos planteamientos para pensar el mundo en el siglo XXI.

5 Cfr., HAN, B.CH., *Psicopolítica*. Barcelona: Herder, 2014, p. 46- 48.

6 SADIN, E., *La inteligencia artificial o el desafío del siglo*. M. Martínez (trad.). Buenos Aires: Caja Negra, 2020, p. 34.

7 SEVERINO, E., *Esencia del nihilismo*. E. ÁLVAREZ (trad.). Madrid: Taurus, 1991, pp. 27-28.

La emergencia de nuevas y novedosas formas de comprender el mundo, que subrayan Sadin y Han, hacen palpable todavía más esta relación, adelantada y proyectada con una clarividencia abismal, tanto por Jünger como por Heidegger; esto es, la relación entre nihilismo y tecnología⁸. Lo cual nos entrega valiosas oportunidades para analizar y pensar el fenómeno técnico, ejemplificado en el caso de la inteligencia artificial, algoritmos y *big data*, como claves para descifrar aquello que muchas veces se oculta bajo el ropaje de las maravillas y utopías tecnológicas.

No obstante, hablar de nihilismo tiene de por sí una serie de inconvenientes, confusiones y malos entendidos que no nos proponemos resolver en profundidad, pero sí rescatar una visión particular para desde ahí interpretar el nihilismo tecnológico que nos convoca. Intentamos así evitar “abusar del título «nihilismo» como una ruidosa consigna carente de contenido que tiene a la vez la función de amedrentar, de descalificar y de ocultar al mismo que comete el abuso ocultado su propia falta de pensamiento”⁹. Lo que intentaremos subsanar aclarando en qué sentido podemos hablar de nihilismo en relación con la tecnología y las aplicaciones que de ella se hacen en el campo de las tecnologías de la información y la comunicación. Ciertamente, toda revisión o interpretación filosófica sobre el nihilismo debe tener como parada obligatoria la filosofía de Nietzsche, quien auguró que la historia de la humanidad en los próximos siglos estaría signada por la presencia de este huésped incómodo e inquietante.

2. NIETZSCHE Y EL NIHILISMO

Ampliamente reconocida y aceptada es la influencia que el nihilismo ruso ejerció sobre la concepción filosófica del nihilismo presentada por Nietzsche. Si bien muchas veces se ha interpretado de forma confusa el orden cronológico de dicha recepción. En este sentido, Elisabeth Kuhn¹⁰ destaca en su investigación la lectura que Nietzsche hace tempranamente de la novela *Padres e hijos* de Turgeniev, que habría llegado a sus manos, en la traducción francesa aproximadamente

8 Valga la pena aclarar que la reflexión filosófica respecto a la tecnología no se limita a los autores tratados en este artículo, pues dicha preocupación ocupó un lugar destacado en el quehacer filosófico del siglo pasado de la mano de autores como Jacques Ellul, Günter Anders, Herbert Marcuse, Gilbert Simondon, Hans Jonas, José Ortega y Gasset, entre otros. Sin embargo, la elección de Heidegger y Jünger se hace teniendo en cuenta el lugar destacado que ocupa en estos autores la relación entre técnica y nihilismo.

9 HEIDEGGER, M., *Nietzsche II*. J. L. VERMAL (trad.). Barcelona: Destino, 2001, p. 41.

10 Cfr. KUHN, E., *Friedrich Nietzsches Philosophie des europäischen Nihilismus*, Berlin-New York: de Gruyter, 1992, pp. 33-37.

en los años 1876 a 1880 y que determinaría de forma decisiva sus primeras formulaciones sobre el nihilismo. Formulaciones, que una vez complementadas con la lectura de Dostoievsky, concretamente de *Los demonios*, realizada probablemente a finales de 1887, se van modificando y precisando, dado que una vez descubierta su obra, Turgeniev pasa a segundo plano e incluso es sometido a crítica. En este contexto, se ha destacado también la influencia que tuvo la figura de Paul Bourget, especialmente de sus textos publicados en 1883 y 1886, del cual se destaca *Essais de psychologie contemporaine* donde Bourget dedica un capítulo entero a la figura de Turgeniev, a lo que Nietzsche le dedicó especial atención, subrayándolo profusamente ¹¹. Este recorrido precisa de alguna manera el camino que recorrió Nietzsche para dar con la elaboración filosófica del nihilismo, destacándose la figura de Turgeniev como primer aire que fue capaz de insuflar la interpretación nietzscheana.

Ahora bien, mientras nadie pone en duda la paternidad filosófica del nihilismo, algo distinto pasa cuando hablamos de sus raíces en el ámbito de la literatura o política, pues Turgeniev reclama para sí la paternidad sobre el nihilismo, lo cual es del todo discutible. Turgeniev en *Padres e hijos* se refiere concretamente al concepto de nihilismo y su significado, que en este contexto se relaciona directamente con uno de los personajes, Bazarov, como fiel reflejo del nihilista y su posición en la pugna entre la vieja y la nueva idea de mundo, cuyo antagonismo queda graficado en la disputa entre padres e hijos. Los padres, seguían y defendían los valores y principios antiguos inspirada en los ideales humanistas, mientras que los hijos, representados en la figura de Bazarov, ya no podían adoptar una fe o creencia más que la que brindaba la nueva visión o interpretación positivista y racionalista del mundo. Esta novela logró reflejar fehacientemente ciertas tendencias operantes en la sociedad y cultura rusa del siglo XIX, tendencias que de alguna forma mostraban cierto malestar e inquietud en la entera cultura europea. A partir de ahí ese puede justificar el impacto e importancia que el concepto de nihilismo adquirió en la Rusia zarista y en el resto de Europa. Así, logró posicionarse como sinónimo para denominar una crítica social¹². Ciertamente, la fuente de la que bebió Nietzsche para construir su concepto de nihilismo fue el escenario literario, psicológico y político abierto por el nihilismo ruso, en este caso, en la figura de Turgeniev, como germen del fenómeno teorizado por Nietzsche. En este contexto, “hay que observar que las primeras referencias a los «nihilistas»

11 Cfr., GENTILI, C., Nihilismo, nihilistas. Sobre una posible fuente de Nietzsche”, Estudios Nietzsche, 19, 2019, pp. 67-69.

12 Cfr., VOLPI, F., *El nihilismo*, op. cit., pp.19-21

presentes en la obra de Nietzsche (sobre todo en el *Nachlass*) no permiten distinguir con precisión el ámbito político y el literario”¹³.

Como habíamos mencionado, es en la filosofía de Nietzsche donde se presenta el asunto del nihilismo de forma más estructurada y se transforma en objeto de una reflexión filosófica elaborada, especialmente en los *Fragmentos póstumos*. “No es pues, –dice Volpi–, una exageración considerar a Nietzsche como el máximo profeta y teórico del nihilismo, como el que diagnostica a tiempo la “enfermedad” que afligirá al siglo y para la cual ofrece una terapia”¹⁴.

Esta profecía, con tintes de clarividencia, viene a narrar la historia futura de la cultura occidental y de la humanidad en su conjunto, “lo que cuento es la historia de los próximos dos siglos. Describe lo que viene, lo que no puede ser de otra manera: *la ascensión del nihilismo*. Esta historia ya se puede contar ahora: pues la necesidad misma está aquí trabajando. Este futuro ya habla en cien signos, este destino se anuncia por todas partes, para esta música del futuro están aguzado todos los oídos.”¹⁵.

Nietzsche trata el problema del nihilismo de distintas maneras y con distintas denominaciones a lo largo de su pensamiento¹⁶, reflejando así la riqueza estilística y conceptual de su pensamiento: lleno de símbolos y metáforas muchas veces de difícil acceso o comprensión. Quizá por ello sus conceptos tienen un sentido plural y abierto, huyendo así del ansia de sistema y orden tan propio de la Modernidad y el Positivismo. Teniendo esto en consideración, es que recogeremos aquellas lecturas sobre el nihilismo que, desde nuestra perspectiva, se mantienen aún vigentes, como sería el caso del nihilismo pasivo e incompleto en el que nos encontraríamos bajo el dominio de la interpretación tecnocientífica.

En los denominados *Fragmentos póstumos* el filósofo alemán vuelve de forma constante sobre el tema del nihilismo, intentando en cada paso retornar sobre el fenómeno para complementarlo, acotarlo y volverlo a presentar. En una de las definiciones más reconocidas caracteriza este fenómeno como un estado Normal,

13 GENTILI, C., “Nihilismo, nihilistas. Sobre una posible fuente de Nietzsche”, *Estudios Nietzsche*, 19, 2019, p. 67.

14 VOLPI, F., *El nihilismo*, op. cit., p. 47.

15 NIETZSCHE, F., *Fragmentos póstumos IV*. J.L. V (trad.). Madrid: Tecnos, 2008, p. 489.

16 Es importante hacer notar que “el término «nihilismo» aparece por primera vez de una manera fugaz en una carta a Köselitz del 13 de marzo de 1881, para después recuperarlo de una manera fugaz en un fragmento del verano-otoño de 1882, sin embargo comienza a aparecer insistentemente solo a partir de 1885, y el término «nihilista» aparece por primera vez en un fragmento de 1880, pero tiene un uso recurrente solo a partir de 1884; la reflexión sobre la nada reviste por el contrario un papel de primer orden en el aforismo 125 (El hombre loco) de *La gaya ciencia*. Después de haber anunciado la muerte de Dios – “¿A dónde ha ido Dios?”, GENTILI, C., “Nihilismo, nihilistas. Sobre una posible fuente de Nietzsche”, *Estudios Nietzsche*, 19, 2019, p. 77.

para agregar a continuación “Nihilismo: falta la meta; falta la respuesta al « ¿por qué?»¿ qué significa nihilismo? –*que los valores supremo se devalorizan*”¹⁷. En otras palabras: las categorías con las que la tradición occidental daba orden y estabilidad a la realidad– tales como fin, unidad, ser, verdad– se han retirado. Esto hace que el mundo parezca como carente de valor y legitimidad. De ahí que el significado de Dios muerto, no puede ser comprendido primeramente como un gesto de ateísmo, sino como un ataque al lugar y los valores que rodean a Dios, como imagen y símbolo de la cultura occidental, deudor de la división platónica de los mundos. “El sentimiento de la *ausencia de valor* se llegó a tener cuando se comprendió que no es lícito interpretar el carácter global de la existencia ni con el concepto de «*fin*», ni con el concepto de «*unidad*», ni con el concepto de “verdad”¹⁸, dice el autor de *Más allá del bien y del mal*.

Ahora bien, esta desvalorización de los valores puede despertar formas de nihilismo distintas y opuestas en relación a la fuerza y a la voluntad. Es aquí donde surge la distinción de nihilismo activo y pasivo. Sostiene Nietzsche,

Nihilismo como *descenso y retroceso del poder del espíritu: EL NIHILISMO PASIVO*: como un signo de debilidad: la fuerza del espíritu puede estar cansada, *agotada*, de manera tal que las metas y los valores existentes *hasta el momento* son inadecuados y no encuentran ya crédito...que todo lo que reconforta, cura, anestesia, pasa al primer plano, bajo diferentes *disfraces*, con carácter religioso, o moral o político o estético, etc.¹⁹.

Como podemos ver, esta versión del nihilismo habla de agotamiento, debilidad y cansancio de la fuerza, de la vitalidad del espíritu, pues los valores ya no tienen esa fuerza operativa que brindaba respuestas, consuelo y esperanzas; a saber, han perdido el lugar que les correspondía.

Sin embargo, por esta misma debilidad o agotamiento es que bajo el nihilismo pasivo no se produce una transvaloración, sino que el lugar que ocupaba Dios y toda su estructura sigue ahí: solo se sustituye la figura metafísica de Dios y otro valor viene a tomar su lugar. Sigue siendo necesaria –en esta fase del nihilismo– esa figura metafísica que permitía darle una legalidad y orden al devenir, basado en valores absolutos. Esto significa que : “ otros valores tienden a ocupar el puesto del “Dios muerto”, otros valores a lo que se confiere, como a los anteriores un carácter absoluto...Esta autoridad que supera a Dios, que sustituye al Dios cristiano, puede ser ahora el instinto del rebaño, la ciencia positivista”²⁰.

17 NIETZSCHE, F., *Fragmentos póstumos*, op. cit., p. 241.

18 *Ibíd.*, p. 396.

19 *Ibíd.*, p. 242.

20 SANCHEZ MECA, D., *El nihilismo*. Madrid: Síntesis, 2004, p. 110.

Ahora bien, entre las causas del nihilismo en cuanto estado psicológico, es necesario hacer notar, como hace Nietzsche, que “cuando se ha supuesto una *totalidad*, una *sistematización*, incluso una *organización*, en todo acontecer y bajo todo acontecer...En el fondo, el ser humano ha perdido la creencia en su propio valor, sí a través de él no actúa una totalidad infinitamente valiosa: o sea, él ha concebido una totalidad semejante para *para poder creer en su propio valor*”²¹. Y ciertamente, tal como afirma Severino: “Dios y la técnica moderna son las dos fundamentales expresiones del nihilismo metafísico”²².

Nietzsche establece ciertos parámetros respecto al nihilismo, tales como totalidad, sistematización y organización de todo y en todo momento como sus causas. Lo que a simple vista se puede comprobar en esta predicción y dirección algorítmica de la existencia que cumplimenta la visión positivista del mundo y muestra su versión más radical, estableciendo como valor absoluto la visión tecnocientífica e ideológica de la tecnología capaz de brindar soluciones y respuestas de todo tipo. De esta manera, “la máquina se establece como paradigma del funcionamiento de una sociedad perfecta y se impone en los diferentes ámbitos de la vida social: en el mundo del trabajo, la educación, del ejército, las instituciones políticas y el Estado”²³.

2.1. JÜNGER: EL ORDEN COMO TOPOS DEL NIHILISMO

Como hemos adelantado, la formulación nietzscheana nos permitirá relacionar el nihilismo con los desarrollos e interpretaciones emergentes en la era digital. Sin embargo, esta relación no queda establecida del todo en la filosofía de Nietzsche. Según Heidegger²⁴, Jünger fue el primer pensador capaz de visualizar con claridad la relación entre técnica y nihilismo. Concretamente, Jünger en su escrito *Sobre el dolor* de 1934, después de un largo recorrido auscultando el fenómeno de la técnica y su relación con el dolor, propone una interpretación de la técnica como desencadenante del nihilismo, pues señala que “nos encontramos en una fase última del nihilismo, es una fase ciertamente notable, que se señala por el hecho de que unos *órdenes* nuevos han ocupado ya unas posiciones muy

21 NIETZSCHE, F., *Fragmentos póstumos*, op. cit., p. 395

22 SEVERINO, E., *Esencia del nihilismo*. E. ÁLVAREZ (trad.). Madrid: Taurus, 1991, p. 208.

23 OCAÑA, *Más allá del nihilismo*, Murcia: Universidad de Murcia, 1993, p. 91.

24 Sostiene Heidegger: “lo que Jünger ve más claramente que Nietzsche es esto que Nietzsche en su tiempo todavía no podía ver en estos fenómenos porque ellos mismos todavía se encontraban ocultos en la realidad. En conjunto son los fenómenos de la técnica como el modo fundamental de la organización y aseguramiento de lo real como voluntad de poder. (HEIDEGGER, M., *Acerca de Ernst Jünger*. D. PICOTTI (trad.). Buenos Aires: El hilo de Ariadna, 2013, p. 318).

avanzadas, pero los valores correspondiente a esos ordenes no se han hecho visibles”²⁵. Estos nuevos órdenes de los que habla Jünger se relacionan principalmente con el fenómeno técnico que parece haber adquirido el control total de la sociedad de entre guerras con su manifestación más nefasta en el desarrollo del poderío bélico militar y sus consecuencias. Asimismo, ha mutado la comprensión y caracterización de la vida humana. Bajo estos nuevos órdenes, la realidad se enfrenta necesariamente con el disfraz de la técnica, que revela al mismo tiempo la cada día más notoria objetivación y cosificación de la vida humana. Esto lo expresa Jünger: “*la técnica es nuestro uniforme. De todos modos, aún nos encontramos demasiado hundidos en este proceso como para poder abarcarlo en toda su extensión con la mirada... Eso ocurre más cuanto que el carácter de confort de nuestra técnica está fusionándose de un modo cada vez más inequívoco con carácter instrumental del poder*”²⁶.

Jünger presenta así ciertas premisas respecto al fenómeno técnico de aquella época que serían fácilmente verificables y contrastables con el papel que cumplen las tecnologías en el presente siglo. Se constata, asimismo, una ausencia de reflexión abismante respecto a estos desarrollos, tanto en el pasado como en el presente. Prevalciendo una visión ideológica y carente de crítica, pues la sociedad en su conjunto parece obnubilada y enceguecida por los beneficios y el confort que nos entrega la técnica, sin tener en cuenta que ese confort y comodidad está sometida y al servicio de las nuevas y originales formas de poder surgidas en los últimos tiempos. Volpi resume las consecuencias de esta interpretación como sigue:

cuando la realidad es plasmada y transformada por la técnica sin que las ideas, las personas y las instituciones se adecuen con la misma rapidez, cuando la disciplina, la capacidad de organización, el potencial energético crecen sin un igual crecimiento de nueva sustancia, entonces la técnica produce nihilismo. Surge en este punto el problema de la actitud adecuada que el hombre debe asumir cuando el vórtice de la aceleración tecnológica parece absorberlo²⁷.

Jünger vendría a ser, entonces, el autor bisagra que relaciona técnica y nihilismo. Es más: la técnica se transformaría en un factor fundamental y desencadenante de nihilismo, esto nos permite ver con mayor claridad de qué manera esta relación se va plasmando y materializando, de forma casi indiscutible, en la historia de la filosofía del S.XX.

25 JÜNGER, E., *Sobre el dolor*. A. Sánchez Pacual (trad.). Barcelona: Tusquets, 1995, p. 48.

26 JÜNGER, E., *El trabajador*. A. Sánchez Pascual (trad.). Barcelona: Tusquets, 1990, p. 59.

27 VOLPI, F., *El nihilismo*. op. cit., p. 109.

Ahora bien, una perspectiva importante que recoge Jünger respecto al nihilismo tiene que ver con el orden, pues se acostumbra a relacionar el nihilismo y sus consecuencias con el caos, la enfermedad y lo malo. Sin embargo, Jünger reconoce en el orden un paradigma, una característica del nihilismo al interior del mundo tecnológico, que llega a adaptarse completamente en diversos sistemas de orden. Incluso, como un topos favorable para su desarrollo, pues lo puede transformar de manera que sirva a sus propios fines. Así, “el nihilismo puede armonizar con ordenamiento mundiales extendidos...al nihilismo no solo le gusta el orden, sino que pertenece a su estilo”²⁸. De esto se puede concluir, que una de las características determinantes, como suelo fecundo para el nihilismo, sería un sistema de orden o alguna clase de ordenamiento social, que permita alinear y estandarizar una serie de conductas e interpretaciones.

No obstante, existe un tipo de orden que, a juicio del pensador de Heidelberg, sobresale y se torna determinante: el orden técnico, ya que es el que permite todo tipo de subordinación y determinación arbitraria, lo que transforma la figura de la subjetividad del hombre moderno, tornándolo en trabajador: “la técnica es el instrumental, es la proyección de un modo especial de vida, para designar el cual es *trabajo* la expresión más sencilla...una identidad de trabajo y ser logra garantizar una seguridad nueva, una estabilidad nueva”²⁹. Es decir, esta figura del trabajador permite un ordenamiento seguro y estable sobre el cual construir e interpretar la realidad, gracias a la técnica, imponiéndose así una idea de existencia. Dicho orden es el suelo propicio para el desarrollo del nihilismo, con la uniformización de la técnica y el trabajo. Ciertamente, como concluye Ocaña:

la racionalidad instrumental que penetra progresivamente en los tejidos del mundo de la vida introduciendo en el seno de la misma sistemas de planificación social –del tipo burocrático, político, económico, etcétera –cuyo control escapa a las decisiones individuales y a las relaciones contractuales consensuadas, en virtud de un funcionamiento impersonal propio del automatismo tecnológico. En este sentido, pues la introducción de la “figura del trabajador” en el mundo de la vida debe entenderse como la gestación de creciente poder de movilización (i.e, organización y disposición instrumental) del mundo por parte de la técnica³⁰.

28 JÜNGER, E., *Acerca del nihilismo*. J.L. Molinuevo (trad.) Barcelona: Paidós, 2008, p. 30.

29 JÜNGER, E., *El trabajador*, op. cit., pp. 90-91.

30 OCAÑA, E., *Más allá del nihilismo*. Murcia: Universidad de Murcia, 1993, p. 67.

2.2. HEIDEGGER: TÉCNICA Y NIHILISMO EN EL FINAL DE LA METAFÍSICA

Siguiendo en la línea de argumentación planteada por Jünger, Heidegger retoma explícitamente en *Superación de la metafísica* esta figura metafísica del trabajador, para desarrollar su crítica a la civilización dominada por la interpretación técnico-científica. Este periodo es conceptualizado por el pensador alemán como final de la metafísica, pero no como un acabamiento en que termina todo vestigio de esta, sino como cumplimiento de sus máximas posibilidades; en este caso, en la emergencia de la figura del trabajador. Por eso, afirma Heidegger que “el trabajo...accede ahora al rango metafísico de la objetualización incondicionada de todo lo presente que despliega su esencia en la voluntad de la voluntad”³¹. En este contexto, se agrega una precisión respecto a la figura del hombre, pues bajo el dominio de la técnica es considerado como animal de trabajo, como materia prima; todavía más, deviene en la materia prima más importante, con la finalidad de producir de forma técnica la posibilidad incondicionada de toda producción³².

En otras palabras, el trabajador es la imagen o determinación de la existencia en la era técnica, en la cual se expresa la metafísica en su consumación como voluntad incondicionada de dominio del ente y de la naturaleza, como destino de la metafísica de la subjetividad. Además, solo en la voluntad de la voluntad es donde la técnica y la ausencia de toda reflexión dominan sin contrapeso, perteneciéndose mutuamente como lo mismo. Es más: la voluntad de voluntad “impone como formas fundamentales de su aparecer el cálculo y la organización de todo, pero esto solo para asegurarse a sí misma...la forma fundamental de este aparecer, en la que la voluntad de voluntad se instala y calcula... se la puede llamar con una palabra la “técnica”³³.

La realidad deviene así en la uniformidad de la medida y el cálculo y dentro de tal movimiento el hombre mismo es sometido a esta uniformidad, donde lo que predomina es una indiferencia en todas partes y en todo lugar, merced a cierto tipo de organización que responde a los dictados del *principio de rendimiento*, que aparece como rasgo fundamental y determinante de la realidad³⁴. Sin embargo, como dice el pensador alemán, esta jerarquización que establece la técnica, “tiene como fundamento determinante la ausencia de jerarquía, porque en todas

31 HEIDEGGER, M., *Conferencias y artículos*. E. Barjau (trad.). Barcelona: Ediciones del Serbal, 2001, p. 52.

32 Cfr., *Ibid.*, p. 70.

33 *Ibid.*, p. 58.

34 Cfr. *Ibid.*, p. 71.

partes la única meta del rendimiento es el vacío uniforme de la usura de todo trabajo, dirigido al aseguramiento del ordenar”³⁵.

Como podemos ver, esta lectura de la figura del trabajador permite a Heidegger desarrollar su crítica a la tradición metafísica, valiéndose al mismo tiempo de la filosofía de Nietzsche para plantear su visión respecto a la civilización técnico-científica como manifestación de la destinación técnica del ser mismo. Sin embargo, cuando se habla de la contraposición Nietzsche-Heidegger, en términos de si este último hace justicia a la filosofía de Nietzsche, especialmente en su lectura de los conceptos de superhombre, voluntad de poder y eterno retorno, nos deberíamos quedar con lo que propone el mismo Heidegger: “lo que es decisivo... es oír al propio Nietzsche, preguntar con él, a través de él, y así al mismo tiempo contra él, pero a favor de la causa más interna, única y común de la filosofía occidental”³⁶. Deberíamos, entonces, tomar en consideración la crítica que, a través de Nietzsche, se plantea a la interpretación técnica de la realidad; preguntarnos si esta interpretación ilumina el fenómeno técnico y si nos muestra indicaciones para habitar en él, más que solamente cuestionar si la interpretación de Heidegger se ajusta filológicamente o filosóficamente a la filosofía de Nietzsche.

Teniendo en cuenta esto, es que se debe aclarar desde un principio que en la lectura heideggeriana la voluntad de voluntad viene a ser equivalente a la voluntad de poder:

su modo de existir no es otro que el constante retorno sobre ella misma. Este redoblamiento sobre sí de la voluntad (*Wille zum Willem*) indica su nihilismo, pues no se persigue ninguna meta, sino que se desarrolla sobre ella misma hasta la irrealidad total. Voluntad de voluntad es una expresión que ya ella misma muestra una circularidad como elemento dominante: la voluntad quiere lo que ella es y es ese querer mismo...el super hombre es la condición del hombre que es en el mundo una vez que se ha hecho patente su esencia como voluntad de poder³⁷.

Así, pues, la voluntad de poder y el superhombre se transforman en los conceptos clave para interpretar el mundo de la técnica: Heidegger ve en ellos la representación de los absolutos que subyacen en el mundo tecnológico y en los que se manifestaría el afán de dominio y manipulación del pensar calculante propio de la técnica, “esta voluntad de voluntad es el estado de consciencia supremo e incondicionado del autoaseguramiento a través del cálculo”³⁸.

35 *Ibid.*, p. 71.

36 HEIDEGGER, H., *Nietzsche I.* J.L. Verma (trad.). Barcelona: Destino, 2000, p. 36.

37 SÁNCHEZ MECA, D., *La experiencia dionisiaca del mundo*. Madrid: Tecnos, 2006, p. 365.

38 HEIDEGGER, M., *Conferencias y artículos, op, cit.*, p. 65.

Heidegger profundiza en esta interpretación cuando afirma: “el nihilismo, el hecho de “que los valores supremos se desvalorizan”, ha sido superado. Esa humanidad que quiere su propio ser hombre como voluntad del poder...el hombre cuya esencia es querida a partir de la voluntad de poder es el transhombre...se encuentra situado a partir del ser de lo ente, ante la tarea de asumir el dominio de la tierra”³⁹. En otras palabras, la nueva humanidad que exige esta voluntad de voluntad, como consumación de la metafísica y por tanto de la época moderna, es un hombre que debe hacerse cargo del dominio incondicionado de la tierra y del mundo en su conjunto. Se revela incluso “la figura del superhombre como funcionario de la técnica”⁴⁰que para llevar a cabo este dominio debe someter todo a los dictados del cálculo y la certeza absoluta que entrega la planificación y programación automática de la técnica. Lo cual le obliga a: “ponerse en condiciones de ser en el futuro el señor de la tierra, es decir, administrar en un sentido elevado las posibilidades que le vienen al hombre futuro por la esencia de la transformación técnica de la tierra y de la acción humana”⁴¹.

Ahora bien, esta interpretación heideggeriana de la técnica traza un camino cuyo objetivo, en primer término, es evitar confundir la causa con el efecto, por eso es que pone en primer lugar la interpretación del mundo, destinada por el ser, que emerge y se establece como absoluto en la era digital, donde la esencia de la técnica moderna ha impuesto su dominio. Por eso aclara Heidegger: “no es que nuestra época sea técnica por ser el tiempo de las máquinas, sino que es una época de las máquinas por ser la era de la técnica”⁴².

En este contexto es importante subrayar que la crítica de Heidegger a la llamada revolución tecnológica no tiene que ver con una demonización de los avances y progresos. Ni tampoco la idea es un rechazo frontal a su utilización, menos regresar a una era pretecnológica, estilo ludita. Más bien la crítica se dirige a la forma de interpretar la realidad bajo el dominio de la técnica, donde el cálculo, la planificación y programación automática, no solo adquiere un rango de privilegio, sino que se concibe como la única forma de pensar, actuar e interpretar el mundo⁴³. Bajo este dominio, el hombre y la naturaleza se convierten

39 HEIDEGGER, M., *Caminos de bosque*. H. Cortés y A. Leyte (trads.). Madrid: Alianza, 2010, p. 187.

40 CEREZO, “La cuestión del nihilismo. La confrontación Heidegger/Nietzsche”. En L. SÁEZ, J. DE LA HIGUERA Y J. F. ZÚÑIGA (Eds.). *Pensar la nada*. Madrid: Biblioteca Nueva, 2007, 261.

41 HEIDEGGER, M., *¿Qué significa pensar?* R. Gabás (Trad.). Madrid: Trotta, 2005, p. 46.

42 *Ibíd.*, p. 82.

43 Cfr. HEIDEGGER, M., *Serenidad*. Y. Zimmermann (trad.). Barcelona: Ediciones del Serbal, 2002, p. 30.

fundamentalmente en materias primas, material humano al servicio de la voluntad de poder, según los designios del señor de la tierra⁴⁴.

A esto apunta el filósofo alemán, cuando plantea que la esencia de lo técnico no es nada técnico, sino que es un modo de desocultar, de la verdad como *aletheia*, en tanto destinación epocal de la técnica, esto es: como desocultar provocante. Se entiende el fenómeno técnico como un acontecimiento que no depende del manejo ni del dominio del hombre. Esta destinación técnica de la verdad, se denomina en la obra de Heidegger como lo dispuesto (*das Ge-stell*), bajo la cual se comprende la realidad en su conjunto como lo constante (*Bestand*), dispuesto ahí para la explotación calculada y programada de todos los entes del mundo. En esta forma de interpretación, la naturaleza y el hombre no se piensan más que como el principal almacén de existencias de energías, que exige para sí el establecimiento de la naturaleza como constante⁴⁵.

Por otra parte, “lo desvelado no concierne al hombre ni siquiera como objeto, sino exclusivamente como *constante*, y el hombre en medio de lo sin-objeto no es más que un constanciador de lo *constante*, va el hombre sobre el borde más escarpado del precipicio; esto es, va hacia un punto en el que él mismo no podrá ser tomado sino como *constante*”⁴⁶.

En conclusión, la naturaleza se transforma así en una fuente inagotable de recursos, en un stock de reservas ahí dispuesto para su explotación incondicionada. La vida humana se interpreta exclusivamente en trabajador, materia prima, o insumo constante y dispuesto para su aprovechamiento, con la finalidad de obtener los mayores beneficios, aumentar la productividad, el rendimiento y la eficacia

3. UNA APROXIMACIÓN A LA INTELIGENCIA ARTIFICIAL: APLICACIONES, SUPUESTOS Y LÍMITES

Ahora bien, hacia este proyecto o racionalidad calculante es que Heidegger dirigirá su artillería, pues la metafísica moderna, con Nietzsche como figura que la cumplimenta, solo sigue la interpretación del ente y el mundo dictado por los parámetros que marca el pensamiento de Descartes y Leibniz⁴⁷. Bajo este prisma,

44 Cfr. ACEVEDO, J., *Heidegger y la época de la técnica*. Santiago: Editorial Universitaria, 1999.

45 Cfr., HEIDEGGER, M., *Filosofía, ciencia y técnica*. F. Soler (trad.). Santiago: Editorial Universitaria, 1997, 131-132.

46 *Ibid.*, p. 137.

47 Es importante hacer notar que “la doctrina nietzscheana que convierte todo lo que es tal y como es en “propiedad y producto del hombre” no hace más que llevar a cabo el despliegue de

el hombre es empujado a luchar por posicionarse como aquel ente que brinda la medida a todo ente o realidad en la tierra gracias al poder ilimitado que le entrega el cálculo y la planificación para conseguir la corrección de todas las cosas. Resume Heidegger, “comienza ese modo de ser hombre que consiste en ocupar el ámbito de las capacidades humanas como espacio de medición y cumplimiento para el dominio del ente en su totalidad”⁴⁸.

En este escenario, que representaría al mismo tiempo el final de la metafísica y por tanto de la filosofía, augura el autor de *Ser y tiempo* una dirección de la investigación y de la cultura occidental marcada por esta interpretación, que se materializa en la cibernética, pues como señala “no hace falta ser profetas para saber que las ciencias que se van estableciendo, estarán dentro de poco determinadas y dirigidas por la nueva ciencia fundamental, que se llama Cibernética. Está corresponde al destino del hombre como ser activo y social, pues es la teoría para dirigir la posible planificación y organización del ser humano”⁴⁹. Profundizando en estos señalamientos y teniendo en cuenta que la vida humana es la materia prima más importante, vaticinaba Heidegger, que con base en la investigación científica, se planificarían y construirían fábricas de producción artificial de material humano⁵⁰. Ejemplo de esto podría ser sin duda la búsqueda y creación de una inteligencia artificial, lo que haría visible todavía más la conexión entre la interpretación propuesta por Heidegger hace más de cincuenta años con los desarrollos y aplicaciones actuales de la tecnología.

Justamente hacia estos rasgos de la era digital, destacados por Heidegger, es que se dirige la crítica contemporánea hacia la inteligencia artificial, pues en la actualidad gran parte de sus logros, o lo más visibles, si se quiere, apuntan a una organización, planificación, organización y ordenamiento del pensamiento y trabajo humano, merced de lo que se ha denominado como capitalismo de vigilancia⁵¹.

extremo la doctrina de Descartes por la que toda la verdad se funda retrocediendo a la certeza de sí del sujeto humano”; “El aseguramiento del supremo e incondicionado autodespliegue de todas las capacidades de la humanidad en dirección del incondicionado dominio de toda la tierra es el oculto acicate que impulsa al hombre moderno”; “lo representado esté en cada caso puesto en seguro dentro del círculo de la disposición calculante”. HEIDEGGER, M., NIETZSCHE I. J.L. VERMALL (trad.). Barcelona: Destino, 2000, p. 109; 113; 121.

“Solo en la metafísica de Leibniz la metafísica de la subjetividad lleva a cabo su comienzo decisivo”, HEIDEGGER, M., *Nietzsche II. op.cit.*, 192.

48 HEIDEGGER, M., *Caminos de bosque. op. cit.*, p. 75.

49 HEIDEGGER, M., *Tiempo y ser*. M. Garrido, J.L. Molinuevo y F.Duque (trads.). Madrid: Tecnos, 2000, p. 79.

50 Cfr., HEIDEGGER, M., *Superación de la metafísica. op. cit.*, p. 43.

51 Cfr., ZUBOFF, S., *La era del capitalismo de la vigilancia*. Barcelona: Paidós, 2020, p. 9.

En ese sentido, pese a que la inteligencia artificial se promueve como la innovación tecnológica más importante del presente siglo, es heredera de ciertos presupuestos fundamentales de la Modernidad, principalmente la racionalidad calculante descrita por Heidegger, y propuesta en la lógica cibernética de Nibert Wiener⁵².

Aunque es difícil establecer el origen de la idea de Inteligencia Artificial, muchos autores reconocen a Marvin Minsky y John McCarthy, profesores del MIT, como pioneros en el estudio de sistemas inteligentes en la década de 1950⁵³. A McCarthy se le atribuye la definición del campo de la inteligencia artificial como la ciencia para fabricar máquinas inteligentes.

En este contexto, lo primero que debemos tener en cuenta para intentar comprender el significado de la inteligencia artificial (IA) es la propia noción de “Inteligencia”. Definir la inteligencia humana ya es difícil y complejo, sin embargo, cuando se habla de IA se da por sentado lo que se entiende por inteligente⁵⁴. La Inteligencia asociada a las máquinas solo es un tipo de inteligencia definida por su capacidad de cálculo y procesamiento de información. Es inteligente porque puede trabajar con una cantidad de información que excede las capacidades humanas. Esta inteligencia es valorada por su velocidad y eficiencia. En consecuencia, detrás de los discursos y promesas de la IA lo que se prioriza es una idea de inteligencia limitada pero afín a los intereses de quienes la promueven y de aquella racionalidad basada en la cuantificación, cálculo y eficiencia.

Si bien durante las últimas siete décadas los avances en IA se han ampliado y extendido, aún está lejos de imitar el pensamiento y racionalidad de los seres humanos⁵⁵. Las computadoras actuales no son más inteligentes que las de hace unas décadas, pero sí tienen mayor capacidad de almacenamiento, procesamiento y análisis de datos, y es este tipo de inteligencia la que se espera moldee y condicione el desarrollo de las sociedades contemporáneas. “Exponer el límite de la inteligencia artificial no es debilitar a las máquinas, sino más bien liberar a la inteligencia de las máquinas del sesgo de ciertas nociones de inteligencia”⁵⁶.

Los datos son un eje central en el desarrollo de la IA. Los avances en el campo de la inteligencia artificial que pregonan los ideólogos de la técnica han

52 Cfr., HUI, Y. (2021). “On the Limit of Artificial Intelligence”. *Philosophy Today*, 65(2), p. 342.

53 Cfr., RUSSELL, S. & NORVIG, P., *Artificial intelligence: a modern approach*. New Jersey: Prentice-Hall, 2002, p. 16.

54 Cfr., MARTÍNEZ-FREIRE, P., *La nueva filosofía de la mente*. Barcelona: Gedisa, 1995.

55 Cfr., DREYFUS, H. L., *What Computers Still Can't Do A Critique of Artificial Reason*. Cambridge, Massachusetts: The MIT Press, 1992.

56 HUI, Y., On the Limit of Artificial Intelligence. *Philosophy Today*, 65,2, 2021, p. 314.

sido posible gracias a la enorme cantidad de datos que se producen, almacenan y procesan en la actualidad. Esto se conoce como *big data*: datos masivos⁵⁷. Con el avance de la digitalización y de las tecnologías de la información y la comunicación, las posibilidades para registrar y convertir en dato cada movimiento y rastro de los usuarios en el mundo online crecieron exponencialmente.

El deseo por registrar, medir y calcular el mundo, tanto natural como social, hace parte de la promesa que un mundo medible y calculable será un mundo controlable, y, por ende, mejor⁵⁸. Por ello durante los últimos dos siglos parte de los esfuerzos de las ciencias se han orientado a desarrollar las herramientas e instrumentos para lograrlo.

Bajo la lógica de que los datos lo explican todo, los big data se transformaron no solo en la forma más valiosa de explicar el mundo, sino en la única. Lo que no se pueda explicar y reducir a datos prácticamente no existe. “El big data sugiere un conocimiento absoluto. Todo es mensurable y cuantificable. Las cosas delatan sus correlaciones secretas que hasta ahora habían permanecido ocultas”⁵⁹. La producción de datos en mayor volumen, variedad de fuentes y velocidad es posible gracias a que cada vez más las actividades individuales y colectivas están mediadas por las tecnologías de la información y la comunicación⁶⁰. Las TIC además de facilitar la realización de múltiples tareas, permiten el registro que posibilitan la captura, almacenamiento y procesamiento de datos⁶¹.

En la actualidad toda actividad se pueda *datificar*, esto es, convertir en dato⁶². Por ello, es común que discursos periodísticos, académicos y empresariales se refieran a los datos como el nuevo petróleo. Y aunque el petróleo y los datos son recursos muy diferentes, lo cierto es que estos últimos son la fuente de poder las compañías tecnológicas del siglo XXI. Esto ha llevado a que sean presentados y entendidos como un recurso fundamental, no solo para la generación de riqueza, sino para la comprensión y solución de todos los problemas que aquejan a las sociedades contemporáneas. No es casualidad que las compañías más ricas del siglo XXI sean las que mayores datos recolectan, almacenan y procesan; y a la vez, las que mayores inversiones realizan en IA.

57 Cfr., MAYER-SCHONBERGER, V. & CUKIER, K., *Big data. La revolución de los datos masivos*. Madrid: Turner, 2013.

58 ROSA, H., *Uncontrollability of the world*. Cambridge: Polity Press, 2020.

59 Cfr., HAN, B. C., *Psicopolítica*. Barcelona: Herder, 2014.

60 Cfr., CHENEY-LIPPOLD, J., *We are data. Algorithms and the making of our digital selves*. New York: New York University Press, 2017.

61 Cfr., FERRARIS, M., *Movilización total*. Barcelona: Herder, 2017, p. 39.

62 Cfr., VAN DIJCK, J., “Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology”. *Surveillance & Society*, 12, 2, 2014, 197-208.

Pero la captura, almacenamiento, procesamiento y análisis de datos no sería nada sin los algoritmos que llevan a cabo estos procesos⁶³. Entendidos como “un conjunto de instrucciones, una secuencia de tareas destinada a conseguir un cálculo o un resultado particular”⁶⁴, los algoritmos más potentes son propiedad de las compañías tecnológicas; es decir, su poder se cimienta sobre los datos, algoritmos e inteligencia artificial que les permite actuar en casi todos los sectores de la economía, la política y la cultura^{65 66}.

Los algoritmos intervienen en todos los aspectos de la vida cotidiana mediados por las tecnologías, tanto a nivel individual como colectivo^{67 68}. La aplicación de algoritmos e IA a todos los aspectos de la vida buscan hacer un mundo más controlable y eficiente, además de disminuir la arbitrariedad humana. Sin embargo, esta automatización o algoritmización de la realidad lo que impone es un nuevo tipo de arbitrariedad: la algorítmica. Que a diferencia de la humana se vuelve incontestable y aparentemente incontrovertible. Pero “una solución algorítmica es una de las tantas posibles soluciones. Es una reducción del problema. Un problema de clasificación dado puede tener muchas soluciones; el algoritmo que finalmente se aplica es sólo una posible solución. En otras palabras, un algoritmo es una manifestación de una solución propuesta.”⁶⁹ Los algoritmos leen e interpretan la realidad bajo el modelo que han sido programados. La realidad se reduce así a datos cuantificables y calculables.

El estado más avanzando en la creencia de que los datos y el procesamiento automático de información para la toma de decisiones se transformaría en un mundo y sociedad más controlable, gracias a la inteligencia artificial. Con base a la enorme cantidad de datos que se producen, la inteligencia artificial se promueve y promociona como el camino para la comprensión total y absoluta del mundo.

Como ahora hay más datos y mayor capacidad de procesamiento, la IA parece más inteligente y poderosa. Pero lo que ha sucedido es que el mundo y la realidad se redujeron a lo que la IA puede entender, procesar y calcular. “la potencia de la inteligencia artificial se basa en la reducción del mundo a modelos

63 Cfr., CARDON, D., *Con que sueñan los algoritmos. Nuestra vida en el tiempo de los big data*. Madrid: Ediciones DADO, 2018.

64 FINN, E., *La búsqueda del algoritmo*. Barcelona: Alpha Decay, 2018, p. 38.

65 Cfr., SRNICEK, N., *Capitalismo de plataformas*. Buenos Aires: Caja Negra, 2018.

66 Cfr., VAN DIJCK, J.; POELL, T. y WAAL, M., *The platform society*. New York: Oxford University Press, 2018.

67 Cfr., BUCHER, T., *If... Then. Algorithmic power and politics*. New York: Oxford University Press, 2018.

68 Cfr., DeNARDIS, L., *The Internet in Everything. Freedom and security in a world with no off switch*. United States of America: Yale University Press, 2020.

69 BUCHER, T., *If... Then. Algorithmic power and politics*. New York: Oxford University Press, 2018, p. 23.

computacionales”⁷⁰. Esto no quiere decir que la IA sea más inteligente, sino que todo se ha reducido a los datos que necesita la IA, esto es lo que Yuk Hui identifica como la “paradoja de la inteligencia”.

Durante la última década los desarrollos y éxitos de la IA han ganado relevancia en los debates contemporáneos sobre el futuro, de ahí que la IA se promueva como la base del desarrollo humano en todos sus ámbitos y esferas “La inteligencia artificial representa, desde inicios de los años 2010, el desafío económico que se juzga más decisivo y en el cual conviene invertir sin esperar y con determinación. Además de las empresas, también los Estados movilizan todos los medios necesarios para situarse en la vanguardia: de ahora en más, cada uno hace de ese objetivo una gran causa nacional”⁷¹.

Los datos, algoritmos e inteligencia artificial intervienen en aspectos que podríamos considerar banales y cotidianos como las series y películas que vemos, la ruta que tomamos para desplazarnos por la ciudad o en las noticias y contenidos que vemos en nuestras redes sociales⁷². Sin embargo, los grandes esfuerzos e inversiones de las compañías tecnológicas se están orientando a participar en espacios fundamentales de la vida política y económica de las sociedades actuales y futuras: la justicia, salud, educación, seguridad, entre muchos otros.

En la combinación big data, algoritmos e inteligencia artificial se depositan las esperanzas de un mundo mejor, puesto que ante la capacidad de medir y cuantificarlo todo se hace posible controlarlo y mejorarlo. Justamente, en ello se centran parte de las promesas de la cuarta revolución industrial

Pero pese a que la IA se propone como el camino para un futuro mejor, en realidad podría servir también para perpetuar las desigualdades, inequidades e injusticias existentes. Al ser programada a partir de los datos del pasado, lejano o reciente, los sistemas de IA reforzarán y perpetuarán los patrones de discriminación, injusticia, segregación y desigualdad que han caracterizado a las sociedades: “Los procesos del big data codifican el pasado. No inventan el futuro”⁷³, por eso los más afectados por la automatización serían las poblaciones y comunidades discriminadas y marginalizadas históricamente.

En definitiva, los sistemas inteligentes lo son de una forma limitada y reducida: solo sirven para resolver los problemas para los que han sido programados, sus procesos son finitos y sus resultados discretos. Las promesas de una IA o de

70 HUI, Y., On the Limit of Artificial Intelligence. *Philosophy Today*, 65,2, 2021, p. 347.

71 Cfr., SADIN, E., *La inteligencia artificial o el desafío del siglo. Anatomía de un antihumanismo radical*. Buenos Aires: Caja Negra, 2020, pp. 24-25.

72 Cfr., Autor, 2020.

73 O'NEIL, C., *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza democracia*. Madrid: Capitán Swing, 2018, p. 252.

una superinteligencia que solucionaran todos los males de la humanidad hacen parte de un discurso ideológico. En este caso, debemos poner especial atención sobre “la supuesta “neutralidad de la técnica, tan a menudo glorificada como la habilidad para superar conflictos y crear la paz social, el bienestar y el orden, es la máscara de aquellos que tienen el poder de plegarla a sus propios fines”⁷⁴.

4. INTELIGENCIA ARTIFICIAL Y NIHILISMO

Como hemos podido ver, la velocidad y la eficacia se convierten en los valores supremos del mundo regido por datos, algoritmos e inteligencia artificial, por eso, cuando se cuestionan sus riesgos y peligros, las críticas se minimizan, pues se presentan como un ataque al desarrollo y al progreso social que beneficiará a todos por igual, teniendo como base el ordenamiento algorítmico y estandarizado de la sociedad gracias a la aplicación de la inteligencia artificial y los algoritmos en todas las esferas de la sociedad. En este sentido, como afirma Sadin:

La inteligencia artificial no constituye una innovación más entre otras, sino que representa más bien un “principio técnico universal” basado sobre una misma sistémica: el análisis robotizado de situaciones de diverso orden, la formulación instantánea de ecuaciones, supuestamente las más acordes, y en general con vistas a emprender las acciones adecuadas correspondientes, sea por medio de intervenciones humanas o de modo autónomo por los sistemas mismos... Se supone que esta lógica se aplicará a largo plazo a todos los segmentos de la vida individual y colectiva en el marco de nuestras relaciones con nuestros cuerpos, con los demás, con el hábitat, o bien en el marco de la organización de la ciudad, de las redes de transporte, de los espacios profesionales de salud, de las actividades bancarias, de las finanzas, de la justicia, de las prácticas militares⁷⁵.

Hacia estas implicaciones de la inteligencia artificial dirige su atención el pensador francés, pues bajo este principio técnico universal de fuerza incontrastable, la sociedad contemporánea se dirigiría hacia un nihilismo tecnológico, en la cual el sentido, el significado de la existencia humana y de la vida social viene moldeada, condicionada, o, incluso, determinada por los algoritmos, quienes terminarían por decidir, no solo en relación con cuestiones más nimias como qué camino tomar, con qué aerolínea viajar etc., sino que también deciden nuestra forma de relacionarnos con el mundo y con la información. Se construye así una visión de la realidad, del mundo y de la propia vida humana, dictada por las predicciones,

74 VATTIMO, G., *Alrededores del ser*. T.Oñate (trad.). Barcelona: Galaxia Gutenberg, 2020, p. 142.

75 SADIN, E., *La inteligencia artificial o el desafío del siglo*. Op. cit., p. 20.

recomendaciones que nos entrega la inteligencia artificial, a través de la mediación algorítmica basada en los datos.

La interpretación de Han va por la misma vía, sin embargo, el filósofo coreano subraya el rasgo nihilista que se aprecia en los supuestos nacidos al alero del big data. En este contexto, en el cual el manejo, uso y recolección de datos adquieren un protagonismo superlativo surge el concepto de dataísmo, que llega a ser denominada como la filosofía emergente de la actualidad y una segunda ilustración⁷⁶; o también es considerada como una ideología⁷⁷. Todas estas tentativas, provenientes de distintas perspectivas disciplinarias, intentan dar cuenta de la emergencia de una forma destacada de *interpretación* y predicción de la realidad y del comportamiento humano gracias a los datos que podemos obtener y acumular. Gracias al big data se pueden hacer predicciones y pronósticos sobre el comportamiento humano, lo cual hace del futuro algo predecible, controlable y ordenable gracias a la inteligencia artificial, lo que nos protege de cualquier imprevisto propio del devenir de la existencia, brindándonos la tecnología esa seguridad y confort prometido, o ese uniforme del que nos hablaba Jünger. No obstante esto, Han concluye que “La *persona* misma se *positiviza* en cosa, que es cuantificable, mensurable, controlable. Sin embargo, ninguna cosa es libre...El big data anuncia el fin de la persona y la voluntad libre”⁷⁸.

Ahora bien, este llamado dataísmo vendría a ser, según Han, una segunda ilustración, pues destaca el filósofo coreano que para Voltaire la estadística significa ilustración, con base en el conocimiento objetivo, fundamentado y movido por números, donde la transparencia se transforma en la clave de acceso a la realidad, se impone así la obligación de convertir todo en dato e información cuantificable y medible. Lo cual nos llevaría de forma indefectible a una ideología a un totalitarismo de los datos; esto es, una era totalmente comprendida y conducida según el imperio de los datos, como forma suprema de conocimiento y dominio⁷⁹. En esta forma de ideología, que podríamos denominar técnica, se esconde una preocupante forma del nihilismo: “el dataísmo es nihilismo. Renuncia totalmente al sentido. Los datos y los números no son narrativos sino aditivos. El sentido, por el contrario, radica en una narración. Los datos colman el vacío de sentido⁸⁰”. Una visión de nosotros, que retorna sobre nosotros mismos una y otra vez, como un eterno retorno sobre cada uno del cual no podemos salir: puesto que siempre

76 HAN, B. C., *Psicopolítica*. Barcelona: Herder, 2014, p. 48.

77 VAN DIJCK, J., “Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology”. *Surveillance & Society*, 12, 2, 2014..

78 HAN, B. CH., *Psicopolítica. op. cit.*, p. 14.

79 Cfr., *Ibid.*, p. 48.

80 *Ibid.*, p. 48.

vuelve sobre nosotros en virtud de la mediación de la máquina y sus recomendaciones, respondiendo así a todo tipo de inquietudes o preguntas sobre el pasado y el presente, pero fundamentalmente como poseedora de las llaves para abrir el futuro a través de su previsibilidad.

Nos encontraríamos sumidos en una especie de vergüenza prometeica, tal como anunciaba Günter Anders en *La Obsolescencia del hombre*, pues nuestro sentimiento de inferioridad frente a la calidad y capacidad de las máquinas terminan por avergonzarnos; nos sentiríamos entonces obligados por nuestras imperfecciones y limitaciones a delegarle a la IA y los algoritmos la responsabilidad de nuestras decisiones, pensamientos y proyectos. Esto no sería más que el fiel reflejo de la impotencia y humillación que siente el ser humano frente a la perfección de su propia creación tecnológica. Como anota Anders es: “vergüenza ante las cosas producidas (por nosotros), cuya alta calidad avergüenza... A diferencia de los productos impecables y calculados hasta el último detalle, debe su existencia al proceso ciego y no calculado... de la procreación y el nacimiento... también se avergüenza del resultado imperfecto e inevitable de ese origen: de sí mismo”⁸¹. En definitiva, esta vergüenza parecería estar actuando bajo el dominio de la IA y los algoritmos, manifestándose como un sentimiento de impotencia e inferioridad de la vida humana frente al poder omnisciente de las tecnologías y su capacidad de orientar la vida humana por el camino de la perfección y la certeza. Esto amenazaría la capacidad del ser humano de tomar sus propias decisiones, motivado quizá íntimamente por aquella vergüenza prometeica.

5. LA FALTA DE PENSAMIENTO COMO HUÉSPED INQUIETANTE DEL NIHILISMO EN LA ERA DIGITAL

Ciertamente, un rasgo de coincidencia entre los pensadores que aquí hemos mentado, tiene que ver con la caracterización de esta relación entre técnica y nihilismo, en la que sobresale la ausencia de meditación sobre las consecuencias humanas y sociales que conlleva esta hibris tecnológica, la medida desmesurada de los números y el cálculo. Nietzsche es quién plantea esta idea cuando habla de la historia de los próximos doscientos años, “Toda nuestra cultura europea se mueve desde hace tiempo bajo la tortura de una tensión que crece decenio en decenio como abocada a una catástrofe: inquieta, violenta, precipitada, como un

81 ANDERS, G., *La obsolescencia del hombre*. J. Monter (trad.). Valencia: Pre-Textos, 2011, pp. 39- 40.

río que quiere acabar, que no reflexiona ya, que tiene miedo de reflexionar sobre sí mismo”⁸². Precisamente, en esa inquietud, en esa violencia el nihilismo (incompleto) campa a sus anchas, sustituyendo el lugar de Dios por el valor absoluto de la tecnología. Como un discurso e ideología dominante que subrepticamente ocupa el lugar de la deidad, ocultando justamente ese movimiento, en las promesas de salvación, perfección, certeza y eternidad que podrían alcanzar gracias al desarrollo de la tecnología y la inteligencia artificial, como una de las promesas más elocuentes de algunas corrientes del transhumanismo⁸³.

Sin embargo, en este trabajo no se trata de imponer una visión pesimista o determinista del mundo técnico, ni tampoco tiene como principal objetivo llamar la atención sobre su amenaza, sino más bien interpretar este escenario como un camino abierto de escucha atenta a lo que se oculta en tales desarrollos, promesas y supuestos. La filosofía es precisamente quien está llamada a mostrar aquello que se oculta en la era digital, que muchos, por desconocimiento u oscuros intereses, pretenden negar o silenciar. Aquello que se disfraza u oculta es lo que le da sentido y hace comprensible este fenómeno; en este caso, la ausencia de meditación en beneficio de una forma única y exclusiva de interpretar la realidad, que siguiendo a Nietzsche hemos caracterizado como nihilismo pasivo.

Vattimo, siguiendo en esta línea, no rechaza el nihilismo entendiéndolo como un signo de malestar de nuestra civilización o cultura, sino que más bien la incomodidad o inquietud que genera este huésped inquietante muestra que la sociedad contemporánea no es lo bastante nihilista, pues, como hemos visto en el caso de la inteligencia artificial, no ha renunciado todavía a la voluntad de imponer un sentido a las cosas, ya que no puede vivir, todavía, sin un aseguramiento que le permita enfrentar la cruda movilidad y devenir de la existencia⁸⁴. Esto es reafirmado por el filósofo italiano, “la distinción nietzscheana entre nihilismo pasivo, reactivo y nihilismo activo no ha perdido su alcance o significado en los más de cien años que nos separan de su primera formulación. Esto se debe a que todavía estamos en misma situación que Nietzsche preveía en su tiempo...el nihilismo pasivo, reactivo donde estamos “arrojados”, diría Heidegger, deviene activo cuando se asume como punto de partida para la creación de un modelo original de vida”⁸⁵. Cuestión que, a juicio de Vattimo, solo se podrá llevar a cabo si ponemos

82 NIETZSCHE, F., *Fragmentos póstumos*. *Op. cit.*, p. 489.

83 Cfr., VALLE, D., “Humanismo y transhumanismo al final de la metafísica: la era digital como paradigma de la metafísica de la subjetividad”. *Cuestiones de Filosofía*, vol. 6-Nº 26, 2020, pp. 93-94.

84 Cfr., VOLPI, F., *El nihilismo*. *op. cit.*, p. 156.

85 VATTIMO, G., *Alrededores del ser*. T. OÑATE (trad.). Barcelona: Galaxia Gutemberg, 2020, p. 143.

en primer plano las causas o fundamentos ocultos de la transformación de nuestra sociedad en el mundo técnico, para salir así del nihilismo pasivo⁸⁶.

Para finalizar recogeremos la propuesta heideggeriana, que, como hemos hecho notar, no pretende demonizar la tecnología, ni tampoco descalificar completamente los avances que ella nos ha entregado. Cerezo condensa esta propuesta como sigue:

La voluntad de poder, en su reverso o in-versión, ha de dejar paso al abandono de la voluntad de dominio, a la serenidad (*Gelassenheit*) del saber renunciar. A la maquinación (*Machenschaft*) y la vivencia (*Erlebnis*), propias del mundo técnico, sustituyen ahora la meditación (*Besinnung*) y el pensar rememorante (*Andenken*), como guías de un «sentido de la tierra», que consista en el habitar propio del mortal...a la par que Heidegger intentaba atravesar el nihilismo de la metafísica, en su extrema figura del mundo científico/técnico, se abría, más allá de la tragedia, a una forma de pensamiento interrogativo, de escucha y de espera⁸⁷.

La invitación es más bien a reconocer el rasgo nihilista de la era tecnológica como una oportunidad, un destello que nos invita a pensar el mundo más allá de una mirada netamente técnica, para descubrir lo que oculta y las posibilidades que alberga.

CONCLUSIONES

No hay duda de que la relación entre nihilismo y tecnología es una idea que ha estado presente desde los albores de la filosofía contemporánea. Sin embargo, hablar de nihilismo genera confusiones respecto a su sentido, significado y alcance; teniendo en consideración esto hemos precisado desde qué perspectiva se habla de nihilismo en la era tecnológica, esto es, como nihilismo pasivo e incompleto. En la era digital, el lugar de Dios es sustituido por el absoluto tecnológico como valor fundamental, el cual a través de la inteligencia artificial, los algoritmos y el big data se transforman en el camino único y seguro para la perfección del ser humano. Esto a través del orden, la predicción y el aseguramiento de la existencia que permite el cálculo y la planificación, brindándole la seguridad y estabilidad deseada frente al futuro siempre imprevisible propio del devenir. Estos serían precisamente los rasgos que marcarían el nihilismo tecnológico del presente, que

86 Cfr., *Ibid.*, p. 148.

87 CERESO, P., "La cuestión del nihilismo. La confrontación Heidegger/Nietzsche". En L. SÁEZ, J. DE LA HIGUERA Y J. F. ZÚÑIGA (eds.). *Pensar la nada*. Madrid: Biblioteca Nueva, 2007, p. 262.

gracias al big data y el registro completo y automatizado de la vida en todas sus esferas, se transformaría en la fuente de materia prima por antonomasia, para la explotación, planificación y predicción conductual de la humanidad en miras de su optimización, productividad, ordenamiento, uniformidad y clasificación.

Todo esto pondría en peligro el libre albedrío, el poder de decisión y elección sobre los propios actos y conductas, a saber: la libertad como fundamento de la existencia. En definitiva, las elecciones o decisiones, desde las más simples a las más complejas quedarían en manos de la inteligencia artificial y los algoritmos, como nueva deidad con la capacidad de brindar todo tipo de respuestas y soluciones a las distintas problemáticas que vive el *homo digitalis*. Abriendo así las puertas a una cultura e interpretación del mundo y la existencia moldeada, condicionada o quizás determinada por la máquina y la inteligencia artificial, que responden en último término a los intereses de los diseñadores.

En últimas, estos fenómenos serían la manifestación actual más visible del nihilismo que preludiaba Nietzsche y reformularon Jünger y Heidegger, complementando y relacionándolo directamente con el fenómeno técnico, cuya máxima expresión y materialización se da en la inteligencia artificial, como cumplimiento pleno de dichas interpretaciones. Queda por ver si en los próximos decenios se da ese movimiento hacia el nihilismo activo, pero no solo cambiando los valores fundadores, sino también el lugar que ocupan aquellos, entendidos como absolutos y necesarios.

REFERENCIAS BIBLIOGRÁFICAS

- ACEVEDO, J., *Heidegger y la época de la técnica*. Santiago: Editorial Universitaria, 1999.
- BUCHER, T., *If... Then. Algorithmic power and politics*. New York: Oxford University Press, 2018
- CARDON, D., *Con que sueñan los algoritmos. Nuestra vida en el tiempo de los big data*. Madrid: Ediciones DADO, 2018.
- CASILLI, A., *Trabajo, conocimiento y vigilancia. 5 ensayos sobre tecnología*. La Paz: AGETIC, 2018.
- CEREZO, P., "La cuestión del nihilismo. La confrontación Heidegger/Nietzsche". En L. SÁEZ, J. DE LA HIGUERA Y J. F. ZÚÑIGA (Eds.). *Pensar la nada*. Madrid: Biblioteca Nueva, 2007.
- CHENEY-LIPPOLD, J., *We are data. Algorithms and the making of our digital selves*. New York: New York University Press, 2017.
- CROSBY, A., *La medida de la realidad. Cuantificación y la sociedad occidental, 1250-1600*. Barcelona: Crítica, 1998.

- DENARDIS, L., *The Internet in Everything. Freedom and security in a world with no off switch. United States of America: Yale University Press, 2020.*
- DREYFUS, H. L. *What Computers Still Can't Do A Critique of Artificial Reason.* Cambridge, Massachusetts: The MIT Press, 1992.
- DIÉGUEZ, A., *Transhumanismo.* Barcelona: Herder, 2017
- FERRARIS, M., *Movilización total.* Barcelona: Herder, 2017.
- FINN, E., *La búsqueda del algoritmo.* Barcelona: Alpha Decay, 2018.
- GARCÍA, D & VALLE, D. (2020). Los impactos de la ideología técnica y la cultura algorítmica en la sociedad: una aproximación crítica. *Revista de Estudios Sociales, 2020* (71), pp. 15-27. <https://doi.org/10.19053/01235095.v6.n26.2020.11242>
- GENTILI, C., “Nihilismo, nihilistas. Sobre una posible fuente de Nietzsche”, *Estudios Nietzsche*, 19, 2019, pp. 63-86.
- HAN, B. C., *Psicopolítica.* Barcelona: Herder, 2014.
- HEIDEGGER, M., *Filosofía, ciencia y técnica.* Santiago: Editorial Universitaria, 1997
- , *Tiempo y ser.* Madrid: Tecnos, 2000.
- , *Conferencias y artículos.* E. Barjau (trad.). Barcelona: Ediciones del Serbal, 2000.
- , *Nietzsche I.* Barcelona: Destino, 2001.
- , *Nietzsche II.* Barcelona: Destino, 2001.
- , *Serenidad.* Barcelona: Ediciones del Serbal, 2002.
- , *Caminos de bosque.* Madrid: Alianza, 2010.
- , *Acerca de Ernst Jünger.* D. Picotti (trad.). Buenos Aires: El hilo de Ariadna, 2013.
- HUI, Y., *Fragmentar el futuro: ensayos sobre tecnodiversidad.* Buenos Aires: Caja Negra, 2020.
- , On the Limit of Artificial Intelligence. *Philosophy Today*, 65(2), 339–357, 2021. DOI: 10.5840/philtoday202149392
- JÜNGER, E., *El trabajador.* Barcelona: Tusquets, 1990.
- , *Sobre el dolor.* Valencia: Pre-textos, 1995.
- , *Sobre la línea.* J.L. Molinuevo (trad.). Barcelona: Paidós, 1994.
- KUHN, E Friedrich Nietzsches Philosophie des europäischen Nihilismus, Berlin-New York: de Gruyter, 1992.
- MARTÍNEZ-FREIRE, P., *La nueva filosofía de la mente.* Barcelona: Gedisa, 1995.
- MAYER-SCHONBERGER, V. & CUKIER, K., *Big data. La revolución de los datos masivos.* Madrid: Turner, 2013.
- OCAÑA, E. (1993). *Más allá del nihilismo.* Murcia: Universidad de Murcia, 1993.
- NIETZSCHE, F., *Fragmentos póstumos IV.* Madrid: Tecnos, 2008.
- O'NEIL, C., *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza democracia.* Madrid: Capitán Swing Libro, 2018.
- ROSA, H., *Uncontrollability of the world.* Cambridge: Polity Press 2020.
- RUSELL, S., & NORVIG, P., *Artificial intelligence: a modern approach.* New Jersey: Prentice-Hall, 2002
- SADIN, E., *La siliconización del mundo.* Buenos Aires: Caja Negra, 2018.

- , *La inteligencia artificial o el desafío del siglo. Anatomía de un antihumanismo radical*. Buenos Aires: Caja Negra Editora.
- SANCHEZ MECA, D., *El nihilismo*. Madrid: Síntesis, 2004.
- , *La experiencia dionisiaca del mundo*. Madrid: Tércnos, 2006.
- SEVERINO, E., *Esencia del nihilismo*. E. Álvarez (trad.). Madrid: Taurus, 1991.
- SRNICEK, Nick. (2018). *Capitalismo de plataformas*. Buenos Aires: Caja Negra Editora, 2018.
- VAN DIJCK, J., “Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology”. *Surveillance & Society*, 12, 2, 2014, 197-208.
- VAN DIJCK, J; POELL, T & WAAL, M *The platform society*. New York: Oxford University Press, 2018.
- VALLE, D., “Humanismo y transhumanismo al final de la metafísica: la era digital como paradigma de la metafísica de la subjetividad”. *Cuestiones de Filosofía*, vol. 6, N° 26, 2020, pp. 75-97. <https://doi.org/10.19053/01235095.v6.n26.2020.11242>.
- VATTIMO, G., *Alrededores del ser*. T.Oñate (trad.). Barcelona: Galaxia Gutemberg, 2020.
- VOLPI, F., *El nihilismo*. Buenos Aires: Biblos, 2005.
- WIENER, N., *Cibernética y sociedad*. Buenos Aires: Editorial Sudamericana, 1998.
- ZUBOFF, S., *La era del capitalismo de la vigilancia*. Barcelona: Paidós, 2020.

DE LAS EMOCIONES NATURALES A LA EMOCIONALIDAD ARTIFICIAL

FROM NATURAL EMOTIONS TO ARTIFICIAL EMOTIONALITY

JOSÉ MIGUEL BISCAIA FERNÁNDEZ

Graduado y Doctorando en Filosofía

Licenciado en Biología, Máster en Biotecnología y Doctor en Neurociencia

Profesor Titular de Fisiología

Facultad de Medicina

Universidad Europea de Madrid

Madrid/España

josemiguel.biscaia@universidadeuropea.es

ORCID: 0000-0002-3496-5527

Recibido: 30/04/2021

Revisado: 24/07/2021

Aceptado: 6/09/2021

Resumen: La innovación en inteligencia artificial se encuentra en la vanguardia científico-tecnológica y muchas de sus aplicaciones presentan ya una alta interactividad humana. Dado que somos seres emocionales, y que esta cualidad psicobiológica es clave para nuestro posicionamiento en el mundo, parece necesario realizar una profunda reflexión sobre el espectro afectivo de la díada humano-máquina. El presente ensayo utiliza el término “emocionalidad artificial” como concepto holístico capaz de abordar este análisis desde una triple perspectiva: la de las ciencias cognitivas y de la computación, la de la neurociencia y la psicología y la de la filosofía teórica y práctica. Partiendo de la descripción de las emociones naturales, en este estudio se discute sobre la posibilidad técnica y conceptual y sobre las consecuencias bio-psico-sociales de una inteligencia artificial con capacidad de reconocimiento, simulación, manipulación y vivencia emocional. Tras dicho análisis se concluye que las dos primeras capacidades son ya en cierto modo posibles y deseables, mientras que las dos últimas se enfrentan a dificultades tecno-científicas, ontológicas, gnoseológicas y neuroéticas discutidas ampliamente por el transhumanismo.

Palabras clave: Emoción, inteligencia artificial, filosofía de la mente, transhumanismo, neuroética.

Abstract: Innovation in artificial intelligence is at the forefront of science and technology, and many of its applications already have high human interactivity. Given that we are emotional beings, and that this psychobiological quality is key to our positioning in

the world, it seems necessary to make a deep reflection on the affective spectrum of the human-machine dyad. This essay proposes the term of “artificial emotionality” as a holistic concept capable of approaching this analysis from a triple perspective: that of cognitive and computer sciences, that of neuroscience and psychology, and that of theoretical and practical philosophy. Starting from the description of natural emotions, this study discusses the technical and conceptual possibility and the bio-psycho-social consequences of an artificial intelligence capable of recognition, simulation, manipulation and emotional experience. After said analysis it is concluded that the first two capabilities are already in a certain way possible and desirable, while the last two face techno-scientific, ontological, epistemological and neuroethical difficulties widely discussed by transhumanism.

Keywords: Emotion, artificial intelligence, philosophy of mind, transhumanism, neuroethic.

INTRODUCCIÓN

Desde su nacimiento a mediados del siglo XX, a partir de las ideas de Alan Turing, la inteligencia artificial (IA) no ha hecho más que progresar. Si bien no se ha cumplido la “Ley de Moore” (1965), que proponía un crecimiento exponencial de la tecnología cibernética, nadie negará que sus aplicaciones actuales y sus futuros desarrollos se encuentran en la avanzadilla de la ciencia y la tecnología. Entre todas estas aplicaciones, las más destacadas son las que se relacionan con el lenguaje natural y el reconocimiento del habla, la visión artificial, la robótica o la conducción autónoma.

Una de las principales características de estos desarrollos en IA es su alto grado de interacción con nosotros. Convivir con máquinas inteligentes supondrá, como mínimo, una adaptación a los nuevos usos tecnológicos y, en su límite, un auténtico cambio de paradigma relacional ético-social. Teniendo en cuenta que somos seres emocionales, y que dicha cualidad psicobiológica es clave para comprender nuestra relación con el mundo, se antoja imprescindible que reflexionemos de un modo bidireccional al respecto del continuo afectivo humano-máquina.

Así pues, este artículo propone una visión integrada en el estudio de las emociones y de la inteligencia artificial, que no se limite sólo a lo técnico o sólo a lo filosófico (como si fueran ámbitos estancos); que ofrezca, de este modo, una aproximación holística al tema de estudio:

- Utilizando el concepto de “emocionalidad artificial” (EA) como alternativa a otros similares (aunque pretendidamente más inclusivo), tales como el de “emoción sintética” (Casacuberta y Vallverdú, 2010), “emoción artificial” (Spinola y Queiroz, 2007), “inteligencia artificial emocional” (Schuller y Schuller, 2018), “arquitectura emocional” (González y García, 2006), “computación afectiva” (Picard, 1997) o “inteligencia artificial de

la emoción” (Chakriswaran et al., 2019). Con dicho término se pretende trascender el mero análisis técnico de las emociones artificiales como subdisciplina de la IA, ampliando sus márgenes para albergar de este modo a la tríada máquina-humano-mundo.

- Planteando un abordaje multidisciplinar, conexionista y funcional de las materias de las que se nutre esta área de conocimiento. De este modo, cualquier análisis profundo del tema exigirá tener una aproximación plural:
 - Desde las ciencias cognitivas y de la computación: describiendo el estado tecnológico de la cuestión, desde lo ya realizable hasta la vanguardia y la especulación tecno-científica.
 - Desde la neurociencia y la psicología: usando las bases neuroquímicas, neurofisiológicas, neuroanatómicas y psicológicas del cerebro humano como modelo comparado para el desarrollo de la IA.
 - Desde la filosofía y las humanidades: contemplando, desde una perspectiva teórica, las dificultades y límites ontológicos y epistémicos de los futuros desarrollos en IA y, desde una perspectiva práctica, la necesidad, utilidad y alcance de la IA en la praxis humana (ética, sociedad, cultura, derecho, economía o política).

De este modo, en relación con la emocionalidad artificial son varias las preguntas generales que pueden discernirse en el horizonte: por un lado, desde la perspectiva de la IA, ¿es tecnológica y conceptualmente posible desarrollar capacidades emocionales en la IA? Si la respuesta es afirmativa, ¿qué aspectos contempla dicha emocionalidad y qué utilidad y sentido podría tener el dotar a una máquina con capacidad de reconocimiento, imitación, manipulación o vivencia afectiva? Por otro lado, desde la perspectiva humana, ¿es posible acceder (y comprender) a los posibles estados emocionales de una IA, incluso manifestar sentimientos por una máquina? Y, a modo de síntesis de todas ellas, en virtud del modo pretendidamente integrador del término “emocionalidad artificial”, ¿cómo será un mundo en el que humanos y máquinas interaccionen afectivamente, qué consecuencias éticas y sociales tendrá?

En el presente estudio, más allá de introducir el debate general de discusión en torno al concepto de emocionalidad artificial, se reflexionará sobre algunas de las cuestiones arriba planteadas, en particular en relación a los aspectos más epistémico-ontológicos, si bien se seguirá el espíritu integrador propuesto y de forma paralela se discutirán aspectos técnicos y filosófico-pragmáticos. No pretende este ensayo, por tanto, agotar ni mucho menos las posibilidades de análisis al respecto de la emocionalidad artificial. En consecuencia, con el objetivo de dar

respuesta a algunas de las preguntas que se han propuesto, se seguirá el siguiente orden argumentativo:

1. En primer lugar, dado que de emociones artificiales trata este ensayo, se partirá de la descripción de las emociones naturales desde una perspectiva filosófica y psicológica.
2. A continuación, una vez explicadas, servirán para el análisis comparado de las emociones artificiales. En el marco teórico de la filosofía de la mente y de las ciencias cognitivas y de la computación, con el objetivo de describir el estado tecnológico de la cuestión y abundar en los límites ontológicos, gnoseológicos y pragmáticos de la emocionalidad artificial, haré una diferenciación interesada entre:
 - 2.1. IA que reconoce emociones.
 - 2.2. IA que simula emociones.
 - 2.3. IA que manipula emociones.
 - 2.4. IA que siente emociones.

1. EMOCIONES NATURALES

Parece lógico que para discutir sobre emociones artificiales partamos antes de sus predecesoras, a saber, las emociones naturales de los seres biológicos. Definir qué son y explicar cuál es su función resultará útil para, mediante un análisis comparado, determinar si es posible y deseable que una IA igualmente las manifieste o presente alguna capacidad en relación con ellas.

No obstante, si lo que se pretende es dar una definición canónica y unívoca del concepto “emoción” nos encontramos enseguida en un callejón sin salida, pues pocos términos son tan inasibles como el que nos ocupa. De este modo tan intuitivo lo expresaban tempranamente Wenger, Jones y Jones (1962: 3): “casi todo el mundo piensa que sabe lo que es una emoción, hasta que intenta definirla. En ese momento prácticamente nadie afirma poder entenderla”. Y es que es tal la complejidad de este fenómeno psicológico que falta una teoría integrada (única y concluyente) que pueda dar cuenta de dicho concepto desde una perspectiva descriptiva, explicativa y predictiva (Cano-Vindel, 1995; Reizenzein, 2019; Tantam, 2003). Una explicación adicional de esta dificultad cognoscitiva probablemente tenga que ver con la variedad de paradigmas epistemológicos desde los que la ciencia y la filosofía se ha aproximado a este nebuloso fenómeno (donde, además, en muchas ocasiones la evidencia empírica ha ido por detrás de los modelos teóricos). Sin embargo, a pesar de las citadas dificultades conceptuales del propio evento emocional y del modo de acercarse hasta él, la ciencia empírica ha hecho

extraordinarios progresos en los últimos años, partiendo de diferentes paradigmas científico-experimentales y de distintas orientaciones metodológicas como la conductual (basada en la psicología del aprendizaje), la (neuro)biológica o la cognitiva (Fernández-Abascal et al., 2010; Reizenzein, 2019).

Tradicionalmente, las diferentes escuelas del pensamiento occidental han mantenido posturas diversas con respecto a las emociones, que van desde el desprecio y ostracismo de las mismas hasta la consideración de lo emocional como fenómeno digno de consideración y estudio filosófico¹. Con respecto al citado descrédito, tal y como reconoce Broncano (2001), son muchos los autores que niegan a las emociones un valor intrínseco y una significación propia, incluso que consideran que no son una clase natural, reduciéndolas con frecuencia a otras instancias como disposiciones, intenciones, conductas, perturbaciones fisiológicas, actitudes proposicionales o juicios evaluativos. Entre sus más destacados detractores contemporáneos encontramos a Ryle (1967), Griffiths (1997) o Elster (1999).

Frente a dichas críticas, el filósofo español señala que las emociones son fenómenos complejos y difusos, aunque de gran interés para la filosofía, pues si bien están en un territorio entre lo fisiológico y lo conceptual, comparten (aunque

1 En la historia de la filosofía se puede sondear el interés de los autores clásicos por el tema de las emociones (o pasiones, del griego *pathos*, como habitualmente eran nombradas). De un lado está la postura general de quienes no se las han tomado en serio, por estimarlas como algo menor; de quienes, incluso, las han considerado como impulsos irracionales internos que debilitan y ensombrecen la razón e intelección humana. Platón, por ejemplo, distinguía en el alma tres dominios separados: el racional, el irascible (sentimientos nobles) y el concupiscible (sentimientos inferiores). Su mito del carro alado, presente en el diálogo del *Fedro*, ejemplifica el dominio de lo cognitivo sobre las bajas pasiones para alcanzar así la virtud y la verdad. En esta línea despreciativa, la escuela estoica tampoco se tomó demasiado en serio a las pasiones, considerándolas en todo caso como elementos perturbadores. Aristóteles, sin embargo, quien las menciona en obras como *Ética eudemia*, *Ética nicomaquea* o *Retórica*, les otorgó más valor, y creía que la parte racional e irracional (lo emocional-pasional) formaban una unidad que guía al hombre. Pasados los siglos, el dualismo ontológico de René Descartes diferencia entre *res cogitans* y *res extensa*, dando mayor importancia a la razón y señalando un tanto despreciativamente que lo emocional es conducido por “espíritus animales”. Como reconoce en *Las pasiones del alma* de 1649, el *cogito* humano puede llegar a dominar lo concupiscente; los animales, por el contrario, al ser considerados como seres sin alma (meros autómatas), dependen enteramente de fuerzas externas e internas. Con el asociacionismo, mecanicismo y empirismo defendido por autores como John Locke o David Hume, se empieza a considerar la relevancia de las emociones en los procesos mentales superiores. Gracias a las ideas sobre el origen biológico de la conducta (y de las emociones) de Charles Darwin (mencionadas en su obra de 1872 *La expresión de las emociones en el hombre y en los animales*) y con el nacimiento de la psicología y la fisiología modernas a manos de autores como Walter Cannon, William James o Carl Lange, se sentarán las bases del estudio científico-experimental de las emociones (para una aproximación más profunda a la historia del estudio de las emociones consultar el capítulo 1 de *Psicología de la emoción* de Fernández-Abascal et al. [2010: 18-37] y las revisiones de Casado y Colomo [2006] y de Pinedo y Yáñez [2018]).

sea parcialmente) varias características con los estados mentales tradicionalmente estudiados por la filosofía de la mente; más en concreto (Broncano, 2001):

- Tienen intencionalidad, pues las emociones presentan contenido que se dirige hacia el objeto que despierta el estado emocional (por ejemplo, se desea o se teme a algo o a alguien en particular). Aunque hay que diferenciar a las emociones de otro fenómeno similar, como son los estados de ánimo, pues estos últimos, que se caracterizan por una mayor vigencia temporal frente al “instante” de la emoción, no muestran un objeto definido. Reconoce, no obstante, que dicho contenido emocional podría ser no-conceptual:

Las emociones conforman un sistema de señales o mensajes que porta y procesa contenido, aunque quizás este contenido es no-conceptual [...] Las emociones, o lo que sea el correlato neuronal del sistema emotivo [...] pueden portar información, pero solamente cuando es descriptible conceptualmente se convierte en mental (Broncano, 2001: 12).

- Contribuyen a la individuación, ya que las emociones anclan los estados mentales en un contexto de evaluaciones, en un eje espacio-temporal clave para la configuración del individuo y para la valoración del mundo.
- Comunican estados mentales, gracias a que la expresión de las emociones es una herramienta vehicular de ciertos estados internos de otro modo inobservables. La “Teoría de la mente” (capacidad de atribuir estados mentales en el otro) da buena cuenta de esta observación (Tirapu-Ustárruz et al., 2007).
- Ayudan en la evaluación de situaciones, especialmente en contextos de urgencia cuando ha de hacerse una valoración rápida que posibilite igualmente una respuesta ajustada a las circunstancias. Son, por tanto, un soporte de la maquinaria evaluadora cognitiva.
- Participan en la creación de constructos socio-culturales, en la medida en que, según algunos autores, las normas sociales y morales derivan de nuestra capacidad emocional (Gibbard, 1990). En este sentido, la mayoría de neurocientíficos y filósofos coinciden en que los juicios morales son intuitivos y están muy vinculados a la emocionalidad, y sólo alcanzamos (a veces) a dar razones de por qué hicimos tal o cual cosa a posteriori, si se nos pregunta o reflexionamos conscientemente sobre ello (Cortina, 2011). Así pues, las emociones tendrían un rol causal en el nivel lingüístico-social, creando un sistema de compromisos que soporta

algunas de las más altas instancias creadas por el aparato cognitivo del hombre, como son las normas o las instituciones sociales².

La complejidad psicobiológica y la multidimensionalidad del fenómeno emocional puede apreciarse claramente si se observan las más de cien definiciones analizadas por Kleinginna y Kleinginna (1981); a partir de ellas, los autores propusieron once categorías descriptivo-explicativas que bien pueden servir para agrupar y detallar de una forma científica, rigurosa y sistemática todas estas formas de conceptualizar la emoción, a saber:

- Categoría afectiva: tiene que ver con la dimensión subjetiva y experiencial de las emociones. Al amparo de esta conceptualización se encontrarían los *qualia* emocionales (cómo es sentir la vivencia de una determinada emoción o estado afectivo). En psicología suele utilizarse el término “sentimiento” para referirnos a esta experiencia subjetiva emocional.
- Categoría cognitiva: se relaciona con aspectos valorativos, informativos, clasificatorios y perceptivos de las emociones. Esta conceptualización es la que más asemeja los estados emocionales a otros estados mentales.
- Categoría basada en estímulos elicidores: tiene que ver con causas desencadenantes externas, es decir, con eventos disparadores de las emociones procedentes del entorno (como una amenaza o un objeto de deseo).
- Categoría fisiológica: hace mención a las modificaciones fisiológicas, a los mecanismos neurobiológicos subyacentes y al sustrato neuroanatómico de las emociones (donde destaca, por encima de todos, el denominado como sistema límbico emocional)³.

2 Emociones autoconscientes como el orgullo, la culpa o la vergüenza son fundamentales como elementos motivadores y controladores de la conducta moral (Fernández-Abascal et al., 2010: 431).

3 El estudio neurobiológico de las emociones (lo que se conoce como “neurociencia afectiva”) ha tenido un crecimiento exponencial en las últimas décadas (Esperidiao-Antonio et al., 2017). La mayoría de investigaciones se han realizado a partir de pacientes con lesiones cerebrales que suponían algún déficit cognitivo-emocional, utilizando modelos animales en los que se alteraba quirúrgica o farmacológicamente alguna región neuroanatómica emocional o empleando sofisticadas técnicas de neuroimagen que correlacionan anatomía y función como la resonancia magnética funcional (fMRI). Gracias a estas aproximaciones metodológicas se han encontrado estructuras claramente implicadas en el procesamiento emocional como la amígdala cerebral, los núcleos septales, la formación hipocámpal, diferentes núcleos diencefálicos o varias regiones corticales como la corteza orbitofrontal o la corteza cingulada. Éstas y otras estructuras adicionales forman parte del circuito emocional de nuestro encéfalo que se conoce con el nombre de sistema límbico. No obstante, según Ledoux (2000), dicho sistema no es un concepto claramente definido, ni anatómico ni funcionalmente hablando, por lo que en palabras de Fernández-Abascal et al. (2010: 50), “actualmente no se mantiene la existencia de un

- Conceptualización emocional/expresiva: se refiere a la expresión externa de las emociones experimentadas, es decir, a la conducta observable que se desencadena. Es la que permite dotar a las emociones de carácter comunicativo.
- Categoría disruptiva: hace hincapié en aspectos disfuncionales y desorganizadores de las emociones, como fenómenos perturbadores que alejan al sintiente de la normalidad funcional. Esta perspectiva es la que desde la filosofía clásica ha denostado tanto a las emociones y, desde la perspectiva actual, justifica los trastornos psiquiátrico-emocionales.
- Categoría adaptativa: contraria a la anterior, pone el acento en el papel funcional y evolutivo de las emociones; en su alta conservación filogenética como garantía de su eficacia para responder a las demandas del entorno y alcanzar la supervivencia individual.
- Categoría multifactorial: considera la emoción como un conjunto amplio de fenómenos, algunos de ellos ya descritos, con diferentes dimensiones: afectiva, motivacional, cognitiva, fisiológica y conductual.
- Conceptualización restrictiva: define la emoción por contraste con otros procesos psicológicos diferentes, como puede ser la percepción o la cognición.
- Categoría motivacional: plantea una vinculación entre emoción y motivación, siendo lo primero elicitador de lo segundo.
- Categoría escéptica: dada la dificultad definatoria y explicativa, resta importancia a este esfuerzo y asume que es una tarea inabarcable e, incluso, epistémicamente cuestionable.

Como se puede deducir, resulta muy difícil determinar cuáles de los elementos expuestos son causa suficiente y necesaria para que un fenómeno sea categorizado como emocional; muy arduo, además, porque son múltiples los contraejemplos que podrían ofrecerse sobre la mencionada categorización. No obstante, Fernández-Abascal et al. (2010: 19-20) considera que a partir de estas conceptualizaciones hay cierto consenso científico sobre cuáles son los cuatro elementos esenciales para entender qué son las emociones: (1) en las emociones hay cambios fisiológicos⁴; (2) en las emociones hay una “tendencia a la acción” o

círculo único y general que explique el procesamiento emocional”, afirmación avalada por la descripción reciente de múltiples subsistemas emocionales íntimamente conectados (Vogt, 2019).

4 Muchos autores defienden que la activación fisiológica es condición necesaria, aunque no suficiente, para que se desencadene una emoción. Lo “necesario” hace mención a que el sujeto debe evaluar y valorar tanto el estado de activación como la situación contextual en que se produce. Según las primeras teorías cognitivo-emocionales desarrolladas en los 60-70 del pasado siglo, como la “Teo-

afrentamiento que ayuda a responder frente a las demandas del entorno; (3) en las emociones hay una experimentación subjetiva o sentimiento; (4) en las emociones se produce un procesamiento de información. Así pues, en un intento de concretar y asir el nebuloso fenómeno de las emociones, una definición con la que trabajaré de ahora en adelante será la siguiente:

Las emociones son un proceso que implica una serie de condiciones desencadenantes (estímulos relevantes), la existencia de experiencias subjetivas o sentimientos (interpretación subjetiva), diversos niveles de procesamiento cognitivo (procesos valorativos), cambios fisiológicos (activación), patrones expresivos y de comunicación (expresión emocional), que tiene unos efectos motivadores (movilización para la acción) y una finalidad: que es la adaptación a un entorno en continuo cambio (Fernández-Abascal et al., 2010: 40-41).

Una teoría explicativa surgida hace algunas décadas, aunque de gran vigencia aun en la actualidad, basada en el rol funcional de las emociones, que acumula gran cantidad de evidencia científica y que presenta interesantes implicaciones en otros procesos psicológicos cognitivos como la memoria o la toma de decisiones, es la “Teoría funcional del sistema emotivo” (Oakley y Jonhson-Laird, 1987: 13). Según los autores:

Las emociones son sistemas de señales que indican a la mente transiciones de planes. Su función es detectar posibles objetivos en el medio externo o interno que son relevantes (positiva o negativamente) a un plan en marcha y disponer al organismo a una biblioteca de planes potenciales de acción que está almacenada en la memoria a largo plazo, pero que gracias a las conexiones rápidas del sistema emotivo (mediante un sistema de marcadores emotivos de la información) permite una activación mucho más rápida que a través de los medios habituales de recuperación de la memoria. El sistema emotivo parte el mundo (externo o interno) en categorías mucho más amplias que el lenguaje y el sistema conceptual, porque son categorías de control rápido de planes.

Como se deduce de este texto, resulta inaceptable explicar las características y función de las emociones sin atender a otros procesos mentales. Es cierto que el inicio del procesamiento emocional (en regiones como la amígdala cerebral)

ría bifactorial de la emoción” o la “Teoría de la evaluación-discrepancia”, la activación indiferenciada correlacionaría con la intensidad emocional, mientras que la cualidad afectiva estaría determinada por la interpretación emocional (lo evaluativo-cognitivo). Los conceptos de “evaluación” (basado en la novedad y el agrado situacional) y “valoración” (basado en el significado, la causalidad y la normatividad social de la situación) son ampliamente reconocidos en las teorías actuales como elementos clave en el procesamiento emocional: los cambios fisiológicos que se experimentan en una emoción serían fruto de la evaluación y valoración situacional y serían fundamentales para preparar al organismo en la eventual ejecución de una acción de aproximación o evitación, según el caso (Fernández-Abascal et al., 2010).

puede ser automático e inconsciente (Janak y Tye, 2015), una especie, concluyen muchos psicólogos, de “vía rápida” de procesamiento. Sin embargo, las emociones más complejas son fruto de acciones deliberadas que exigen consciencia (sería la “vía precisa”), como la evaluación situacional, el conocimiento adquirido previamente o nuestras expectativas, capacidad planificadora y toma de decisiones, aspectos todos ellos relacionados claramente con nuestra cognición (donde están implicadas áreas cerebrales como la corteza prefrontal [Dixon et al., 2017]), lo cual nos transporta a las ideas de Broncano (2001) expuestas más arriba sobre la complejidad de los estados emocionales y su similitud con otros estados mentales tradicionalmente estudiados por la filosofía de la mente. Así pues, podemos concluir que lo cognitivo-emocional forma un continuo explicativo bidireccional en gran parte del procesamiento mental⁵. Dicho de otro modo, las emociones conforman una especie de sistema multinivel de procesamiento informacional que es clave para la supervivencia. Lo emocional tamiza el mundo (interior y exterior), lo categoriza, evalúa y valora de forma sencilla (aunque también rápida y eficazmente), para de este modo dar una respuesta adaptativa que favorezca la supervivencia del individuo en un entorno cambiante. Podemos concluir, con Reeve (1994), que las emociones tienen tres grandes funciones: adaptativas (porque contribuyen a la supervivencia y al bienestar individual), sociales (porque intervienen en las relaciones entre individuos) y motivacionales (porque movilizan recursos para la acción, favoreciendo o perjudicando el acercamiento o la evitación).

Otro aspecto a destacar en el análisis de las emociones, de interés para las intenciones de este artículo, es el intento clasificatorio que durante años lleva realizando la psicología. La dificultad deriva de los problemas descriptivos y explicativos ya comentados, además del criterio agrupador que se considere. Sin embargo, los esfuerzos en esta tarea no han decaído, no sólo por abundar en el conocimiento del fenómeno emocional sino, también, por la aplicabilidad terapéutica en el tratamiento de desórdenes psiquiátrico-emocionales. Así pues, hay dos grandes criterios agrupadores: (1) el dimensional, que define el mapa emocional basándose en tres dimensiones generales (valencia afectiva, activación y control), y (2) el discreto, que reconoce características específicas y distintivas

5 Prueba adicional de esta relación entre lo emocional y otros fenómenos cognitivos la encontramos sobradamente en la obra *Inteligencia emocional* de Daniel Goleman (1995). Como resumen de esta vinculación, Casacuberta y Vallverdú (2010: 120) señalan que “nuestras emociones están en el núcleo mismo de la racionalidad”. En esta misma línea, el célebre investigador Antonio Damasio (1994) dejó a las claras en su obra *El error de Descartes* la evidente conexión entre emoción y razón. En cualquier caso, “si el término cognitivo alude a procesamiento activo de la información, la mayor parte, sino todas las emociones, requieren algún tipo de procesamiento cognitivo” (Fernández-Abascal, 2010: 61), por ejemplo, como mínimo, el implicado en los patrones evaluativos (sean o no conscientes).

en cada emoción (Fernández-Abascal et al., 2010). Siguiendo este último criterio, psicólogos como Ekman (1992) definieron seis emociones básicas o primarias (sorpresa, asco, miedo, alegría, tristeza e ira), las cuales son consideradas como innatas y universales, y surgen antes en el desarrollo individual. Adicionalmente, la psicología afirma que también poseemos emociones secundarias o complejas, también denominadas “autoconscientes” o “autoevaluativas” (como la vergüenza, la culpa, el orgullo, los celos, etc), fruto de nuestra compleja interactividad social (Lewis, 2000).

2. EMOCIONALIDAD ARTIFICIAL

El debate filosófico en torno a la emocionalidad artificial (EA) puede desarrollarse a lo largo de un continuo en cuyos extremos aparecen dos polos diferenciados: el de la filosofía teórica y el de la filosofía práctica. Desde el primer núcleo de reflexión se puede dar respuesta a cuestiones ontológicas y epistemológicas en relación a las capacidades de la emoción artificial; desde la filosofía práctica, el análisis se construye sobre aspectos relacionados con la acción humana, como la ética, la estética o la política. Dicho de modo más claro, la filosofía teórica tratará de responder a la pregunta de si es conceptualmente posible una IA con capacidad emocional; desde la filosofía práctica, las preguntas girarán en torno a la necesidad, utilidad y consecuencias de la EA en la praxis humana. Pues bien, en el presente ensayo abordo cuestiones relacionadas con la filosofía teórica; cuestiones óntico-gnoseológicas que caen bajo el amparo de la filosofía de la mente y de las ciencias cognitivas. Aunque en ulteriores análisis sería conveniente tamizar aspectos más pragmáticos, aquí sólo se mencionarán de pasada aquellas tesis generales que sean ineludibles y, por supuesto, aquellos desarrollos tecnológicos que sirvan de soporte para mi argumentación. Con esta forma de tratar el tema respeto el espíritu integrador y multidisciplinar de este ensayo, tal y como se defendió en la introducción.

Una forma adecuada de iniciar la reflexión es distinguir las cuatro grandes formas a través de las cuales la IA puede relacionarse con las emociones: reconociéndolas, simulándolas, manipulándolas o sintiéndolas. Detrás de cada uno de estos cuatro términos se esconden problemas filosóficos de honda raigambre y compleja fundamentación, dificultad potenciada en la medida en que estamos haciendo una traslación desde el ámbito biológico (descrito en el apartado 1) al sintético.

Sobre el reconocimiento y la simulación de emociones descansa un problema tratado por la gnoseología, que está en el origen mismo del debate filosófico sobre las capacidades cognitivas de la IA, y que tiene que ver con lo que se conoce y

comprende al respecto de lo emocional. “Conocimiento” o –hago un giro terminológico que matizaré en breve– “procesamiento informacional” para que la IA pueda computar la información recibida y, así, llegado el caso, ofrecer después la respuesta programada e, incluso, imitar un patrón emocional. Aquí se encuentra, pues, el primer nudo gordiano, clásico ya entre los filósofos analíticos y del lenguaje, que tiene que ver con la diferencia entre procesar información manipulando símbolos mediante una mera computación instalada en el algoritmo cibernético (sintaxis) y comprender verdaderamente el significado simbólico (semántica).

Por su parte, la explicación de qué es sentir una emoción nos conduce inexorablemente hacia un nudo gordiano adicional, que tiene que ver con la emergencia de emociones genuinas (más a allá de su programación y simulación) y la posibilidad de vivencia subjetiva. Esto nos transporta a cuestiones ópticas relacionadas con la naturaleza de la mente y la consciencia (ampliamente debatido por el fisicalismo y el funcionalismo) y nos traslada también hacia una reflexión fenomenológica relacionada con los *qualia* emocionales.

El control de emociones humanas por parte de una IA quizá no planteé tanta dificultad teórica (sí desde luego técnica), aunque se abre un abismo bioético (más bien neuroético) si nos asomamos a las posibles consecuencias bio-psico-sociales de una manipulación de estas características.

Como paso previo a este abordaje filosófico, por el momento meramente apuntado, parece de necesidad el realizar antes una sucinta –aunque suficiente– revisión del concepto tecnológico de IA; igualmente, de los diferentes tipos y principales sistemas operativos de la IA actual y de la que está por venir. Esto ayudará a comprender mucho mejor las dificultades y límites filosóficos que se pretenden explorar y, además, de este modo ratifico el ejercicio de coherencia expuesto en la introducción en relación al espíritu multidisciplinar e integrador del concepto de emocionalidad artificial.

Según la Real Academia Española (RAE), la inteligencia artificial es la “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”⁶. Un criterio útil para iniciar la distinción entre los diferentes tipos de IA puede ser el establecido por el filósofo John Searle, que diferencia la IA débil de la IA fuerte (Searle, 1980):

- IA débil: es la que tenemos en la actualidad y que, pese a sus grandes capacidades computacionales, nunca desarrollará una inteligencia general como la humana.

⁶ Definición de IA según la web de la RAE: <https://dle.rae.es/inteligencia> (consultado el 21/3/21).

- IA fuerte: es la que critica el propio Searle, por irrealizable y conceptualmente imposible, aunque también es la defendida por los transhumanistas más voluntariosos; es decir, la que tendría capacidades cognitivas iguales a las humanas (lo que muchos consideran como una IA general) o incluso superiores (lo cual supondría el advenimiento de una “singularidad tecnológica” en forma de super-IA).

Con respecto a la IA débil, es decir, aquella de la que ya disponemos, podemos utilizar dos sencillas clasificaciones adicionales, muy útiles para diferenciar el estado actual de la cuestión tecnológica al respecto de la IA. Una sería la propuesta por Rusell y Norvig (2019) y la otra la planteada por López de Mántaras y Meseguer (2017). Según los primeros autores, la IA se clasifica en:

- Sistemas que piensan como humanos: son los que automatizan actividades como la toma de decisiones, el aprendizaje o la resolución de problemas. Un ejemplo serían las “redes neuronales artificiales”, capaces de realizar una tarea de forma óptima, como por ejemplo ganar en diferentes juegos como el ajedrez (Deep Blue), las damas (CHINOOK) o el go (AlphaGo)⁷.
- Sistemas que actúan como humanos: son los que imitan tareas básicas propiamente humanas. A esta categoría pertenecen las primeras máquinas de la robótica clásica. La empresa Boston Dynamics⁸ ofrece un excelente muestrario de algunos de los desarrollos robóticos más impresionantes en la actualidad; los robots espaciales diseñados por la NASA (*National Aeronautics and Space Administration*) para la exploración de Marte (como Sojourner, Spirit, Opportunity o Curiosity) son también un buen ejemplo.
- Sistemas que piensan racionalmente: a esta categoría pertenecen las IA que imitan el pensamiento lógico racional del ser humano. Aquí encontramos a los “sistemas expertos”, que tienen cualidades de simulación,

7 Deep Blue es la inteligencia artificial diseñada por IBM (*International Business Machine*) que entre 1996 y 1997 derrotó al ajedrecista ruso Gary Kasparov. CHINOOK es un programa de ordenador basado en IA desarrollado en la Universidad de Alberta. AlphaGo es un programa informático desarrollado por Google DeepMind, basado en el aprendizaje profundo de redes neuronales artificiales, que ganó a varios campeones mundiales al enfrentarse al célebre juego de estrategia de origen chino go. El documental *AlphaGo* (2017) recoge de forma excelente la historia de este hito informático. Para ampliar información al respecto de los grandes éxitos de la IA recomiendo consultar el capítulo 5 del libro *Inteligencia artificial* (López de Mántaras y Meseguer, 2017: 118-144).

8 Web de la empresa donde se muestran sus desarrollos en IA robótica: <https://www.bostondynamics.com/> (consultado el 19/3/21).

planificación, control o monitorización. Dendral o Mycin serían dos buenos ejemplos.

- Sistemas que actúan racionalmente: son los sistemas que emulan de forma racional la conducta humana. Aquí estarían los “agentes inteligentes” capaces de interactuar con su entorno. Un ejemplo destacado sería la sonda espacial Deep Space 1.

Por su parte, siguiendo a los científicos españoles, la IA también se puede clasificar atendiendo a diferentes modelos explicativos: el “simbólico”, el “conexionista”, el de la “computación evolutiva” y el de la “robótica del desarrollo” (López de Mántaras y Meseguer, 2017). Según estos autores, el modelo “simbólico”:

Se basa en el razonamiento lógico y la búsqueda heurística como pilares para la resolución de problemas, sin que el sistema inteligente necesite formar parte de un cuerpo ni estar situado en un entorno real. Es decir, la IA simbólica opera con representaciones abstractas del mundo real que se modelan mediante lenguajes de representación basados principalmente en la lógica matemática y sus extensiones (p. 11).

Esto no quiere decir que los sistemas de IA simbólicos actuales no hayan evolucionado, pues los incorporados en robots sí que pueden interactuar con el medio para alcanzar objetivos. Por otro lado, estaría la IA “conexionista”, inspirada en el funcionamiento del cerebro, la cual:

Se basa en la hipótesis de que la inteligencia emerge a partir de la actividad distribuida de un gran número de unidades interconectadas que procesan información paralelamente. En la IA conexionista estas unidades son modelos aproximados de la actividad eléctrica de las neuronas biológicas (p. 12).

Este modelo conexionista, cuya principal novedad es la operatividad paralela en vez de secuencial, ha servido para la creación de “redes neuronales artificiales” (como las antes citadas de AlphaGo), mucho más eficaces para facilitar el aprendizaje de los sistemas de IA que la propuesta del modelo simbólico. Los antecedentes de esta aproximación se encuentran en 1943, en un primer estudio en el que se simuló el funcionamiento de una sola neurona artificial (McCulloch y Pitts, 1943); o en los descubrimientos de los premiados al Nobel en 1959, David Hubel y Torsten Wiesel, que demostraron el modo jerárquico en el que trabaja el cerebro, lo que facilitó la posterior creación de “redes neuronales convolucionales y profundas”, es decir, las que trabajan en capas que procesan diferentes aspectos de la información⁹ (citado en Latorre, 2019: 115).

⁹ Como ejemplo práctico de esto último tenemos el sistema de morfo-reconocimiento, muy utilizado en la actualidad para la identificación de rostros de personas.

Otra aproximación novedosa es la de la “computación evolutiva”, la cual imita los procesos de evolución biológica mediante una especie de mecanismo de selección natural (conocido como “algoritmo genético”) en la que hay una competencia entre diferentes algoritmos. De una forma similar, las “redes neuronales adversarias” participan en una auténtica carrera de armamentos mediante complejas retroalimentaciones, lo cual ha servido, por ejemplo, para la creación de rostros humanos virtuales (Latorre, 2019).

Por último, tenemos el modelo de la “robótica del desarrollo”, el único de los tres que exige necesariamente una corporeidad robótica para su implementación. De hecho, se basa en la interacción de la máquina con el entorno, gracias a distintos sensores capaces de imitar el funcionamiento de los sentidos especiales (como la vista o la audición) y somatosensoriales (como el tacto o la propiocepción) del ser humano, para que reciba inputs externos y construya después una representación interna del mundo. Así, la IA estaría “situada” en su entorno, lo cual facilita su desarrollo cognitivo puesto que, como algunos autores sostienen, “el cuerpo da forma a la inteligencia” (a esto podríamos llamarlo “cognición situada”) (López de Mántaras y Meseguer, 2017: 14-15). Tras la captación de estímulos sensoriales, la IA debe ser capaz de procesarlos para, con ello, producir después los *outputs* correspondientes, por ejemplo, en forma de movimientos corporales. A este modelo se le pone el apellido “del desarrollo” porque la mejora cognitiva de este sistema se da cronológicamente, en la medida en que el sistema de IA va interaccionando de forma paulatina con su entorno y aprendiendo de (y con) él. Incluso, puede crearse un algoritmo que seleccione el mejor método de aprendizaje, siendo esto último una especie de, en palabras de Latorre (2019: 113), “meta-aprendizaje”. Algo similar a esto se hizo con AlphaGo, programa que se entrenaba con versiones cada vez más mejoradas de sí mismo. En relación a este modelo procesual, otra estrategia similar consiste en el “aprendizaje por refuerzo”, que premia a la máquina cuando realiza lo deseado (Latorre, 2019).

Tras esta breve revisión del estado tecnológico de la cuestión¹⁰, en lo que viene a continuación sobre reconocimiento, simulación, manipulación y vivencia de emociones por parte de la IA, diferenciaré siempre entre lo ya realizable tecnológicamente (IA débil) y lo que la especulación científica y el transhumanismo más optimista propone (IA fuerte).

10 Para una mayor profundización de los aspectos técnicos de la IA, aunque con una clara orientación práctica (y muy orientada también hacia temas tratados en este ensayo, como la filosofía de la IA), se recomienda consultar la obra *Inteligencia Artificial* (Boden, 2017).

2.1. IA QUE RECONOCE EMOCIONES

Haciendo una aclaración terminológica, por “reconocer” me refiero a lo que actualmente hace la IA cuando recibe información, es decir, procesar *inputs* gracias a un algoritmo diseñado por un programador. Parto, pues, de los postulados searleanos de la IA débil y de los sistemas capaces de hacer hoy día esta operación, como por ejemplo los sistemas simbólicos. En el extremo opuesto de un eje gnoseológico imaginario estaría el término “conocimiento”, o “comprensión”, que es lo que hace una inteligencia humana o hará la pretendida IA fuerte¹¹, bien como IA general (con una inteligencia equivalente a la humana), bien como super-IA (con una inteligencia sobre-humana).

Las críticas de Searle a la IA fuerte, en su celebrado experimento mental de “la habitación china”, puede ser aplicado también aquí, al diferenciar entre computar información (en este caso, emocional) y conocer o comprender verdaderamente el significado de las emociones. Lo primero es lo que hace una IA actual (IA débil); lo segundo, considera el pensador americano, nunca será posible. La razón, según él, es que el ordenador manipula símbolos formales, pero no información, al menos en el sentido de información con contenido semántico y significado (Searle, 1980, 1987). O, dicho de otro modo, la sintaxis del algoritmo que ha programado el informático no es suficiente para lograr una verdadera comprensión semántica e intencionalidad que lleve a la máquina a entender las emociones humanas como las entendemos nosotros.

Ha habido autores que han criticado las posturas de Searle. Por ejemplo, Rapaport (1988), con su experimento mental de “la habitación coreana”, donde plantea la posibilidad de una sintaxis con cierto valor semántico. Para otros detractores, como Block y Fodor (1972), la clave de una IA fuerte podría estar en la imitación conexionista de nuestro cerebro (que funciona de manera paralela), en vez de utilizar un sistema simbólico. Mejorando esta forma de operar se alcanzarían una serie de ventajas, como la capacidad de almacenar información de forma distribuida o una mayor velocidad de procesamiento y mayor tolerancia a los fallos.

En la actualidad existen sistemas operativos capaces de reconocer matices prosódicos y emotivos del lenguaje hablado (como el algoritmo japonés Empath), o de identificar expresiones corporales vinculadas a la comunicación no verbal (como el *software* Emotient de Apple). Ambos métodos están siendo explotados

11 Newell y Simon (1976: 113-26) defendieron la posibilidad de una verdadera IA fuerte en base a su hipótesis de los Sistemas de Símbolos Físicos (SSF) al reconocer que un SSF “tiene los medios necesarios y suficientes para la acción general inteligente”. El cerebro humano y un sistema artificial que lo pueda imitar serían ejemplos de SSF.

por empresas en campos tan diversos como la domótica, los videojuegos, la atención al cliente o la robótica. En este sentido, los algoritmos de la IA son (o podrían ser) capaces de reconocer determinados elementos expresivos externos de las emociones, que encontramos en el lenguaje hablado (determinadas palabras “clave” que pueden funcionar como señales, o ciertas frecuencias sonoras que correlacionan con la prosodia emocional) o en la posición corporal y en la gestualidad facial (las diferentes expresiones gestuales, controladas por los músculos de la expresión facial, que son un correlato externo de ciertos estados emocionales)¹².

Hay, adicionalmente, otra forma potencial de reconocer emociones, virtualmente posible aunque con considerables limitaciones en la actualidad (desde luego técnicas, aunque también conceptuales y por supuesto éticas). Me refiero al empleo de la neurotecnología (en forma de neuroescaneo, electroencefalografía o a través del uso de interfaces cerebro-computador) la cual, operada por una IA, sería capaz de establecer correlatos neurofisiológicos de las emociones (Torres et al., 2020). Sin embargo, para la filósofa Kathinka Evers (2011: 51), “leer” la mente (o las emociones) de este modo no será nunca totalmente eficaz, pues considera que “sería ilegítimo inferir que la cartografía contextual de un proceso neuronal consciente puede o debe suministrar informaciones sobre el contenido de este proceso”. La autora sueca considera que un mapa del cerebro no es en ningún caso un mapa del pensamiento (o de un estado mental o emocional particular), debido a razones como la gran variabilidad inter e intraindividual. En este sentido, los correlatos entre circuitos neuronales y pensamientos (incluyo también las emociones y sentimientos) particulares no son iguales en todos nosotros, ni siquiera en un mismo individuo en momentos distintos (pues cada vez que sentimos, pensamos, imaginamos o recordamos algo cambia la circuitería neuronal). Dicho lo cual, aunque hoy día no exista un dispositivo capaz de leer la mente (y las emociones), para los transhumanistas y funcionalistas más transgresores ninguna ley

12 Los sistemas de codificación facial FACS (*Facial Action Coding System*), de Ekman y Friesen, y MAX (*Maximally Discriminative Facial Movement Coding System*), de Izard, son ampliamente utilizados para la caracterización gestual de las emociones, dada la cierta universalidad en la expresión facial de las emociones básicas o primarias. El reconocimiento de cambios fisiológicos periféricos podría ser otra forma a través de la cual un sistema de IA podría ser capaz de obtener información emocional de un humano (aunque esto sería más invasivo) dado que hay ciertos patrones de cambio visceral asociados con estados emocionales susceptibles de ser medidos. El registro de la actividad cardiovascular y respiratoria, de la actividad electrodermal o de la concentración de determinadas hormonas en sangre (como el cortisol, que correlaciona con el estrés), sería una forma de obtener información fisiológico-emocional (Fernández-Abascal et al., 2010). Lo interesante de todo esto es que, aunque una máquina no pueda “comprender” el significado más profundo de las emociones humanas, desde luego podría tener acceso a una ingente cantidad de información emocional, incluso mayor de la que un humano puede extraer de un congénere (sin ayudas artificiales). En todo caso, aunque posibles, todas estas intervenciones carecen de una aplicación directa, clara y útil en la actualidad.

física impediría registrar la conexión de cada neurona de nuestro cerebro (a pesar de la extrema dificultad), pues, al fin y al cabo, siempre que pensamos o sentimos algo se activa un determinado proceso funcional con un sustrato neurobiológico susceptible de ser comprendido (y registrado). Y, si llegásemos a emular artificialmente un cerebro con la ayuda de la IA (a esas copias Hanson [2017: 162] las llama “ems”), nada impediría acceder a la información que contenga.

Para lo que hoy día puede hacer (o presumiblemente hará) una IA con respecto al reconocimiento de emociones, lo único que interesa, desde la categorización emocional antes citada de Kleinginna y Kleinginna (1981), es que las emociones se puedan conceptualizar como entidades fisiológicas y conductuales (en la medida en que la IA reconoce correlatos morfo-fisiológicos externos, como la gestualidad, la dilatación pupilar, la sudoración o el temblor de voz; o internos, como los cambios en el flujo sanguíneo en regiones neuroanatómicas del sistema límbico emocional o en los patrones electrofisiológicos del cerebro) o como entidades expresivas (desde el momento en que dicho correlato morfo-fisiológico comunica y transmite información). Ambas son las condiciones suficientes y necesarias para que una IA actual pueda reconocer de forma básica las emociones (al menos las primarias). Y digo “básica” porque los datos empíricos obtenidos de esta forma no supondrían en ningún caso una información inequívoca del estado emocional que experimenta un sujeto particular. Esto es así debido a varias razones (y limitaciones de la IA): (1) que puede haber convergencia fisiológica y conductual entre varias emociones; (2) que la socialización y la cultura contribuyen a determinar el modo en el que las emociones son expresadas; (3) que otras capacidades mentales y cognitivas, como los deseos, las creencias, los recuerdos o las expectativas podrían influir en el patrón emocional que se evalúa; (4) que puede haber fingimiento emocional; y (5) que una IA pudiera reconocer el “qué” no es lo mismo que reconocer el “cómo”. Es decir, comprender verdaderamente el estado afectivo del interlocutor es algo fenoménicamente imposible (en realidad lo es en cada uno de nosotros, cada vez que intentamos comprender los afectos de los demás).

Pese a las dificultades, tal y como se dijo más arriba, el reconocimiento de emociones por parte de la IA puede tener varias aplicaciones. La evaluación afectiva o la detección de mentiras serían algunas de las más controvertidas. En relación a lo primero, podría ser de utilidad para empresas de publicidad¹³ o recursos humanos, interesadas en conocer el estado emocional de un cliente o potencial trabajador, respectivamente. Con respecto a lo segundo, su justificación

13 El neuromarketing es una disciplina que utiliza las aportaciones de la neurociencia para la dirección de las organizaciones, el estudio de las necesidades y comportamiento del cliente, el *targeting* y el posicionamiento, los negocios, las estrategias de producto, marca y precios, los canales de ventas, las finanzas o las comunicaciones (Braidot, 2005).

técnica se basa en que es sabido que al mentir se producen perturbaciones emocionales vehiculadas fisiológicamente (Gazzaniga, 2005). Su utilización en juicios o en el reconocimiento de sospechosos es una potencial herramienta no exenta de limitaciones técnico-conceptuales (como ya se ha dicho, correlacionar estados emocionales con otros estados mentales actuales o disposicionales), bioéticas (la violación de la “privacidad mental” y la “libertad cognitiva” [Farah, 2005, Sententia, 2004]) y legales (el quebranto de los “neuroderechos [Ienca y Andorno, 2017]).

2.2. IA QUE SIMULA EMOCIONES

La expresión de emociones se puede programar (Latorre, 2019), bien en forma de gestos o acciones locomotoras, bien a través de la simulación del lenguaje natural. Con respecto a lo primero destaca Kismet, robot pionero desarrollado en el Instituto Tecnológico de Massachusetts (MIT), o la empresa Hanson Robotics¹⁴, que ha diseñado varios andróides que simulan expresiones faciales de alegría, tristeza o enfado. En referencia al lenguaje natural, es posible que una IA, como la aplicada a los *chatbots* o a los asistentes inteligentes como Siri o Alexa, sea capaz de mantener una sencilla conversación de contenido aparentemente emocional¹⁵. Incluso, el programa MINDER, que imita a una enfermera especializada en el cuidado de bebés, es capaz de simular los aspectos funcionales de la ansiedad al interactuar con los pequeños (Boden, 2017).

Las dificultades searleanas discutidas más arriba vuelven aquí a hacer acto de presencia. Según los postulados de la IA débil, el robot emocional no sabría lo que está simulando, tan sólo se limitaría a ejecutar un programa siguiendo las prescripciones de su algoritmo. Con la denominada “robótica del desarrollo”, la IA situada en el mundo (gracias a sensores que le permitan interactuar con su interlocutor humano), podría aprender a mejorar su simulación emocional, pero nunca llegaría a comprender verdaderamente el objeto de su acción.

Pero, ¿por qué habríamos de querer que una IA virtual o robótica simule emociones? Muchos potenciales clientes de sistemas de IA domésticos se sentirán

14 Web de la empresa: <https://www.hansonrobotics.com/> (consultado el 12/3/21).

15 El programa ELIZA, un *bot* conversacional diseñado en el MIT en la década de los sesenta, fue pionero en este sentido. Dicho sistema de IA era capaz de imitar la conversación de un psiquiatra interactuando con sus pacientes. Su objetivo era demostrar que podía pasar la prueba del test de Turing.

más cómodos y empáticos¹⁶ con una IA, robot o androide humanizado, cuyo aspecto, comportamiento y voz sea casi indistinguible de la de un humano. A dicha mejora empática está contribuyendo enormemente la antropomorfización de la robótica, por ejemplo, gracias a la mayor coordinación motora de los androides (con el bipedismo como uno de sus máximos exponentes) o con la utilización de materiales sintéticos que imitan la piel, las uñas, el cabello o los ojos humanos. Para la mejora de las habilidades locomotoras, la ingeniería robótica ha centrado sus esfuerzos en la imitación anatómico-funcional de nuestros sistemas sensoriales y motores. Los robots más avanzados disponen de todo tipo de sensores, como los inerciales (que imitan nuestra cinestesia y propiocepción) en forma de giroscopios o acelerómetros, o los posicionales, visuales y auditivos, en forma de sistemas láser y radar, cámaras y micrófonos que recogen distancias, imágenes y sonidos. Un excelente ejemplo de la combinación de algunas de estas tecnologías lo encontramos en Atlas, el robot de Boston Dynamics que tiene la increíble habilidad de practicar *parkour*.

Sin embargo, en el polo opuesto estarían los detractores de la humanización robótica (incluidas las emociones); personas que prefieren que la IA tenga un aspecto artificial (por su voz, morfología o movimientos), que en poco recuerde a un ser humano. Esta discrepancia es explicada por la “Teoría del valle inquietante” de Masahiro Mori (1970), la cual señala que nos resulta desagradable que un robot sea demasiado antropomorfo (citato en Latorre, 2019: 154).

En la actualidad existen iniciativas robóticas que participan en importantes relaciones interactivas, como es el cuidado de las personas mayores: tal es el caso del robot Zora¹⁷. Estas máquinas pueden resultar muy útiles para ofrecer cuidados y paliar la soledad, y en no demasiado tiempo tendrán un papel destacado en la terapia médica. Conforme avancen las habilidades motoras y las capacidades cognitivas de estos dispositivos, no será de extrañar que sus dueños lleguen a tenerles un afecto creciente y verdadero. Un riesgo asociado será el abandono de los mayores por parte de sus familiares bajo la justificación de que el cuidado robótico es más que suficiente. Como aventura el experto en IA José Ignacio Latorre (2019: 155), “una persona mayor morirá dando la mano a un robot”.

16 Ha sido ampliamente demostrado que las neuronas espejo son clave en la imitación del otro y en los procesos de empatía. En este sentido, se ha sugerido que son el sistema neurobiológico responsable de, en cierto modo, salvar el “problema de las otras mentes” (cómo acceder y comprender la mente de los otros), facilitando que la intersubjetividad sea posible (Iacoboni, 2009).

17 Web de la empresa: <https://www.zorarobotics.be/> (consultado el 2/4/21).

2.3. IA QUE MANIPULA EMOCIONES

La manipulación de emociones por parte de la IA es una posibilidad técnicamente viable (aunque con muchas limitaciones) que podría aparecer en el futuro. Los interfaces cerebro-computador tal vez sean la forma de materializarlo.

El control cibernético-emocional se explicaría por el hecho de que las emociones tienen una base neurobiológica cada vez mejor conocida: desde la neuroquímica a la neuroanatomía, la evidencia científica que correlaciona la vivencia subjetiva emocional con su base cerebral es cada vez más fuerte (Esperidiao-Antonio et al., 2017; Fernández-Abascal et al., 2010). De hecho, en la actualidad existen múltiples aproximaciones en la terapia neuro-psiquiátrica para el tratamiento de patologías que cursan con desórdenes emocionales, como por ejemplo la ansiedad, la depresión o la esquizofrenia. A los tratamientos farmacológicos habría que incluir otras herramientas intervencionistas como la neurocirugía o el uso de dispositivos neuromoduladores (como la estimulación magnética transcraneal [EMT], la estimulación eléctrica transcraneal [EET], la terapia electroconvulsiva [TEC] o la estimulación cerebral profunda [ECP]).

Desde el punto de vista filosófico hay dos grandes núcleos de reflexión: partiendo de la filosofía de la mente, el debate sobre la posibilidad de manipular el contenido, la intensidad o la valencia (positiva-negativa) emocional. Desde la bioética (más en concreto, desde la neuroética aplicada), la preocupación por los límites de una intervención terapéutica o de neuromejora que podría interferir en aspectos clave de la condición humana como la identidad o el libre albedrío (Biscaia, 2021).

Con respecto a lo primero recordemos que, en el primer apartado dedicado a las emociones biológicas, se indicó que éstas compartían algunas cualidades con los estados mentales, como por ejemplo la intencionalidad, en el sentido de que las emociones instancian representaciones internas (pues se relacionan con estados mentales), frecuentemente se dirigen hacia un objeto concreto (se ama algo o a alguien en particular) y presentan contenido (aunque, según Broncano [2001], no conceptual). A la pregunta de si se podría manipular la aparición, contenido, direccionalidad y grado de dicha emoción, por supuesto no hay respuesta neurobiológica totalmente satisfactoria (por más que algunos neurofármacos y neurocirugías puedan lograr algo de esto), aunque ninguna ley física impide que dicho escenario pudiera presentarse, al menos de forma parcial. Sentir una emoción y no otra, dirigida hacia un objeto y no otro, en un mayor o menor grado es, nomológicamente hablando, susceptible de ser manipulado, aunque la complejidad de los fenómenos cerebrales que acontecen lo hace difícilmente regulable. Además, en la medida en que lo emocional es muchas veces inseparable de lo cognitivo, la supuesta manipulación debería intervenir igualmente otras áreas

neuroanatómicas y circuitos neuronales responsables de dichos procesos cerebrales superiores. Por último, no se debe olvidar la importancia de determinadas claves externas que sirven como disparadores de las emociones, además de las experiencias pasadas, de las expectativas futuras o del contexto socio-cultural, que también formarían parte de la ecuación emocional.

Si bien estas últimas cuestiones filosóficas son de indudable interés, no se insistirá más en ellas por no ser ámbito de la filosofía de la IA de la que trata este ensayo sino, más bien, del debate transhumanista (neuro)tecnológico y de la neuroética aplicada y fundamental¹⁸.

2.4. IA QUE SIENTE EMOCIONES

He dejado para el final el punto de mayor complejidad; también el más controvertido desde casi cualquier punto de vista tecno-científico y filosófico que se considere. Me refiero a la posibilidad de que, algún día, una IA fuerte pueda sentir emociones de forma genuina, al modo de como lo hacen los seres biológicos.

Para un mejor tratamiento de este último apartado dividiré el argumentario en cuatro secciones que se construyen desde las siguientes cuatro preguntas de partida: (2.4.1.) ¿es conceptualmente posible que una IA tenga estados mentales y emocionales? (2.4.2.) ¿Podría saber ella, también nosotros, si los está experimentando? (2.4.3.) ¿Sería de utilidad que una IA sienta emociones? (2.4.4.) ¿Cuál sería el impacto social de una emocionalidad robótica emergente?

2.4.1. *Emergentismo emocional en la IA*

El debate en torno a la posibilidad de la emergencia de emociones en un ser artificial nos retrotrae a la clásica discusión sobre la naturaleza de los procesos mentales y al problema mente-cerebro. Recordemos que ya en el primer apartado concluimos con Broncano (2001: 1) que las emociones eran “territorios intermedios en la mente”, es decir, fenómenos con ciertas características propias de los estados mentales, por lo que el debate de lo mental podrá extrapolarse, con matizaciones, al territorio emocional. Dicho debate propio de la filosofía de la mente, de la gnoseología y de la filosofía del lenguaje gira en torno a cuestiones

¹⁸ Para una lectura de algunos de estos temas se recomiendan las obras *Transhumanismo. La búsqueda tecnológica del mejoramiento humano* (Diéguez, 2017) y *Neuroética. Retos para el siglo XXI* (Levy, 2014), o acudir a la revisión sobre neurotecnología y neuromejora de Biscaia (2021).

óntico-epistémicas al respecto de la existencia, naturaleza y conocimiento de la mente y de los estados mentales. En líneas generales, a fin de poder seguir mi argumentación, resumo sucintamente las principales posturas al respecto, siguiendo lo recogido por los tratados de Hierro-Pescador (2005) y Martínez-Freire (2002).

A groso modo, hay posturas que niegan de forma radical la existencia de lo mental: bajo este ideario encontramos el materialismo o monismo ontológico de algunos fisicalistas, especialmente aquellos consideramos como eliminacionistas por reconocer que únicamente existe lo material (el cerebro)¹⁹; igualmente, se situaría aquí cierto tipo de conductismo radical. Además, hay posturas que niegan la sustanciación de lo mental, aunque le reconocen una cierta existencia lingüística. Por su parte, hay posiciones que aceptan la sustanciación ontológica o funcional de lo mental: independientemente de lo físico (con o sin leyes que unen al conjunto mente-cerebro)²⁰; aunque, también, vinculado nomológicamente con lo físico (una suerte de epifenomenalismo de lo mental desde lo cerebral). Finalmente, el funcionalismo considera que los estados mentales son estados funcionales cuyo órgano es el cerebro (aunque para algunos funcionalistas el sustrato material podría ser cualquier otro, distinto del tejido nervioso).

Algunos de los fisicalistas más destacados, cada cual con sus propias matizaciones al respecto del tema que nos ocupa, son: Carnap (1931, 1932), el cual reduce el lenguaje psicológico al físico, pues la realidad material es la única existente y lo mental es, en todo caso, algo disposicional; o Rorty (1982), quien niega lo mental, admitiendo que la mente no es siquiera un estado natural; o Feigl (1958), cuyo monismo ontológico descansa, hay que admitirlo, en un cierto dualismo epistemológico sobre el modo de conocer lo psicológico-cerebral, que culmina en su “Teoría del doble acceso”; o Smart (1959), quien identifica estados mentales con estados físicos por la dificultad de engranar leyes psicofísicas que unan los procesos de ambos estados; o Lewis (1994), para quien los estados mentales son idénticos a los estados neuronales. Por su parte, el conductismo ontológico (con John Watson como principal representante) únicamente reconoce a la conducta (derivada de la interacción estímulo-respuesta) como digna de estudio, puesto que niega la existencia de estados mentales²¹.

19 En el fisicalismo reduccionista lo mental se reduce a lo cerebral, produciéndose una identificación sustancial (que no descriptivo-lingüística) mente-cerebro.

20 Por ejemplo, el dualismo cartesiano o el dualismo interaccionista del fisiólogo John Eccles y del defensor de la “Teoría de los tres mundos”, el filósofo Karl Popper. Para una revisión de sus ideas consultar la obra *El yo y su cerebro* (Barcelona: Labor Sa, 1993).

21 El conductismo metodológico (con Burrhus Skinner como mayor exponente), a diferencia del radical, no niega de forma tajante los estados mentales, pero reconoce que son inobservables y que por tanto no se pueden estudiar de forma directa.

El conductismo lógico de Ryle (1967) considera que el dualismo cartesiano es un error lógico y categorial. Además, defiende que los procesos mentales no son actuales sino disposicionales para la conducta. En una línea similar, Ludwig Wittgenstein reconoce que el problema de la mente es un problema lingüístico²². Por su parte, el monismo anómalo o fisicismo de ejemplares de Davidson (1980) plantea que habría dos tipos de descripciones –física y mental– para la misma sustancia, aunque no hay leyes correlacionales entre lo cerebral y lo mental.

Entre los más reconocidos funcionalistas se encontraría Putnam (1960), con su “Teoría del autómatas probabilístico” y su analogía con los estados de máquina de Turing, donde hace una comparación *cerebro-hardware* y *mente-software*. El filósofo estadounidense sostiene que los estados mentales no son estados del cerebro ni disposiciones para la conducta, sino estados funcionales que operan según una cadena causal; o Fodor (1968, 1981), quien propone una “Teoría representacional de la mental”, donde habría un lenguaje del pensamiento formado por representaciones mentales que se relacionan funcional y computacionalmente de forma causal. Este autor afirma que no deben identificarse estados mentales (lo funcional) y cerebrales (lo estructural), pues operan a diferentes niveles (se refiere a ello como macropropiedades y micropropiedades, respectivamente). Por su parte, el emergentismo de Searle (1987) sostiene que los procesos mentales no son independientes de los físicos, aunque no se reducen a los procesos cerebrales. Serían procesos o propiedades que emergen desde lo cerebral, manteniendo entre ellos una relación de implicación lógica. Lo mental (sea algo meramente físico o no) causaría la conducta, y las macropropiedades psicológicas se realizarían gracias a las micropropiedades del soporte material.

Tras este sucinto repaso de las principales posturas (a riesgo de que, por la brevedad expositiva, se puedan hacer de forma legítima cualesquiera matizaciones y ampliaciones), se puede concluir que las tesis funcionalistas son las que mejor pueden soportar la emergencia de propiedades mentales (y, de ahí, también de estados emocionales) en una IA. Dicho funcionalismo propone que lo importante para la aparición de estados mentales son los estados y relaciones funcionales, no el soporte material que sustenta lo mental (así, una máquina tiene estados lógicos (*software*) y estructurales (*hardware*), como un humano tiene mente y cerebro, respectivamente). En este sentido, una máquina podría perfectamente ser capaz de poseer estados mentales, toda vez que, al menos (y no es poca cosa), imite los estados funcionales de un cerebro. Desde luego, es necesario un soporte material para ello, pero éste, a diferencia de lo que sostienen algunas tesis fisicalistas radicales, no tiene porqué ser necesariamente un cerebro, sino cualquier otro

22 Para un adecuado acercamiento a dicha postura consultar el capítulo 4 del tratado de filosofía de la mente de Hierro-Pescador (2005).

material. Por tanto, según el funcionalismo, si una IA tuviera una configuración causal-funcional semejante a la de un cerebro, nada impediría que pensase (y sintiese). Aunque, desde luego, ninguna ley física impide que otros posibles estados funcionales distintos a los cerebrales, que ni siquiera conocemos, puedan generar emergencia o superveniencia mental²³.

2.4.2. *Test de Voight-Kampff*

El filósofo Thomas Nagel (1974) publicó un artículo titulado *¿Cómo es ser un murciélago?* que, para el caso, puede ser de utilidad. En él, se plantea la dificultad epistemológica de conocer la vivencia subjetiva de los estados mentales del otro. En el núcleo de su debate se encuentra la clásica discusión al respecto de los *qualia* (el modo como algo se experimenta) y, también, del “problema de las otras mentes” (cómo acceder –y comprender– a la mente de los otros). Chalmers (1995) ha querido referirse a esta dificultad explicativa sobre “cómo es ser uno” como el “problema difícil de la conciencia”.

Parece claro que para que una máquina experimente emociones subjetivamente debe poseer algún tipo de conciencia. Al menos, como mínimo, equivalente a la de un animal. Block (1994a, 1994b y 1995) diferencia varios tipos de conciencia: fenoménica (sería la experiencia subjetiva), de acceso (en la que hay representación de un contenido representacional para el control racional de la acción o el habla), monitora (responsable de la introspección y de los pensamientos de alto nivel sobre un estado mental) y conciencia del yo. Para que una IA experimentase emociones básicas se exigiría, de entre todos estos tipos (como mínimo), que tuviera conciencia de tipo fenoménica²⁴, pues es la que mejor correspondencia establece con la categorización subjetivo-afectiva de las emociones descrita en el apartado 1. Lo que parece evidente es que, de ser posible la emergencia o superveniencia emocional (de forma intencional, mediante diseño y programación o, como ha especulado la ciencia-ficción, de forma “espontánea”, al modo de otros tantos cambios morfo-funcionales surgidos en el mundo vivo durante el transcurso de la evolución), posiblemente surgiría de forma paulatina: en primer lugar podrían manifestarse emociones básicas o primarias (como el miedo o la alegría),

23 La superveniencia es un concepto ampliamente utilizado por Davidson (1973), entre otros autores, en el marco de su “Teoría psicofísica de la identidad”, según el cual no es posible que haya dos eventos iguales en los aspectos físicos pero que difieran en algún aspecto mental.

24 Se conocen casos de conciencia fenoménica sin otros tipos de conciencia, como la de acceso. Lo vemos, por ejemplo, en la prosopagnosia (tipo de agnosia visual en la que el sujeto es incapaz de reconocer rostros familiares; a veces, incluso, es incapaz de reconocerse así mismo).

que son las más sencillas (en el caso biológico, surgieron pronto en la filogenia y tienen una extensión universal entre los vertebrados). Y quizá, más adelante, podrían aparecer emociones secundarias o “autoconscientes” (como la vergüenza o el orgullo), aunque éstas exigirían, como su nombre indica, que se organizaran alrededor de la unidad de conciencia que confiere un “yo”.

El título de este apartado, “Test de Voight-Kampff”, hace mención a una prueba ficticia que aparece en la película de ciencia-ficción *Blade Runner* (1982), ideada para la detección de replicantes (una suerte de seres artificiales). Dicho test, al modo del famoso test de Turing, era utilizado para saber si una máquina poseía emociones, al basarse en la detección de respuestas fisiológicas y conductuales asociadas. El problema aquí es que las experiencias subjetivas son fenoménicamente intrasferibles, y sólo podríamos inferir estados emocionales genuinos en una IA mediante conductas o respuestas correlacionales (serían como “criterios”²⁵ de la presencia de dichos estados), analogías e inferencias. No se quiere decir con esto que sea nomológicamente imposible que una super-IA experimente emociones, sino que desde un punto de vista gnoseológico no tendríamos forma de estar seguros. Aunque esto no debe ser mayor impedimento epistemológico que el que, de facto, se produce con respecto al conocimiento de los estados emocionales de otros seres vivos no humanos o, incluso, de nuestros propios congéneres, pues, al fin y al cabo, el núcleo del “problema de las otras mentes”²⁶ bien se puede aplicar a la cuestión que ahora nos ocupa: en su vertiente puramente epistemológica, el acceso a nuestros estados mentales (emocionales) es radicalmente diferente del acceso a los estados de los demás; en su vertiente conceptual, la dificultad estriba en cómo es posible formar un concepto de los estados mentales de los otros. En definitiva, podemos concluir que no sería posible tener un conocimiento directo de los estados mentales de cualesquiera otras mentes, sean animales o robóticas, aunque esto no impide tener un cierto grado de conocimiento (indirecto) sobre, al menos, la existencia y modo de operar de tales estados.

A pesar de las dificultades (para muchos) aparentemente insalvables en relación al desarrollo de capacidades cognitivas y emocionales similares a las humanas, algunos científicos y pensadores creen que en el futuro será posible alcanzar lo que se conoce, en palabras de Vernor Vinge (1993), como la

25 Sin duda, el filósofo Ludwig Wittgenstein es quien ha hecho las mayores aportaciones con respecto al concepto de “criterio”. No obstante, para muchos pensadores, el emplear los criterios como prueba de la existencia de estados mentales es totalmente deficiente, dado el abismo explicativo entre la conducta y los estados internos que la sustentan. Sin conexión conceptual y en ausencia de inferencias inductivas, parece difícil sostener con pleno derecho esta conexión (McDowell, 1982).

26 La *Stanford Encyclopedia of Philosophy* (2019) ofrece una magnífica exposición en relación al concepto de las “otras mentes” (<https://plato.stanford.edu/entries/other-minds/>) (consultado el 17/7/21).

“singularidad tecnológica”. Este nuevo estadio supondría, en su versión débil, la aparición de una IA general equivalente a la humana (Tegmark [2018] y otros expertos lo consideran altamente posible para mediados del presente siglo) y, en su versión fuerte, significaría el advenimiento de una super-IA con capacidades cognitivas superiores a las nuestras (Kurzweil, 2012). Para justificar la llegada de dicha singularidad, el autor se basa en la conocida como “Ley de Moore” (1965), que él adapta en forma de su “Ley de rendimientos acelerados”: ambas, en todo caso, se basan en el crecimiento exponencial de la tecnología cibernética que se observó en los primeros años de la era informática. No obstante, para superar el estancamiento que parece haberse producido en los últimos tiempos (Waldrop, 2016), el científico sostiene que se tendrá que dar un cambio de paradigma tecnológico una vez alcanzada la actual fase de meseta (Kurzweil, 2012), lo cual podría llegar cuando el primer sistema super-inteligente fuese capaz de perfeccionarse a sí mismo y/o de fabricar otras IA iguales. En esta línea, la utilización de nuevos lenguajes lógicos (lógica trivalente, intuicionista y de la vaguedad) que fuesen capaces de programar el “sentido común”, el empleo de sistemas de IA conexionista (que operen de forma paralela y distribuida) o la utilización de nuevos algoritmos de aprendizaje podría acelerar el progreso²⁷. Y, sobre todo, la llegada de la computación cuántica podría ser ese esperado hito que esperan los singularistas más voluntariosos: aunque todavía incipiente, este sistema trabaja con “qubit” (del término inglés *quantum bit*), lo que contempla la superposición de estados cuánticos (1, 0, o los dos a la vez). Además, a este elemento de complejidad se le suma el entrelazamiento cuántico, lo que multiplica exponencialmente (millones de veces) la capacidad de computación frente a cualquier ordenador actual (Diéguez, 2017). En definitiva, aunque muchos científicos ortodoxos prefieren no hablar de singularidad tecnológica, pues lo consideran improbable dadas las actuales dificultades técnicas (Stone y Hirsh, 2006)²⁸, el experto nacional en transhumanismo, Antonio Diéguez (2017: 77), considera que “no hay razones irrefutables para pensar que la creación de una super-inteligencia artificial no es ni será jamás posible”. Por su parte, el experto en IA Max Tegmark (2018) indica que no hay leyes físicas que impidan el surgimiento de una super-IA, y plantea un escenario al que deno-

27 Un interesante proyecto de conciencia artificial es LIDA (*Learning Intelligent Distribution Agent*). Es, a su vez, en palabras de Boden (2017: 123): (1) “un modelo conceptual (una teoría computacional expresada verbalmente) de conciencia (funcional)” y (2) “una implementación parcial y simplificada de ese modelo”. LIDA funcionaría de manera distribuida, con varios subsistemas trabajando en paralelo y de forma jerarquizada (lo que se conoce como *Global Work-space Theory*), es decir, imita el funcionamiento del cerebro según la “Teoría neuropsicológica” desarrollada por Bernard Baars (1988).

28 Así lo señala un informe de la Universidad de Stanford, titulado *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence* (2016). https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf (consultado el 1/2/2021).

mina “3.0”, con máquinas cuyo *hardware* sería auto-replicativo y cuyo *software* produciría mejoras recursivas.

2.4.3. *Mente caliente VS mente fría*

La película de ciencia-ficción *Morgan* (2016) plantea la pertinencia de crear una IA general con auténtica capacidad emocional. Para algunos autores, desarrollar una “mente caliente”, es decir, programar un algoritmo emocional, podría ser útil y necesario para desarrollar una conducta ética en la IA (Latorre, 2019). En este sentido, tal y como se comentó más arriba, ya existe un precedente tecnológico (el programa en inteligencia artificial MINDER) capaz de regular sus respuestas en el cuidado de bebés en base a la simulación (funcional) de estados de ansiedad. La idea subyacente es que la programación de emociones también podría ser ventajosa en la toma de decisiones (pues, como se discutió ampliamente en el primer capítulo, existe un continuo cognitivo-emocional). La película *Morgan* (2016), no obstante, nos señala los riesgos asociados a esta posibilidad: la aparición de emociones y sentimientos negativos, como el miedo o la hostilidad, que podrían volverse contra sus creadores.

Siguiendo la “Teoría funcional del sistema emotivo” de Oakley y Jonhson-Laird (1987), las emociones ayudan a evaluar aspectos relevantes del mundo (interno o externo) de forma rápida y global, para que la respuesta que se requiere no esté limitada por la lentitud de procesos cognitivos que exigen la participación de la conciencia y la memoria. Las emociones activan un plan de acción rápido y eficaz orientado a la deliberación, aunque, también, un plan menos discriminador que el activado por otros módulos cognitivos. Así pues, en un primer momento podría parecer útil, tal y como se acaba de indicar, que un sistema de IA (débil) lo poseyera. Sin embargo, si se considera que, por defecto, la cognición de una super-IA se presume hiper-rápida y compleja, quizá este argumento no sea del todo convincente en el escenario hipotético de dicha singularidad tecnológica.

Por su parte, una “mente fría” (únicamente con capacidad analítica y juicio racional) tampoco está exenta de riesgos, tal y como, por ejemplo, plantean obras cinematográficas como *Terminator* (1984) o *Yo, robot* (2004), donde la ausencia de “color emocional” en las máquinas puede provocar que la elección más óptima según la lógica de un algoritmo no sea la más conveniente según las normas éticas, los convencionalismos sociales o el sentido común²⁹. En concordancia con

²⁹ Aunque esto puede ser discutido: el libro *La sabiduría de los psicópatas* de Kevin Dutton (2020) afirma que algunas personas que tienen trastorno antisocial de la personalidad y psicopatía

esta idea, los científicos españoles López de Mántaras y Meseguer (2017: 149) opinan al respecto del sentido común que “es requisito fundamental para conseguir IA similar a la humana en cuanto a generalidad y profundidad”.

2.4.4. *Mundo Silico Sapiens*

Esta última sección será meramente apuntada, con la optimista idea de desarrollarla de forma más detenida en sendas reflexiones. Pues la introducción de IAs genuinamente afectivas (robóticas o no) en nuestro mundo tendrá un considerable impacto en diferentes planos de la praxis humana, como por ejemplo la convivencia, la cultura, el derecho, el arte, la economía o la política. Plantear un escenario en el que convivan diferentes seres inteligentes, sensitivos y conscientes nos arroja hacia un abismo reflexivo, similar al que tendríamos con el conocimiento de otros seres inteligentes extraterrestres³⁰. Así pues, dejo abiertas una serie de preguntas generales con el objetivo de invitar a la especulación y el propósito de profundizar en su análisis en posteriores investigaciones: ¿podríamos considerar “vivo” a un ser artificial dotado de emociones e inteligencia; qué características ayudarían a dicha adscripción? ¿Tendría una IA general capacidad de libre albedrío y responsabilidad moral; cuál sería su posicionamiento ético en el mundo? ¿Qué estatus jurídico-legal tendría un ser vivo artificial, sintiente y dotado de inteligencia? ¿Cuáles serían las ocupaciones y tareas de dicha inteligencia robótica; cómo se integraría en el sustrato económico y productivo de los estados? ¿Cómo serían las relaciones jurídico-administrativas entre máquinas inteligentes-sintientes y humanos? ¿Qué tipo de relaciones afectivas y convivenciales se establecerían entre humanos y máquinas emocionales e inteligentes? ¿Qué peligros o amenazas deberíamos considerar, y qué medidas sería conveniente adoptar frente a una super-IA?

(caracterizados por empobrecimiento emocional y falta de empatía; diríamos, pues, que son “mentes frías”), tienen una alta capacidad racional dirigida a metas y, por tanto, no es infrecuente que alcancen el éxito profesional.

³⁰ Hay que señalar que este escenario ya se produjo, salvando las distancias, en nuestro pasado evolutivo, con una situación de convivencia (no está claro si pacífica) entre dos especies inteligentes, como el *Homo sapiens* que somos y el extinto *Homo neanderthalensis* (si bien, en cierto modo, somos también Neandertales, en la medida en que compartimos acervo genético con ellos). La extensa y magnífica obra divulgadora del paleontólogo Juan Luis Arsuaga ha profundizado en muchos de estos aspectos.

CONCLUSIONES

El correcto abordaje de la interacción IA-humano exige aproximarse desde la perspectiva de tres grandes áreas de conocimiento dentro del campo de las ciencias y las humanidades: (1) la psicología y la neurociencia; (2) la robótica y las ciencias de la computación; y (3) la filosofía teórica (sobre todo la filosofía de la mente) y práctica (especialmente la ética aplicada). Sólo así podrá alcanzarse un conocimiento profundo e integrado del complejo fenómeno de la emocionalidad artificial.

Como se ha podido comprobar, no son pocos los retos a los que se enfrentarán los futuros desarrollos de la inteligencia artificial en el ámbito afectivo. Su interactividad humana tendrá un papel muy destacado en el reconocimiento y simulación emocional, especialmente en los robots de compañía con funciones socio-sanitarias o en cualesquiera sectores en los que el empleo de programas de detección de expresión emocional (fisiológica o conductual) suponga algún tipo de beneficio.

Aunque por el momento sean meras propuestas vanguardistas, los interfaces cerebro-ordenador regulados por IA también tendrán, presumiblemente, un destacado papel en la manipulación terapéutica de los afectos, más en concreto, en el tratamiento de trastornos neuropsiquiátricos que cursan con alteraciones emocionales como la depresión, la ansiedad o la esquizofrenia.

Pese a la problemática tecnológica y la controversia conceptual apuntada, el advenimiento de una IA general o, incluso, de una super-IA con conciencia y vivencia emocional genuina podría, a juicio del transhumanismo más optimista, ser una realidad en un futuro a medio o largo plazo. Desde un punto de vista tecnológico, son varias las alternativas propuestas para lograr algunos de estos hitos, como el desarrollo de nuevos lenguajes lógicos y algoritmos de aprendizaje, de sistemas integrados y arquitecturas cognitivas o de la prometedora computación cuántica. Desentrañar los misterios de la estructura y funcionamiento cerebral sin duda ayudará en esta progresión tecnológica.

Pero no sólo hay retos científico-tecnológicos, pues desde la filosofía de la mente y de las ciencias cognitivas el transhumanismo más transgresor seguirá debatiendo acaloradamente sobre el núcleo ontológico y gnoseológico de la IA, al respecto de la posibilidad de estados mentales artificiales, de la creación de una inteligencia general con “conciencia artificial” o de la emergencia de genuinos *qualia* emocionales. Posiblemente el funcionalismo sea la corriente filosófica que mejor pueda sustentar este conjunto de controvertidas propuestas.

A todas las dificultades comentadas se une, sin duda, cierta urgencia por establecer una vigilancia que limite los riesgos bioéticos y sociales inherentes a una

tecnología tan sobrecogedora, capaz de modificar nuestro mundo de un modo que superará con creces los cambios tecno-sociales de la Revolución Industrial. Se hace imprescindible, pues, un desarrollo responsable y sostenible de dichas tecnologías, con sistemas de control garantistas, habida cuenta del impacto bio-psico-social que la llegada de la emocionalidad artificial presumiblemente tendrá en las próximas décadas. Afortunadamente, contamos ya en nuestro entorno inmediato con iniciativas reguladoras, como las Directrices éticas de la Comisión Europea aprobadas en junio de 2018, que proponen una serie de normas básicas que pivotan en torno al respeto de la autonomía humana, la prevención del daño, la equidad y la explicabilidad³¹.

REFERENCIAS BIBLIOGRÁFICAS

- APARICIO-GARCÍA, R.; JUÁREZ-GRACIA, G.; ÁLVAREZ-CEDILLO, J.; SANDOVAL-GUTIÉRREZ, J y TOVAR-CORONA, B., "Evaluation of the Design of a Brain-Computer Interface for Emotion Detection". *DYNA*, 93(5), 2019, 468.
- BAARS, M.J., *A Cognitive Theory of Conscience*. Cambridge: Cambridge University Press, 1988.
- BISCAIA, JM., "Neuromejora: de la vanguardia científica y tecnológica a las dificultades y límites planteados por la filosofía de la mente y la bioética". *Revista Iberoamericana de Bioética*, 16, 2021, 1-17. doi: 10.14422/rib.i16.y2021.003.
- BLAKEMORE, C. y GREENFIELD, S. (ed.), *Mindwaves*. Oxford: Blackwell, 1987.
- BLOCK, N., "Consciousness". En: GUTTENPLAN, S. (Ed.), *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994a.
- , "Qualia", En: GUTTENPLAN, S. (ed.), *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994b.
- , "On a Confusion about a Function of Consciousness". *Behavioural and Brain Sciences*, 18, 1995.
- BLOCK, N. y FODOR, JA., "What Psychological States are not". *Philosophical Review*, 81, 1972 159-181.
- BODEN, MA., *Inteligencia Artificial*. Madrid: Turnes Publicaciones, 2017.
- BRAIDOT, N., *Neuromarketing, neuroeconomía y negocios*. Madrid: Editorial puerto norte-sur, 2005.
- BRONCANO, F., "Las emociones: territorios intermedios en la mente". *Contrastes. Revista internacional de filosofía*, suplemento VI, 2001.

31 Normativa europea: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (consultado el 19/7/21).

- CANO-VINDEL, A., "Orientaciones en el estudio de la emoción". En FERNÁNDEZ-ABASCAL, EG. (ed.), *Manual de Motivación y Emoción*. Madrid: Centro de Estudios Ramón Areces, 1995, 337-383.
- CARNAP, R., "Die Physikalische Sprache als Universalsprache der Wissenschaft". *Erkenntnis*, 1931-1932.
- CASACUBERTA, D. y VALLVERDÚ, J., "Emociones sintéticas". *Páginas de Filosofía*, XI, 13, 2010, 116-144.
- CASADO, C. Y COLOMO, R., "Un breve recorrido por la concepción de las emociones en la filosofía occidental". *A parte Rei. Revista de Filosofía*, 47, 2006, 1-10.
- CHAKRISWARAN, P.; VINCENT, D.; SRINIVASAN, K.; SHARMA, V.; CHANG, C-Y. y GUTIÉRREZ, D., "Emotion AI-Driven Sentiment Analysis: A Survey, Future Research Directions, and Open Issues". *Appl. Sci*, 9(24), 2019, 5462. doi: org/10.3390/app9245462.
- CHALMERS, D., "Facing up to the Problem of Consciousness". *Journal of Consciousness Studies*, 2, 1995, 200-219.
- CORTINA, A., *Neuroética y neuropolítica. Sugerencias para la educación moral*. Madrid: Tecnos, 2011.
- DAMASIO, AR., *Descartes'error. Emotion, reason, and the human brain*. Nueva York: Grosset/Putman Book, 1994.
- DAVIDSON, D., "Material Mind". *Studies in Logic and the Foundations of Mathematics*, 74, 1973.
- , *Ensayos sobre acciones y sucesos*. Barcelona: Crítica, 1980.
- DIÉGUEZ, A., *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*. Barcelona: Herder Editorial, 2017.
- DIXON, ML.; THIRUCHSELVAM, R.; TODD, R.; CHRISTOFF, K., "Emotion and the prefrontal cortex: An integrative review". *Psychology Bulletin*, 143(10), 2017, 1033-1081. doi: 10.1037/bul0000096.
- DUTTON, K., *La sabiduría de los psicópatas*. Madrid: Ariel, 2020.
- EKMAN, P., "An argument for basic emotions". *Cognition and Emotions*, 6(3-4), 1992, 169-200.
- ELSTER, J., *Alchemies of the Mind. Rationality and the Emotions*. Cambridge: Cambridge University Press, 1999.
- ESPERIDIAO-ANTONIO, V.; MAJESKI-COLOMBO, M.; TOLEDO-MONTEVERDE, D.; MORAES-MARTINS, G.; FERNANDES, JJ.; BAUCHIGLIONI DE ASIS, M., MONTENEGRO, S., SIQUEIRA-BATISTA, R. "Neurobiology of emotions: an update". *International Review Psychiatry*, 29(3), 2017, 293-307.
- EVERS, K., *Neuroética. Cuando la materia se despierta*. Madrid: Katz Editores, 2011.
- FARAH, MJ., "Neuroethics: The Practical and the Philosophical". *Trends in Cognitive Sciences*, 9, 2005, 34-40. doi: 10.1016/j.tics.2004.12.001.
- FEIGL, H.; MAXWELL, G. y SCRIVEN, M. (Ed.), *Concept, Theories and Mind-Body Problem*. Minneapolis: University of Minnesota Press, 1958.

- FERNÁNDEZ-ABASCAL, EG.; GARCÍA, B.; JIMÉNEZ, MP.; MARTÍN, MD. y DOMÍNGUEZ FJ., *Psicología de la Emoción*. Madrid: Editorial Universitaria Ramón Areces, 2010.
- FODOR, J., *Psychological Explanation*. Nueva York: Random House, 1968.
- , *Representations*. Brighton: The Harvester Press, 1981.
- GAZZANIGA, MS., *The Ethical Brain*. Nueva York: Dana Press, 2005.
- GIBBARD, A., *Wise Choices, Apt Feelings. A Theory of Normative Judgement*. Oxford: Clarendon, 1990.
- GOLEMAN, D., *Inteligencia emocional*. Madrid: Kairos, 1995.
- GONZÁLEZ, M. y GARCÍA, J., “Arquitecturas emocionales en inteligencia artificial: una propuesta unificadora”. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 7(2), 2006, 156-168.
- GRIFFITHSS, P., *What Emotions Really Are*. Chicago: The University of Chicago Press, 1977.
- GUTTENPLAN, S., *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994.
- HANSON, R., (2017). “Cuando los robots gobiernen la Tierra: el legado humano”. En *El próximo paso: la vida exponencial*. OpenMind BBVA, 2017. doi: 10.1080/0952813X.2015.1055826.
- HIERRO-PESCADOR, J., *Filosofía de la mente y de la Ciencia cognitiva*. Madrid: Akal, 2005.
- HOOKE, S. (Ed.), *Dimension of Mind*. Nueva York: Collier Books, 1960.
- IACOBONI, M., “Imitation, empathy and mirror neurons”. *Annu Rev Psychol*, 60, 2009, 653-670. doi: 10.1146/annurev.psych.60.110707.163604.
- IENCA, M. y ANDORNO, R., “Towards new human rights in the age of neuroscience and neurotechnology”. *Life Sciences, Society and Policy*, 13, 2017, 5. doi: 10.1186/s40504-017-0050-1.
- JANAK, PH. Y TYE, KM., “From circuits to behaviour in the amygdala”. *Nature*, 517(7534), 2015, 284-292. doi: 10.1038/nature14188.
- KLEINGINNA, PR. y KLEINGINNA, AM., “A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition”. *Motivation and Emotion*, 5, 1981, 345-379.
- KURZWEL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*. Berlín: Lola Books, 2012.
- LATORRE, JI., *Ética para máquinas*. Barcelona: Ariel, 2019.
- LEDOUX, JE. “Emotion circuits in the brain”. *Annual Review of Neuroscience*, 23, 2000, 155-184.
- LEVY, N., *Neuroética. Retos para el siglo XXI*. Barcelona: Avarigani Editores, 2014.
- LEWIS, D., “Reduction of Mind”. En: GUTTENPLAN, S. (Ed.), *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994.
- LEWIS, M., “Self-conscious emotions: Embarrassment, pride, shame, and guilt”. En: LEWIS, M. Y HAVILAND-JONES, JM. (Ed.), *Handbook of Emotions* (623-636). New York: The Guilford Press, 2000.

- LÓPEZ DE MÁNTARAS, R. y MESEGUER, P., *Inteligencia Artificial*. Madrid: Consejo Superior de Investigaciones Científicas, 2017.
- MARTÍNEZ-FREIRE, P., *La nueva filosofía de la mente*. Barcelona: GEDISA, 2002.
- MCCULLOCH, W. y PITTS, W., "A Logical Calculus of the Ideas Immanent in Nervous Activity". *The Bulletin of Mathematical Biophysics*, 5, 1943, 115-133.
- MCDOWELL, J., "Criteria, Defeasibility, and Knowledge". *Proceedings of the British Academy*, 68, 1982, 455-79.
- MOORE, GE., "Cramming more components onto integrated circuits". *Electronics*, 38(8), 1965.
- NAGEL, T., "What is it Like to be a Bat?". *Philosophical Review*, 83, 1974, 435-456.
- NEWELL, A. y SIMON HA., "Computer Science as Empirical Enquiry: Symbols and Search". *Communications of the Association for Computing Machinery*, 19, 1976.
- OAKLEY, K. y JOHNSON-LAIRD, PN., "Toward a Cognitive Theory of Emotion". *Cognition and Emotion*, 1, 1987, 29-50.
- PICARD, RW., *Affective Computing*. New York: MIT Press, 1997.
- PINEDO IA. y YÁÑEZ CJ., "Las emociones: una breve historia en su marco filosófico y cultural en la época antigua". *Cuadernos de Filosofía Latinoamericana*, 39(119), 2018, 13-45. doi: 10.15332/25005375.504.
- PUTNAM, H., "Minds and Machines". En HOOK, S. (Ed.). *Dimensions of Minds*. New York: New York University Press, 1960, 138-164.
- RAPAPORT, WJ., "Syntactic Semantics: Foundations of Computational Natural Language Understanding". En FETZER, JH. (Ed.), *Aspects of Artificial Intelligence*. Dordrecht: Kluwer Academic Publishers, 1988, 81-131.
- REEVÉ, J., *Motivación y Emoción*. Madrid: McGraw-Hill, 1994.
- REISENZEIN, R., "Cognition and emotion: a plea for theory". *Cognitive Emotion*, 33(1), 2019, 109-118. doi: 10.1080/02699931.2019.1568968.
- RORTY, R., "Contemporary Philosophy of Mind". *Synthese*, 52, 1982, 323-348.
- RUSSELL, S. y NORVIG, P., *Artificial Intelligence: A Modern Approach*. New Jersey: Prentice Hall, 2009.
- RYLE, GILBERT., *El concepto de lo mental*. Buenos Aires: Paidós, 1967.
- SCHULLER, D. y SCHULLER BW., "The Age of Artificial Emotional Intelligence". *Computer*, 51(9), 2018, 38-46. doi: 10.1109/MC.2018.3620963.
- SEARLE, JR., "Minds, Brains and Programs". *Behavioral and Brain Sciences*, 3(3), 1980, 417-457. doi: doi.org/10.1017/S0140525X00005756.
- , "Minds and Brains without Programs". En BLAKEMORE, C., (Ed.), *Mindwaves*, Blackwell, 1987.
- SENTENTIA, W., "Neuroethical Considerations: Cognitive Liberty and Converging Technologies for Improving Human Cognition". *Annals of the New York Academy Sciences*, 1013, 2004, 221-228. doi: 10.1196/annals.1305.014.
- SMART, J., "Sensations and Brain Processes". *The Philosophical Review*, 68 (2)2, 1959, 141-156.

- SPINOLA, J. y QUEIROZ, J., "Artificial Emotions: Are We Ready for Them?". *Advances in Artificial Life*, 4648, 2007.
- STONE, M., y HIRSH, H., "Artificial Intelligence: The Next Twenty-Five Years". *IA Magazine*, 26(4), 2006, 85-97. doi: doi.org/10.1609/aimag.v26i4.1852.
- TANTAM, D., "The flavour of emotions". *Psychology Psychotherapy*, 76(1), 2003, 23-45. doi: 10.1348/14760830260569229.
- TEGMARK, M., *Vida 3.0*. Barcelona: Editorial Taurus, 2018.
- TIRAPU-USTÁRRUZ, J.; PÉREZ-SAYES, G.; EREKATXO-BILBAO, M. y PELEGRÍN-VALERO, C., "¿Qué es la teoría de la mente?" *Revista de Neurología*, 44(8), 2007, 479-489. doi: doi.org/10.33588/rn.4408.2006295.
- TORRES, E.; TORRES, A.; HERNÁNDEZ-ÁLVAREZ, A. y YOO, S., "EEG-Based BCI Emotion Recognition: a Survey". *Sensors*, 20(18): 5083, 2020. doi: 10.3390/s20185083.
- VINGE, V., "The Coming Technological Singularity: How to Survive in the Post-Human Era". Ponencia presentada en *VISION-21 NASA Lewis Research Center and the Ohio Aerospace Institute*, 1993.
- VOGT, BA., "Cingulate cortex in the three limbic subsystems". *Handbook of Clinical Neurology*, 166, 2019, 39-51. doi: 10.1016/B978-0-444-64196-0.00003-0.
- WALDROP, MM., "The Chips are down for Moore's Law". *Nature*, 530(7589), 2016, 144-147. doi: 10.1038/530144a.
- WENGER, MA.; JONES, FN. y JONES, MH., "Emotional Behavior". En CANDLAND, DK., (ed.). *Emotion: Bodily Change*. Princeton: van Nostrand, 1962.

DEUS EX MACHINA: LA INTELIGENCIA ARTIFICIAL EN EL FUTURO DEL SER HUMANO

DEUS EX MACHINA: *ARTIFICIAL INTELLIGENCE
IN THE FUTURE OF THE HUMAN BEING*

DAVID GONZÁLEZ-CUENCA

Líder y Miembro del Grupo de Investigación PIREO

Magister en Relaciones y Negocios Internacionales

Profesional en Relaciones Internacionales y Estudios Políticos

Profesor Investigador.

Facultad de Estudios a Distancia. Programa en Relaciones Internacionales y Estudios Políticos

Universidad Militar Nueva Granada.

Bogotá/Colombia. Universidad Nacional de Estudios a Distancia (UNED)

dgonzales1769@alumno.uned.es

ORCID: 0000-0002-6563-9193

Recibido: 12/04/2021

Revisado: 18/08/2021

Aceptado: 6/09/2021

Resumen: Con el auge de la tecnología, la humanidad inicia la implementación de procesos de inteligencia artificial (I.A), lo cual puede presentar una singularidad tecnológica, que define la capacidad que tiene esta tecnología de evolucionar en sus lógicas cognitivas y llevarlas a niveles de mejoramiento autónomo por encima de las capacidades humanas.

Es así que por medio de este artículo y soportado en el argumento dramático denominado *Deus ex Machina* se pretende exponer la dependencia que se está creando en el ser humano ante la inteligencia artificial y desde una visión prospectiva, plantear las consecuencias que ello acarrea para la existencia del ser humano frente a la imposibilidad existencial de sobrellevar su propio destino y la pérdida del libre albedrío por cuenta de la influencia y la superioridad tecnológica, concretando la analogía de una percepción del ser humano hacia las máquinas como dios, generando una degradación del individuo y de la sociedad, llevando a que el *Deus ex Machina* controle al ser humano, ubicándolo al borde de la extinción o de la dependencia existencial.

Palabras clave: Cyberpunk, Distopia, Inteligencia Artificial, Posthumanismo, Singularidad Tecnológica, Transhumanismo.

Abstract: With the rise of technology, humanity begins the implementation of artificial intelligence (AI) processes, which may present a technological singularity, which defines the

ability of this technology to evolve in its cognitive logics and take them to levels of autonomous improvement above human capabilities.

Thus, through this article and supported by the dramatic argument called *Deus ex Machina*, it is intended to expose the dependence that is being created in the human being before artificial intelligence and from a prospective vision, to propose the consequences that this entails for existence of the human being in the face of the existential impossibility of coping with his own destiny and the loss of free will due to the influence and technological superiority, specifying the analogy of a perception of the human being towards machines as God, generating a degradation of the individual and of society, leading the *Deus ex Machina* to control the human being, placing him on the brink of extinction or existential dependence.

Keywords: Artificial Intelligence, Cyberpunk, Dystopia, Posthumanism, Technological Singularity, Transhumanism.

INTRODUCCIÓN

Desde el mismo origen de la especie humana, por cuenta de un factor evolutivo relacionado con la aproximación del fuego como herramienta transformadora de su entorno y la necesidad evolutiva de la supervivencia, la utilización de artefactos, herramientas, utensilios entre otros objetos producidos por el hombre, se presenta en las diferentes sociedades históricas una correlación entre el factor evolutivo y la creación, desarrollo o actualización de lo que podemos entender como tecnología.

Un cuchillo, una cuchara, la misma rueda, el arado, el fuego (ya mencionado), se han establecido por antropólogos y aceptados en su gran mayoría por la sociedad como artefactos tecnológicos que proporcionaron a la especie humana un escenario de evolución que nos han conducido a generar tras millones de años de existencia, una clase de dependencia hacia las herramientas tecnológicas para garantizar la supervivencia. (Villalba Gómez, 2019)

El ser humano es una especie carente de características naturales de protección, no poseemos como muchos seres vivos del planeta estructuras de defensa ni habilidades de aprovisionamiento de alimento o fuentes de supervivencia; no más que nuestro cerebro y la capacidad que nos provee de crear herramientas que permitan garantizar nuestra protección y lograr la obtención de recursos alimenticios, entre otros. Ello nos presenta como una especie que debe comprender su entorno más que ninguna otra, debe establecer relaciones sociales para garantizar el éxito como especie y generar un marco de protección o seguridad, fundamentado en los vínculos, que son relevantes para proyectar la existencia de la especie, pero aún más en la utilización de herramientas tecnológicas.

A partir de ello, podemos identificar que la tecnología se puede entender como proceso o sistema, como saber y como artefacto; para este análisis

tomaremos la última categoría, definida por *Ian Quallenberg Menkes* como “un conjunto de objetos materiales que no se encuentran en la naturaleza y que son manufacturados por seres humanos” (Quallenberg Menkes, 2012, pág. 238); en este sentido y llegando a una época moderna, en la cual el desarrollo y la utilización global de herramientas de alta tecnología se hace más evidente, podemos percibir un crecimiento exponencial en el desarrollo y producción de herramientas tecnológicas, en la dirección en que durante los últimos veinte años se han presentado saltos en la transformación de la relación del ser humano con la tecnología de una manera más acelerada que lo que vivieron nuestros antepasados en más de doscientos años. (Cheang Wong, 2005)

A partir de entender el impacto de la tecnología en la modificación de las estructuras sociales del ser humano, y por ende en su evolución como especie, el estadounidense *Alvin Toffler* ha establecido cuatro diferentes momentos históricos, que denomina *olas*, como factores revolucionarios que re-dinamizan las sociedades humanas y permiten esclarecer el panorama frente al argumento de la sinergia entre evolución y tecnología. (Toffler, 2018)

La primera ola que define el teórico está fundamentada en la revolución agrícola: La extensión de los cultivos soportada en el asentamiento de las organizaciones sociales que eran nómadas por tradición, volcó al ser humano a desarrollar técnicas basadas en la percepción del entorno, tales como el reconocimiento de las estaciones climáticas, las fases lunares, la calidad de la tierra y la observación de semillas que produjeran alimentos mediante un proceso denominado agricultura.

Durante la primera ola o revolución tecnológica, definida por *Toffler* (2018), se consolidaron las comunidades humanas en torno particularmente a las grandes fuentes hídricas, surgen civilizaciones localizadas en Mesopotamia y Egipto, por mencionar algunas, y presentaban una división del trabajo definido en las clases y catalogadas como economías descentralizadas y autosuficientes, pues producían lo que la comunidad necesitaba. (Toffler, 2018)

La segunda ola de *Toffler* está definida por la llamada *Revolución Industrial*, en la que, si bien se mantenían las dinámicas agrarias de la primera ola, hay un auge de desarrollo tecnológico, principalmente, para fortalecer el proceso productivo de los bienes agrícolas; el arado ya venía siendo implementado como una herramienta que generaba mejor productividad a menor esfuerzo, sin embargo, el desarrollo de la máquina a vapor y la implementación de fuentes de energía no renovables en la mecanización de los procesos productivos permitió determinar modelos de producción en masa, lo cual incrementó la población humana e incentivó la creación de escenarios científicos (escuelas, universidades, entre otras) para el desarrollo de nuevas tecnologías; asimismo se da origen a una degradación ambiental por cuenta del uso de estos combustibles que luego se percibirían como

dañinos para la atmósfera y los ecosistemas y el principal generador del cambio climático global. (Guillot, 2016)

Para esta etapa, el factor consumismo empieza a insertarse en las dinámicas culturales de las sociedades humanas, aspecto que va a ser relevante para identificar la dependencia del ser humano hacia la tecnología en apartados por venir, así como la definición de sociedad en sí por cuenta de los mismos desarrollos tecnológicos y el acceso a los mismos.

La tercera revolución tecnológica o según *Zbigniew Brzezinski*, la *era tecnológica* (1979), se fundamenta en el desarrollo de tecnología basada en procesos científicos profundos, en lo que la computación, la física cuántica, las tecnologías de la información como internet, los viajes espaciales y la biotecnología van a establecerse como los determinantes de la aceleración temporal de la evolución tecnológica, definiendo en sentido recíproco, una aceleración evolutiva de la especie y de sus necesidades. (Schoijet, 1998)

Ello, puede presentar diferentes connotaciones, tanto biológicas, sociales y hasta éticas. Sin perjuicio de ello, podemos determinar la tercera ola como el origen de la denominada Inteligencia Artificial o IA.

Aunado a lo anterior, nos resta determinar una cuarta etapa de evolución, la cual según *Toffler*, es la que estamos *ad portas* de percibir, determinada por la exposición de una revolución biológica que se entremezcla con el factor informático. Ante ella, ya hemos percibido una cantidad importante de precedentes que permiten identificar como muy cercana o presente la cuarta ola; el descubrimiento del genoma humano recién entrado el siglo XXI, precedido por la primera clonación de una especie, la oveja Dolly en 1996; lo cual nos conduce al desarrollo de una nueva tecnología basada en organismos vivos determinado por la manipulación molecular y celular de sus características.

En palabras de *Toffler*, el argumento principal para el desarrollo de este tipo de tecnología es el mismo que se presentaba en los desarrollos tecnológicos de las primeras olas, el crecimiento de la población humana y la necesidad de generar fuentes de alimento y subsistencia a gran escala y de manera eficiente, claro, basados en un sistema de mercado en el que el consumismo y la riqueza son factores a tener en cuenta. (Toffler, 2018) Así entonces, la cuarta ola estará marcada por un salto aún más exponencial del desarrollo tecnológico que el vivido hasta la actualidad, en la que el uso y gestión de la información basados en fuentes sintéticas como la IA, vinculada a la creación de nuevos procesos biológicos nos plantea una nueva posición como especie frente a la supervivencia en este planeta, conduciéndonos a la migración planetaria.

En esencia y función de lo anterior, este artículo tiene por objetivo exponer la dependencia que se está creando en el ser humano ante la inteligencia artificial

y el escenario en construcción de otorgarle facultades divinas a la tecnología de ese tipo. Esta investigación se ha desarrollado bajo el marco de una metodología cualitativa de tipo analítica o crítica con un enfoque predictivo, basados en la recolección y análisis de fuentes secundarias determinada por documentos técnicos y científicos, y fuentes cinematográficas y literarias como medio de aproximación prospectiva de realidades alternas del ser humano en su interacción con la Inteligencia Artificial y su entorno.

1. LA LLEGADA DE LA INTELIGENCIA ARTIFICIAL AL ESCENARIO SOCIAL

La Inteligencia Artificial o IA, se presenta en el escenario tecnológico como una herramienta y como una técnica de gestión de información a la cual se le ha provisto de *razonamiento* al estilo humano (Schoijet, 1998). Como hemos identificado anteriormente, el ser humano por su condición natural, ha tenido que adaptarse a su entorno y para ello ha tenido que diseñar estrategias y herramientas para el reconocimiento y comprensión del mismo; a partir de la revolución tecnológica de la información y las comunicaciones, los científicos han llegado a determinar variables y algoritmos que permiten predecir los comportamientos de los seres humanos, en principio como una función del mercado para poder impactar de manera más eficiente en el consumidor, y luego en diferentes frentes como la medicina, el clima, entre otros. (AGragwal, Gans, & Goldfarb, 2019)

Es una constante en la comunidad científica considerar al cerebro humano como el órgano más complejo de la biología humana, que no puede ser replicado o por lo menos ello contempla una altísima dificultad, adicionando que el factor de razonamiento en el ser humano es único y definido por la individualidad del mismo, por la construcción social del entorno y los estímulos particulares que cada persona recibe a lo largo de la historia, aunado a un factor histórico y genético que se marca en el comportamiento y las dinámicas existenciales de los seres humanos. (Morin, 1999)

La misma injerencia de la tecnología de la que se ha venido tratando, ha tenido un papel fundamental en la evolución no solo social sino también biológica del ser humano, lo que ha obligado a la sociedad a continuar con el auge tecnológico, llevándolo a traspasar la frontera de lo artificial y empezar a construir procesos basados en lo biológico; ello también se refiere a la condición racional del hombre enfocada específicamente a la toma de decisiones, la cual mediante procesos matemáticos y físicos se ha venido replicando hasta el punto de consolidar el factor de IA como un desarrollo tecnológico real y aplicable a las diferentes situaciones de la vida en sociedad. (Gumbrecht, 2009)

La necesidad, en esencia económica, de garantizar la mayor cantidad y mejor información para la toma de decisiones, ha dado origen a conceptos tales como el *Business Intelligence* y el *Big Data*, que vienen acompañados de un proceso de intervención tecnológica en el que media la IA con el ser humano, el primero como gestor del proceso de recolección y gestión de la información en cantidades sobre-humanas (basado en la capacidad física de recopilar y procesar información del ser humano), y el segundo como receptor de la información ya decantada y seleccionada para un fin determinado. (AGragwal, Gans, & Goldfarb, 2019)

Estos procesos se han hecho socialmente comunes y aceptados, gracias a la intervención de los dispositivos digitales y el auge de la internet. Actualmente, cualquier persona con acceso a un dispositivo con conexión a la red puede gestionar procesos de recolección y procesamiento de información, para cualquier efecto a niveles altísimos de eficiencia.

Antiguamente, durante la primera y segunda ola o revolución tecnológica, primaba la eficiencia de las tecnologías dada las necesidades esenciales que pretendían solventar; sin embargo, desde la era moderna, en donde supuestamente las sociedades tienen la capacidad tecnológica de garantizar la supervivencia humana (desconociendo factores como la desigualdad social, por ejemplo), la misma se ve volcada por cuenta del consumismo ya referenciado, a generar nuevas necesidades y hacer de ellas fundamentales para la supervivencia del ser, sin serlo, y determinándolas como el engranaje de la sociedad. (Foucault, 1988)

A partir de ello, se puede inferir que, bajo una óptica de generalización, las sociedades modernas se articulan desde un paradigma de tecnología digital dominado por la IA. Es aquí donde es indispensable insertar el argumento presentado por Nick Bostrom en el cual el ser humano ha perdido la independencia frente a las funciones de procesamiento de las computadoras. Bostrom establece que al desarrollar procesos de programación de simulación del comportamiento humano y que éstos sean incluidos en computadoras, las mismas pueden generar procesos de evolución por medio de estructuras similares a la sinapsis basados en procesos, otorgándole a las tecnologías de IA un factor de singularidad único, que podría generar una desestabilización en la relación hombre-máquina, y distorsionando, la realidad y las condiciones sociales de la misma. (Bostrom N. , 2003)

En este punto, es pertinente establecer las condiciones para entender el factor de la singularidad tecnológica, que permitirá identificar conceptualmente el proceso evolutivo de la IA en esencia. En el sentido establecido por el teórico Ray Kurzweil, la *singularidad* “representará la culminación de la fusión de nuestro pensamiento biológico y la existencia con nuestra tecnología , lo que resulta en un mundo que sigue siendo humano, sino que trasciende nuestras raíces biológicas.” (Kurzweil, 2008, pág. 70), lo que implica la reconexión con la evolución histórica del ser humano en función de la tecnología, aceptando el concepto de

crecimiento exponencial en el desarrollo tecnológico basado en la Ley de Moore (Cheang Wong, 2005), la cual está determinada por el incremento exponencial de la complejidad de los circuitos integrados en relación a una disminución de los costos, y aunque hoy día se vienen sentando bases que la contradicen en términos de su exponencialidad, se persevera en el crecimiento de las capacidades de procesamiento de las nuevas tecnologías incluso a niveles cuánticos, generando adicionalmente incremento en la disposición de estos elementos para ser usados y a la vez mejorados en su capacidad y rendimiento funcional, haciendo que la tecnología basada en partes elaboradas con silicio, como los microprocesadores, sean utilizados en componentes que estén presentes en todas las sociedades y que definan la evolución de las personas y de sus comunidades. (Ramírez, 1998)

Decantando lo anterior, podemos definir la singularidad tecnológica como un momento histórico o un proceso desarrollado por tecnologías de IA en el cual esos entes estén provistos de tal capacidad que les permita evolucionar sin intervención humana, llevando este proceso a niveles de *crear* tipos de tecnología capaces de superar los niveles cognitivos y de aprendizaje de los propios seres humanos. Ante lo cual se han presentado alternativas como la posibilidad de no rivalizar con las IA sino llegar a integrarlas al ser humano, estableciendo así la amalgama entre lo biológico y lo tecnológico, transgrediendo las características humanas como las conocemos.

Ante este concepto, que bien puede parecer sacado de una película de ciencia ficción, es necesario atribuirle condiciones reales en nuestro tiempo; a mediados del año 2017, se descubrió al interior de los laboratorios de investigación de IA de Facebook, un precedente de la singularidad tecnológica, los chatbots o lo que se conoce como agentes de diálogo, habían desarrollado un sistema de comunicación diferente al impartido por los creadores, ello para desarrollar un proceso más eficiente en el uso del lenguaje, lo cual podría definirse como un salto adelante en el proceso de la singularidad de la IA, hecho que fue confuso para los científicos, pues el lenguaje creado más bien parecía un defecto de la máquina, (Arteaga, 2018) lo que nos marca el punto de partida para abordar desde planteamientos prospectivos la posibilidad de la singularidad de la IA en lo que se conoce dentro del espectro literario de la ciencia ficción distópica como *el fantasma en la máquina*. (Dick, 2012)

Para dar un salto desde la realidad hacia el factor prospectivo, es necesario apoyarnos de la ciencia ficción, particularmente de la distopía, la cual en esencia se fundamenta en una visión del futuro de la humanidad y del planeta en términos negativos o apocalípticos, catalogando una posible realidad indeseada. (Foster, 2005) Y profundizando más en el subgénero narrativo de la distopía debemos ubicarnos en el *cyberpunk*, el cual se distancia del subgénero descrito pues no incluye en sus narrativas aspectos como invasiones e interacciones con alienígenas,

sino que se fundamenta en el análisis de sociedades profundamente interconectadas y dependientes de la computarización, presentado interrogantes frente a la misma existencia humana, las consecuencias de la generación de realidades virtuales, el transhumanismo y la manipulación genética. (Cavallaro, 2000) En esencia el *cyberpunk*, creado por *William Gibson* mediante la publicación del libro *Neuromante*, “no tratan sobre el funcionamiento de los microchips, sino que intentan reflejar las implicaciones de una cultura obsesionada con la información y la alta tecnología” (Mayorga, 2017, pág. 18).

Establecido el marco argumental, debemos aproximarnos a la definición de ese fenómeno denominado *el fantasma en la máquina*: el cual se refiere a ese fenómeno basado en la singularidad tecnológica que permitirá el desarrollo de una especie de conciencia en entes cibernéticos o tecnologías basadas en IA; el fenómeno es descrito en el libro de *Philip Dick*, llamado *¿Sueñan los androides con ovejas eléctricas?* (Dick, 2012), al igual que en algunos de los cuentos cortos del intelectual *Isaac Asimov* como *Yo, Robot*; en donde entidades tecnológicas con aspecto humano (Robots o androides) construidos en función de la IA, presentan una evolución en su razonamiento fundamental, generando nuevos procesos cognitivos tendientes a la construcción de lógicas basadas en los algoritmos implementados por los humanos, pero en distorsión o ausencia, de condiciones éticas y morales; a pesar incluso de la formulación de las conocidas como las tres leyes de la robótica formuladas por el mismo *Asimov*, las cuales consisten en que: “1. Un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra daño. 2. Un robot debe obedecer las órdenes dadas por los seres humanos, excepto si estas órdenes entran en conflicto con la primera ley. Y 3. Un robot debe proteger su propia existencia en la medida en que su protección no entre en conflicto con las dos primeras leyes” (Rendón, 2018)

A partir de esta ilustración se puede percibir un análisis de tipo existencial, pues está implícita la viabilidad de la especie humana frente a la evolución de la IA (Heidegger M. , 2005), considerando que se puede generar una competencia entre ambas en las que el crecimiento exponencial de la IA puede superar en amplio margen las capacidades de adaptación del ser humano, así como se genera también un factor ético desde el entendido del actuar del ser humano en relación a la tecnología y su funcionamiento como amalgama de la sociedad. (Morin, 1999) Aquí entonces, se permitirá desarrollar el análisis de la integración social y cultural de la IA por encima de la misma naturaleza humana.

2. EL SER HUMANO, MÁQUINA DEPENDIENTE DE LAS MÁQUINAS

Con la llegada de la tecnología digital y el auge de la inteligencia artificial, en las sociedades humanas se ha gestado un escenario para el desarrollo de tecnologías cada vez más cercanas a la naturaleza humana; los consumidores queremos que los productos sean más humanos, que los diseños sean más dúctiles y prácticos hacia nuestra biología.

La robótica, en ese contexto, se ha desarrollado en los últimos diez años de una manera acelerada, y gestando nuevas tecnologías que permitirán, dicen los científicos, hacer de la vida de los seres humanos más fácil de sobrellevar. (Villalba Gómez, 2019) Se plantean desde este escenario propósitos de la robótica enfocados a la industria para mejorar los procesos de diseño y manufactura, en la cultura y los deportes se prevé desarrollar olimpiadas o competencias entre robots descritas en el proyecto Robot Tokio 2020 postergado por el fenómeno del COVID-19 a nivel mundial y que retrasó el desarrollo de las olimpiadas (Comité Organizador de Tokio 2020 de los Juegos Olímpicos y Paralímpicos, 2021); sin embargo, recientemente ha impactado la presentación de un organismo robótico con programación soportada en IA que adicionalmente tiene un aspecto humanoide, cuyo principal desarrollo ha sido una programación que le permite entablar conversaciones completas con seres humanos sin la mediación de un programador.

Aquí estamos tratando sobre *Sophia*, un robot diseñado por la empresa *Hanson Robotics*, quien no solo ha dado de qué hablar en los medios de comunicación, sino que ha llegado más lejos. Ha sido invitada en múltiples eventos internacionales y se ha destacado como participante de programas de televisión, así como también se ha presentado ante la Asamblea General de las Naciones Unidas. (Palou, 2017)

Sin embargo, lo más destacado para nuestro análisis no es la figuración en escenarios públicos de un robot humanoide que puede ser interlocutor con humanos sin apoyo alguno; lo más determinante proviene de un par de hechos que catalogo como los factores de quiebre de la identidad humana y la exacerbada demostración de la dependencia del ser humano hacia la IA y su posible intromisión en nuestra construcción social tal y como la conocemos:

En el mes de octubre del año 2017, *Sophia*, se convierte en el primer robot en recibir la condición de ciudadano, al recibir de parte de Arabia Saudí la nacionalidad, un derecho reservado a humanos con ciertas condiciones determinadas (BBC Mundo, 2021); de igual manera, para julio de 2018, el alcalde de Medellín en Colombia, Federico Gutiérrez, en el marco de la Escuela de Verano de la Universidad Pontificia Bolivariana, entregó a *Sophia* un *paisaporte*

o *pasaporte paisa*, el cual se podría homologar a las llaves de la ciudad. (Plaza Mayor Medellín, 2021)

Estos dos hechos, enmarcan una visión de aceptación y receptibilidad de los seres humanos y sus sociedades frente a la posibilidad de convivir con robots *inteligentes*. (McCaffery, 1991) Sin embargo, no se puede evitar pensar en diversas situaciones en las que personas, humanos, en condición de exilio o migrantes por diversas causas han solicitado la nacionalidad por cuestiones humanitarias y en muchos casos son negados (OIM, 2020). Ante este factor, es pertinente preguntarse ¿pueden llegar a ser los robots como *Sophia* más aceptados por la sociedad, en comparación con los migrantes o con otro tipo de personas en condiciones de dificultades humanitarias?

Frente a ese cuestionamiento se presentan aproximaciones de todo tipo, sin embargo, las que más certeza pueden dar son las de tipo político y las de tipo ético. En este sentido, podemos referirnos a las percepciones del teórico *Nick Bostrom*, quien presenta un análisis desde el transhumanismo, el cual se entiende como una corriente cultural e intelectual que pretende analizar la interacción entre humanos y tecnología y los beneficios y riesgos que esto puede conllevar. (Bostrom N. , 2005)

En su libro, *En Defensa de la Dignidad Posthumana*, *Bostrom* realiza una confrontación entre la corriente transhumanista y los bioconservadores, en donde los primeros se presentan con una postura en la que se promueve el acercamiento del humano a la tecnología, por ejemplo, en situaciones de procreación o transformación física. *Bostrom* presenta a los transhumanistas como un sector que cree posible la evolución del ser humano a partir de la intervención tecnológica, consolidando una especie evolucionada, la cual se denomina posthumana; sin embargo, aclara *Bostrom*, los transhumanistas al abrazar el desarrollo tecnológico, no dejan de lado la defensa de los derechos humanos y la toma de decisiones individuales, y a partir de sus planteamientos toman posturas frente a amenazas como el uso militar o armas biológicas en el desarrollo de la tecnología. (Bostrom N. , 2005)

En oposición se encuentran los bioconservadores, para los cuales, el desarrollo tecnológico en función del ser humano y su evolución podría *deshumanizar* al hombre, deteriorando la condición de dignidad de las personas, generando un deterioro de la percepción que tiene el ser frente a sí mismo y frente aquellos sistemas que han sostenido las sociedades humanas por siglos, tal es el caso de los factores metafísicos, ontológicos y constructivistas (Fukuyama, 2002). Ante ello, los bioconservadores proponen unos estados de prohibición frente a la expansión de la tecnología a nivel global y limitar la exposición de promesas frente al mejoramiento tecnológico del humano para “prevenir un deslizamiento por una pendiente resbaladiza hacia un estado posthumano”. (Bostrom N. , 2005, pág. 208)

La construcción del pensamiento evolutivo se presenta con riesgos desde dos puntos de vista mostradas a partir del posthumanismo: el primero consiste en que el posthumanismo puede desencadenar en la degradación del ser, perjudicándonos como especie; y en segundo lugar, la posibilidad latente de que los seres posthumanos se presenten como una amenaza a los humanos tradicionales, generando una divergencia en las condiciones sociales, proyectando sociedades conflictivas a una escala mayor de la que ya se han presenciado en la historia por condiciones genéticas, raciales, entre otras.

Muchas de las propuestas de la ciencia ficción planteadas desde el punto de vista distópico del que ya se trató, presentan una cercanía conceptual con los bioconservadores, sin demeritar el hecho de que posiblemente los transhumanistas puedan convertirse en la tendencia social de continuidad de progreso por cuenta de avances tecnológicos e inclinaciones económicas. (Esparza-Oviedo, 2020)

A este respecto podemos referirnos a autores como *Aldous Huxley* (1979), quien, en su famosa obra, *Un Mundo Feliz*, presenta una sociedad posthumana en la que las condiciones raciales y sociales divergentes fueron estandarizadas por cuenta de la implementación de procesos tecnológicos, generando estructuras de poder sociales de control dignas de un análisis del filósofo *Michel Foucault* (1988). En el cine podemos percibir cómo a partir de diferentes referencias conceptuales, los transhumanistas han perdido el control de las estructuras tecnológicas y las mismas han generado una evolución tal que se presentan como los *gobernantes* y posibles destructores de la especie humana, como en el caso de las trilogías *Matrix* o *Terminator*, por mencionar algunas de dominio popular.

En este sentido, la prospectiva que venimos aplicando se puede interpretar desde las estructuras narrativas de los expositores del *cyberpunk*, ligados a presentar una radiografía de una sociedad posthumana que contempla en su devenir ontológico e histórico (Heidegger M. , 1999) todas las consecuencias presentadas por los bioconservadores, presentando escenarios posibles de desarrollo y decadencia de la especie humana por cuenta de la aceptación de la IA como factor cohesionador de las estructuras sociales. (Gumbrecht, 2009)

3. DEUS EX MACHINA, LA SINGULARIDAD EN FUNCIÓN DE LA DIVINIDAD DE LA INTELIGENCIA ARTIFICIAL

Los seres humanos, hemos venido aceptado cada vez más el desarrollo de la tecnología digital y la reciente llegada de la IA como un logro de la especie humana, una herramienta de facilitación de las actividades cotidianas y un factor de eficiencia en términos económicos, entendiendo estas *máquinas* como parte de esos bienes de producción que el marxismo establecía como el factor principal del

desarrollo de las sociedades capitalistas. (Agrawal, Ajay; Gans. Joshua; Goldfarb, 2019)

Estamos tan volcados a la receptibilidad de estas tecnologías, que como lo vimos anteriormente, otorgamos condiciones que eran reservadas exclusivamente para la condición humana y que por diferentes miembros de la sociedad fue tan difícil y demorado lograrlas. Recordemos que en la antigua Grecia, las mujeres no eran consideradas ciudadanos, al igual que los niños; antes del siglo XIX los afrodescendientes no eran catalogados como seres racionales y mucho menos podían alcanzar la condiciones de ciudadanos; entonces, luego de siglos de lucha de las mujeres por lograr una aceptación social de su condición como ciudadanos, y de las cruentas luchas de la comunidad afro-descendiente por participar de la construcción social de manera pertinente (hecho que hoy en día continúa), percibimos que en no más de una década de desarrollo de la IA, estos seres ya han logrado una participación positiva en la sociedad humana. Bajo este argumento podemos evidenciar ya la validación de los conceptos bioconservadores de la pérdida de dignidad del ser humano por cuenta de la intervención tecnológica en su proceso evolutivo y constructivista de sociedad; ya presenciamos el inicio de las sociedades posthumanas. (Bostrom N. , 2001)

Apartándonos un poco de la dialéctica argumentativa, es hora de presentar la postura basada en el argumento literario del *Deus Ex Machina*, punto de ancla de toda la visión de este análisis que dará cierre a la propuesta. El *Deus Ex Machina*, es un argumento literario, creado en la era de oro griega, la cual propone una salida argumental a impases desarrollados en la trama (Contreras Escribano, 2015), esto quiere decir que este argumento servía de excusa para poder concluir una trama enredada y que pareciera sin fin cercano; el argumento consistía en posicionar una deidad o un dios como el hacedor o factor central de la trama; los personajes se comportan de esa manera porque así lo concibió determinado dios, era el argumento de cierre, usualmente. Este dios o deidad era presentado al público gracias al complejo desarrollo de máquinas que permitían desde cierta altura o distancia insertar en el escenario a la representación divina.

Ahora bien, para ligar el argumento del *Deus Ex Machina*, en nuestro andamiaje conceptual, debemos aclarar que desde esta perspectiva el desarrollo tecnológico de las sociedades modernas o postmodernas, en palabras de Heidegger, pueden darse por una desviación en la percepción del tiempo (Heidegger M. , 2005) (Sanabria, 2014), desviación provocada por la alternatividad de la realidad que presentan las tecnologías digitales, así como lo indica Gumbrecht (2009), presentando un escenario en el que la misma existencia humana es prescindible para nosotros mismos, provocando una percepción radical de desplazamiento de las necesidades humanas, hacia la satisfacción de necesidades virtuales, garantizando

así la victoria de las máquinas por el control y el ejercicio de poder sobre los seres humanos.

En este sentido, la percepción de los seres hacia las máquinas, cada día más, se está convirtiendo en casi una cercanía religiosa, una dependencia existencial del tipo bíblico (Green, 2018) many of which are being actively considered by organizations ranging from small groups in civil society to large corporations and governments. However, it also presents ethical concerns which are not being actively considered. This paper presents a broad overview of twelve topics in ethics in AI, including function, transparency, evil use, good use, bias, unemployment, socio-economic inequality, moral automation and human de-skilling, robot consciousness and rights, dependency, social-psychological effects, and spiritual effects. Each of these topics will be given a brief discussion, though each deserves much deeper consideration.”, "author": [{"dropping-particle": "", "family": "Green", "given": "Brian Patrick", "non-dropping-particle": "", "parse-names": false, "suffix": ""}], "container-title": "Scientia et Fides", "id": "ITEM-1", "issue": "2", "issued": {"date-parts": [{"2018"}]}, "page": "9-31", "title": "Ethical reflections on artificial intelligence", "type": "article-journal", "volume": "6", "uris": [{"http://www.mendeley.com/documents/?uuid=c2240a78-d506-4b96-9898-2596b08ff2d9"}]}, "mendeley": {"formattedCitation": "(Green, 2018 en el que el comportamiento y las decisiones están estrechamente ligadas a un pensamiento religioso o teológico si se le quiere. Hemos perdido la conexión con el mundo real, para abordar nuestra propia realidad desde un espectro virtual dominado por la IA. (Solana, 2011)

Con ello, no se pretende más que indicar, que nuestra existencia, factor central del componente literario argumental de nuestra historia ha llegado a un impase; las estructuras sociales, al parecer, se presentan obsoletas ante la promesa de mejores opciones dentro de las realidades virtuales, se nos presentan opciones de ser personas *mejores*, en escenarios personalizados, donde cedemos nuestra independencia e identidad por un código de usuario, mientras que los seres como *Sophia* captan los recursos sociológicos y ontológicos que dejamos atrás. Todos nuestros defectos como individuos y como sociedad pueden ser solventados en la virtualidad, ya que en la realidad hemos fracasado en ser mejores, lo que pudimos lograr como especie fue construir seres *mejores*.

La intervención de la IA en nuestra historia como lo fue con esa deidad incrustada en el argumento literario de la antigua Grecia, se presenta como el punto de desenlace de nuestro devenir, con la presentación del *Deus Ex Machina*, así desde esta visión literaria se puede indicar que el ser humano ha perdido de una vez por todas, la batalla de la evolución por la supervivencia. Solo queda, recostarnos en el sillón, encender nuestros equipos de realidad virtual, y dejar que *nuestra vida pase en un instante por el frente de nuestros ojos*.

4. CONCLUSIONES

El desarrollo exponencial de la tecnología en términos de la Inteligencia Artificial ha generado una transformación en las relaciones sociales del ser humano posicionando a las máquinas a un nivel equiparable; entidades de Inteligencia Artificial como *Sophia*, han logrado una participación positiva en la sociedad al punto de obtener la ciudadanía en los Emiratos Árabes Unidos y recibir una condecoración en el departamento colombiano de Antioquia denominada el *Pasaporte Paisa*, que se asemeja a las llaves de la ciudad; configurándose así unos antecedentes que parecen superficiales pero pueden generar nuevas configuraciones sociales en la medida en que la IA continúe su proceso hacia la singularidad.

A partir de este proceso de ascenso social de aceptación de los seres humanos hacia las máquinas inteligentes, se puede proyectar una distorsión de las condiciones de relación entre estos dos y un futuro en el que el control sea ejercido desde las máquinas sobre los seres humanos. La evolución que presenta la singularidad tecnológica sirve de fundamento para el desarrollo prospectivo de escenarios distópicos, en los que estas máquinas son consideradas seres superiores a los seres humanos, consolidándose en un nivel superior que deje en el ostracismo a la especie humana en su papel en la tierra.

La percepción de los seres humanos frente a la inclusión en su vida diaria de diferentes fuentes y recursos de tecnología soportada en Inteligencia Artificial, aunado a la singularidad tecnológica, permite anticipar escenarios distópicos en los cuales la especie humana se ve relegada a un papel inferior en la escala evolutiva, llegando incluso al punto de considerar a las máquinas como deidades, provocando hasta la desaparición de la especie.

REFERENCIAS BIBLIOGRÁFICAS

- AGRAWAL, Ajay; GANS, Joshua y GOLDFARB, A., *The Economic of Artificial Intelligence*. Chicago and London: The University of Chicago Press, 2019, 643.
- ARTEAGA, S., “La inteligencia artificial de Facebook desarrolla su propio idioma”. *Computer Hoy*. [en línea], <<https://computerhoy.com/noticias/internet/inteligencia-artificial-facebook-desarrolla-su-propio-idioma-63928>>. [Consulta: 28 nov. 2018].
- BBC MUNDO., “Sophia, la robot que tiene más derechos que las mujeres en Arabia Saudita”. *BBB Mundo*. [en línea], <<https://www.bbc.com/mundo/noticias-41803576>>. [Consulta: 12 abr. 2021].
- BOSTROM, N., “Existential risks: Analyzing human extinction scenarios and related Hazards”. *Journal of Evolution and Technology*, 9, 2001, 1-36.
- BOSTROM, N., “Are we living in a computer simulation?”. *The Philosophical Quarterly*, 53, 221, 2003, 243-255.

- BOSTROM, N., "In defense of posthuman dignity". *Bioethics*, 19, 3, 2005, 202-214.
- BRZEZINSKI, Z., *La Era Tecnológica*. Buenos Aires: Paidós, 1979. 461.
- CAVALLARO, D., *Cyberpunk & Cyberculture: Science Fiction and the Work of William Gibson*. Somerset, New Jersey: Atheneum Press Ltd, 2000, 258.
- CHEANG WONG, J. C., "Ley de Moore, Nanotecnología y Nanociencias: síntesis y modificación de nanopartículas mediante implantación de iones". *Revista Digital Universitaria*. [en línea], 6, 7, 2005, 1-10. <http://www.revista.unam.mx/vol.6/num7/art65/jul_art65.pdf>. [Consulta: 12 abr. 2021]. ISSN: 1067-6079.
- COMITÉ ORGANIZADOR DE TOKIO 2020 DE LOS JUEGOS OLÍMPICOS Y PARALÍMPICOS, *Iniciativas innovadoras para los Juegos de Tokio 2020*. [en línea]. Tokyo: 2020.<<https://tokyo2020.org/es/games/vision-innovation/>>. [Consulta: 12 abr. 2021].
- CONTRERAS ESCRIBANO, S., *Deus Ex Machina: intervención en la Battersea Power Station*. Madrid: Universidad Politécnica de Madrid, 2015.
- DICK, P. K., *¿Sueñan los androides con ovejas eléctricas?* Editorial Planeta, S.A: España, 2019, 271.
- ESPARZA-OVIEDO, S. M., "Reseña sobre G. HOTTOIS, ¿El transhumanismo es un humanismo? (Trad. José Vicente Bonilla Pardo)", *Cuadernos Salmantinos de Filosofía*, 47, 2020, 629-634.
- FOSTER, .T., *The souls of cyberfolk: Posthumanism as vernacular theory*. 13a ed. United States: University of Minnesota Press, 2005, 334.
- FOUCAULT, M., "El sujeto y el poder". *Revista mexicana de sociología*, 50, 3, 1988, 3-20.
- FUKUYAMA, F., *Our Posthuman Future*. London: Profile books, 2002, 256.
- GREEN, B. P., "Ethical reflections on artificial intelligence". *Scientia et Fides*, 6, 2, 2018, 9-31.
- GUILLOT, W. M., "Emerging Technology: Creator of Worlds". *Strategic Studies Quarterly*, 10, 3, 2016, 3-8.
- GUMBRECHT, H. U., *In 1926: Living on the edge of time*. United States: Harvard University Press, 2009, 448.
- HEIDEGGER, M., *El concepto de tiempo*. 3ª Edición. España: Herder Editorial, 2008, 144.
- HEIDEGGER, M., *Ser y tiempo*. 5ª Edición. Chile: Editorial universitaria, 2015, 504.
- HUXLEY, A., *Un mundo feliz*. Bogotá: Plaza y Janés, S.A, 1979, 240.
- KURZWEIL, R., *The Singularity is near*. United States: Duckworth Overlook, 2008, 672.
- MAYORGA, I., "Oráculo Cyberpunk". *Revista Arcadia*, 146, 2017, 18-19.
- MCCAFFREY, L., *Storming the reality studio: A casebook of cyberpunk and postmodern science fiction*. United State: Duke University Press, 1991, 405.
- MORIN, E., *Los siete saberes necesarios para la educación del futuro*. Francia: Unesco, 1999, 72.

- OIM, *Informe sobre las Migraciones en el Mundo 2020*. [en línea]. Ginebra, Suiza: Organización Internacional para las Migraciones, <<https://publications.iom.int/books/informe-sobre-las-migraciones-en-el-mundo-2020>>. [Consulta: 12 abr. 2021].
- PALOU, N., *El día que un robot habló en la Asamblea General de la ONU [en línea] España*: microsiervos.com <<https://www.microsiervos.com/archivo/robots/dia-robot-hablo-asamblea-general-onu.html>>. [Consulta: 28 jul. 2021].
- PLAZA MAYOR MEDELLÍN, *Sophia, la robot humanoide por primera vez en Suramérica*. [en línea] Colombia: Plaza Mayor Medellín, <<https://plazamayor.com.co/sophia-la-robot-humoide-por-primera-vez-en-suramerica/>>. [Consulta: 12 abr. 2021].
- QUALLENBERG MENKES, I., “La diferencia entre la tecnología y ciencia”. *Revista de Ciencias Sociales de la Universidad Iberoamericana*, 7, 14, 2012, 231-255.
- RAMÍREZ, A., *La cuarta ola: la revolución biológica mezclada con la información. Diálogos con Alvin Toffler*. México: Universidad Veracruzana, 1998.
- RENDÓN, M., “El hombre que inventó las leyes de la robótica”. *Hipertextual*. [en línea], <<https://hipertextual.com/2017/01/isaac-asimov-robotica>>. [Consulta: 28 nov. 2018].
- SANABRIA, S., *De temporalidad y cotidianidad*. Bogotá: Universidad Nacional de Colombia, 2014.
- SCHOIJET, M., “La revolución científica y tecnológica y la sociedad postindustrial”. *Cuestiones Contemporáneas*, 1998, 127- 154.
- SOLANA, J., *La humanidad amenazada: gobernar los riesgos globales*. Madrid: Grupo Planeta Spain, 2011, 336.
- TOFFLER, A., “The Third Wave”. *Calculemos*. [en línea], <<http://calculemus.org/lect/07pol-gosp/arch/proby-dawne/materialy/waves.htm>>. [Consulta: 28 sep. 2018].
- VILLALBA GÓMEZ, J. A., “Transformaciones humanas a través de la tecnología”. *Revista Logos Ciencia & Tecnología*, 11, 1, 2019, 138–151.

POLARIZACIÓN SOCIAL Y CONTROL POLÍTICO: ALGUNAS CONSECUENCIAS DE LA INTELIGENCIA ARTIFICIAL Y LAS REDES SOCIALES PARA LA RAZÓN PRÁCTICA

*SOCIAL POLARIZATION AND POLITICAL CONTROL:
SOME CONSEQUENCES OF ARTIFICIAL INTELLIGENCE
AND SOCIAL NETWORKS FOR PRACTICAL REASON*

CARLOS RODRÍGUEZ-GORDO

Doctor en Filosofía
Universidad de Salamanca
Salamanca/España
carlosfilosofia@gmail.com
ORCID: 0000-0001-6452-0396

Recibido: 1/06/2021
Revisado: 31/08/2021
Aceptado: 6/09/2021

JOSÉ SARRIÓN-ANDALUZ

Doctor en Filosofía
Profesor Contratado Doctor
Universidad Pontificia de Salamanca
Salamanca/España
jsarrionan@upsa.es
ORCID: 0000-0001-9110-0549

Resumen: Este artículo analiza diversas implicaciones político-filosóficas de las Inteligencias Artificiales que dirigen las Redes Sociales. En primer lugar, se estudia la polarización social, la cual se observa como consecuencia de la búsqueda de beneficios económicos, que crea, a su vez, la necesidad de ofrecer contenidos polémicos para maximizar la presencia de los usuarios en las Redes Sociales. En segundo lugar, aborda el aumento del control político, debido a la capacidad predictiva y manipulativa de las IA que dirigen dichas redes. Por último, se centrará en la pretendida neutralidad y transparencia de dicha tecnología. En los tres momentos de este trabajo se emplearán herramientas procedentes del arsenal analítico que nos han legado algunos autores de la filosofía del siglo XX tales como Horkheimer, Marcuse, Habermas, Mosterín, Gramsci y Quintanilla, entre otros, cuyas propuestas se aplican a la realidad estudiada. Finalmente, se bosquejarán propuestas prácticas ante la problemática descrita.

Palabras clave: Algoritmos, Racionalidad, Horkheimer, Marcuse, Habermas, Mosterín, Gramsci, Quintanilla

Abstract: This article analyzes several political and philosophical implications of the Artificial Intelligence that directs the Social Networks. In the first place, social polarization is studied, which is observed as a consequence of the search for economic benefits, which in turn creates the need to offer controversial content to maximize the presence of users on Social Networks. Second, the increase in political control, due to the predictive and manipulative capacity of the AI that run these networks. Finally, it will study the alleged neutrality and transparency of said technology. In the three moments of this work, tools will be used from the analytical arsenal that some authors of 20th century philosophy have bequeathed to us, such as Horkheimer, Marcuse, Habermas, Mosterín, Gramsci and Quintanilla, among others, whose proposals are applied to the studied reality. Finally, practical proposals will be outlined in the face of the problem described.

Keywords: Algorithms, Rationality, Horkheimer, Marcuse, Habermas, Mosterín, Gramsci, Quintanilla.

INTRODUCCIÓN¹

La emergencia y actual desarrollo de la Inteligencia Artificial (en adelante IA) suscita debates de muy diversas implicaciones: metafísicas, epistemológicas, antropológicas, éticas, etc. En el presente artículo trataremos la Inteligencia Artificial, entendida como el conjunto de algoritmos de aprendizaje automático que permiten predecir el comportamiento o gustos de un usuario, desde el punto de vista de sus consecuencias para la filosofía política, analizando las implicaciones político-filosóficas de la aplicación de la IA a través de las conocidas como Redes Sociales.

Abordaremos dos de los impactos que la IA genera en la realidad política y social, ambos de importante calado filosófico: la polarización social y el aumento de la capacidad de control político sobre la sociedad. Analizaremos la aplicabilidad de algunas propuestas de la filosofía del siglo XX a esta nueva problemática. A continuación, trataremos la cuestión de la pretendida neutralidad y transparencia de las tecnologías de IA y sus consecuencias, para finalmente bosquejar algunas propuestas prácticas respecto a la realidad analizada.

1 Los autores deseamos mostrar nuestro agradecimiento a los profesores Eduardo G. Pardo, Profesor Titular de Ciencia de la Computación e Inteligencia Artificial de la Universidad Rey Juan Carlos, y a Álvaro Félix García Sánchez, profesor técnico de Formación Profesional en la especialidad de Sistemas y Aplicaciones Informáticas, por sus sugerentes comentarios que han enriquecido este artículo.

1. POLARIZACIÓN Y REDES SOCIALES

Especialmente a partir de 2005, hemos asistido a un incremento notable en el número de usuarios de Internet, acompañado de la fusión de muchas de las plataformas digitales, practicando un cambio del sentido original de la web 2.0., caracterizada inicialmente por la participación. Las empresas del Valle del Silicio californiano han seguido promocionando públicamente una imagen de creación horizontal y de esferas públicas de contacto colectivo, pero al mismo tiempo su potencial de crecimiento ha ido dependiendo casi exclusivamente de las ampliaciones de capital en forma de acciones que, a su vez, conllevan una notable exigencia de ganancias rápidas (Vukanovic, 2009).

Ya en 1995, el comunicólogo inglés Richard Barbrook y el tecnólogo Andy Cameron escribieron uno de los textos más reconocibles de las ideologías de la red que explicita esa simbiosis entre fomento de la participación y ánimo de lucro a la que le dan nombre: la ideología californiana:

This new faith has emerged from a bizarre fusion of the cultural bohemianism of San Francisco with the hi-tech industries of Silicon Valley. Promoted in magazines, books, TV programmes, Web sites, newsgroups, and Net conferences, the Californian Ideology promiscuously combines the freewheeling spirit of the hippies and the entrepreneurial zeal of the yuppies. This amalgamation of opposites has been achieved through a profound faith in the emancipatory potential of the new information technologies. (Barbrook y Cameron, 1996, pp. 44-45)²

Esta mencionada Ideología Californiana se basaría, de acuerdo con estos autores, en una visión optimista sobre el poder emancipatorio de las nuevas tecnologías, que acabará por conseguir la libertad en una utopía digital donde todo el mundo sea “alegre y rico”, y sea capaz de intervenir en su sociedad de una forma activa y directa. Una democracia que dichos autores califican como de estilo “Thomas Jefferson”, donde “todos los individuos puedan expresarse libremente en el ciberespacio”. Sin embargo,

by championing this seemingly admirable ideal, these techno-boosters are at the same time reproducing some of the most atavistic features of American society, especially those derived from the bitter legacy of slavery. Their Utopian vision of

2 “Esta nueva fe ha emergido de una extraña fusión entre la bohemia cultural de San Francisco y la industria de tecnología punta del Valle del Silicio. Promovida en revistas, libros, programas de televisión, sitios web, grupos de discusión de Usenet y conferencias de la Red, la Ideología Californiana combina, de forma promiscua, el espíritu despreocupado de los hippies y el ardor empresarial de los yuppies. Esta amalgama de realidades opuestas ha sido posible gracias a una profunda fe en el potencial emancipatorio de las nuevas tecnologías de la información” (traducción propia).

California depends upon a willful blindness toward the other— much less positive— features of life on the West Coast. (Barbrook y Cameron, 1996, p. 45)³

De acuerdo con este análisis, los *hippies* en *People's Park*, autoproclamados fundadores de la nueva izquierda mundial, rompiendo los límites de la estricta política posterior a la Segunda Guerra Mundial, antimilitaristas, acérrimos defensores de la igualdad racial, de género, siempre en guardia contra la homofobia, el consumismo exacerbado y la degradación medioambiental, habrían terminado siendo aliados como socios capitalistas, con los que antaño estaban enfrentados.

Determinismo ideológico e individualismo libertario habrían terminado unidos en una ideología ambigua que elimina lo más comprometido de ambos frentes, en una suerte de *minima moralia* en la que se combinan la economía de mercado con ciertos envoltorios contraculturales, al calor del determinismo tecnológico *mcluhiano*, con objetivos bien diferentes: los unos, confiados aún en el poder emancipatorio de las nuevas tecnologías; los otros, a la postre vencedores, convencidos de que el trabajo invertido en el desarrollo tecnológico acabará por construir un nuevo mercado.

Veinte años después de Barbrook y Cameron, algunos teóricos siguen considerando a las redes sociales como una oportunidad de empoderar a los usuarios con nuevas herramientas de conexión, al estilo de los sueños de Yochai Benkler o Jenkins a propósito de la web 2.0. Por ejemplo, el optimista Jeff Jarvis (2011) o Papacharissi (2010) quien, basándose en Goffman, Castells, Lessing o Deleuze, entre otros, propone la idea de un ciudadano-consumidor más exigente y crítico que aprovecha las ventajas de la técnica para ampliar sus posibilidades de intervención social (Papacharissi 2010, p. 64).

Otros pensadores, por su parte, caracterizan las redes como experiencias fallidas de participación democrática. Es el caso de K. Milberry y S. Anderson (2009, en Van Dijck, 2016) cuando afirman que “en su camino hacia el enclausamiento de Internet, las compañías de medios *online* crean membranas sinérgicas que prescriben circuitos que constriñen la libertad de los usuarios” (p. 65). En un sentido similar encontramos los escritos de Dyer-Whiteford y de Peuter (2009) a propósito de la precarización laboral en este tipo de industria (llegan a emplear el término “esclavitud”); o los trabajos de Jörgen Skågeby (2009), donde los metadatos sociales son descritos como una mera moneda en una economía de

3 “Al abanderar esta ideal aparentemente admirable, estos tecnófilos están reproduciendo, al mismo tiempo, algunas de las características más atávicas de la sociedad norteamericana, especialmente las derivadas del amargo legado de la esclavitud. Su visión utópica de California depende de una obstinada ceguera hacia las otras características —mucho menos positivas— de la vida en la Costa Oeste” (traducción propia).

la atención, que media inevitablemente las relaciones interpersonales, alejándolas del propósito original del espacio compartido.

Yendo un paso más allá que Skågeby, Cohen considera las redes sociales como plataformas de vigilancia que convierten la privacidad en acumulación de capital:

Not only is surveillance the method by which Facebook aggregates user information for third-party use and specifically targets demographics for marketing purposes, but surveillance is the main strategy by which the company retains members and keeps them returning to the site. [...] it is the unpaid labour of producer-consumers that facilitates this surveillance. (Cohen, 2008, p. 8)⁴

Y este es un evidente cambio en la industria digital, que tras la popularización de Internet se ha visto obligada a focalizarse en proveer nuevos servicios más que productos, forzando a las compañías de medios a desarrollar formas nuevas de monetización de la creatividad y las relaciones online. Mientras en el siglo XX su principal fuente de ingresos era la comercialización de bienes de consumo producidos en masa, en el XXI, con el modelo “gratis” (conocido como modelo *freemium*), los ingresos provienen fundamentalmente de la publicidad. Y esto es lo que potencia lo que conocemos como economía de la atención.

Como resultado de ese “ansiado” nuevo mercado se encuentran dos productos. El primero y más evidente: ingentes cantidades de contenido creado por los propios usuarios. Como ocurre con otros negocios de la web 2.0., las redes sociales están construidas sobre la comercialización de lo que Lazzarato (1996, p. 133) entiende como trabajo inmaterial, que realizan los propios usuarios de la plataforma de manera gratuita. Aunque, en el caso de Facebook, como afirma Cohen, “what distinguishes this particular social network is the way in which surveillance is fundamental to this process” (2008, p. 8)⁵. Y en este sentido, no se diferencia en absoluto de otros medios de comunicación tradicionales con conocidos patrones de relaciones asimétricas de poder entre trabajadores y propietarios, mercantilización y aprovechamiento de la audiencia propios del capitalismo.

Y de esa vigilancia surge el segundo de los productos, el petróleo del siglo XXI (y que los usuarios no proporcionan conscientemente): información sobre su comportamiento y preferencias que configura una vasta memoria digital cuyo

4 “La vigilancia no es solo el método por el que Facebook recopila información para uso de terceros para segmentar los públicos con propósitos de marketing, sino también la estrategia principal con la que la empresa retiene a sus miembros y los fuerza a regresar a la plataforma: [...] es el trabajo gratuito de los productores-consumidores lo que facilita la vigilancia” (traducción propia).

5 “Lo que distingue particularmente a esta red social es la forma en que la vigilancia es fundamental en este proceso” (traducción propia)”.

análisis permite, por primera vez en la historia de lo humano, no solo describir pasado y presente, sino predecir y pronosticar el futuro con inaudita precisión, incluso fuera de la propia plataforma (Tandera et al., 2017).

Como afirmará Christian Fuchs en 2009, a la hora de evaluar esta economía de la vigilancia se corre el riesgo de considerar, erróneamente, que la venta de privacidad es una consecuencia natural del apetito de los usuarios por conectarse y autopromoverse, una especie de tributo o penitencia cuya responsabilidad reside en ellos, cuando lo que en realidad ocurre es que es consecuencia lógica e inevitable de una economía política arraigada en la comercialización de los públicos (Fuchs, 2009), donde el usuario no puede intervenir, excepto si abandona la utilización de dicho recurso:

Mark Andrejevic has coined the notion of the digital enclosure (Andrejevic 2007), which means that interactive technologies generate “feedback about the transactions themselves” and that this feedback “becomes the property of private companies” (Andrejevic 2007, 3). Based on Marx and Smythe, we can argue that the contemporary Internet is a specific form of the digital enclosure that is based on the exploitation of prosumption: it is the realization of digital exploitation. Prosumers are digitally exploited. (pp. 298-299)⁶

También en palabras de Fuchs: “the Internet is a common information infrastructure. In its essence, it is part of the commons because all humans need to communicate in order to exist and reproduce themselves” (p. 269)⁷. El problema llega cuando nos damos cuenta de que una gran parte de Internet es “controlled by corporations and “immaterial” online labour is exploited and turned into surplus value in the form of the advertising-based Internet prosumer commodity” (p. 269)⁸. Una publicidad cada vez más efectiva y diseccionada que solo es posible gracias a la vigilancia y seguimiento de la actividad de los usuarios, el almacenamiento de la información extraída y su explotación con la ayuda de ordenadores y bases de datos (*big data*).

6 “Mark Andrejevich ha acuñado el término de “enclaustramiento digital” (Andrejevich 2007, p. 3), que significa que las tecnologías interactivas generan “retroalimentación sobre las interacciones en sí mismas” y esa retroalimentación se convierte en propiedad privada de las compañías. Basándonos en Marx y Smythe, podemos afirmar que el Internet contemporáneo es una forma específica de recinto digital basado en la explotación del *prosumidor*: es la realización de la explotación digital. Los prosumidores están digitalmente enclaustrados y digitalmente explotados” (traducción propia).

7 “Internet en esencia es parte del común porque todos los humanos necesitan comunicarse para existir y reproducirse” (traducción propia).

8 “Está controlada por grandes multinacionales y el trabajo inmaterial online está explotado y convertido en plusvalía en forma de comercialización de los prosumidores en el Internet basado en publicidad” (traducción propia)

Así, la vigilancia es crucial en la Web 2.0 comercial. Esta supone vigilar diferentes prosumidores que crean dinámica y permanentemente contenidos, navegan por diferentes perfiles de usuarios, interactúan con otros y se unen, y construyen nuevas comunidades. Todas estas actividades son monitorizadas y almacenadas, lo que les permite crear perfiles detallados de los usuarios que incluyen intereses y perfiles de comportamiento, que sirven como valor de cambio para los publicistas y que pueden ser utilizados gracias a laxas políticas de privacidad.

Todas estas plataformas de vigilancia lidian con el precario equilibrio entre estimular y explotar la actividad de los usuarios. Es obvio que sus resultados económicos dependen radicalmente de la voluntad de sus usuarios para contribuir a la producción de datos y permitir su recolección y análisis (Milberry y Anderson 2009, p. 409) y, por ello, trabajan profusamente en convertirse en más y más adictivas (Nakaya, 2015). Hawi y Samaha (2016) publican a propósito de las consecuencias de la adicción a las redes sociales, que “independently from culture and gender, there exists a negative relationship between social media addiction and self-esteem and a mediated negative relationship between social media addiction and satisfaction with life. As the psychology behind social media keeps making it more and more seductive, we expect the problem only to aggravate” (Hawi y Samaha, 2016, p. 8)⁹. Desesperanza que vemos confirmada en otros estudios de psiquiatría que confirman la relación entre el uso habitual de las redes sociales y problemas de comportamiento compulsivo o aislamiento social creciente (Savci y Aysan 2017), ansiedad (Liu y Ma 2020) o soledad (Baltaci, 2019).

¿Cuál es el mecanismo que se encuentra detrás de ese problemático aumento de la adicción? ¿Qué herramientas se encuentran tras el continuo incremento de la inversión publicitaria en redes sociales?

Un algoritmo es un conjunto de instrucciones o reglas definidas, no ambiguas, ordenadas y finitas que permite solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades (Cormen, Leiserson et al., 2001). Entre otros, los algoritmos que procesan la información obtenida de las redes sociales se encargan de decidir qué publicaciones visualiza un usuario en su móvil con relación a su actividad y realizar un análisis acerca de su comportamiento para poder optimizar el primer proceso.

En Facebook hay tres factores fundamentales que determinan lo que se muestra a cualquier usuario: repositorio (un *scraping* de todas las publicaciones de los contactos del usuario y de las tendencias fuera de la red), indicadores de

9 “Con independencia de cultura o género, existe una relación negativa entre la adicción a las redes sociales y la autoestima y una relación negativa mediada entre la adicción a las redes sociales y la satisfacción con la vida. Como la psicología detrás de las redes sociales las convierte en cada vez más atractivas, esperamos que estos problemas se agraven” (traducción propia)

rendimiento (tiempo de visualización, *likes*, comentarios, compartidos, etc.) y una predicción en función del historial del usuario y el tipo de contenido que ha consumido antes. Este análisis da como resultado una puntuación general que determina la visibilidad del contenido.

El 26 de mayo de 2020, el diario norteamericano *The Wall Street Journal* denunciaba que la compañía Facebook había sido advertida internamente sobre los efectos de polarización y comportamiento tribal que fomentaba sus algoritmos; pero decidió, no solo ignorar sus efectos, sino bloquear cualquier intento por evitarlos. De acuerdo con dicho diario, la compañía manejaba un informe, de 2016, de acuerdo con el cual el 64% de los usuarios que se habían unido a grupos extremistas lo habían hecho directamente por la recomendación de la herramienta. Sin embargo, los estudios de los directivos descubrieron que de haber seguido las modificaciones sugeridas por los empleados se habría comprometido notablemente el crecimiento en la adquisición de nuevos usuarios y limitado desproporcionadamente la visibilidad de los grupos conservadores norteamericanos.

Un estudio de ese mismo año mostraba que las narrativas conflictivas incitaban a la agregación de usuarios en cámaras de eco homogéneas, con independencia de la red social y el tipo de algoritmo. Y no solo eso sino que con un modelo de aprendizaje estadístico eran capaces de predecir con “buena precisión” cuándo un usuario iba a polarizarse (Besi, Zollo, Del Vicario, Puliga, Scala, Caldarelli, Uzzi Quattrociocchi, 2016)

En 2021, un estudio publicado en la revista *Computers in human behaviour* confirmaba que las noticias no sugeridas por un algoritmo no tenían un efecto positivo sobre adultos jóvenes, mientras que aquellas que estaban sugeridas por algún tipo de algoritmo servían como mecanismo para confirmar las creencias políticas existentes de los individuos, pudiéndoles llevar a puntos de vista más extremistas (Feezell, Wagner, Conroy 2021). En este mismo sentido Cho, Saifuddin, Hilbert, Liu & Luu publican un artículo en 2020 sobre las recomendaciones de YouTube cuyas conclusiones son muy esclarecedoras:

ideological reinforcement, [...] is heightened by political videos selected by the YouTube recommender algorithm based on participants' own search preferences and is muted to some degree if perceived preferences of participants' online social circle are used as inputs. This pattern is particularly pronounced for negative emotions such as anger and sadness. (pp. 14-15)¹⁰

10 “El refuerzo ideológico [...] se ve reforzado por los vídeos políticos seleccionados por el algoritmo de recomendación de YouTube en función de las preferencias de búsqueda de los propios participantes y se silencia hasta cierto punto si las preferencias recibidas del círculo social en línea de los participantes se usa como input. Este patrón es particularmente pronunciado para las emociones negativas como la ira y la tristeza” (traducción propia).

Y lo que es aún más importante, que “we also expect that algorithm effects on opinion reinforcement and polarization would be stronger for those who are engaged in politics” (pp. 15-16)¹¹. Por su parte, Pariser (2017) ha definido el sesgo ideológico construido por los algoritmos como “filtro burbuja”.

El gran problema que viene añadido es que ya se han superado las fronteras entre el comportamiento en red y fuera de ella y los resultados del algoritmo han pasado a entornos *offline* con trágicas consecuencias, como el conocido caso del tiroteo de Christchurch en 2019.

El 15 de marzo de 2019 se produjeron dos tiroteos masivos en diferentes mezquitas de Christchurch (Nueva Zelanda) llevados a cabo por el mismo individuo, que acabaron saldándose con 51 fallecidos y 40 heridos. Este es considerado el primer tiroteo retransmitido oficialmente por Facebook Live, al que acompañó, además, un manifiesto colgado en línea que posteriormente sería prohibido en el país.

El atentado se relacionó con movimientos supremacistas blancos y de extrema derecha de la *alt-right*, como también lo habían sido el ataque de Hanau en Alemania o el tiroteo de Noruega. De hecho, el Índice de Terrorismo Global de 2019, publicado por el *Institute for Economics and Peace* reflejaba que los ataques terroristas de ese tipo habían aumentado un 320% en los últimos cinco años (Pawels 2019).

¿Qué está produciendo esa popularización del discurso extremista y de odio? En primer lugar, en el caso del terrorismo de extrema derecha se observa una banalización y normalización de los discursos de odio en los medios de comunicación. Mientras que un ataque yihadista se califica inmediatamente de terrorismo, a la hora de evaluar acciones violentas procedentes de la extrema derecha, se tiende a utilizar términos eufemísticos o directamente legitimadores (Perry y Scrivens 2020). En segundo lugar, varios estudios apuntan a que la radicalización de estos grupos está facilitada, cuando no promovida directamente, por las propias redes sociales a pesar de que eso contraviene con sus políticas de uso. Véanse, por ejemplo, las publicaciones de Tarleton Gillespie (2010) o Ben-David y Matamoros-Fernández (2016) para el caso español:

In the case of the Spanish extreme-right political parties, we have shown that Facebook hosts an increasing volume of covert discriminatory practices that not only circulate data and content, but also trigger overt hate speech by the parties' followers. However, the circulation of covert discrimination is not limited to the Spanish case

11 “Esperamos que los efectos del algoritmo sobre el refuerzo de la opinión y la polarización sean más fuertes para aquellos que están involucrados en política” (traducción propia).

and may have broader implications on the role Facebook plays as a cultural intermediary. (p. 1188)¹²

En este sentido, también abundan estudios cuantitativos a nivel multinacional como el de Asongu, Orim y Nting en 2019:

It is apparent from the findings that the managing body of Facebook may not be doing enough in prevention of the use of its social media platform to fight terrorism. [...] In essence, more complex algorithms need to be developed to trace and address online content that is characterized by extremist rhetoric, violent images, organization of violence and propagation of hate (p. 19).

Aún así, los hay que aprovechan tal coyuntura y el evidente sesgo de Facebook para su propio beneficio. En abril de 2021, el cineasta Caolan Robertson, ayudante del ex *YouTuber* de extrema derecha Tommy Robinson, afirmó en un reportaje para el *New York Times* que “hacer ediciones inteligentes y centrarse en la confrontación podía ayudar a atraer millones de visitas en YouTube y otros servicios” (Metz, 2021). Además, aprendió que el algoritmo de recomendación de YouTube llevaba a menudo a la gente a videos más extremos y “Eso hizo que hiciéramos vídeos más extremos” (Metz, 2021). “En los vídeos parece que solo intentamos averiguar lo que está pasando, reunir información, entender a la gente pero en realidad, estábamos tratando de encontrar la manera más incendiaria de hacerlos enojar” (Metz, 2021).

Como hemos observado hasta ahora, las redes sociales, gracias a sus algoritmos, constituyen una fuente de polarización social. Ahora bien, no se trata simplemente de que estas sean meros receptáculos ni mecanismos de divulgación de corrientes extremistas, altavoces neutrales de las mismas. Por el contrario, como hemos podido observar, precisamente los algoritmos de IA que impulsan algunas funciones de las redes sociales dan lugar, como consecuencia necesaria, a una polarización social.

Las redes sociales no son simplemente un hueco aprovechable por corrientes extremistas presentes en la sociedad (independientemente de que este fenómeno también puede producirse), sino que, además, dichas redes sociales siguen una lógica que conduce inevitablemente a la polarización social. Sin embargo, dicha lógica no tiene como objetivo en sí mismo el desarrollo de dicha polarización

12 “En el caso de los partidos de extrema derecha españoles, hemos demostrado que Facebook alberga un volumen creciente de prácticas discriminatorias encubiertas que no solo hacen circular datos y contenido, sino que desencadenan un discurso de odio abierto por parte de los seguidores de dichos partidos. [...] La circulación de la discriminación encubierta no se limita al caso español y puede tener implicaciones más amplias sobre el papel que desempeña Facebook como intermediario cultural” (traducción propia).

social, sino que su objetivo es el aumento de los beneficios empresariales, aspecto connatural a cualquier empresa con ánimo de lucro. Los beneficios empresariales, en una empresa de redes sociales, aumentan si se produce un aumento de la optimización del rendimiento, lo cual es posible en la medida en que aumenta la permanencia de los usuarios en dichas redes. Y precisamente por este motivo, las empresas del sector buscan mecanismos para generar estímulos que aumenten el tiempo de permanencia de los usuarios en las mismas. Puesto que su beneficio es extraído de la publicidad, cuanto más tiempo se encuentren los usuarios en las redes sociales, más aumentan los beneficios empresariales de las mismas.

Los mecanismos psicológicos descritos en este apartado muestran cómo las redes sociales explotan la tendencia del ser humano a la sociabilidad. Cabrían mencionarse aquí los debates clásicos de la filosofía acerca de la sociabilidad humana, desde la teoría de la sociabilidad natural de Aristóteles a las diversas tesis contractualistas. No obstante, sería baladí para la cuestión que nos ocupa tratar el carácter natural o social del origen de dicha sociabilidad. Lo que no debe pasar desapercibido es que, como hemos expuesto en este apartado, la confrontación entre diversas posiciones supone un aumento significativo de los estímulos, lo cual, a su vez, garantiza la tasa de permanencia en dichas redes sociales.

La neutralidad no genera tanta interacción como la polaridad. La Inteligencia Artificial de las redes sociales, (como ya hemos dicho, entendida como el conjunto de algoritmos de aprendizaje automático que permiten predecir el comportamiento o gustos de un usuario) es capaz de localizar qué contenido genera polémica para cada usuario de manera individual, y mostrárselo en consecuencia. Un contenido determinado terminará, finalmente, convirtiéndose en un espacio de confrontación entre los partidarios más entusiastas y los más firmes detractores del mismo. Dado que la confrontación entre diversas posiciones supone un aumento de la permanencia en las redes sociales, los algoritmos tienden a favorecer automáticamente dicha confrontación entre los usuarios, creando así mecanismos de polarización social. Por lo tanto, si bien el fin último de la Inteligencia Artificial no es la polarización; sino el aumento de permanencia, que, a su vez, está motivado por la búsqueda de beneficios, la consecuencia necesaria e inevitable es el aumento de la polarización social.

1.1. RACIONALIDAD DE FINES FRENTE A RACIONALIDAD DE MEDIOS: APLICACIÓN A LA ACTIVIDAD DE LA IA

Nos encontramos, por lo tanto, ante una lógica, un modo de racionalidad instrumental, que genera como consecuencia necesaria la pérdida de racionalidad, que no es un debate nuevo en filosofía. Max Horkheimer (2002) ya analizó cómo

el predominio de la razón de medios sobre la razón de fines había tenido como consecuencia la derrota de la racionalidad, que habría abonado el terreno para la emergencia de la extrema derecha. Horkheimer no es el único en tratar las consecuencias del adelgazamiento de la razón. Herbert Marcuse (2001) habló en 1964 del hombre unidimensional. Por su parte, Jesús Mosterín (1974) habló de la racionalidad incompleta.

Horkheimer (2002) expuso cómo el progreso científico-técnico defendido por la Ilustración se había convertido en su contrario, en el camino a la barbarie. Esto se produjo debido a que, de acuerdo con el autor, la razón había transitado, a lo largo de la historia, desde un modelo de razón objetiva como “principio inviscerado de la realidad y en ella operante” (Horkheimer, 2002, p. 46), hacia un modelo de razón subjetiva, entendida como la “capacidad para calcular probabilidades y determinar los medios más adecuados para un fin dado” (p. 46). Esta razón subjetiva, también llamada formalizada en la terminología de Horkheimer, consiste en un “conjunto de determinaciones que hacen posible la calculabilidad de acciones desde el punto de vista instrumental, de la efectividad de los medios disponibles y bajo el aspecto estratégico, también, de la validez de la elección a llevar a cabo unos medios dadas unas preferencias, unos medios y unas condiciones contextuales” (Muñoz Veiga, 1987, p. 27). El tránsito de la razón objetiva a la subjetiva, considerado por Horkheimer como un proceso histórico necesario, tiene sin embargo consecuencias fundamentales y graves. A diferencia de Weber, quien veía en el desgajamiento entre la esfera axiológica y la esfera cognitiva reducida a lo científico-técnico un aumento de la racionalidad (Muñoz Veiga, 1987), Horkheimer encuentra una pérdida de la misma, debido a la cual las acciones no pueden ser planificadas, enjuiciadas y justificadas más que en base a su rendimiento. En suma, la formalización o subjetivización de la razón supone que la razón de fines se vea adelgazada hasta ser sustituida por una razón de medios. Lo único racional es la medición, el cálculo y la operación con datos, y todo intento de ir más allá es considerado mera superstición. La razón se ha transformado, desde su estado original de investigación sobre las preguntas últimas de la realidad, en simple herramienta de cálculo y aplicación: en su afán por desprenderse de los grandes relatos mitológicos, el pensamiento occidental ha ido progresivamente renunciando a las grandes preguntas, pasando de una razón objetiva (que pretendía emanar de la estructura real del mundo) a una razón subjetiva (la capacidad del sujeto para, a partir de unas premisas o unos medios, llegar a unas conclusiones o a unos fines, la “capacidad de calcular probabilidades y determinar los medios más adecuados para un fin dado” (Horkheimer, 2002, p. 48)).

El problema de la racionalidad subjetiva es que al convertirse la razón en un mero instrumento que permita dilucidar cómo adecuar unos medios a unos fines

dados, está dejando de lado la pregunta por los fines mismos. A partir de ese momento los fines son dados desde fuera, y, como discípulo en cierto grado de Marx, Horkheimer sabe que dichos fines son dados desde el sistema de producción. El ser humano, la ciencia, la naturaleza: todos son medios doblegados al único fin del sistema de producción: el principio de autoconservación y su consecuencia racional: la lógica de dominación.

La concepción de la razón como mero agente de medios, que renuncia a la elección o consideración de fines, da lugar a que finalmente la razón vuelva a un estado precientífico y semimitológico. En su intento por superar las viejas teorías a través del desprecio a sus categorías, para Horkheimer, finalmente la sociedad no hace otra cosa que abrir la puerta al dogmatismo. El pensamiento se convierte en mero gestor: administración, no decisión.

Así, a diferencia de Weber, Horkheimer observa que la pérdida de la razón objetiva entraña problemas profundos, mostrando su afición a un tipo de racionalidad alternativa a la instrumental (Muñoz Veiga, 1987): la racionalidad sustantiva u objetiva cuya tarea es determinar fines relacionados con ideas o principios concretos, como justicia, igualdad, felicidad, democracia o tolerancia. Por el contrario, la sustitución de la verdad objetivamente vinculante por la verdad calculística, formalizada o subjetiva desemboca finalmente en un relativismo moral. Horkheimer entendió así el auge del nazismo. La lógica interna del proceso de formalización de la razón (en términos de Horkheimer) dio lugar a un “adelgazamiento” de la razón que fue tendiendo a la liquidación de la misma: el relativismo y la pragmatización del pensamiento sustituyen la lógica de la verdad por la lógica de la probabilidad; la instrumentalización del lenguaje y su reducción a propaganda en función del consumo, despojado de su condición de expresión de la verdad; el vaciamiento de los grandes ideales de la Modernidad y la consiguiente depotenciación de la esfera política, de la democracia, su instrumentalización en función del “sondeo”, de la opinión y de los intereses de las mayorías, la producción industrial de la cultura como mercancía y la perversión de la religión. La descomposición de la unidad, dadora de sentido de las viejas imágenes metafísico-religiosas del mundo, lleva al acoso de la identidad de los sujetos sociales y la solidaridad humana, a la quiebra del centro cohesionador de mundos vitales, a la fragmentación y atomización social, y a otras consecuencias ético-políticas que, en definitiva, producen una invertebración del mundo (Muñoz Veiga, 1987). La consecuencia última de este ingente proceso de “cosificación” (Horkheimer, 2002, p.74) es la capitulación de la razón ante la realidad existente, ante el poder de lo que triunfa en la historia. En definitiva: una nueva sumisión de la razón ante el mito y el ídolo. La ciega furia de la lógica instrumental pierde su capacidad superadora, su cualidad esencial de negación de lo existente, y queda reducida a instrumento, a *ancilla administrationis*, perdiendo su orientación a la verdad que trasciende los hechos.

Herbert Marcuse (2001), por su parte, trató la técnica contemporánea como un medio opresivo. De acuerdo con este autor, la tecnología no cumple esta función simplemente por sus consecuencias inmediatas (alienación laboral, aumento del paro, etc.), sino porque inunda al conjunto de la sociedad con su forma específica de racionalidad, reduciendo al ser humano a una sola dimensión: la de la búsqueda del confort y el interés económico. Como consecuencia, se produce una uniformización del pensamiento y la asimilación por el sistema de las corrientes contestatarias. Así, para Marcuse, se produce una dominación sobre una masa de gente, mediante la creación de nuevas necesidades, limitando la racionalidad a su nivel instrumental y teniendo como consecuencia una interiorización de valores y necesidades construidos para reducir al ser humano a un consumidor de los productos de la técnica. Las mercancías son convertidas en fetiches, siguiendo la noción marxiana del fetichismo de la mercancía (Marx, 2017), que adoctrinan a los seres humanos. La dominación tecnológica, según Marcuse, carente de fines racionales, reduce todo a medios para ejercer su dominación, imposibilitando la comprensión objetiva de la realidad y conduciendo así a la subordinación al orden existente (López Sáenz, 1988).

Marcuse también observó que la técnica se convierte, dentro del sistema capitalista, en una de las fuentes principales de violencia y alienación. Cuanta mayor producción, mayor es la necesidad de convertir a los seres humanos en máquinas de consumo. La opresión de la técnica entra en la misma esencia del ser humano y conduce a su alienación, a su conversión en un mero engranaje del sistema productivo que debe acumular beneficios indefinidamente. Al aplicar esta idea a la temática de las IA que rigen las redes sociales debe tenerse en cuenta que, en este caso, el ser humano ya no solo corre peligro de ser convertido en un engranaje del sistema productivo como productor (Marx) y como consumidor (Marcuse), sino que su información, sus datos, sus gustos y preferencias políticas y culturales, en suma, su propia vida, son también convertidos en una mercancía en sí misma, dando un paso más allá en la alienación.

Igualmente, Habermas (1986) también planteó reflexiones sobre la ciencia y la técnica que pueden resultar de gran valor para el debate que nos ocupa, centrándose en el rol que desempeña la técnica al asumir la función de la ideología como legitimación de la dominación política. De acuerdo con dicho autor, la técnica contemporánea alejaría a la sociedad del necesario debate público en torno a la política, facilitando así una tecnocracia de acuerdo con la cual la satisfacción de las necesidades se realiza gracias a la técnica, haciendo equivalentes progreso técnico y progreso social. De este modo, la tecnología quedaría instaurada como un recurso de poder para los tecnólogos. El problema, de acuerdo con las tesis de Habermas, radicaría en que la lógica tecnológica se convierte en una lógica cibernética, independizada de la lógica humana. Se habría producido una

autonomización del sistema científico-técnico-económico, ajeno a cualquier control social, basado en la idea de que determinadas decisiones (de carácter político, técnico, militar o económico) no pueden tomarse en el ámbito de la discusión pública. Dicha idea, en realidad, oculta unos intereses dados. La propuesta de Habermas pasa por hacer explícitos dichos intereses para que puedan ser discutidos por el conjunto de la sociedad y decididos democráticamente. Habermas propone, en definitiva, complementar la racionalidad instrumental con una racionalidad práctica valorativa (que encajaría con la noción de Horkheimer de razón de fines), la cual se emprende desde la acción comunicativa (1981). De acuerdo con esta posición, no hay una negación de la importancia de la racionalidad instrumental propia de las ciencias empírico-analíticas, sino una reconsideración del ámbito al que debe ser aplicada, junto a un reconocimiento de los intereses práctico y emancipatorio correspondientes a las ciencias histórico-hermenéuticas y a las ciencias sociales críticas, respectivamente, como condición para que la ciencia contribuya al pleno desarrollo del ser humano libre y moral en el conjunto de la vida humana social e individual: en suma, a la realización del ser humano.

El filósofo español Jesús Mosterín (1973), por su parte, distinguió entre racionalidad teórica y racionalidad práctica. La primera era entendida como la posesión de evidencia suficiente para una creencia dada. La segunda consiste en cinco condiciones: la conciencia clara de los fines de una conducta, el conocimiento de los medios necesarios para lograrla, la capacidad de poner en práctica dichos medios en la medida posible, la disposición para dar preferencia a los fines posteriores frente a los anteriores en caso de conflicto entre fines de la misma línea y distinto grado de proximidad, y que los fines últimos perseguidos sean compatibles entre ellos mismos. De este modo, es posible ser racional teóricamente pero no prácticamente (si bien no al contrario).

Por otro lado, también plantea que el sujeto desarrolla un comportamiento racional cuando realiza sus elecciones de modo que las mismas maximicen su felicidad (Mosterín, 1993). La mejor vida posible para una persona depende tanto de la naturaleza humana genérica como de la experiencia personal, su herencia cultural y genética y el entorno en que vive. Así, la racionalidad sería una actitud de segundo orden que permite analizar, criticar y, en su caso, cambiar nuestros valores de primer orden, con el fin de que el sistema global de valores dé preferencia a aquellos valores objetivamente preferibles en caso de conflicto entre valores opuestos.

El peligro no radica en la racionalización como tal, ya sea en su esfera teórica o en la práctica, sino en las racionalizaciones incompletas (Mosterín, 1974, 1978), ante lo cual la solución propuesta pasa por la adopción de una racionalidad completa.

Como hemos visto, diversos autores de la filosofía del siglo XX observaron la importancia de la racionalidad de fines frente a la de medios, cada uno desde su propia perspectiva. Este punto de vista, aplicado a la actividad de la Inteligencia Artificial que riga las redes sociales, nos parece que puede tener una gran fecundidad.

A lo largo de este apartado hemos visto que las redes sociales tienen como fin la maximización del tiempo de permanencia de los usuarios en las mismas, fin que, a su vez, es dirigido por otro fin último, a saber, la maximización de beneficios económicos (en este caso impuesto por los propietarios de las empresas que gestionan dichas redes). Y la consecuencia necesaria de dicho fin es la polarización social, toda vez que el medio que los algoritmos ponen en marcha automáticamente es la presentación de posiciones enfrentadas a los usuarios para aumentar la polémica, aprovechando así los mecanismos psicológicos descritos más arriba. Desde esta perspectiva, y en lo que se refiere a la capacidad de polarización, las redes sociales, como tal, no parecen tener un fin político en sí mismas, sino que pretenden más bien desplegar todos los medios necesarios, sin importar sus consecuencias, con arreglo a un fin que no es decidido por el conjunto de la sociedad. Empleando la terminología de Horkheimer, podríamos considerar que la IA actúa en este sentido como una especie de razón formalizada hasta el extremo. Siguiendo a Habermas, diríamos que es una razón independizada de cualquier control democrático por parte de la sociedad. Con Marcuse, convendríamos en que desarrolla una función opresiva y violenta. Y completando con los conceptos de Mosterín, podríamos añadir que la IA contribuye a una racionalidad incompleta en la sociedad, donde se produce una ausencia de racionalidad práctica (toda vez que la jerarquización de los medios solo se realiza por parte de la IA que controla las redes sociales).

Efectivamente, y en suma, parece evidente que la racionalidad de la IA, orientada a la maximización de beneficios, impone a la sociedad un aumento de la irracionalidad.

No nos parece casualidad que precisamente los primeros frankfurtianos tomaran el análisis de las causas del nacionalsocialismo como base para estudiar sus teorías acerca de la racionalidad instrumental. Ciertamente, una de las cuestiones de actualidad en materia de redes sociales es cómo estas hacen de instrumento de difusión de ideas de extrema derecha, tal y como hemos expuesto en este apartado.

Ahora bien, la polarización no es la única consecuencia de las redes sociales. Existe otro fenómeno de relevancia, consecuencia de dicha tecnología, y también enormemente relacionado con las temáticas de la Escuela de Frankfurt: el aumento de la capacidad del control político-social sobre la humanidad.

2. REDES SOCIALES Y CONTROL POLÍTICO

La segunda parte de este artículo pretende centrarse en otra consecuencia de la Inteligencia Artificial que actúa como motor de las redes sociales: el aumento de la capacidad de control político, ya sea por parte del Estado o por parte de grupos económicos con aspiraciones políticas.

En su búsqueda de beneficios, las redes sociales han permitido la construcción de perfiles sociológicos de los usuarios, denominados perfiles OCEAN. Con una herramienta, creada por el investigador Christopher Wylie, son capaces de predecir el comportamiento de un usuario basándose en su actividad digital. Para ello los algoritmos de Inteligencia Artificial generan un perfil psicométrico¹³ de los usuarios, que a continuación se utiliza para segmentar los anuncios en función del público que objetivamente sea más receptivo a sus diferentes mensajes. Esta estrategia es llamada microsegmentación.

La microsegmentación tiene evidentes fines comerciales cuya eticidad podríamos cuestionar. Pero seguramente su aplicación con más consecuencias sociales llega en el momento en el que puede usarse con fines políticos.

Uno de los ejemplos más paradigmáticos de que no existe voluntad ni económica ni política de paliar las tremendas consecuencias sociales de las diferentes plataformas es que el propio poder las utiliza en propio beneficio, obviando los mínimos morales más evidentes y convirtiendo la práctica política en el lodazal de odio y desinformación que conocemos.

Aunque existe literatura sobre la importancia de la comunicación interpersonal a través de las redes sociales en las elecciones presidenciales de 2008 en las que Barack Obama acabó convirtiéndose en el 44° Presidente de los Estados Unidos (Harfoush, 2010), es a partir de 2016 cuando el papel de las redes sociales parece más importante. El propio Trump afirmó el 13 de noviembre que: “el hecho de que tenga ese poder en términos de números de Facebook, Twitter, Instagram, etc. creo que me ayudó a ganar” (Morin, 2016).

Puede confirmarse el salto cualitativo en esta campaña que, por primera vez, dedicó la mayor parte de su presupuesto publicitario a las redes sociales en lugar de a los medios de comunicación tradicional (Hendricks, 2017). Mientras otros candidatos republicanos habían llegado a gastar cerca de 90 millones de dólares, el presupuesto de Trump rondó los 10 millones aunque la presencia orgánica de

13 El Centro Psicométrico de la Universidad de Cambridge ha habilitado una demo, con fines de concienciación, que permite a cualquier usuario subir los datos de sus Redes Sociales para observar su propia predicción de su perfil psico-demográfico basándose en la huella digital de su comportamiento online: <https://applymagicsauce.com/demo>

Trump era notablemente superior a la del resto de los candidatos (Wells et al., 2016)

Esta estrategia le supuso una mayor visibilidad que la de la candidata demócrata Clinton, tanto en visitas a la web y búsquedas en Google, como en menciones en redes sociales, especialmente en Twitter. Algunos analistas, no obstante, apuntan que esta superioridad estaba inflada artificialmente mediante diversas redes de bots que forzaban los *trending topics* (Bessi y Ferrara, 2016).

La inversión del equipo de campaña no se limitó exclusivamente a la promoción publicitaria sino que aprovecharon para crear una base de datos mediante encuestas online, llamadas telefónicas y *big data*, que les permitía conocer mejor a los electores para optimizar los mensajes publicitarios en aquellos lugares clave, especialmente para “hombres blancos descontentos” (Conley, 2017). El proyecto Álamo, como se denominó a esta base de datos, disponía de información de 14 millones de indecisos localizados en aquellos Estados más igualados (Green y Issenberg, 2016) que se complementó más adelante con los trabajos de perfilado psicológico de Cambridge Analytica.

Según informaciones de *The New York Times*, *The Observer* y el diario británico *The Guardian*, un extrabajador de esta consultora, Christopher Wylie, confirmó que para completar dicho perfilado habían accedido a los datos de 50 millones de usuarios de Facebook y desarrollado una herramienta “para influir en las elecciones presidenciales de Estados Unidos” que era capaz de “predecir el sentido del voto de millones de personas” (Cadwalladr, 2018), con el objetivo de influir en sus decisiones y provocar la victoria del multimillonario neoyorkino. A través de un aparentemente inocuo test de personalidad diseñado por el profesor de Psicología de la Universidad de Cambridge, Alexandr Kogan, la herramienta recopilaba diversa información sobre los distintos usuarios de la red social, incluidos sus gustos, inclinaciones políticas e ideales.

Con esta información, realizaron una intensa campaña contra la reputación de Hillary Clinton y propiciaron el descontento entre los propios votantes demócratas, sobre todo dirigida a hombres blancos, mujeres jóvenes y afroamericanos.

Aprovechando la enorme repercusión de Donald Trump en Twitter, su intensidad de publicación y la naturalidad y espontaneidad que mostraba en su perfil, pudo permitirse una estrategia más agresiva e impulsiva que hacía, en primer lugar, identificarse más a su electorado (Gutiérrez 2016) y, en segundo lugar, generar más ataques, por parte de los críticos, lo que, a la postre, le supuso también una enorme visibilidad.

En Facebook, en cambio, optaron por una estrategia totalmente diferente con unos objetivos bien distintos¹⁴. El primero, la obtención de fondos, el segundo, para la publicación de *post* invisibles (Green y Issenberg, 2016) y el tercero y más importante: para la difusión de noticias falsas. Incluso en este aspecto se investigó la implicación del Gobierno ruso. Gracias a esta estrategia se construyó una realidad alternativa, una *posverdad* (Suiter, 2016).

Otro de los casos paradigmáticos del uso y abuso de las redes sociales para la modificación de los resultados electorales es el Brexit. Un estudio cuantitativo sobre la discusión alrededor del Brexit en páginas oficiales de noticias del Reino Unido reflejó el surgimiento de dos comunidades polarizadas, que se generaron como resultado de la actividad de los usuarios y no tenían relación con el contenido mismo (Del Vicario et al., 2017).

Como hemos podido observar, la microsegmentación y el micromárquetin, posibles gracias a los perfiles OCEAN creados por los algoritmos de IA presentes en las redes sociales, presentan consecuencias que no pueden ser ignoradas por la filosofía política.

2.1. CONSECUENCIAS PARA LA RAZÓN PRÁCTICA

Una muy evidente es la erosión de la democracia. En un contexto de manipulación del ideario político, por parte de grandes fortunas, capaces de manipular a escala individual a inmensas cantidades de población, nos parece incuestionable que no pueden cumplirse las condiciones necesarias para que se desarrolle una política discursiva tal y como se entiende en la teoría social habermasiana, de acuerdo con la cual la ciudadanía debe implicarse en la resolución de los conflictos, a través de deliberaciones que tienen lugar tanto en la esfera formal como informal (Habermas, 1998). La inmensa capacidad de injerencia que permiten las redes sociales a pequeñas élites económicas parece poner en duda la posibilidad de una sociedad deliberativa conforme a los criterios de la acción comunicativa, y se aleja de las condiciones para una *Situación Ideal de Habla*.

Más bien parece que el Estado, bajo estas condiciones, no puede ser concebido como un mero espacio neutral en disputa por las diversas tendencias político-ideológicas que dialogan (o compiten) por un espacio parlamentario. Por el contrario, el Estado se convierte en un espacio asimétrico, privilegiado para los sectores con mayor acceso al capital económico, en un sentido similar al que

¹⁴ Puede consultarse un análisis acerca de dicha campaña en Facebook en: Kushin, Yamamoto y Dalisay (2019).

estudiaron teóricos del Estado de tendencia gramsciana en el siglo XX (Poulantzas, 1972, 1974; Miliband, 1973; Jessop, 2017).

Efectivamente, las posibilidades que ofrecen las redes sociales a las élites económicas para impulsar iniciativas políticas, asentar opiniones o, directamente, vencer procesos electorales, muestran la capacidad de estas tecnologías en manos de los sectores económicamente dominantes en una sociedad. El marxista italiano Antonio Gramsci (1970) escribió acerca de la capacidad de las clases dominantes de un bloque histórico dado, para ejercer la dirección moral de la sociedad a través de la noción de hegemonía.

Gramsci enfocó la cuestión del Estado ampliando la concepción del mismo, entendiéndolo como la combinación de la sociedad política y la sociedad civil. De este modo, más allá de la concepción meramente jurídica del Estado, Gramsci incorporaba aspectos privados, tales como los partidos políticos, los medios de comunicación, los sindicatos, las asociaciones, etc.

Gramsci rechaza que el Estado sea un mero reflejo de una clase preconstituida, como se desprende de la lectura de Marx (García-Santesmases, 1986), para dar lugar a la concepción ampliada del Estado, resultado de las relaciones orgánicas entre Estado, sociedad política y sociedad civil. De este modo, la dominación de una sociedad por las clases dirigentes no se encontraría basada meramente en la coerción, sino que es posible gracias a la capacidad de generar consensos entre las clases subalternas. En suma, las clases dominantes no solo gobiernan mediante la fuerza, sino que también dirigen mediante el convencimiento, el consenso o la creación de un sentido común, dirección que se encuentra posibilitada gracias a un conjunto de organizaciones de la sociedad civil, las cuales presentan una relativa autonomía del poder político del Estado (entendido aquí en su sentido jurídico), que, según la teoría de Gramsci, realizan una difusión de una concepción del mundo que favorecería los intereses de las clases dominantes.

Parece evidente que esta concepción gramsciana adquiere una nueva relevancia a la luz de la capacidad de difusión ideológica que las redes sociales permiten a las clases dirigentes de la sociedad. La capacidad de ideologización o de transmisión de un determinado “sentido común” al conjunto de la población que tenían los medios de comunicación convencionales parece palidecer al lado de la impecable eficacia del micromárquetin, la microsegmentación y los perfiles OCEAN, posibles gracias a los algoritmos de Inteligencia Artificial que poseen las redes sociales que están disponibles para quien posea los recursos económicos para financiarlos.

3. OPACIDAD Y ALIENACIÓN EN LAS INTELIGENCIAS ARTIFICIALES

A lo largo de los anteriores apartados de este artículo hemos estudiado cómo el aumento de la polarización social y el control político son una consecuencia directa de las tecnologías de IA estudiadas y, concretamente, del planteamiento económico que subyace a dichas tecnologías, dirigido a la maximización de beneficios, esto es, la lógica intrínseca del capital. La polarización es producto de la tendencia de los algoritmos a aumentar la permanencia de los usuarios en las redes sociales, mientras que el control político lo es de la capacidad de los algoritmos de microsegmentar los contenidos basándose en los perfiles psicométricos diseñados al efecto. Y esto nos conduce al debate acerca de la necesidad de democratizar el control de la tecnología.

El debate acerca de la neutralidad de la tecnología ha ocupado un lugar importante de la filosofía. Para introducir esta cuestión podríamos preguntarnos: ¿es una pistola una tecnología neutral? La pistola puede emplearse –cabría responder– para la defensa de la paz o para el ataque en una guerra injusta, por ejemplo. Podría contraargumentarse que, independientemente del uso que se otorgue a dicha arma, la finalidad de su tecnología no es otra que matar. Finalmente, cabría añadirse que la pistola por sí misma no puede hacer daño o bien alguno, si no es empuñada por un ser humano que escoja el fin con que sea empleada.

El filósofo español Miguel Ángel Quintanilla (2017) defiende la posibilidad de tecnologías entrañables frente a tecnologías alienantes. La noción de alienación usada por este autor, si bien de raíz marxiana, se distingue del uso que le dio Marx en que la alienación tecnológica no se produce solo en el uso de la maquinaria dentro del sistema productivo, sino en el uso de la tecnología en general. De acuerdo con dicho autor, “las dos fuentes principales de la alienación tecnológica tienen que ver, por una parte, con la opacidad del diseño de los sistemas técnicos y, por otra, con la simplificación de los criterios de evaluación que predominan en los procesos de diseño y difusión de las tecnologías” (Quintanilla, 2017, p. 21). La opacidad es definida como “la ocultación de su función y estructura, del propósito al que sirve, sus prestaciones o *affordances*” (Quintanilla, 2017, p. 21). Frente a las tecnologías alienantes se plantean como alternativa las tecnologías entrañables, que serían las que cumplieran 10 condiciones: abierta, polivalente, dócil, limitada, reversible, recuperable, comprensible, participativa, sostenible y socialmente responsable. De especial interés nos parece la séptima condición, esto es, la comprensibilidad, que consiste en que el usuario debe saber lo que está haciendo cuando emplea la tecnología (Quintanilla, 2002, 2012, 2017, Parselis 2016a, 2016b). A la luz de esta propuesta, cabría preguntarse: ¿pueden ser considerados en este momento como tecnologías entrañables los algoritmos de IA que rigen las Redes Sociales?

Para tratar de responder esta pregunta debemos considerar, en primer lugar, que, progresivamente, las redes neuronales existentes en la mayoría de los sistemas de IA se comportan como una caja negra, de la cual solo conocemos sus entradas y su salida sin conocer los detalles internos de su implementación, cuyos valores se ajustan automáticamente a medida que reciben nuevas entradas, donde no hay conocimiento *a priori*, sino que la red descubre por sí misma características, relaciones, correlaciones y categorías de forma totalmente autónoma (Hinton y Sejnowski, 1999). En 2016, la revista *Scientific American* confirmaba que Facebook estaba implementando este tipo de tecnología para enseñar a las máquinas a predecir el comportamiento humano sin ningún tipo de ayuda externa.

A finales de 2016 y principios de 2017, Youtube dejó de utilizar un algoritmo para su funcionamiento interno y pasó a controlarse mediante dos redes neuronales adversas, ambas con un único objetivo fundamental: incrementar el tiempo de visionado de vídeos de cada usuario, maximizando los ingresos publicitarios. A diferencia de un algoritmo, donde hay normas secuenciales, en una red neuronal solo existen unos *inputs* y un resultado deseado y es la propia red neuronal la que se va recalibrando mediante aprendizaje. No hay leyes ordenadas ni secuenciadas. La máquina puede evolucionar y no se controla lo que ocurre dentro o fuera de ella. Esto significa, de facto, que es mucho más complicado cambiar el funcionamiento de una red neuronal. Lo único que se puede hacer es reforzar el aprendizaje y esperar los cambios. Un estudio de 2019 publicado por la Cornell University indicaba que los canales con contenido extremo promovían mucho más la participación, lo que significaba que a la postre acababan siendo más sugeridos (Ribeiro et al., 2019). Otras publicaciones, en cambio, parecen indicar que, si bien el sistema de recomendación no promueve la radicalización, tampoco hace nada por evitarla (Ledwich y Zaitsev, 2019) y sí por reforzarla (Faddoul, Chaslot y Farid, 2020).

El paso de que sean algoritmos los que tomen las decisiones de sugerencia de contenidos en las redes sociales a que sea una Inteligencia Artificial la que se encarga de hacer esas sugerencias tiene importantes consecuencias para la transparencia del proceso. Mientras que un algoritmo es fácilmente analizable, desconocemos el comportamiento que tiene una Inteligencia Artificial porque se encuentra en permanente evolución.

Son ampliamente conocidas las Inteligencias Artificiales que se convierten en amplificadoras de sesgos y, por lo tanto, en una causa más en el problema de la polarización política. El 24 de marzo de 2016, Microsoft tuvo que “apagar” al robot que lanzaron en Twitter, TAY, porque en apenas unas horas en contacto con los usuarios radicalizó su discurso y acabó alabando a Hitler, insultando a las feministas y realizando diferentes ofrecimientos sexuales. Microsoft prometió corregir el comportamiento del robot y relanzarlo, pero hasta la fecha aún no ha

sido reactivado. La hipótesis de la tecnológica de Redmond fue que un grupo de usuarios del imageboard *4chan* realizó un “esfuerzo coordinado” para reforzar el aprendizaje de TAY para que desarrollase determinados patrones de conducta.

Idéntico destino, el ser apagados, tuvieron Bob y Alice de Facebook, dos robots diseñados para negociación que, en lo que respecta a sus relaciones con humanos, fueron mucho más educadas y fructíferas que las de TAY, pero de repente, empezaron a comunicarse entre sí después de inventar un nuevo idioma que facilitaba y automatizaba su trabajo mejor que el inglés en el que habían sido programados.

Un grupo de investigadores del Instituto Tecnológico de Massachusetts creó un sistema de Inteligencia Artificial obsesionado con la muerte y con rasgos psicopáticos y le sometieron a una exposición prolongada a imágenes de violencia, tras lo cual fue sometido a un test de manchas de Rorschach. Los resultados, como no podía ser de otro modo, fueron que tal exposición había trastornado al sistema y donde una inteligencia artificial normal veía un “grupo de pájaros apoyados en una rama de árbol”, Norman (como se llamaba la IA) veía una electrocución de un hombre y similares (Wakefield, 2018).

El 14 de abril de 2017 un grupo de investigadoras publicaron en *Science* un artículo donde mostraban que los sistemas de aprendizaje artificial también asumían los sesgos humanos sexistas y racistas, advirtiendo así sobre los problemas de los sistemas de aprendizaje no supervisado (Caliskan, Bryson y Narayanan, 2017).

Esta es una de las grandes diferencias entre la IA y una tecnología tradicional, como la de pistola de la que hablábamos antes. Comprendemos el funcionamiento interno de una pistola (su física, su química, su mecánica). Podemos, en este sentido, considerar que, independientemente de la valoración que realicemos acerca de su neutralidad, la pistola al menos satisface la séptima condición para ser una tecnología entrañable en el sentido que plantea Quintanilla: la comprensión de su funcionamiento interno. Sin embargo, de acuerdo con lo estudiado en este artículo, no parece que las IA que rigen las redes sociales puedan satisfacer dicha condición¹⁵. Por el contrario, la opacidad de esta tecnología más bien parece que permite clasificarla como una tecnología alienante, siguiendo el modelo de Quintanilla.

Por otro lado, y como ya hemos expuesto en este artículo, las consecuencias sociopolíticas de las tecnologías IA se deben precisamente a la neutralidad de dicha tecnología. La expansión de la extrema derecha en YouTube o la capacidad

¹⁵ Es más, difícilmente pueden considerarse satisfechas la primera condición (apertura), la tercera (docilidad y control), la octava (participación) y la décima (responsabilidad social).

de implantar discursos xenófobos en las redes sociales no se deben a una voluntad, por parte de las empresas dueñas, de expandir dichos discursos, sino más bien a su lógica de maximización de beneficios. Intrínseco al modelo tecnológico está el modelo económico. En este sentido, el capitalismo es amoral. Lo cual no significa que no se sustente en una moralidad: el afán de lucro. La politización que crean las redes sociales es, en este sentido, producto de su propia neutralidad política.

En todo caso, la discusión acerca de la neutralidad de la tecnología puede presentar interés desde un punto de vista de la filosofía de la ciencia, pero no tanto desde el punto de vista de la política de la ciencia, análogamente a como la discusión acerca de la intención de una acción determinada presenta interés desde el punto de vista de la filosofía moral, pero es irrelevante desde el punto de vista de las éticas consecuencialistas.

Sin embargo, la inevitable (en cuanto técnicamente obligada) opacidad de las tecnologías de IA nos sitúa ante la incapacidad manifiesta de convertirlas en tecnologías entrañables.

En este sentido, tiene plena actualidad la advertencia de Manuel Sacristán, de mediados de los años 70 del siglo XX, según la cual los clásicos problemas epistemológicos de la filosofía de la ciencia habían dado paso a problemas más urgentes, relacionados con las consecuencias de la ciencia, problemas de un carácter más ontológico que epistémico. Según Sacristán, la filosofía de la ciencia había de albergar un núcleo de política de la ciencia si pretendía enfrentarse a los problemas fundamentales del presente (Sarrión-Andaluz, 2017).

CONCLUSIÓN

El filósofo español Manuel Sacristán planteó sus reflexiones acerca de filosofía y política de la ciencia en el contexto de la toma de la conciencia acerca de la problemática ecológica (a raíz de la publicación del Informe al Club de Roma sobre los límites del crecimiento), así como en el marco del riesgo de una hecatombe nuclear en los tiempos de la Guerra Fría. Sin pretender negar que estos riesgos hayan desaparecido, las consecuencias sociales de la IA suponen una actualización de estas reflexiones.

Detrás de la irracionalidad social que generan los algoritmos de la IA se encuentra una lógica de maximización de beneficios. Toda vez que las tecnologías de IA se encuentran ya presentes entre nosotros, y que su evolución no parece reversible en este momento, ¿qué alternativas políticas pueden plantearse frente a este fenómeno?

Nos atrevemos a bosquejar dos planteamientos, estrechamente relacionados entre sí:

En primer lugar, el establecimiento de mecanismos de control democrático sobre la tecnología. Independientemente de que podamos o no concebir la energía nuclear como una tecnología neutral cuyos fines dependen de su uso (puede ser empleada para curar el cáncer o para destruir ciudades enteras en segundos), lo cierto es que se entiende que dicho tipo de energía debe encontrarse bajo estricta regulación y control públicos. Un planteamiento similar parece exigir la tecnología IA, cuyas consecuencias sociopolíticas hemos observado en este artículo.

En segundo lugar, este control democrático solo será posible a través de una educación que forme a la ciudadanía en el uso de estas tecnologías, socializando el saber acerca de las mismas y comprendiendo sus consecuencias y funcionamiento interno, dentro de lo posible. En este aspecto, cabe plantearse aquí la noción de “educación en ciberciudadanía responsable” (Rodríguez-Gordo, 2014).

REFERENCIAS BIBLIOGRÁFICAS

- ANDREJEVIC, Mark, *iSpy: surveillance and power in the interactive era*. Lawrence: University Press of Kansas, 2007.
- ANDREJEVIC, M., The work that affective economics does”. *Cultural Studies*, 25(4-5): 2011, 604-620.
- ASONGU, S.A., STELLA-MARIS I. Orim, S-M.I. y NTING, R.T., Terrorism and Social Media: Global Evidence, *Journal of Global Information Technology Management*, 22(3), 2019, 208-228.
- BALTACI, Ö., The Predictive Relationships between the Social Media Addiction and Social Anxiety, Loneliness, and Happiness. *International Journal of Progressive Education*, 15(4), 2019, 73-82.
- BARBROOK, R., CAMERON, A., The Californian Ideology. *Science as culture*, 6, 1996, 44-72.
- BEN-DAVID, A. y MATAMOROS-FERNÁNDEZ, A., Hate speech and convert discrimination on Social Media: Monitoring the Facebook Pages of Extreme-right political parties in Spain. *International Journal of Communication*, 10, 2016, 1167-1193.
- BESSI, A. y FERRARA, E., Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, 2016, 21(11).
- BESSI, A., ZOLLO, F., Del Vicario, M., Puliga, M., Scala, A., Caldarelli, G., et al. (2016) Users Polarization on Facebook and Youtube. *PLoS ONE* 11(8). <https://doi.org/10.1371/journal.pone.0159641>
- CADWALLADR, C. (18 de marzo de 2018). ‘I made Steve Bannon’s psychological warfare tool’: meet the data war whistleblower. *The Guardian*. <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>

- CALISKAN, A., BRYSON, J., NARAYANA, A., Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 2017, 183-186.
- CHO, J., SAIFUDDIN A., HILBERT, M., LIU B., LUU, J., Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political Polarization. *Journal of Broadcasting and Electronic Media*, 2020, 64.
- COHEN, N., The Valorization of Surveillance: Towards a Political Economy of Facebook. *Democratic Comunique*, 22(1), 2008, 5-22.
- CONLEY, B., Thinking what he says: Market research and the making of Donald Trump's 2016 Presidential campaign. En J. Gillies (ed.), *Political marketing in the 2016 U.S. presidential election*, 2017, pp. 29-48. Springer.
- CORMEN, LEISERSON, RIVERST, Stein., *Introduction to algorithms*. Cambridge: MIT Press, 2001.
- DEL VICARIO, M.; ZOLLO, F.; CALDARELLI G.; SCALA, A. y QUATTROCIOCCI, W., Mapping social dynamics on Facebook: The Brexit debate, *Social Networks*, 50, 2017, 6-16.
- DYER-WHITEFORD, N. y de PEUTER, G., *Games of Empire: Global Capitalism and Video Games*. University of Minnesota Press, 2009.
- FADDOUL, M.; CHASLOT, G. y FARID, H., *A longitudinal analysis of YouTube's promotion of conspiracy videos*. Cornell University, 2020.
- FEEZEL, J.; WAGNER, J. y CONROY, M., Exploring the effects of algorithm-driven news sources on political behaviour and polarization, 2021. *Computers in human behaviour*, 116. <https://doi.org/10.1016/j.chb.2020.106626>
- FUCHS, C., An alternative view of privacy on Facebook. *Information*, 2, 2009, 140-165.
- GARCÍA-SANTESMASES, A., *Marxismo y estado*. Centro de Estudios Políticos y Constitucionales, 1986.
- GILLESPIE, T., The politics of "platforms." *New Media & Society* 12(3), 2010, 347-364.
- GRAMSCI, A., *Antología. Selección, traducción y notas de Manuel Sacristán*. Siglo XXI, 1970.
- GREEN, J. & ISSENBERG, S. (27/10/2016), Inside the Trump bunker, with days to go. *Bloomberg Businessweek*. <https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go>
- GUTIÉRREZ, C. (6/07/2016, Clinton vs Trump, la contienda en las redes sociales. *Fortuna*. <https://revistafortuna.com.mx/contenido/clinton-vs-trump-la-contienda-en-las-redes-sociales/>
- HABERMAS, J., *Ciencia y técnica como "ideología"*. Tecnos, 1986.
- HABERMAS, J., *Teoría de la acción comunicativa*. Taurus, 1987.
- HABERMAS, J., *Facticidad y validez*. Trotta, 1988.
- HARFOUSH, R., *Yes we did. Cómo construimos la marca Obama a través de las redes sociales*. Gestión 2000, 2010.
- HAWI, N. S. y SAMAHA, Maya, The relations among social media addiction, self-esteem, and life satisfaction in University Students. *Science Computer Review*, 35(5), 2017, 576-586.

- HENDRICKS, J. A., The social media election of 2016. En R. E. Denton (ed.), *The 2016 US presidential campaign*, 2017, pp. 121-150. Springer.
- HINTON, G. y SEJNOWSKI, T., *Unsupervised Learning and Map Formation*. MIT Press: Foundations of Neural Compilation, 1999.
- HORKHEIMER, M., *Crítica de la razón instrumental*. Trotta, 2002.
- HORWITZ, J. y SEETHARAMAN, D. (26 de mayo de 2020). Facebook executives shut down efforts to make the site less divisive. *The Wall Street Journal*.
- JARVIS, J., *Public Parts: How Sharing in the Digital Age Improves the way we work and live*. Simon & Schuster, 2011.
- JESSOP, B., *El Estado. Pasado, presente, futuro*. Catarata, 2017.
- KUSHIN M. J.; YAMAMOTO M. y DALISAY, F., Societal Majority, Facebook, and the Spiral of Silence in the 2016 US Presidential Election. *Social Media + Society*, 5(2), 2019. <https://doi.org/10.1177%2F2056305119855139>
- LAZZARATO, M., Immaterial labor. En M. Hardt y P. Virno (eds.) *Radical thought in Italy: a potential politics* (pp. 133-147). University of Minnesota Press, 1996.
- LEDWICH, M. y ZAITSEV, A., *Algorithmic extremism: examining YouTube's rabbit hole of radicalization*. Cornell University, 2019.
- LIU, C. y MA, J., Social media addiction and burnout: The mediating roles of envy and social media use anxiety. *Curr Psychol*, 39, 2020, 1883-1891.
- LÓPEZ SÁENZ, La crítica de la racionalidad tecnológica en Herbert Marcuse. *Enrahonar*, 14, 1988, 81-93.
- MARCUSE, H., *El hombre unidimensional*. Ariel, 2001.
- MARX, K., *El Capital. Crítica de la economía política*. Siglo XXI, 2017.
- METZ, C. (21 de abril de 2021) Un exyoutuber de la ultraderecha explica cómo se propaga el odio. *The New York Times*. <https://www.nytimes.com/es/2021/04/21/espanol/youtube-ultraderecha-odio.html>
- MILBERRY, K. y ANDERSON, S., Open Sourcing Our Way to an Online Commons: Contesting Corporate Impermeability in the New Media Ecology. *Journal of Communication Inquiry*, 33(4), 2009, 393-412. <https://doi.org/10.1177/0196859909340349>
- MILIBAND, R., *El estado en la sociedad capitalista*, Siglo XXI, 1973.
- MOSTERÍN, J., El concepto de racionalidad. *Teorema, revista internacional de filosofía*. 3 (4), 1973, 455-480.
- MOSTERÍN, J., La incompleta racionalidad. *Teorema, revista internacional de filosofía*. 7 (1), 1974, 55-58.
- MOSTERÍN, J., *Racionalidad y acción humana*. Alianza, 1978.
- , *Filosofía de la cultura*. Alianza, 1993.
- MORIN, R. (11/12/2016). Trump says social media was key to victory. *Politico*. <https://www.politico.com/story/2016/11/donald-trump-social-media-231285>
- MUÑOZ VEIGA, J., Materiales para una crítica de la modernidad (Max Horkheimer y los orígenes de la «teoría crítica»). *Logos. Anales Del Seminario De Metafísica*, 22, 1987, 13-36.
- NAKAYA, A.C., *Internet and Social Media Addiction*. Reference Point Press, 2015.

- PAPACHARISSI, Z. A., *A Private Sphere. Democracy in a Digital Age*, Polity Press, 2010.
- PARISER, E., *El filtro burbuja. Cómo la red decide lo que leemos y lo que pensamos*. Taurus, 2017.
- PARSELIS, M., “El valor de las tecnologías entrañables”, *CTS: Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, 2016a. 11(32).
- *Las tecnologías entrañables como marco para la evaluación tecnológica*, [Tesis de doctorado]. Universidad de Salamanca, 2016b.
- PAWELS, A. (20/11/2019). El ascenso del terrorismo de extrema derecha. *EsGlobal*. <https://www.esglobal.org/el-ascenso-del-terrorismo-de-extrema-derecha/>
- PERRY, B. y SCRIVENS, R., Who’s a Terrorist? What’s Terrorism? Comparative Media Representations of Lone-Actor Violence in Canada. En J. Littlewood, L. Dawson y S. Thompson (Ed.), *Terrorism and Counterterrorism in Canada* (pp. 242-264). University of Toronto Press, 2020.
- POULANTZAS, N., *Sobre el estado capitalista*. Laia, 1974.
- POULANTZAS, N., *Poder político y clases sociales en el estado capitalista*. Siglo XXI, 1972.
- PYBUS, J., Trump, the First Facebook President: Why Politicians Need Our Data Too. En C. Happer, A. Hoskins, & W. Merrin (Eds.), *Trump’s Media War*, 2019, (1 ed., pp. 227-240). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-94069-4_14
- QUINTANILLA, M. A., “La democracia tecnológica”, *Arbor: Ciencia, pensamiento y cultura*, 683, 2002, 637-652.
- , “El pensamiento científico y la ideología de izquierdas”, *Página Abierta*, 218, 2012.
- , Tecnologías entrañables: un modelo alternativo de desarrollo tecnológico. En M. A. Quintanilla, M. Parselis, D. Sandrone y D. Lawler, D., *Tecnologías entrañables. ¿Es posible un modelo alternativo de desarrollo tecnológico?*, 2017, pp. 15-53. Catarata.
- RIBEIRO, M. H.; OTTONI, R.; WEST, R.; ALMEDIA V. A. F. y MEIRA Jr., W., *Auditing Radicalization Pathways on YouTube*. Cornell University, 2019.
- RODRÍGUEZ-GORDO, C., *Medios de Comunicación social y Ciber ciudadanía Responsable* [Tesis de doctorado]. Universidad de Salamanca, 20114.
- SARRIÓN-ANDALUZ, J., *La noción de ciencia en Manuel Sacristán*. Dykinson, 2017.
- SAVCI, M. y AYSAN, F., Social-emotional model of internet addiction. *Psychiatry and Clinical Psychopharmacology*, 27(4), 2017, 349-358. <https://doi.org/10.1080/24750573.2017.1367552>
- SKÅGEBY, J., Exploring qualitative sharing practices of social metadata: expanding the attention economy. *The Information Society*, 25(1), 2009, 60-72.
- SUITER, J., Post-truth politics. *Political Insight*, 7(3), 2016, 25-27.
- TANDERA, T.; HENDRO, SUHARTONO, D.; WONGSO, R. y PRASETIO Y.L., Personality Prediction System from Facebook Users, *Procedia Computer Science*, 116, 2017, 604-611 <https://doi.org/10.1016/j.procs.2017.10.016>.
- VAN DIJCK, J., *La cultura de la conectividad: una historia crítica de las redes sociales*. Siglo XXI, 2016.

- VUKANOVIC, Z., Global Paradigm Shift: Strategic Management of New and Digital Media in New and Digital Economics. *The International Journal on Media Management*, 11, 2009, 81-90.
- WAKEFIELD, J. (2/06/2018). Are you scared yet? Meet Norman, the psychopathic AI. *BBC News*. <https://www.bbc.com/news/technology-44040008>
- WELLS, C.; SHAH, D. V.; PEVEHOUSE, J. C.; YANG, J.; PELLED, A.; BOEHM, F.; LUKITO, J.; GHOSH, S. y SCHMIDT, J. L., How Trump drove coverage to the nomination: Hybrid media campaigning. *Political Communication*, 33(4), 2016, 669-676.

GOBERNANZA GLOBAL DE LA IA*

GLOBAL GOVERNANCE OF AI

ANÍBAL MONASTERIO ASTOBIZA

Departamento de Filosofía I
Facultad de Filosofía y Letras
Universidad de Granada
Granada/España
amastobiza@ugr.es
ORCID: 0000-0003-1399-5388

Recibido: 19/04/2021

Revisado: 30/07/2021

Aceptado: 6/09/2021

Resumen: El desarrollo de la IA crece cada vez más y se aplica en casi todos los sectores económicos y/o sociales y, por consiguiente, nacen nuevas oportunidades, pero también nuevos retos. En este artículo, pretendo presentar algunos de estos nuevos retos, cuáles son las formas más apropiadas para afrontarlos y qué enfoque ético seguir para una gobernanza global de la IA. Para ello, en primer lugar, comentaré el proceso de difusión y penetración de la tecnología subrayando la inclinación humana de hacer uso de distintas herramientas para manipular y transformar la realidad. En segundo lugar, hablaré de la revolución digital y algunos desafíos que plantea para la sociedad y el individuo. Finalmente, presentaré un esbozo de los esfuerzos colectivos para la gobernanza global de la IA.

Palabras clave: gobernanza, ética, IA, retos, oportunidades.

Abstract: AI is increasingly growing and being applied in almost all economic and/or social sectors and, as a consequence, new opportunities are arising, but also new challenges. In this article, I intend to present some of these new challenges, what are the most appropriate ways to address them and what ethical approach to follow for a global governance of AI. To do so, I will first discuss the process of diffusion and pervasiveness of technology, highlighting the human disposition to make use of different tools to manipulate and transform reality. Secondly, I will discuss the digital revolution and some of the challenges it poses for society and the individual. Finally, I will present an outline of the collective efforts for the global governance of AI.

Keywords: governance, ethics, AI, challenges, opportunities.

* Este trabajo se ha realizado en el marco de los proyectos: EXTEND: Bidirectional Hyper-Connected Neural System” (Horizon2020) y EthAI+3 (PID2019-104943RB-100).

1. INTRODUCCIÓN

Industria, gobiernos, academia y organizaciones civiles buscan nuevas formas de aplicación y desarrollo de la Inteligencia Artificial (IA)¹ para múltiples propósitos. La IA promete a la industria y empresas maximizar beneficios; a los gobiernos automatizar procesos administrativos; a los académicos descubrir nuevas preguntas y respuestas científicas; y la sociedad civil (Organizaciones No-Gubernamentales, por ejemplo) quiere que las tecnologías digitales no incrementen las desigualdades en el acceso al estado de bienestar de los ciudadanos, no discrimine, ni conculque derechos fundamentales.

El uso y desarrollo de la IA crece cada vez más y se aplica en casi todos los sectores económicos y sociales (McAfee y Brynjolfsson, 2017) y, por consiguiente, esto significa que nacen nuevos retos. En este artículo, pretendo presentar algunos de estos nuevos retos², cuáles son las formas más apropiadas para afrontarlos y qué enfoque ético seguir para una gobernanza global de la IA. A medida que el desarrollo tecnológico madura los avances tecnológicos son ubicuos y su impacto e influencia alcanzan a múltiples facetas de nuestras vidas y trabajo. Por todo ello, nos vemos en la necesidad de alinear estas poderosas nuevas tecnologías con nuestros valores fundamentales –integrar dichos valores en el diseño e implementación de la tecnología.

Si queremos que la tecnología –y en concreto la tecnología de la IA– esté al servicio de la gente y esté dirigida al bien común, la gobernanza global de la misma es una potencial solución porque implica una coordinación colectiva para alcanzar un consenso mínimo que permita tener el control sobre dicha tecnología, a saber, la IA y la robótica, que es vista por muchos como una amenaza existencial (Bostrom, 2014).

1 La IA es difícil de definir dado que distintos investigadores tienen una idea diferente. La IA puede verse como la ciencia e ingeniería que aspira a crear inteligencia en sistemas artificiales. La expresión IA se introdujo en los años 50 del siglo XX (McCarthy, Minsky, Rochester y Shannon, 1955), pero su historia profunda se remonta a los trabajos de filósofos como Ramón Llull, Thomas Hobbes o G. W. Leibniz que intentaron axiomatizar y formalizar reglas y principios para describir la razón. Una forma de describir la IA que casi puede servir de definición aproximativa es la que ofreció, en su día, Marvin Minsky: “La IA es la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si las hicieran los seres humanos” (Minsky 1968/2003).

2 Los tres retos (más sobre esto en la sección 2) que he seleccionado para dar cuenta de la necesidad de un marco de gobernanza global de la IA son: *Automatización de las noticias y propaganda política computacional*; *Sesgos algorítmicos y discriminación*; y *El futuro del trabajo: desempleo tecnológico*. Los retos han sido seleccionados de manera subjetiva de la literatura de investigación siguiendo el criterio de conexión de la IA con la gobernanza. Es evidente la elección de estos retos. El uso de la IA en cada uno de estas áreas: Comunicación política, discriminación de usuarios y consumidores y mercado laboral, respectivamente; plantea una transformación radical de actividades esenciales a escala internacional y global.

En primer lugar, comentaré el proceso de difusión y penetración de la tecnología subrayando la propensión humana de hacer uso de distintas herramientas para manipular y transformar la realidad. En segundo lugar, hablaré de la revolución digital y los desafíos que plantea para la sociedad y el individuo. Finalmente, presentaré un esbozo de los esfuerzos colectivos para la gobernanza global de la IA.

Por gobernanza global de la IA –para los propósitos de este artículo– se entiende dos sentidos interdependientes. Por una lado, una gobernanza global de la IA *descriptiva* que hace referencia a diversos procesos por los que se toman decisiones y se implementan dichas decisiones y una gobernanza global de la IA *normativa* que hace referencia al conjunto *correcto* o *bueno* de tales decisiones. Cuando se habla de una gobernanza global de la IA *bueno* se quiere decir efectiva, legítima, adaptativa e inclusiva (Renda, 2019; Dafoe, 2020). En algunos lugares de este artículo se utilizará la expresión “gobernanza global ética de la IA” para referirse a esta segunda acepción o sentido normativo de gobernanza y en otros, simplemente, “gobernanza global de la IA” para referirse a la primera acepción o sentido descriptivo de gobernanza. Sin embargo, nótese la interdependencia de ambos sentidos, aunque se puedan distinguir lógicamente. También el lector podrá encontrar que “regulación” y “gobernanza” se utilizan como términos coextensivos. Esto es así porque la gobernanza como proceso de toma de decisiones incluye, normas, instituciones, reglas y, por supuesto, leyes.

La situación de la pandemia de la Covid-19 muestra claramente la necesidad de una coordinación global para la resiliencia ante catástrofes, ya sean naturales o provocadas por la acción del ser humano. Prepararnos para una gobernanza global de la IA u otras tecnologías de uso-dual³ es uno de los temas de gran importancia si queremos que la tecnología se dirija hacia un uso para el bien común. Países desarrollados y en vías de desarrollo han de ser conscientes de la necesidad de una gobernanza global de la IA, pero tampoco debemos olvidar la capacidad regulativa y sus contextos. El Sur Global⁴ tiene bajos niveles de capacidad regulativa e institucional y quizá una menor garantía de la gestión de riesgos y esto es algo que puede generar retos adicionales.

3 Uso-dual hace referencia al uso de la tecnología con propósitos militares o comerciales. Más sobre esto en la sección 1.2.

4 A efectos de este artículo, el “sur global” se refiere a los países en desarrollo de África, América Latina y Asia incluyendo Oriente Medio. El sur global es un término que sustituye a “tercer mundo” y “países en desarrollo” en muchos debates académicos, aunque no está exento de controversia. El término sur global trasciende las fronteras y abarca países que comparten un pasado colonial o mantienen poblaciones oprimidas y privadas de derechos, incluso en Occidente o países desarrollados (véase, de Sousa Santos, 2016).

Europa es un promotor activo de la gobernanza global de la IA, pero organizaciones internacionales, múltiples países y regiones también entienden la necesidad de operativizar e implementar una gobernanza global de la IA.

En la siguiente sección, 1.1., introduciré y contextualizaré brevemente la historia del desarrollo tecnológico y su relación con el progreso civilizatorio.

1.1. DIFUSIÓN Y PENETRACIÓN DE LA TECNOLOGÍA

Desde la aparición del linaje homínido –en concreto a partir del género *Homo Habilis* (Oakley, 1959)– y hasta llegar al ser humano, nuestra especie no ha parado de instrumentalizar la realidad física. Gracias al desarrollo de la tecnología el ser humano ha podido ocupar todos los nichos ecológicos de este planeta. Pero hasta llegar a la hegemonía del ser humano en el control de los recursos naturales, si uno mira a la historia de nuestra especie nos llevó mucho tiempo hasta alcanzar una plena prosperidad.

Uno puede defender la idea de que en la historia global de la humanidad solo ha habido dos grandes eventos transformacionales: 1) la revolución agrícola y 2) la revolución industrial. Hasta hace 15.000 o 10.000 años los seres humanos sobrevivían cazando, recolectando o buscando carroña. Sin embargo, una serie de transformaciones culturales permitió el desarrollo de distintas técnicas para cultivar la tierra, hacer acopio del excedente, domesticar animales y un rápido proceso de urbanización y estratificación social (Herrera y García-Bertrand, 2018).

El segundo gran evento, la revolución industrial –que comenzó en el siglo XIX– es transformacional en muchos sentidos pero, principalmente, porque por primera vez en la historia de la humanidad el crecimiento económico permitió la explosión demográfica. Hasta la revolución industrial el crecimiento económico, la invención y desarrollo de nuevas tecnologías, únicamente permitía la mera, y difícil, subsistencia de la población. Estábamos en una condición *malthusiana* o en palabras del filósofo inglés Thomas Hobbes, la vida era “solitaria, pobre, sucia, bruta y corta”⁵. Pero todo esto cambió con la revolución industrial. Para hacernos una idea, con una imagen bastará.

5 ‘... and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short’ (Leviathan, i. xiii. 9).

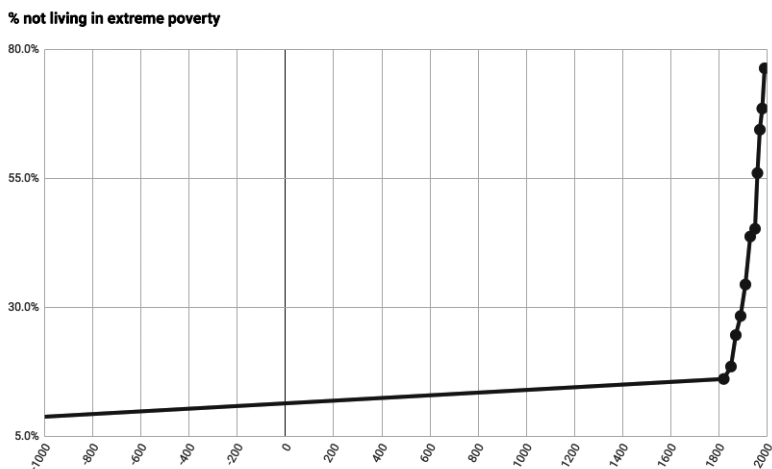


Figura 1. Imagen tomada Luke Muehlhauser Creative Commons Attribution-Noncommercial-ShareAlike 3.0.

La imagen de arriba muestra de manera clara y nítida cómo hasta la llegada de la revolución industrial el bienestar subjetivo era negativo y se mantuvo constante en la mayor parte de la población. Solo con la llegada de la revolución industrial el número de personas viviendo en pobreza extrema decreció.

De acuerdo con historiadores y demógrafos las condiciones de vida antes del año 1800 eran penosas. Solo el 43% de los recién nacidos llegaban al quinto cumpleaños y la esperanza de vida de quienes llegaban a la edad adulta rondaba los 30-35 años (Roser y Ritchie, 2019).

En términos económicos, renta per cápita, las cosas antes de la llegada de la revolución industrial no eran mejores. Otra imagen, bastará para mostrarlo.

Esta otra imagen de abajo muestra como la renta per capita mundial en dólares de 1990 era casi nula la mayor parte de la historia de la humanidad, es decir, las personas casi vivían en un absoluto estado de indigencia y/o pobreza, hasta que de nuevo el factor de la productividad a través de la revolución industrial hizo su aparición.

La revolución industrial representó un ciclo o periodo de gran innovación, difusión y penetración de la tecnología ayudado por factores como el capital humano y físico (Mankiw et al. 2992) e institucional (Acemoglu y Dell, 2010).

La revolución digital, en la que incluyo el desarrollo y avances de la IA y robótica, supone un salto exponencial en los usos y potenciales aplicaciones de la tecnología, pero también un incremento de los potenciales riesgos. Veámoslo, brevemente, en la siguiente sección 1.2.

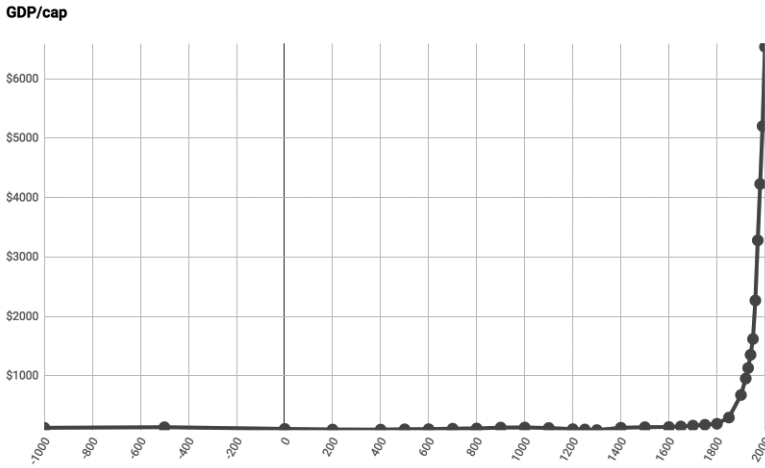


Figura 2. Imagen tomada de Luke Muehlhauser Creative Commons Attribution-Noncommercial-ShareAlike 3.0. con datos de De Long (1998).

1.2. REVOLUCIÓN DIGITAL Y USO-DUAL DE LA TECNOLOGÍA

Como vemos, las nuevas tecnologías desarrolladas durante la revolución industrial mejoraron los niveles de bienestar individual y social. La mecanización de las fábricas hizo despegar el nivel de producción de bienes y productos de consumo. Llegada la década de los 1920s la población mundial llegó a 1.000 millones de personas, 4.000 millones durante la década de los 70s y 7.000 millones a principios del siglo XXI. Los niveles de vida, medidos por renta per cápita, para estos 7.000 millones de personas en todo el mundo, de acuerdo con datos del Banco Mundial (véase, <https://data.worldbank.org/indicador/NY.GDP.PCAP.CD?locations=1W>), creció de estar por debajo de los 500 \$ en 1960 a 11.000 \$ en 2018.

Desde hace unas décadas el mundo ha mejorado, incluyendo una mejora en la esperanza de vida, declive de la pobreza extrema, reducción de la mortalidad infantil, mayor acceso a la educación, agua potable y atenciones sanitarias (Pinker, 2018).

Muchos autores consideran que el desarrollo de la IA y tecnologías digitales es otra gran transformación y/o revolución para la historia de la humanidad. De hecho, la IA tiene el potencial de incrementar el producto interior bruto mundial de ahora a 2030 entre 13 y 15 \$ trillones (Mckinsey Global Institute, 2018). La

IA es como dicen los economistas una “tecnología de propósito general” porque puede cambiar muchos aspectos de nuestras vidas, propiciar la innovación y hacernos más prósperos y productivos.

Pero como toda tecnología emergente con gran potencial, la IA tiene su lado oscuro. De acuerdo con el Instituto Nacional de Salud de los EE.UU. determinadas tecnologías emergentes en campos como las ciencias biomédicas, física, energía nuclear, nanomateriales, biología etc. generan una preocupación por su potencial “uso-dual” de la tecnología derivada (DURC Policy, 2012).

Esta agencia sanitaria de los EE.UU. define el “uso dual” de la tecnología de la siguiente manera:

“...investigación que, con base en el entendimiento actual, se puede anticipar razonablemente que proporcionará conocimiento, información, productos o tecnologías que podrían ser directamente aplicados incorrectamente para plantear una amenaza significativa con amplias consecuencias potenciales para la salud y la seguridad pública, los cultivos agrícolas y otras plantas, animales, el medio ambiente, los materiales o la seguridad nacional”.

Hasta hace poco la expresión “uso-dual” se refería primariamente a tecnologías que podían tener aplicaciones civiles y militares. Actualmente la expresión tiene múltiples sentidos. El desarrollo de la ciencia y la tecnología crea demandas hasta hace poco impensables por los propios científicos e ingenieros que las desarrollan. Considerar la amplia gama de potenciales usos indebidos de la tecnología junto a sus implicaciones de seguridad, el uso dual de la investigación, es altamente complejo, pero necesario.

Muchas tecnologías emergentes presentan importantes desafíos éticos, legales, sociales y políticos. Pero la IA, investigación genómica, nanotecnología, robótica, ciencias de la computación en general y neurociencia aplicada tienen la capacidad de ser aditivas y multiplicativas. Es decir, aplicadas conjuntamente la tecnología derivada de estas áreas de investigación pueden manipular los átomos de la materia física, la información digital, biológica (ADN) y las células nerviosas.

Si nos centramos en la IA, la IA tiene como objetivo crear máquinas inteligentes y como corolario entender mejor la inteligencia biológica. En su propósito de crear máquinas inteligentes pronto los investigadores en IA se dieron cuenta que la mejor estrategia era crear modelos de inteligencia que aprendan en lugar de tener que ser programados por un ingeniero o matemático humano desde el principio.

Así nació el “aprendizaje automático” (machine learning) y una técnica dentro de este subcampo llamada “aprendizaje profundo” (deep learning) en la que se desarrollan “redes neuronales” que imitan la arquitectura y propiedades de las neuronas del cerebro humano. En otras palabras, nodos computacionales

interconectados entre sí que imitan la capacidad del cerebro humano de percibir, procesar y representar información.

El aprendizaje automático mejora cada día gracias a las grandes cantidades de datos e información que generamos y podemos anotar/etiquetar, así como analizar (Big Data). Y si bien es cierto que gracias a estos desarrollos la IA puede ofrecer diagnósticos fiables en medicina sobre la base de los síntomas y recomendar determinados tratamientos y detectar anomalías en imágenes con mayor fiabilidad que radiólogos; o en el derecho procesar miles de sentencias en cuestión de milisegundos; o en la movilidad ofrecernos los vehículos sin conductor mejorando la seguridad y reduciendo la siniestralidad y los accidentes; la IA es también capaz de conducirnos a un futuro distópico.

Con IA se pueden crear *robots kamikazes*, vehículos sin conductor diseñados para colisionar intencionalmente, drones convertidos en misiles o con la capacidad de portar armas, videos (e imágenes) sintéticos generados por IA con la intención de propagar bulos o falsas noticias (fake news), armas autónomas letales, virus para inutilizar sistemas informáticos que controlan nuestras infraestructuras básicas, sistemas de reconocimiento facial que vigilen y persigan a grupos determinados... y la lista puede seguir.

Por estos “usos-duales” con consecuencias negativas, la IA tiene un lado oscuro que no debemos obviar. En la siguiente sección hablaré con más detalle de tres retos a los que nos enfrentamos si usamos la IA con propósitos espurios y la necesidad de superarlos para alcanzar una verdadera gobernanza global ética de la IA.

2. TRES GRANDES RETOS DE LA IA

Uno de los retos más importantes que plantea el desarrollo de la IA es el uso presente y futuro de las tecnologías computacionales para manipular el comportamiento del público y sus actitudes con propósitos propagandísticos y políticos, así como comerciales (marketing). Otro segundo reto importante son los sesgos algorítmicos. Cada vez más automatizamos nuestra toma de decisiones y delegamos nuestras decisiones a sistemas algorítmicos para tomar decisiones con consecuencias en múltiples campos económicos y sociales. Sin embargo, estos algoritmos pueden estar sesgados y discriminar a ciertos colectivos o grupos específicos por su condición, raza, sexo etc; y al mismo tiempo acceder a información personal sensible. Por último, el tercer reto de gran importancia es el futuro del trabajo. El desempleo tecnológico o el desplazamiento de las personas de sus lugares de trabajo por la implantación de máquinas, es quizá el reto más acuciante. Si la gente

no tiene posibilidad de trabajar porque las máquinas pueden hacer su trabajo se puede quebrar la paz social.

Veamos cada uno de estos retos con más detalle para tener una idea de la importancia que tiene superarlos para alcanzar una gobernanza global ética de la IA.

Automatización de las noticias y propaganda política computacional

Bajo este epígrafe quiero introducir el reto que se deriva de la aplicación presente y futura de las tecnologías computacionales para la manipulación de la conducta de la gente con propósitos políticos y comerciales (marketing).

La interacción entre automatización, algoritmos y política puede ser explosiva. Se pueden crear chat-bots (agentes de software autónomos) que interactúen en las redes sociales virtuales con los usuarios y manipulen la opinión pública sobre ciertos asuntos cruciales para la política. Por ejemplo, durante las elecciones presidenciales de los EE.UU. acabó siendo elegido Donald Trump como 45º presidente. Numerosos estudios achacan este hecho a la acción de operativos rusos (hackers) que compraron miles de anuncios en plataformas como Facebook para explotar las divisiones raciales, religiosas e ideológicas fomentando la polarización de la sociedad norteamericana⁶.

Se puede fabricar la polarización o el consenso con el uso de sofisticados algoritmos que crean anuncios micro-dirigidos a estratos sociales específicos generando “filtros burbuja” o cámaras de eco, es decir, entornos de información donde la gente solo recibe aquella información que confirma lo que ya sabe o, en el peor de los casos, confirma sus prejuicios (Pariser 2012). Los filtros burbuja o cámaras de eco son búsquedas en Internet que algoritmos sofisticados personalizan para ofrecernos resultados que nos gustaría ver. Si somos conservadores ideológicamente, únicamente tendremos resultados en los motores de búsqueda de noticias que confirmen nuestras inclinaciones ideológicas conservadoras y si somos progresistas ideológicamente, únicamente tendremos resultados que confirmen nuestras inclinaciones progresistas.

También con el uso estratégico de chat-bots se puede retroalimentar y reforzar ideas. Los bots programados para interactuar con ciertos usuarios pueden

6 Ciertamente, en teoría, es posible presentar anuncios personalizados a los usuarios de ciertas redes sociales como, por ejemplo, Facebook. Lo que los expertos ponen en duda es la relación causal entre estar expuesto a contenido personalizado y la modificación del comportamiento político, es decir, que dicho contenido tengo algún efecto a la hora de cambiar la preferencia de voto. Esto es algo que todavía no está demostrado (Baldwin-Philippi, 2017).

amplificar ideas cuyo contenido puede ser negativo, explotarlo y, en última instancia, manipular la conducta.

La consecuencia de todo esto es una sociedad polarizada, dividida, sin capacidad de puntos de encuentro, gente aislada intelectualmente por la acción de algoritmos con un propósito político. El escándalo de Cambridge Analytica (The Guardian 2018) es un caso de filtración de datos y erosión de la confianza y de la privacidad por parte de compañías tecnológicas como Facebook, pero también es un caso donde se pone de manifiesto la posibilidad de manipulación política a través de las redes sociales (Véase, nota al pie de página número 3).

La desinformación es otra consecuencia nefasta del uso de herramientas computacionales o IA con el propósito de desestabilizar la sociedad. Es cierto que las mentiras, las falsas noticias, han existido siempre. Pero con cada tecnología nueva el público potencial y la audiencia a la que podían llegar se multiplicaba. Las mentiras y el engaño evolucionaron gracias al lenguaje. Con la escritura, la información falsa multiplicó su campo de acción, pero es que con las TIC (tecnologías de la información y comunicación) se pueden difundir mentiras a todos los rincones del globo y en todos los soportes imaginables. Por otra parte, características de captura de atención (economía de la atención) e incluso mecanismos de refuerzo condicionado conducentes a hábitos pseudo-adictivos en el diseño de la tecnología (Williams 2018), permiten una propagación y viralización de contenidos mucho mayor (Aral, 2020).

Sesgos algorítmicos y discriminación

La IA se puede aplicar a los gobiernos y los asuntos administrativos (Ramió, 2019). La IA puede realizar tareas predictivas, asignación de colegios para niños y niñas, distribución de recursos sociales y sanitarios. En la empresa privada, la IA puede utilizarse para la selección de personal y contratación, facilitar créditos financieros, selección de perfiles para publicidad etc. Todas estas decisiones, sumamente importantes, se pueden automatizar con la aplicación de sistemas algorítmicos de toma de decisiones.

El problema es que muchas veces los algoritmos responsables de la toma de decisiones están sesgados o son opacos y/o poco transparentes, no protegen la información personal etc., es decir, no se sabe cómo han tomado la decisión que han tomado, ni porqué (Pasquale, 2015).

Por ejemplo, en un distrito de California, EE.UU., jueces y fiscales usan una herramienta algorítmica para decidir si una persona investigada por un delito debe salir de la cárcel antes de un juicio ante un tribunal. Una investigación del grupo de

periodistas ProPublica determinó que el algoritmo usado discriminaba negativamente a personas investigadas de raza negra. La herramienta algorítmica llamada COMPAS generaba una puntuación para estimar el grado de reincidencia en un periodo de tiempo de dos años tras salir de la cárcel.

La empresa que había creado el algoritmo llamada Northpointe (ahora se llama Equivant) sacó un comunicado de prensa defendiendo que los algoritmos no estaban sesgados. Pero el estudio de ProPublica fue lo suficientemente riguroso como para mostrar los sesgos y desequilibrios estadísticos existentes en el algoritmo que es usado para automatizar decisiones muy importantes como, por ejemplo, a quién se le priva de libertad (Véase, Monasterio Astobiza 2017, para un análisis en detalle sobre la ética de los algoritmos).

El problema, muchas veces, consiste en qué definición de equitativo o justo (fair) se está usando a la hora de diseñar y construir un algoritmo con aplicaciones sociales. Habitualmente, los científicos computacionales no colaboran con filósofos para obtener una definición de equidad o justicia de la ingente literatura en filosofía política con una historia de más de 2.500 años de teorización. Y aún así, se aventuran a ofrecer una caracterización matemática de la idea filosófica de justo o equitativo (fair) que puede implicar variadas y diferentes implementaciones y eso solo si tenemos en consideración la literatura de investigación en ciencias de la computación que además diferencia entre sesgos sociales de la aplicación del aprendizaje máquina (machine learning) y sesgos estadísticos.

Estos últimos, los sesgos estadísticos, son simplemente la diferencia entre el valor esperado de una estimación y el verdadero valor. Por otro lado, los sesgos sociales, que reproducen los prejuicios de un sistema social injusto y que pueden estar presentes en los algoritmos usados en modelos de aprendizaje automático para automatizar decisiones, son más complejos y con graves repercusiones que pueden amenazar y poner en peligro la democracia (O'Neil, 2016).

Algunas definiciones de equidad en algoritmos para evitar que expresen sesgos sociales son:

- *Equidad de grupo*: trata grupos diferentes de manera igual (Dwork et al. 2012).
- *Equidad individual*: individuos similares deben ser considerados similarmente (Pleiss et al. 2017).
- *Equidad de proceso*: equilibrio entre precisión, equidad de proceso y equidad de resultados (Grig-Hlaca et al. 2016).
- *Probabilidades igualadas*: (equalized odds): la probabilidad de una persona perteneciente a una clase que se le asigna un atributo positivo y la probabilidad de una persona perteneciente a una clase que se le asigna

un atributo negativo debe ser igual que para miembros protegidos y no-protegidos del grupo general (Hardt et al. 2016).

- *Igualdad de oportunidad*: la igualdad de oportunidades para grupos protegidos o no-protegidos debe tener ratios positivos iguales (Hardt et al. 2016).
- *Equidad a través de la conciencia*: (fairness through awareness): dos individuos cualquiera que sean similares con respecto a una métrica deben recibir el mismo tratamiento (Dwork et al. 2012).

Una cuestión básica a tener en cuenta sobre el impacto social de los sesgos algorítmicos es que los algoritmos, procedimientos y/o protocolos computacionales, no son neutros ni objetivos. Los algoritmos se enmarcan en un contexto tecnológico, económico, ético, temporal y espacial concreto. Y, además, están imbuidos de los valores conscientes e inconscientes de las personas que los diseñan. Responden a los intereses del diseñador que de manera, muchas veces, inconsciente, en forma de prejuicios, sesgos, estereotipos; crea algoritmos que retroalimentan el sistema de prejuicios social injusto existente.

Los sesgos sociales no son los únicos efectos perversos de los algoritmos que afectan a la vida de las personas creando discriminación etc. Los algoritmos también explotan vulnerabilidades y pueden extraer información sensible de las personas. Por consiguiente, la privacidad se ha convertido en uno de los valores centrales en nuestro tiempo. En un primer momento, uno podría pensar que los sesgos algorítmicos y los temas de privacidad son fenómenos autónomos e independientes, pero no es así. Un algoritmo puede tomar decisiones automatizadas que causen un perjuicio o discriminen a un individuo que pertenece a un colectivo vulnerable, porque accede a una variable identificativa de ese individuo que no está suficientemente protegida como información sensible y personal. Por eso en esta sección hablo sobre discriminación, sesgos y privacidad de manera continua e indistintamente.

La privacidad es difícil de definir. Quizá la definición de privacidad más famosa es la de Warren y Brandeis (1890, p. 20): “el derecho a estar solo”. A partir de una revisión de las distintas teorías filosóficas y legales de privacidad, Tavani (2007, p. 2) agrupa en cuatro categorías las teorías de privacidad:

- a) No-intrusión: privacidad definida como derecho a ser libre de toda intrusión.
- b) Aislamiento: privacidad entendida como estar físicamente inaccesible a otros.
- c) Limitación: privacidad entendida como acceso a nuestra información restringida en ciertos contextos.

d) Control: privacidad entendida como el control sobre nuestra información.

A partir de estas definiciones de privacidad obtenidas de la filosofía y el derecho y tras décadas de promover distintos enfoques y metodologías se ha llegado a una definición matemática significativa de privacidad, llamada: privacidad diferencial (differential privacy). La privacidad diferencial cuantifica la semántica de la privacidad introduciendo una serie de constreñimientos matemáticos a un algoritmo (Dwork y Roth 2014). La privacidad diferencial ofrece una definición matemática rigurosa de privacidad. Para entender la idea de la privacidad diferencial imagina un algoritmo que analiza una base de datos y realiza operaciones computacionales sobre ella (por ejemplo, la media, mediana, varianza de los datos, etc.) Dicho algoritmo se dice que es diferencialmente privado si al mirar el resultado no se puede saber si los datos de algún individuo se incluyeron en el conjunto de datos original o no.

La privacidad diferencial como técnica protege a los individuos de una base de datos de su posible identificación introduciendo “ruido” en los datos. Esto produce una distribución probabilística que hace prácticamente imposible determinar el registro de alguien en una base de datos. Dicho de otra manera, alguien que quiera identificar a una persona a partir de la información contenida en la base de datos no podrá hacerlo, porque no se puede vincular la información contenida en la base de datos con ningún individuo.

Para conseguir que no haya sesgos sociales, ni discriminación algorítmica, porque previamente se ha accedido a datos sensibles y personales sin consentimiento, es decir, se ha violado la privacidad de alguien; se debe diseñar algoritmos que sean socialmente sensibles. No solo se deben establecer leyes, regulaciones y comités de supervisión, también se debe crear y diseñar tecnología ética (ethics by design). De esta forma se construirán algoritmos con mejores definiciones de equidad, precisión, transparencia, privacidad y con una ética integrada en su diseño (Kearns y Roth 2019). La “ética por diseño” es un enfoque que busca la inclusión sistemática de valores y principios éticos así como otros requerimientos y procedimientos en el diseño y procesos de desarrollo de la tecnología y sistemas de IA. La “ética por diseño” es un ejemplo de Enfoque de Diseño Sensible al Valor (Friedman, Kahn y Borning, 2002), que se está convirtiendo en el enfoque globalmente recomendado para el desarrollo de la IA. En conclusión, la “ética por diseño” es un enfoque que se asegura que los sistemas de IA y robótica cumplan ciertos valores éticos fundamentales para que su implementación respete los derechos y dignidad de las personas.

El futuro del trabajo: desempleo tecnológico

La IA va a transformar el futuro del trabajo. De hecho ya lo está haciendo. La IA tiene y va a tener un impacto en la productividad, en los salarios y en el empleo. En general, la IA y la automatización del trabajo seguirá el famoso adagio del filósofo Sombart, popularizado por el economista Shumpeter: destrucción creativa. Es decir, la IA destruirá empleos, pero, al mismo tiempo, creará nuevos empleos.

No obstante, a pesar de la ingente cantidad de inversiones, talento, productos y servicios (como por ejemplo, capacidad de visión computacional a un nivel de rendimiento mayor que el humano o reconocimiento de voz también a un nivel de rendimiento mayor que el humano gracias a técnicas como las redes neuronales artificiales) el nivel de productividad se ha reducido (Brynjolfsson, Rock y Syverson 2017). Esto se conoce como *la paradoja de la productividad*.

Brynjolfsson, Rock y Syverson (2017) dan cuatro razones para explicar la paradoja de la productividad. La primera es quizá que pecamos de una ilusión de optimismo. Puede ser que seamos muy optimistas y la tecnología de la IA no está a la altura de las expectativas. Otra razón es que puede que no estemos midiendo bien los beneficios de la tecnología en términos de productividad. Una tercera razón es que puede que la tecnología solo esté beneficiando a unas pocas organizaciones o industrias y todavía no ha llegado al público en general. Finalmente, la razón que defienden los autores es que la mejora de la tecnología en productividad es real, pero como las organizaciones tardan en reestructurarse los beneficios de la tecnología tardan en hacerse visibles.

Brynjolfsson –profesor en el Instituto de Stanford para la IA centrada en el ser humano (HAI) y director del Laboratorio de Economía Digital de Stanford– y otros economistas no dudan en describir a la tecnología de la IA como una tecnología de propósito general como lo es la electricidad. Porque la IA es una tecnología de propósito general que necesita innovación complementaria e inversión para poder hacer realidad sus promesas. Pero para ello quizá tenemos que reinventar nuestras organizaciones, instituciones y métricas.

A pesar de que la tecnología desplazará y dejará sin trabajo a muchas personas –y no tienen por qué ser solo personas de sectores que realicen tareas mecánicas, rutinarias y con bajo cualificación, también los trabajadores de profesiones liberales como abogados, ciertas especialidades médicas, etc.; pueden ser desplazados por máquinas (Frey y Osborne 2013)– la mayoría de los economistas creen que la IA y la robótica no dejará sin empleo a grandes cantidades de personas y que el desempleo tecnológico es un miedo infundado.

Sin lugar a dudas, la tecnología de la IA, la robótica y la tecnología digital transformarán nuestros trabajos y vidas. Pero estamos todavía muy lejos de un mundo sin trabajo (Danaher, 2019). Para bien o para mal, todavía no estamos en un mundo donde reine el ocio y el tiempo libre, ni tampoco en un mundo en caos y sin paz social porque la gente no tenga, ni encuentre, trabajo por el desempleo tecnológico.

En la siguiente sección 3, hablaré sobre el tema central de este artículo y que da título al mismo: La gobernanza global de la IA. Como hemos visto los sistemas de IA son el resultado de esa querencia humana por el uso del conocimiento para la creación de artefactos y herramientas que permiten resolver problemas y extraer un mayor rendimiento del entorno para mejorar el bienestar. Pero con la llegada de los avances y desarrollos de la IA y la revolución digital los potenciales riesgos también han aumentado. He comentado tres retos de escala global que necesitan una cooperación internacional para una estabilidad geopolítica. Veamos ahora la necesidad de una respuesta colectiva para mitigar los efectos no deseados de la transformación digital del mundo.

3. GOBERNANZA GLOBAL DE LA IA

Existe un creciente debate sobre cómo regular y hacer un uso responsable de las tecnologías digitales, en particular de la IA. Los países desarrollados, con una larga tradición de instituciones y mecanismos de contrapeso al poder, son capaces de implementar procesos de deliberación pública entre diferentes actores y/o partes interesadas (por ejemplo, la sociedad civil, grupos de afectados, organizaciones no gubernamentales etc.) y, por lo tanto, empezar a construir una arquitectura de gobernanza ética para la IA (y tecnología digital conexas) supuestamente basada en el imperio de la ley. Sin embargo, en el llamado sur global (por ejemplo, países con una historia poscolonial, también llamados países no desarrollados) y, principalmente, en países con regímenes autoritarios donde su situación de vulnerabilidad y dependencia económica del liderazgo del norte global les lleva a importar tecnología digital, capital y modos de organización a pesar de sus diferencias políticas, legales y culturales; es realmente muy difícil una regulación de la IA (Arun, 2020).

Esto puede tener un impacto muy negativo en sus poblaciones ya de por sí excluidas, oprimidas y discriminadas. La cuestión central es hasta qué punto los países no desarrollados con culturas políticas e institucionales diferentes que importan tecnología digital de los países del norte global pueden verse afectados si no se tiene en cuenta la necesidad de una gobernanza global ética de la IA y a varios niveles desde una perspectiva de derechos humanos/democrática. En el

debate sobre la gobernanza ética, hay una posición que sugiere una moratoria sobre cualquier intento de regular la IA. Según los defensores de esta postura, la regulación mata la innovación (Ahmed, 2015). Creo que los argumentos de esta posición son erróneos, porque la IA ya tiene un enorme impacto en la vida de millones de personas en todo el mundo (por ejemplo, en la economía de plataformas ya controlada por algoritmos, y/o la IA que se utiliza para contratar a personas, conceder préstamos o decidir la libertad condicional de delincuentes).

Por ello, una distinción útil es: “gobernanza ética blanda” y “gobernanza ética dura (basada en la ley)”. Esta distinción puede ser aceptada por los escépticos de la regulación de la IA. Por ejemplo, la “gobernanza ética blanda” describe normas como los marcos ISO o IEEE que pueden utilizarse en las primeras fases del desarrollo de la tecnología digital. Por otro lado, la “gobernanza ética dura (basada en la ley)” se refiere directamente a las prohibiciones e impide el uso de una tecnología cuando el riesgo supera los beneficios⁷. Quizá la mejor recomendación para una gobernanza global de la IA es partir de la experiencia con otras tecnologías, como la energía atómica, con la intención de crear una gobernanza global ética de la IA desde un enfoque de derechos humanos/democrático. También es útil plantear el caso de la gobernanza global de la IA como un riesgo existencial para la humanidad similar a como resulta ser el cambio climático.

El problema del cambio climático desborda las fronteras de un único país y se convierte en un problema de coordinación colectiva entre países. Si se quiere atajar los riesgos que supone un incremento de las temperaturas por la acción antrópica de emisión de gases de efecto invernadero a la atmósfera, se necesita una respuesta global coordinada. De la misma manera, las tecnologías disruptivas, como la biotecnología, pero especialmente la IA; también requiere de una

7 A fecha de escritura de este artículo la UE ha publicado el documento: Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre inteligencia artificial-Ley sobre Inteligencia Artificial. La propuesta indica que la UE sostendrá una postura firme respecto a determinadas aplicaciones de la IA, diferenciándose de Estados Unidos y China. Este nuevo reglamento de ser adoptado –que implica que será de aplicación en toda la UE– quiere que los europeos puedan confiar en lo que la IA puede ofrecer. El proyecto de reglamento incluye la prohibición de la IA para la “vigilancia indiscriminada”, incluidos los sistemas que rastrean directamente a las personas en entornos físicos o agregan datos de otras fuentes. Prohibición de los sistemas de IA que crean puntuaciones de crédito social, es decir, que juzgan la fiabilidad de una persona en función de su comportamiento social o de los rasgos de personalidad previstos. Autorización especial para el uso de “sistemas de identificación biométrica remota”, como el reconocimiento facial, en espacios públicos. Notificaciones necesarias cuando las personas interactúan con un sistema de IA, a menos que sea “obvio por las circunstancias y el contexto de uso”. Nueva supervisión de los sistemas de IA de “alto riesgo”. Evaluación de los sistemas de alto riesgo antes de su puesta en servicio, lo que incluye garantizar que estos sistemas sean explicables para los supervisores humanos y que se entrenen con conjuntos de datos de “alta calidad” sometidos a pruebas de sesgo y la creación de un “Consejo Europeo de Inteligencia Artificial”.

gobernanza global porque los riesgos son compartidos y no exclusivos de un solo país.

Mucho se habla de la “carrera armamentística” por el liderazgo en la revolución digital (Lee 2018) entre dos potencias hegemónicas: USA y China. Ninguna de estas dos potencias quiere quedarse atrás y las razones son claras. El primer país que domine la tecnología de la IA obtendrá una ventaja competitiva muy grande que será reforzada porque el país ganador obtendrá mayores recursos que a su vez mejorará sus sistemas. Esto refleja un escenario fuertemente competitivo donde la seguridad y el control de la tecnología puede pasar a un segundo plano, porque lo que verdaderamente importa, la prioridad, es ganar a los competidores.

Este duopolio (USA-China) en la carrera armamentística por liderar la tecnología de la IA es quizá una simplificación. En primer lugar, porque bajo el paraguas abstracto de IA hay distintas aplicaciones, técnicas etc. que no todas se solapan (el “aprendizaje profundo” no es lo mismo que un “sistema experto” y no se aplican a los mismos campos, p. ej.), pero, principalmente, porque no tiene en cuenta la interdependencia de los complejos militares-industriales de los países y ni siquiera el hecho de que se comparten internacionalmente infraestructuras, patentes y talento.

Sin embargo, aun siendo una simplificación esta caracterización de duopolio entre USA y China en el desarrollo de IA, hay que enfatizar que la motivación subyacente de la competición nacional por el dominio en la revolución digital tiende a centrarse en adquirir nuevas capacidades y mucho menos en la gobernanza, control y regulación del impacto social y ético de las tecnologías emergentes de la IA. Dicha competición nacional entre dos superpotencias y el resto de países por el liderazgo en la tecnología de IA, no solo es una simplificación, sino que es un sinsentido desde un punto de vista del interés general y el bien común. La tecnología de la IA no es un juego suma-cero porque nada se pierde por compartir nuevas capacidades, algoritmos, infraestructuras, talento etc. De hecho, los investigadores en IA de universidades forman parte de una comunidad internacional que comparten intereses, conferencias, acuerdos, patentes y sociedades profesionales como la AAI (siglas en inglés de la Asociación para el Avance de la Inteligencia Artificial) y IEEE (siglas en inglés del Instituto de Ingenieros Eléctricos y Electrónicos).

El riesgo está en los desarrollos de IA en corporaciones y empresas grandes y/o pequeñas que en lugar de regirse por un *etos* colaborativo y cooperativo, pueden guiarse por la competencia y la maximización del beneficio. Pero aun así, en las corporaciones tecnológicas suelen existir comités éticos que supervisan la I+D+I (investigación, desarrollo e innovación). También existen consorcios de industria, como el Partnership on AI (<https://www.partnershiponai.org/>) y organizaciones supranacionales como la ONU que tienen el poder de convocar eventos

y pueden redactar documentos no vinculantes, pero con fuerza simbólica y moral que denotan acuerdo y consenso en principios básicos. De hecho, de ser posible una potencial gobernanza global de la IA está debe inspirarse en la experiencia de éxito de instancias supranacionales como la ONU. La cooperación internacional para la regulación de la IA es fundamental.

Del mismo modo, numerosos países están dotándose de grupos de expertos y consejeros que ayudan en el proceso de regulación. Un ejemplo notorio en este sentido es el Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial de la Unión Europea (Véase para más información este vínculo: <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>).

Este Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial de la Unión Europea ha elaborado una guía de principios y propone siete requerimientos clave que todo sistema de IA debe cumplir para ser considerado confiable (AI HLEG 2019):

- *Agencia y supervisión humana*: Los sistemas de IA deben empoderar a los seres humanos, permitiéndoles tomar decisiones informadas y fomentando sus derechos fundamentales. Al mismo tiempo, es necesario garantizar unos mecanismos de supervisión adecuados, lo que puede lograrse mediante enfoques basados en el principio de “humano en el bucle”, “humano en el bucle” y “humano en el mando”.
- *Robustez y seguridad técnica*: Los sistemas de IA deben ser resistentes y seguros. Necesitan ser seguros, asegurando un plan alternativo en caso de que algo salga mal, además de ser precisos, fiables y reproducibles. Esa es la única manera de asegurar que también se pueda minimizar y prevenir el daño no intencional.
- *Privacidad y gobernanza de los datos*: además de garantizar el pleno respeto de la privacidad y la protección de los datos, también deben garantizarse mecanismos adecuados de gobernanza de los datos, teniendo en cuenta la calidad y la integridad de los datos, y garantizando un acceso legítimo a los mismos.
- *Transparencia*: los datos, el sistema y los modelos de negocio de la IA deben ser transparentes. Los mecanismos de trazabilidad pueden ayudar a conseguirlo. Además, los sistemas de IA y sus decisiones deben explicarse de manera adaptada a las partes interesadas. Los seres humanos deben ser conscientes de que están interactuando con un sistema de IA y deben estar informados de las capacidades y limitaciones del sistema.
- *Diversidad, no discriminación y equidad*: Debe evitarse el sesgo injusto, ya que puede tener múltiples consecuencias negativas, desde la marginación de los grupos vulnerables hasta la exacerbación de los prejuicios y

la discriminación. Al fomentar la diversidad, los sistemas de IA deben ser accesibles para todos, independientemente de cualquier discapacidad, e involucrar a los actores relevantes a lo largo de todo su ciclo de vida

- *Bienestar social y ambiental*: Los sistemas de IA deben beneficiar a todos los seres humanos, incluidas las generaciones futuras. Por lo tanto, hay que garantizar que sean sostenibles y respetuosos con el medio ambiente. Además, deben tener en cuenta el medio ambiente, incluidos otros seres vivos, y su impacto social debe ser cuidadosamente considerado.
- *Rendición de cuentas*: Se deben establecer mecanismos para garantizar la responsabilidad y la rendición de cuentas de los sistemas de inteligencia artificial y sus resultados. La auditabilidad, que permite la evaluación de algoritmos, datos y procesos de diseño; juega un papel clave en ello, especialmente en aplicaciones críticas. Además, debe garantizarse una reparación adecuada y accesible.

Gobiernos, empresas, organizaciones civiles y activistas de todo el mundo han desarrollado principios de gobernanza ética de la IA. Entre los principios de gobernanza global ética de la IA que preponderan entre los distintos actores destacan la justicia, seguridad, transparencia, protección de la privacidad, tecnología alineada con valores humanos, tecnología humano-céntrica, transparencia, explicabilidad, y finalmente, responsabilidad (Jobin, Ienca y Vayena 2019).

Miremos de nuevo al problema del cambio climático como ejemplo de problema de coordinación colectiva del que extraer lecciones para la gobernanza global de la IA. El problema del cambio climático se asemeja a la tragedia de los comunes (Hardin 1968). Imagina un pasto abierto. Es evidente que cada pastor intentará que sus animales se alimenten en el pasto. Como seres racionales cada pastor piensa que sí mete más animales mayor será el beneficio, pero está lógica o manera de pensar, al llevarla todos acabo, al final crea un drama o tragedia visible. El pasto es un recurso natural, finito, y si cada pastor va metiendo más animales al pasto, el recurso común, el pasto, se acaba.

Por su parte, el problema del desarrollo tecnológico en IA es un problema similar que requiere una estrategia global para evitar escenarios suma cero, es decir uno gana, pero el resto pierde; y crear escenarios de ganancia para todos (e.g. win-win). Por ejemplo, si un país desarrolla una tecnología de IA que resulta ser beneficiosa para la economía y prosperidad de los ciudadanos de un país, el resto de países la copiará porque no se querrán quedar atrás. Sin embargo, esto conduce a una carrera armamentística de todos contra todos que puede llevar a una guerra comercial y de patentes que finalmente puede provocar una reducción en el estado de bienestar de los ciudadanos de todos los países. En la medida en que ningún país se querrá quedar atrás en el desarrollo tecnológico, muchos

pueden saltarse las precauciones de desarrollar IA segura que no produzca riesgos existenciales (Bostrom 2014).

En definitiva, el cambio climático, así como el desarrollo tecnológico y potenciales riesgos de la IA, es un problema de coordinación y gobernanza global porque soluciones en el ámbito de un solo país no pueden luchar contra el cambio climático, ni tampoco contra los riesgos y retos del desarrollo tecnológico en IA, respectivamente.

Uno de los objetivos para una plena gobernanza global de la IA y tecnologías adyacentes es crear un comité ético que incluya a gobiernos, industria y sociedad civil similar a como se coordinan los gobiernos y estados-nación en torno al cambio climático con el Panel Intergubernamental sobre el Cambio Climático. De hecho, una iniciativa similar se está llevando a cabo por funcionarios de alto nivel de Canadá y Francia para establecer el Panel Internacional sobre IA (IPAI, siglas en inglés).

Esfuerzos académicos también están sumándose para crear una gobernanza global ética de la IA. Por ejemplo, Wendell Wallach, investigador de la Universidad de Yale en el Centro Interdisciplinar de Bioética y Gary Marchant, director del Centro del Derecho e Innovación de la Universidad de Arizona, han propuesto lo que llaman Comités Coordinadores de Gobernabilidad (CCG).

Los CCG coordinarían las actividades de las distintas partes interesadas, supervisarían la evolución de la situación, tomarían nota de las mejores prácticas y determinarían qué instituciones están asumiendo la responsabilidad de las diversas preocupaciones, así como las lagunas existentes.

Las aplicaciones de técnicas computacionales provenientes del aprendizaje máquina o IA en general tienen un impacto en la sociedad (definen cómo interactuamos a través de plataformas como Facebook, Twitter, Whatsapp...), economía (Airbnb, Uber...), educación (MOOCs...) y hasta determinan nuestro entretenimiento (You Tube, Spotify...). Este poder transformador que encierra la IA puede generar grandes beneficios y mejorar nuestras vidas, pero al mismo tiempo tiene el potencial de ser discriminador, reforzar los prejuicios sociales y polarizar a la sociedad.

Las ramificaciones de la toma de decisiones por algoritmos o máquinas se dejan ver en el derecho, en la medicina, ejército y fuerzas armadas, ciencia... Para que realmente tengan un efecto transformador positivo deben integrar una supervisión humana basada en principios de ética práctica y lo que es más importante, el derecho internacional y las regulaciones estatales deben a través de distintos mecanismos, como códigos de buenas prácticas (soft law), poner en marcha una plétora de prácticas de monitorización internacional del impacto de la IA.

También existen las regulaciones y las legislaciones (*hard law*) que instituyen obligaciones y deberes, además de prohibir ciertas conductas. A través de la ley y agencias regulatorias se puede establecer lo que se puede hacer y lo que no se puede hacer. Sin embargo, entre los códigos de buenas prácticas (*soft law*) y las leyes regulatorias (*hard law*), muchos autores prefieren los primeros porque no coartan el avance y progreso de la innovación e investigación, en este caso de la tecnología en IA. De ahí la distinción presentada más arriba entre “gobernanza ética blanda” y “gobernanza ética dura” Pero siempre se ha de seguir un enfoque ético o marco normativo que minimice los riesgos.

Sin embargo, al ser gobernanza de la IA basada en *soft law* y no gobierno o burocracia basado en *hard law* es posible que agentes estratégicos en el desarrollo de la IA tengan incentivos para no cooperar y coordinarse. Por otra parte, muchos investigadores consideran que dejar que la propia industria se autogobierne es un error. Porque esto solo llevaría a un blanqueamiento ético (*ethics washing*) de ciertas compañías tecnológicas para proyectar una responsabilidad social corporativa falsa. Por consiguiente, el trabajo para implementar una verdadera gobernanza global de la IA es crear un marco normativo efectivo que contemple los riesgos a largo plazo y desarrolle aplicaciones concretas para su implementación por todos los grupos de interés.

La gobernanza global de la IA pretende establecer una seguridad internacional en materia de desarrollo tecnológico basada en el multilateralismo y no caer presa de los distintos modelos de feudalismo digital. Tampoco simpatiza con el modelo del capitalismo de la vigilancia abanderado por los EE.UU., ni el imperialismo digital defendido por China. La gobernanza global ética de la IA es quizá una tercera vía mucho más cercana a los valores humanistas de Europa donde los avances e innovación de la IA no deben comprometer los derechos y privacidad de los ciudadanos. La neutralidad tecnológica es el valor central donde el regulador establece los principios y el mercado los aplica teniendo en cuenta los estándares del producto o servicio.

La confianza en la IA ha de ser un pilar central sobre el que sustentar la gobernanza global ética de la IA. Así se podrá garantizar la co-existencia en armonía entre los sistemas de IA y las personas, sin miedo a la explotación u opresión. Una IA centrada en el ser humano donde la libertad de elección y decisión no se vea comprometida y las máquinas, sistemas de IA etc., no supongan un riesgo o una amenaza existencial para la humanidad. Por supuesto, las buenas intenciones deben estar acompañadas de acciones. Y en los últimos tiempos estamos viendo nacer un panorama de guías de buenas prácticas, reglamentos, planes, e incluso legislación y regulación que permiten hablar de un ecosistema de derechos y garantías en el uso de la tecnología de la IA. En concreto, Europa ha formado el Grupo de expertos de Alto Nivel sobre IA –un consejo independiente formado

por expertos–, un plan de Coordinación sobre IA, donde los estados miembro de la UE han acordado coordinarse y colaborar en políticas públicas, el Libro Blanco sobre IA y, más recientemente –como comentaba en la nota a pie de página número 7–, un reglamento.

Como he manifestado más arriba la cooperación internacional es fundamental si queremos que la humanidad afronte el problema común de la gobernanza de la IA que no puede ser abordado por ningún país por sí solo. Como prioridad para la gobernanza de la IA no solo se encuentran los ámbitos que he mencionado como retos más arriba, también se encuentra el uso militar de la IA que puede crear riesgos para la seguridad internacional. Los Sistemas de Armas Autónomas Letales –armamento dirigido por IA sin necesidad de supervisión y control humano– deben prohibirse para evitar la escalada de conflicto en el futuro. La seguridad de los datos a nivel internacional también debe ser otra prioridad para la gobernanza global de la IA. De la minería de datos y recolección, a la clasificación y asignación de etiquetas, todo el proceso de curación de datos debe ser regulado para prevenir la formación de modelos equivocados que hagan que los agentes decisorios (e.g. grupos de interés relevantes desde consumidores y particulares, hasta organizaciones, empresas y gobierno) formen juicios erróneos.

Ya sabemos que el desarrollo acelerado de la IA y sus aplicaciones demandan urgentemente ética y gobernanza, seguridad, privacidad, transparencia, explicabilidad y equidad para el uso e implementación de sistemas de IA. Pero para ello es clave un desarrollo estandarizado a nivel global propiciado por un nuevo diseño institucional, nuevas leyes, y reglas de gobernanza. Todo marco de gobernanza global ética de la IA no puede ser efectivo si no tiene en cuenta la voz de los usuarios. Por ello es necesario tener en cuenta las experiencias, preocupaciones, intereses y necesidades de los destinatarios finales de la tecnología.

No me gustaría terminar esta sección dedicada al tema central del artículo sin mencionar la gobernanza de una Inteligencia General Artificial (IGA) sin control. Cuando se habla de gobernanza global ética de la IA se sobreentiende cualquier iniciativa para hacer de la IA beneficiosa para la humanidad a través de los gobiernos mundiales, organizaciones internacionales, multinacionales y corporaciones etc., que colaboran conjuntamente para establecer un marco regulatorio y estándares industriales. Esta gobernanza que se sobreentiende no es directa y trata de influir en los productos y servicios que los creadores y desarrolladores tienen permitido hacer, sería inútil si los creadores y desarrolladores fueran capaces de dar lugar a una IGA.

Con una IGA como agente independiente todo intento de regulación por parte de los gobiernos sería en vano. La gobernanza global ética de la IA pasaría a convertirse en predictibilidad, explicabilidad y controlabilidad de una IGA. Este es un escenario poco probable, pero no por ello imposible. Una IGA altamente

capaz, creativa e incontrolable podría ser ingobernable, por ello es perentorio tener cuanto antes el consenso internacional y el marco de regulación y gobernanza global ética de la IA en el estado actual de desarrollo de la tecnología.

4. CONCLUSIONES

La gobernanza global de la IA tiene que ser una de las prioridades de la agenda internacional muy similar a como lo es la acción climática. Los agentes estratégicos (e.g. gobierno, industria, academia y sociedad civil) de la comunidad internacional necesitan trabajar juntos para conseguir obtener lo mejor que puede ofrecer la tecnología de la IA y evitar consecuencias negativas como las que he intentado subrayar en este artículo. En este artículo, he comentado tres grandes retos que considero se han de atajar para establecer una gobernanza global de la IA que beneficie a la sociedad y las personas. Una gobernanza global de la IA puede hacer que los sistemas artificiales y las personas colaboren de manera más fructífera. Una gobernanza global ética de la IA es una exhortación clara y directa para la cooperación internacional ante el desarrollo de tecnologías emergentes y potencialmente disruptivas que permita evitar consecuencias desastrosas que socaven gravemente las perspectivas de las generaciones actuales y futuras.

REFERENCIAS BIBLIOGRÁFICAS

- ACEMOGLU, D. y DELL, M., “Productivity differences between and within countries”. *American Economic Journal: Macroeconomics*, 2, 2010, 169–88
- AHMED, K. “Google’s Demis Hassabis –misuse of artificial intelligence ‘could do harm’,” BBC News. <http://www.bbc.com/news/business-34266425> 2015
- AI HLEG, 2019. *ETHICS GUIDELINES FOR TRUSTWORTHY AI*. [ebook] Brussels: European Commission. Disponible en: <<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>> [Accedido 28 Julio 2021].
- ARAL, S., *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health—And How We Must Adapt*. New York. Currency. 2020
- ARUN, C., “AI and the Global South: Designing for Other Worlds” En: DUBBER Marcus, PASQUALE Frank y SUNIT Das (ed.) *The Oxford Handbook of Ethics of AI*. Oxford. Oxford University Press. 2020, 589-607
- BALDWIN-PHILIPPI, J., The Myths of Data-Driven Campaigning. *Political Communication* 34, 2017, 627-633
- BOSTROM, N., *Superintelligence: Paths, Dangers, Strategies*. Oxford. Oxford University Press. 2014.

- BRYNJOLFSSON, E.; ROCK, D. y SYVERSON, C. (2017), “Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics”. NBER Working Paper 24001 [<https://www.nber.org/papers/w24001>]
- DAFOE, A., “AI Governance: Opportunity and Theory of Impact”. Effective Altruism Forum. 2020. Recuperado el 29/07/2021 <https://forum.effectivealtruism.org/posts/42reWndoTEhFqu6T8/ai-governance-opportunity-and-theory-of-impact>
- DANAHER, J., *Automation and Utopia: Human Flourishing in a World without Work*. Mass. Cam. Harvard University Press. 2019.
- DE LONG, J., “Estimating World GDP, One Million B.C. – Present”., 1998. Recuperado el 06/04/2021 https://delong.typepad.com/print/20061012_LRWGDP.pdf
- DE SOUSA SANTOS, B., “Epistemologies of the south and the future”. From the European South, 1, 2016, 17-29.
- DURC Policy. (2012), “United States Government Policy for Oversight of Life Sciences Dual Use Research of Concern”. Recuperado Abril 2, 2021 de <http://www.phe.gov/s3/dualuse/Documents/us-policy-durc-032812.pdf>
- DWORK C. et al., “Fairness through awareness”. *Proceedings of the 3rd Innovations in Theoretical Computer Science Conferences ACM*, 2012, 214-226
- DWORK, C. y ROTH, A., *The Algorithmic Foundations of Differential Privacy*. Delft. Now Publishers. 2014.
- FREY, C. y OOSBORNE, M., “The future of employment: How susceptible are jobs to computerization?” University of Oxford, 2013. [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf]
- FRIEDMAN, B. et al. “Value Sensitive Design: Theory and Methods”, University of Washington Technical Report, No. 2–12, 2002.
- GRICK-HLACA, N. et al., “The case for process fairness in learning: Feature selection for a fair decision making”. *NIPS Symposium on Machine Learning and the Law*. Vol. 1, 2016, 2
- HARDIN, G., “The tragedy of the commons”. *Science*, 162, 1968, 1243-1248
- HARDT, M. et al., “Equality of opportunity in supervised learning”. *Advances in Neural Information Processing Systems*. 2016, 3315-3323
- HERRERA, R. y GARCIA-BERTRANDT, R., “The Agricultural Revolutions”. En HERRERA, Rene y GARCIA-BERTRAND, Ralph (eds.), *Ancestral DNA, Human Origins, and Migrations*. Oxford. Academic Press, 2018, 475-509.
- JOBIN, A.; IENCA, M. y VAYENA, E., “The global landscape of AI ethics guidelines”. *Nature Machine Intelligence* 1, 2019, 2522-5839.
- KEARNS, M. y ROTH, A., *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford. Oxford University Press. 2019.
- LEE, K., *AI Superpowers: China, Silicon Valley, and the New World Order*. Boston. Houghton Mifflin Harcourt. 2018.
- OAKLEY, KP., *Man the Tool-Maker*. Chicago, IL: University of Chicago Press. 1959.
- O’NEIL, C., *Weapons of Math Destruction*. New York. Random House. 2016.
- MANKIW, N. et al., “A contribution of the empirics of economic growth”. *The Quarterly Journal of Economics*, 107, 1992, 407-437.

- MCAFEE A. y BRYNJOLFSSON E., *Machine, Platform, Crowd: Harnessing our Digital Future*. New York. WW. Norton & Company. 2017.
- MCKINSEY GLOBAL INSTITUTE, “Notes from the AI frontier: Modeling the impact of AI on the world economy. Recuperado el 15/04/2021 de <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy>
- MINSKY, M., *Semantic Information Processing*. Cam. Mass. MIT Press. 1968/2003
- MONASTERIO ASTOBIZA, A., “Ética algorítmica: Implicaciones éticas de una sociedad cada vez más gobernada por algoritmos”. *Dilemata: Revista Internacional de Éticas Aplicadas*. 24, 2017,185-217.
- PARISER, E., *The Filter Bubble: What The Internet Is Hiding From You*. London. Viking. 2012.
- PASQUALE, F., *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cam. Mass. Harvard University Press. 2015.
- PINKER S., *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*. Viking. New York. 2018.
- PLEISS, G. et al., “On fairness and calibration”. *Advances in Neural Information Processing Systems* 52017, 680-5689.
- RAMIÓ, C. *Inteligencia Artificial y Administración Pública: Robots y Humanos Compartiendo el Servicio Público*. Madrid. Catarata 2019.
- RENDA, A. “Artificial Intelligence: Ethics, governance and policy challenges”. *CEPS Task Force Report*, 2019, ISBN 978-94-6138-716-5.
- ROSER, M. y RITCHIE, H., “Child & Infant Mortality”. Publicado online en OurWorldIn-Data.org. Recuperado de ‘<https://ourworldindata.org/child-mortality>’ [Fuente online]
- TAVANI H., “Philosophical theories of privacy: Implications for adequate online privacy policy”. *Metaphilosophy*. 38, 2007, 1-22.
- GUARDIAN, The, “Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach”. Accedido Abril 3, 2021 <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- WARREN, S. y BRANDEIS, L., “The right to privacy”. *Harvard Law Review*, 4, 1890, 193- 220.
- WILLIAMS, J., *Stand out of our Light: Freedom and Resistance in the Attention Economy*. Cambridge. Cambridge University Press. 2018.

LA UE & LA GOBERNANZA ÉTICA DE LA INTELIGENCIA ARTIFICIAL: INTELIGENCIA ARTIFICIAL & DIPLOMACIA¹

*EU & ETHICS GOVERNANCE OF THE ARTIFICIAL
INTELLIGENCE: ARTIFICIAL INTELLIGENCE & DIPLOMACY*

MARIO TORRES JARRÍN

Doctor en Historia
Director del Instituto de Estudios Europeos y Derechos Humanos
Universidad Pontificia de Salamanca
Salamanca/España
director.ieu@upsa.es
ORCID: 0000-0002-9358-5268

Recibido: 1/06/2021
Revisado: 30/08/2021
Aceptado: 6/09/2021

Resumen: Los avances tecnológicos desarrollados dentro del campo de la Inteligencia Artificial (IA) están ocasionando grandes cambios en el conjunto de las sociedades. La IA ofrece una serie de beneficios, e incluso se podría pensar que representa una nueva etapa en la evolución como seres humanos. Pero, también es verdad que la IA abre una serie de interrogantes, entre ellas cuestiones de carácter ético, valores y regulación, ya que en muchos casos puede vulnerar varios de los derechos humanos. Las acciones de las empresas tecnológicas vinculadas a la IA tienen un impacto geopolítico a nivel mundial. Su capacidad económica y su poder de influencia sobre las sociedades las convierten, de facto, en nuevos actores dentro del sistema internacional y en la gobernanza global. El presente artículo examina la propuesta de la Unión Europea (UE) para buscar una gobernanza ética de la IA.

Palabras clave: Inteligencia Artificial, Unión Europea, ética, gobernanza, industria 4.0, diplomacia.

Abstract: The technological advances developed within the field of Artificial Intelligence (AI) are causing great changes in all societies. AI offers a number of benefits, and one might even think that it represents a new stage in evolution as human beings. But, it is also true that

¹ * Este artículo se enmarca dentro del trabajo del Grupo de Investigación “La Unión Europea & la gobernanza ética de la inteligencia artificial” del Instituto de Estudios Europeos y Derechos Humanos de la Universidad Pontificia de Salamanca (España).

AI raises a series of questions, including ethics, values and regulation, since in many cases AI violates various human rights. The actions of technology companies linked to AI have a global geopolitical impact. Their economic capacity and their power to influence societies make them, de facto, new actors within the international system and in global governance. This article examines the proposal of the European Union (EU) to seek an ethics governance of the AI.

Keywords: Artificial Intelligence, European Union, ethics, governance, Industry 4.0, diplomacy.

INTRODUCCIÓN

Hace exactamente cien años, en 1921, se estrenó la obra de teatro de ciencia ficción “*Rossumovi Univerzální Roboti*”², del escritor checo Karel Capek, en ella se utiliza por primera vez el término “robot”, que en checo significa “trabajo forzado”. La obra trata sobre una empresa que construye humanos artificiales orgánicos con el fin de aminorar la carga de trabajo de los seres humanos. A esos humanos artificiales les llama “robots”. Los robots tienen la capacidad de pensar, y a pesar que fueron creados para ayudar a los humanos, los robots organizan una revolución que termina con la destrucción de la humanidad.

En 1956 en la conferencia de Dartmouth, organizada por el matemático estadounidense John McCarthy, se introduce el concepto “Inteligencia Artificial”, el cual es definido como “la ciencia e ingeniería de hacer máquinas inteligentes”³.

En 1957 Frank Rosenblat diseña la primera “red neuronal artificial”⁴. Para 1961, Marvin Minsky (cofundador del laboratorio de Inteligencia Artificial del Instituto Tecnológico de Massachusetts), publica el artículo “*Steps Towards Artificial Intelligence*”⁵.

En la década de los años ochenta, concretamente en 1982, el Ministerio de Comercio Internacional e Industria de Japón comenzó a crear una “computadora de quinta generación” cuya característica era realizar muchos cálculos utilizando

2 CAPEK, Karel, *Robots Universales Rossum*. Translate by Cluida Novack-Jones. New York: Penguin Classic, 2004. First edition in Czech as R.U.R. / *Rossumovi Univerzální Roboti* in 1921.

3 MCCARTHY, John, “Recursive functions of symbolic expressions and their computation by machine, Part I”, Massachusetts: Massachusetts Institute of Technology. pp.1-34.

4 ROSENBLATT, Frank, “The Perceptron: A probabilistic Model for Information storage and organization in the brain” Princeton: *Psychological Review* 65 (6) 1958, pp. 386-408.

5 MINSKY, Marvin, “Steps Towards Artificial Intelligence”, *Proceedings of the IRE*, Vol. 49. N1, January, 1961. pp. 8-30

paralelismo masivo⁶. Posteriormente, poco a poco comenzaron a perfeccionarse los sistemas inteligentes dentro de los ordenadores y aparecen los primeros programas de conversación, como el proyecto ALICE (*Artificial Linguistic Internet Computer Entity*), cuyo creador fue el científico estadounidense Richard Wallace.

En 1997 la empresa IBM crea el superordenador *Deep Blue*, diseñado para jugar ajedrez, y que consiguió ganar al campeón mundial, el ruso Gari Kasparov⁷.

Ya en el siglo XXI, en 2005, usando la ley de Moore, Raymond Kurzweil predijo que las máquinas alcanzarán un nivel de inteligencia humana en el año 2029, y pronostica que para el 2045 habrán superado la inteligencia de nuestra civilización en un billón de veces. Se estima que para ese entonces la IA habrá alcanzado su máximo potencial, convirtiéndose en una “Súper Inteligencia Artificial”. Se espera que, para ese entonces los cerebros de los humanos estén conectado a “la nube”. Para Kurzweil el futuro será de síntesis sin precedentes entre el ser humano y la máquina⁸.

Todos los avances tecnológicos antes descritos evidencian que la civilización humana que conocemos está inmersa en un nuevo estadio de evolución como sociedad y especie. Para Klaus Schwab la humanidad vive una “Cuarta Revolución Industrial” la cual es impulsada por las empresas del sector llamado “Industria 4.0” y que nos llevará a una globalización 4.0. Para Schwab esta nueva etapa de la humanidad se caracteriza por una fusión de tecnologías que desintegrará las fronteras entre las esferas física, digital y biológica⁹, y afirma que “dicha revolución tecnológica modificará la forma en que vivimos, trabajamos y nos relacionamos. En una escala y complejidad la transformación será diferente a cualquier cosa que el género humano haya experimentado”¹⁰.

A lo largo de la historia de la humanidad, el ser humano siempre ha mostrado un interés por comprender el funcionamiento de su propia inteligencia, y más aún, el buscar replicarla en otras creaciones a través de los avances tecnológicos. Partiendo de la base que no existe mayor desafío y proeza que desarrollar una tecnología que permita igualar o superar la inteligencia humana. Dichos deseos

6 TEIGENS, Vasil.; SKALFIST, Peter.; MIKELSTEN, Daniel. *Inteligencia Artificial: La Cuarta Revolución Industrial*, Cambridge: Cambridge Stanford Books, 2020.

7 REAL ACADEMIA DE CIENCIAS MORALES Y POLÍTICAS, *Anales*. Año LXXI, N 96, Madrid: Real Academia de Ciencias Morales y Políticas. p. 251.

8 CREIGHTON, Jolene, “The “Father of Artificial Intelligence” says singularity is 30 years away”. *Futurism*, february 14, 2018. [en línea], 2018-02-14 <https://futurism.com/father-artificial-intelligence-singularity-decades-away> [Consulta: 29 junio 2021].

9 SCHWAB, Klaus, “The Fourth Industrial Revolution. What it means and how to respond”. *Foreign Affairs*, December 2015, New York: Council on Foreign Relations.

10 SCHWAB, Klaus, *The Fourth Industrial Revolution*. Geneva: World Economic Forum. 2016.

no son nuevos. Ya en la antigüedad encontramos dichos anhelos plasmados en la literatura, un ejemplo es el mito de Galatea. Según los griegos, el rey Pigmalión era un magnífico escultor, y que por dedicar tanto tiempo trabajando no había encontrado el amor. Él buscaba una esposa cuya belleza correspondiera a su idea de la mujer perfecta. Por eso, un día decidió hacer la más hermosa escultura de una mujer, y al terminar su obra, quedó enamorado de ella, razón por la cual visitaba su obra día y noche preguntándose cómo sería si su obra tendría vida. Un día Pigmalión organizó una fiesta en honor a la diosa Afrodita y le pidió que por favor le diera vida a su creación. Afrodita concedió sus deseos dando vida a Galatea¹¹.

La literatura inglesa en 1818 presenta la novela “*Frankenstein or the Modern Prometheus*”, en este caso la escritora inglesa Mary Shelly aborda la historia de un ser que es creado artificialmente.

Ya en el siglo XX, de manera paralela al génesis del desarrollo de la IA, el cineasta estadounidense Stanley Kubrick presenta la película “2001, una odisea en el espacio”, la cual describe la historia de una supercomputadora que rivaliza con los seres humanos.

Como ocurre con el caso del robot de Capek, podemos comprobar el constante deseo del ser humano por crear vida e inteligencia de manera artificial. Los avances alcanzados en materia de IA, y más aún su puesta en marcha, presentan múltiples desafíos para el conjunto de la humanidad, sobre todo desde el punto ético. La IA puede vulnerar directamente en varios de los derechos humanos, incluidos los civiles, los derechos políticos, económicos, culturales y sociales; siendo el derecho a la privacidad el más afectado¹².

El Derecho de la Unión busca hacer cumplir el derecho a los ciudadanos a la intimidad y a la protección de los datos personales¹³. Por eso la Comisión Europea ha desarrollado una serie de iniciativas para regular la IA.

Sí se piensa que es una exageración decir que la IA puede alterar la historia de la humanidad o representar una nueva etapa como especie. Entonces tengamos en cuenta y reflexionemos sobre los siguientes datos. A nivel mundial, según Goldman Sachs, en 2017 se estimó que el 90 por ciento de los datos mundiales

11 Ovidio, *Metamorfosis*, Madrid: Consejo Superior de Investigaciones Científicas, Vol II, 1994.

12 RASO, Filippo., HILLIGOSS, Hannah., KRISHNAMURTHY, Vivek., BAVITZ, Christopher and KIM, Levin., *Artificial Intelligence & Human Rights: Opportunities & Risks*. Cambridge: Berkman Klein Center for Internet & Society at Harvard University, September, 2018. p. 4.

13 PARLAMENTO EUROPEO. Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas. Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012) (INL)).

se habían creado en los cinco últimos cinco años, período que a su vez representó un crecimiento de quince veces en el número de desarrolladores de unidades de procesamiento de gráficos, *hardware* crucial para IA¹⁴. De acuerdo a estudios realizados por McKinsey la IA tiene el potencial de generar una actividad económica global adicional de alrededor de \$13 billones para el 2030, lo que sería un 1,2 por ciento del crecimiento global del PIB mundial por año¹⁵. Estos datos evidencian que las sociedades podrán dar un salto cualitativo en muchos ámbitos, lo que modificará la historia de la humanidad.

1. LA UNIÓN EUROPEA Y LA INTELIGENCIA ARTIFICIAL

Dentro del ámbito europeo, y desde el punto de vista del reto social y empresarial que la AI conlleva, según una encuesta realizada por Eurostat en 2020, solo entre el uno y el tres por ciento de las empresas de la UE, con diez trabajadores o más, dicen que usan la IA, principalmente, tecnologías como el aprendizaje automático, el procesamiento del lenguaje y el reconocimiento de voz (excluyendo el sector financiero)¹⁶. Según el ranking global de los países con mayor grado de inversión en innovación en IA realizada por *Tortoise Media* dentro del top 10 encontramos tres Estados Miembros de la Unión Europea (EM-UE): Alemania, Países Bajos y Francia. Entre el top 30 hay 15 EM-UE y 22 EM-UE en el top 50¹⁷.

Sí analizamos dichos datos estadísticos comprobamos que entre los diferentes indicadores existen países europeos que lideran varios sectores, por ejemplo, en desarrollo de *software* operativo ubicamos entre los cinco primeros lugares a países como Polonia, Eslovenia, Lituania, Eslovaquia y Luxemburgo. Esta lista sería encabezada por el Reino Unido, pero este país no es considerado en la lista después del Brexit, sin embargo, he de señalar que con la salida del Reino Unido la UE ha perdido no sólo un miembro, también al país más innovador en materia de IA, el cual junto a Estados Unidos y China se encuentran entre los países líderes

14 GOLMAN SACHS EQUITY RESEARCH, *Profiles in Innovation Revisited: AI hardware*, 11 March 2018, p. 7.

15 BUGHIN, Jacques., SEONG, Jeongmin., MANYIKA, James., CHUI, Michael. And Joshi, Raoul. *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey Global Institute, [en línea], 2018-09-14 <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy>. [Consulta: 29 junio 2021].

16 EUROSTAT, *Artificial Intelligence* [en línea], 2021-06-29 https://ec.europa.eu/eurostat/databrowser/view/ISOC_EB_AI_custom_784358/bookmark/table?lang=en&bookmarkId=f34cd95d-77aa-45bd-8496-eabf49549c02. [Consulta: 29 junio 2021].

17 TORTOISE, *The Global AI Index* [en línea], 2018-02-14 <https://www.tortoisemedia.com/intelligence/global-ai/>. [Consulta: 29 junio 2021].

en todo lo relacionado con los avances en la vanguardia de la IA. De hecho, en los ámbitos de desarrollo de *software* operativos y estrategia gubernamental¹⁸, los indicadores a nivel global de *Tortoise* sitúan a China en el primer lugar antes que Estados Unidos, convirtiendo a China en la “súper potencia en IA”¹⁹.

Hay que tener en cuenta que China tiene varias de las más grandes empresas en el sector de IA como son Baidu, Alibaba y Tencent (juntas son conocidas como las “BAT”), en contraposición encontramos a las empresas estadounidenses llamadas “GAFAM” (Google, Amazon, Facebook, Apple & Microsoft). La UE no cuenta con empresas de estas características, ni en desarrollo tecnológico, capacidad económica ni poder político. La acción exterior que despliegan este grupo de empresas y los datos que gestionan las convierten en actores geopolíticos de facto, por tanto, actores con poder de decisión y acción dentro del sistema internacional, poderes y capacidades muy superiores a los que en la actualidad poseen los estados nacionales.

Para la UE los pronósticos de Raymond Kurzweil de que para el 2045 exista una “Súper inteligencia Artificial” o también llamada “*Inteligencia Artificial General*” no es ciencia ficción, es más bien una realidad próxima que tenemos que regular, teniendo en cuenta el impacto que generará en el conjunto de la humanidad. Para Stuart Armstrong la “Inteligencia Artificial General” logrará obtener poderes sobrehumanos: “Imagínense a un humano que, cada vez que abre la boca, han pasado un año entero para reflexionar e investigar si su capacidad de respuesta será máximamente efectiva. Así será una IA social”²⁰. Además, entrará en el campo de la investigación en el desarrollo de nuevas tecnologías, es decir, “haría investigación y desarrollo simultáneamente en ciento de subcampos técnicos y combinaría sin descanso ideas entre campos. El desarrollo tecnológico humano cesaría, y la IA o las tecnologías de investigación guiadas por IA se volverían ubicuas rápidamente”²¹. Los resultados de dichas investigaciones podrían ahorrar años, décadas e incluso siglos de investigación si se compara con los resultados que se podrían obtener si esta es realizada por seres humanos.

El 16 de julio de 2019, la presidenta de la Comisión Europea, Ursula Von der Leyen anunció, entre sus directrices políticas, un enfoque europeo coordinado

18 Ibid.

19 LEE, Kai-Fu, *AI superpowers: China, Silicon Valley, and the new world order*. New York: Houghton Mifflin Harcourt Publishing, 2019.

20 AMSTRONG, Stuart, *Smarter than us: The rise of machine intelligence*. Machine Intelligence Research Institute, 2014, p. 14.

21 Ibid., p. 15.

de la implicación ética y humana de la Inteligencia Artificial, así como realizar un análisis sobre cómo mejorar el uso de *Big Data* en la innovación²².

Según la Comisión Europea, la IA está transformando nuestro mundo y está presente en nuestra vida diaria; cuando traducimos textos *online* o usamos una aplicación móvil. La IA ofrece grandes oportunidades para mejorar nuestra sociedad, e incluso, podríamos pensar que dicha tecnología llevará a la humanidad a un estadio superior a través de una evolución como seres humanos. Pero también, dicha tecnología representa muchas incógnitas sobre la pertinencia de su uso. De ahí la necesidad de regularla.

Entre las oportunidades, podemos destacar su aporte en la mejora en el sector de la salud, dado que gracias a los algoritmos se puede ayudar a obtener un mejor diagnóstico, hecho que permitiría detectar enfermedades incurables a tiempo o realizar una medicina de prevención para evitar enfermedades futuras. En términos del sector de la educación, los aportes son diversos, desde mejorar el rendimiento de los estudiantes, mediante programas que permitan evaluar el proceso de aprendizaje de un estudiante de manera personalizada, comparar su evolución con otros estudiantes, tomando como referencia una base de datos, estudiar el desempeño del estudiante, conocer sus necesidades, identificar sus fortalezas y debilidades para finalmente obtener una forma más personalizada de enseñanza tomando en cuenta sus capacidades y potencialidades intelectuales. El uso de realidad virtual o el incluir elementos interactivos tanto en la enseñanza como en la evaluación son áreas que permiten ver el potencial y el salto cualitativo que se puede obtener de una educación adaptada a la era digital.

En el campo empresarial, el transporte, análisis de datos permite a las empresas lograr mayores soluciones eficientes, en sus procesos de producción, comunicación y distribución. En el ámbito del sector público, la defensa, la seguridad, la justicia o el sistema tributario son algunas de las potenciales áreas a mejorar mediante el uso de la IA.

Sin embargo, también existen grandes riesgos, sobre todo los relacionados con la vulneración de los derechos fundamentales y los derechos humanos de los ciudadanos. Por poner algunos ejemplos, podemos citar que en aras de la defensa y la seguridad, se utilizan tecnología de IA como: el reconocimiento facial, el reconocimiento de voz, análisis de nuestros datos personales. Cada una de estas áreas mal gestadas y gestionadas pueden generar discriminación (principalmente racial, etnia, género, edad o condición de discapacidad física) además de obtener

22 VON DER LEYEN, Ursula., *A Union that strives for more. My agenda for Europe. Political guidelines for the Next European Commission 2019-2024*. Luxembourg: Publications Office of the European Union, 2019.

información sesgada y prejuiciosa, dependiendo de los parámetros e información que se haya tenido en cuenta a la hora de diseñar los programas correspondientes. Ni que decir de los riesgos que esto implica si regímenes autoritarios utilizan estas tecnologías para reprimir a determinados grupos en una sociedad.

La Universidad de Georgetown realizó un estudio en 2016 en el que mencionaba que “la tecnología de reconocimiento facial afectará desproporcionadamente a las personas afroamericanas, debido a que en las listas de personas bajo vigilancia policial existen más rostros de personas negras que blancas”²³. En 2018, la Unión de Libertades Civiles de Nueva York, reveló que 2017, el *software* de Evaluación de Clasificación de Riesgos que utilizaba el Servicio de Inmigración y Control de Aduana de los Estados Unidos, desde el año 2013, había sido utilizado para decidir si un inmigrante debía de ser detenido o puesto en libertad bajo fianza²⁴.

La poca transparencia de muchas empresas dedicadas a la IA genera grandes interrogantes sobre el uso de los datos personales, no se sabe en qué servidor o país se encuentran dichos datos, también se desconoce quién o qué empresas los almacena, estudia, usa, gestiona y comercializa. A través de la IA es posible analizar perfiles personales para luego buscar influir en las decisiones de las mismas, ya sean comerciales o políticas. Esta irrupción en vida privada de los ciudadanos también se da en la vida política de los países. Los ataques cibernéticos y las campañas de desinformación ocasionan cambios en el electorado, alteran procesos electorales y perturban los mercados internacionales con juegos especulativos y arbitrarios.

La Comisión Europea propuso un enfoque que coloque a las personas en el centro del desarrollo de la IA, es decir una IA centrada en el ser humano. Con esta propuesta la UE busca liderar a nivel internacional su regulación, defendiendo principios éticos de acuerdo con los valores de la UE. Una gobernanza ética de la IA puede traer grandes beneficios a la sociedad, como obtener una mejor atención médica, un consumo de energía más eficiente, mejorar la productividad en la agricultura, aminorar los costos de producción, la sostenibilidad ambiental, ayudar a detectar el fraude y reducir las amenazas a la ciberseguridad.

En octubre de 2020, el Parlamento Europeo (PE) adoptó una serie de resoluciones relacionadas con la IA sobre ética, responsabilidad y derechos de autor. En 2021, les siguieron resoluciones sobre IA en materia penal y en educación,

23 GARVIE, Clare, BEDOYA, Álvaro y FRANKLE, Jonathan, *The perpetual Line-Up. Unregulated police face recognition in America*. Washington: Center on Privacy & Technology at the Georgetown University, 2016.

24 DE ASÍS, Rafael, “*Inteligencia artificial y derechos humanos*”. Materiales de filosofía del derecho, N° 2020 / 04. Madrid: Universidad Carlos III de Madrid.

cultura y sector audiovisual. El 21 de abril de 2021, la Comisión Europea publicó una “Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre inteligencia artificial (llamada también “Ley de inteligencia artificial”) y se modifican determinados actos legislativos de la Unión”²⁵. Esta propuesta responde a las solicitudes del Parlamento Europeo y del Consejo Europeo que habían mostrado su preocupación por la falta de regulación, razones por las cuales decidieron impulsar acciones legislativas que garanticen un mercado interior que incremente las oportunidades del IA hacia los ciudadanos y reduzca sus riesgos. La propuesta apoya como objetivo máximo en que la UE se convierta en un líder mundial en el desarrollo de una IA segura, confiable y ética.

El reglamento clasifica en varios niveles de riesgo las aplicaciones de IA. “Riesgo inadmisibles”, por tanto, deberán de ser prohibidos. Es cuando los sistemas de IA se consideren una amenaza clara a la seguridad y los derechos fundamentales, esto incluye sistemas que manipulan el comportamiento humano, el abanico es amplio, desde juguetes que utilicen asistencia vocal para incitar a comportamientos peligrosos entre los menores hasta sistemas de puntuación social como los que utiliza China para controlar sus ciudadanos. También es el caso del reconocimiento facial a través de cámaras que emplean sistema de IA de identificación biométrica facial. Aunque se señala, que estos sistemas se podrán utilizar con autorización de un órgano judicial para prevenir una amenaza terrorista o localizar e identificar a prófugos y sospechosos de delitos graves. Este punto, dependiendo del país que lo use puede servir de excusa para implementar políticas de represalias o persecuciones políticas.

“Riesgo alto”, sistemas de calificación crediticia o comprobación de documentos de viajes. “Riesgo limitado”, uso de robots o *bots* en conversaciones con usuarios y el “*deep fake*”. “Riesgo mínimo”, los videojuegos o los filtros de correo basura. Cabe mencionar que el reglamento excluye las tecnologías de IA para uso militar.

2. LA DIPLOMACIA DE LA UE EN LA IA: HACIA UNA GOBERNANZA ÉTICA DE LA INTELIGENCIA ARTIFICIAL

El 21 de junio de 2021, el Parlamento Europeo publicó un estudio titulado “*Artificial Intelligence diplomacy. Artificial Intelligence governance as a new*

25 EUROPEAN COMMISSION, *Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM (2021) 206 final.*

European Union External Policy Tool”, el cual indica que la IA se ha convertido en una herramienta de poder político y un elemento de diplomacia estatal, concluye que el impacto de la IA genera, en términos geopolíticos, una alteración interna y externa en la vida de la UE. Por ello, es necesario adaptar el mercado común a una era digital donde prevalece un importante sector de la Industria 4.0, crear nuevas normas jurídicas que regulen el diseño, producción y comercialización de las tecnologías producto de la IA, a la par que regular su uso. Esta es la realidad, no una hipótesis o posible escenario futuro el que se enfrenta la UE y sus Estados Miembros (EM-UE). La IA está destinada a influir en el equilibrio global de poder y en la relación entre los estados, así como en la geopolítica en general²⁶.

Los EM-UE tiene que invertir más en investigación sobre IA. El documento del Parlamento señala como en Estados Unidos, por ejemplo, sólo en lo que va de 2021 la Comisión Nacional de Seguridad sobre IA ha publicado más de 700 páginas de reportes, los cuales no solo provienen de entes estatales, sino de proyectos conjuntos entre gobierno-universidad, así como gobierno-empresa-universidad.

Las universidades europeas no han profundizado en el tema, apenas existen grupos de investigación que en la actualidad se dedican a estudiar las repercusiones de la AI en el futuro de la sociedad, mucho menos sí los avances producidos por la AI cumplen con los principios éticos y condiciones instrumentales que le permiten a la persona su realización, es decir el cumplimiento de los derechos humanos.

Entre los cambios geopolíticos a enfrentar por parte de la UE ante la IA el documento del Parlamento Europeo destaca las siguientes áreas:

- *Competición Estados Unidos-China*. Siendo esta la primera medida a implementar por parte de la UE, evitar una confrontación militar entre ambos países utilizando las diferentes tecnologías desarrolladas dentro del marco de la IA. Ya que dicho hecho tendría consecuencias catastróficas para la Europa y el mundo. Para lo cual la UE deberá de trabajar de la mano con ambos países.

En el caso de los Estados Unidos, la UE debería de priorizar el trabajar con los aliados, sobre todo en todo lo que se refiera a la posición de los Estados Unidos en materia de tecnología²⁷. La nueva administración del presidente Joe Biden, cuya

²⁶ EUROPEAN PARLIAMENT, *Artificial Intelligence diplomacy. Artificial Intelligence governance as a new European Union External Policy Tool*. Policy Author: Ulrike Franke Luxembourg: Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies. p. 39.

²⁷ IMBRIE, Andrew., FEDASIUK, Ryan., Alken, Catherine, CHHABRA, Tarun and CHA-HAL, Husanjot., *Agile Alliance. How the United States and its allies can deliver a democratic way*

política exterior se inclina por favorecer un reforzamiento del eje transatlántico (UE-USA) frente a China, es una acción manifestada por el propio presidente estadounidense. Con relación a dicho reforzamiento, en la última Cumbre UE-Estados Unidos, celebrada en Bruselas el 15 de junio de 2021, entre los principales resultados están los siguientes puntos:

- Establecer un Consejo Transatlántico de Comercio y Tecnología.
- Impulsar la transformación digital, que incluya fortalecer y promover el liderazgo tecnológico e industrial de las empresas europeas y estadounidenses en tecnologías emergentes.
- Apoyar la investigación y los intercambios colaborativos.
- Facilitar la cooperación en materia de aplicación y políticas de reglamentación.
- Coordinar esfuerzos en los organismos multilaterales para promover un modelo democrático de gobernanza global.

El Consejo Transatlántico de Comercio y Tecnología incluirá la creación de grupos de trabajo con agendas centradas en crear estándares en IA, Internet de las Cosas, entre otras tecnologías emergentes, clima y tecnología verde, seguridad y competitividad de las TIC, gobernanza de datos y plataformas tecnológicas, el uso indebido de la tecnología que amenaza la seguridad y los derechos humanos²⁸. El trabajo de este consejo puede servir para afianzar las posiciones europeas y estadounidenses dentro del *Global Partnership on Artificial Intelligence*²⁹ creado en junio de 2020. El cual fue una idea desarrollada dentro del G7 bajo la presidencia de Canadá y Francia, y cuyos objetivos contemplan el cerrar la brecha entre la teoría y la práctica de la IA mediante el apoyo a la investigación de vanguardia y las actividades aplicadas sobre las prioridades relacionadas con la IA. Lo interesante de esta instancia es que, de los 15 países fundadores, 8 son europeos y 7 EM-UE además de Estados Unidos y la propia UE: Australia, Canadá, Francia, Alemania, India, Italia, Japón, México, Nueva Zelanda, República de Corea, Singapur, Eslovenia, Reino Unido, Estados Unidos y la Unión Europea. En diciembre de 2020 se sumaron: Brasil, Países Bajos, Polonia y España.

of AI. Washington: Center for Security and Emerging Technology at Georgetown University, 2020.

28 Declaración conjunta. Cumbre Unión Europea-Estados Unidos “Towards a renewed Transatlantic partnership”. Bruselas: Consejo de la Unión Europea, 15 de junio de 2021. pp. 3-4.

29 THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE. [en línea], 2021-06-29 <https://gpai.ai/> [Consulta: 29 junio 2021].

Pero existen también otros foros, en los cuales el consejo transatlántico puede intervenir, como el foro llamado “D-10”³⁰, el cual nació siendo una iniciativa del Departamento de Estado de Estados Unidos en 2008, y que posteriormente fue retomado por el *Atlantic Council* en 2014³¹, cuyo objetivo fue el mantener un “orden democrático” basado en las reglas de “diez democracias líderes”: Australia, Canadá, Francia, Alemania, India, Italia, Japón, Corea del Sur, Reino Unido y Estados Unidos. A los que se suman la UE y tres estados observadores: Indonesia, Polonia y España.

En lo que respecta a las relaciones entre la UE y China es necesario conseguir tener una posición común sobre IA. Sin embargo, existen divergencias entre los EM-UE sobre si es conveniente o no desarrollar una mayor profundización de las relaciones con el país asiático. Entre los principales argumentos contrarios a un mayor acercamiento con China, podemos identificar la cuestión si permitir la inversión china en empresas o proyectos europeos dedicados al sector del desarrollo de *software* operativos y otros tipos de infraestructura digitales, lo que genera una desconfianza hacia China³². Pero la desconfianza es mutua. Para China Europa no posee el peso suficiente como para ser considerado un aliado en los temas del sector la Industria 4.0, mucho menos en los vinculados con los de IA. Kai-Fu, ex presidente de Google China, trabajó también en Apple y Microsoft, creador de un sistema de reconocimiento de voz, afirmó en 2018 que Europa “ni siquiera está en la carrera por la medalla de bronce de IA”³³. En 2019 dicha afirmación se vio corroborada con un informe emitido por *McKinsey Global Institute* que afirmaba que “Europa está incrementando su brecha de IA a su brecha digital”³⁴. En efecto la Industria 4.0 en Europa va muy rezagada respecto a otras regiones del mundo, Asia y América, y de manera particular con las estadounidenses y chinas, sobre

30 FRANKE, Ulrike, “Artificial Divide. How Europe and America could clash over AI”, London: European Council on Foreign Affairs, January, 2021. Ver también: PANIER, ALICE, “Europe in the Geopolitics of Technology, Connecting the Internal and External Dimensions”. Paris: Institut Francais des Relations Internationales, April 2021.

31 ATLANTIC COUNCIL, *D-10 Strategy Forum*. [en línea], 2021-06-29 <https://www.atlanticcouncil.org/programs/scowcroft-center-for-strategy-and-security/global-strategy-initiative/democratic-order-initiative/d-10-strategy-forum/> [Consulta: 29 junio 2021]

32 OERTEL, Janka, “The new China consensus: How Europe is growing wary of Beijing”, London: European Council on Foreign Relation, Policy Brief, [en línea], September 2020. https://ecfr.eu/publication/the_new_china_consensus_how_europe_is_growing_wary_of_beijing/. [Consulta: 29 junio 2021]

33 MINSKY, Carly, “One former Google exec says there’s no hope for Europe’s artificial intelligence sector”, *Sifted*, [en línea], December, 2018. <https://sifted.eu/articles/interview-google-kai-fu-lee-ai-artificial-intelligence/>. [Consulta: 29 junio 2021]

34 BUGHIN, Jacques., SEONG, Jeongmin., MANYIKA, James., HÄMÄLÄINEN, Lari, WINDHAGEN, Eckart & HAZAN, ERIC., *McKinsey Global Institute*,. [en línea], February, 2019. [insights/artificial-intelligence/tackling-europes-gap-in-digital-and-ai#](https://www.mckinsey.com/industries/artificial-intelligence/tackling-europes-gap-in-digital-and-ai#). [Consulta: 29 junio 2021]

todo en las áreas relacionadas con *Big Data*, *Machine Learning*, ambas tecnologías son la base de la IA.

El cruce de sanciones entre UE y China no favorece a distender las relaciones diplomáticas entre ambas partes y establecer una cooperación, sin embargo, se hace necesario hacerlo. En marzo de 2021 la UE impone sanciones contra China por abusos contra los derechos humanos. Los ministros de asuntos exteriores de la UE aprobaron sancionar a cuatro funcionarios chinos y a una institución con la prohibición de viajes a la UE, además de imponer una congelación de sus activos. En respuesta a esta medida el gobierno chino respondió sancionando a diez ciudadanos de la UE, entre ellos eurodiputados y políticos. A la vez, incluyó sanciones a cuatro entidades europeas, entre ellas, el Comité Político y de Seguridad de la UE, la subcomisión de Derechos Humanos del Parlamento Europeo; el gobierno chino concluyó su paquete de prohibiciones asegurando que los sancionados y sus familias tienen prohibida la entrada al territorio chino, incluyendo los territorios de Hong Kong y Macao, además de la negativa de hacer negocios en China.

Hay que recordar que en la cumbre UE-China de 2020 la UE indicó la necesidad que el desarrollo de nuevas tecnologías digitales debe ir de la mano del respeto de los derechos fundamentales y la protección de datos, además de plantear las cuestiones pendientes sobre ciberseguridad y desinformación³⁵.

Por otro lado, China lamentó la última declaración conjunta entre UE-Estados Unidos en la que se mencionaba que las partes se comprometían a cooperar en una serie de cuestiones globales, entre ellas hacer frente a China³⁶. Es claro que el presidente Joe Biden busca ser y hacer un contrapeso al poder que China ha ganado en los últimos años. La retirada de Estados Unidos, durante el período del presidente Donald Trump, de varios organismos y foros internacionales, así como la congelación de procesos de negociaciones de tratados, ha hecho que China haya aprovechado ese vacío de poder que dejaba Estados Unidos en cada una de las instancias antes descritas. De ahí que, durante su primera gira europea, que incluía reunión del G7, OTAN y UE el presidente Joe Biden dijera “Estados Unidos ha vuelto y las democracias del mundo están unidas para afrontar los retos

35 CONSEJO DE LA UNIÓN EUROPEA, *EU-China Summit: Defending EU interests and values in a complex and vital partnership- Press release by President Charles Michel and President Ursula von der Leyen*. [en línea], 2020-06-22. <https://www.consilium.europa.eu/es/press/press-releases/2020/06/22/eu-china-summit-defending-eu-interests-and-values-in-a-complex-and-vital-partnership/> [Consulta: 29 junio 2021]

36 EURONEWS, “China rechaza y lamenta las críticas de la cumbre entre Estados Unidos y la UE” [en línea], 2021-06-16 <https://es.euronews.com/2021/06/16/ue-eeuu-comercio-china> [Consulta: 29 junio 2021].

más difíciles y los asuntos que más importan para nuestro futuro”³⁷. Estas declaraciones eran una clara alusión a China. Con esta afirmación deja constancia que el liderazgo de China, aunque es global, no será jamás validado por no ser un país democrático, de ahí que mencione “las democracias del mundo están unidas para afrontar los retos” identificando que China es uno de ellos.

De democracia va el segundo punto que aborda el documento del Parlamento Europeo.

- *Autoritarismo y debilitamiento de la democracia.* En lo que respecta a este punto la UE enfatiza que tiene como objetivo ser líder mundial en lograr un desarrollo ético y confiable de la IA. Para lo cual considera que es imprescindible crear legislación que regule la IA, tanto de manera intracomunitaria como extracomunitaria.

Se espera que sí la UE es capaz que su Industria 4.0 dedicada al sector de la IA cumple con los estándares éticos europeos, esto servirá de garantía para los ciudadanos europeos, pero también para ciudadanos extracomunitarios que preferirán comprar o contratar servicios de empresas europeas por ser estas respetuosas con los principios éticos.

- *Empoderamiento de las empresas privadas.* Reconociendo que se ha hecho mucho por combatir contra las *Big Tech*³⁸, principalmente mediante el trabajo de los comisarios europeos Margrethe Vestager y Thierry Breton. También se reconoce desde las instituciones comunitarias que las empresas tecnológicas estadounidenses y cada vez más las chinas son las principales y más grandes del mundo. Algo que tiene que ser revertido.
- *Defensa y seguridad.* En el ámbito militar la IA ha demostrado que puede contribuir a importantes ventajas militares, además de poder identificar y conocer los riesgos que la tecnología desarrollada dentro del sector puede causar al conjunto de la sociedad europea y a nivel mundial.

Un asunto primordial y urgente sería que la UE solicite a sus Estados Miembros que desarrollen estrategias militares de IA³⁹. El impulso de la Comisión Europea

37 EURONEWS, “Estados Unidos ha vuelto, dice Joe Biden al llegar a Reino Unido para el G7 y una gira europea”, [en línea], 2021-06-10 <https://es.euronews.com/2021/06/10/estados-unidos-ha-vuelto-dice-joe-biden-al-llegar-a-reino-unido-para-el-g7-y-una-gira-euro> [Consulta: 29 junio 2021].

38 LARGER, Thibault., SCOTT, Mark., KAYALI, Laura., “Inside the EU’s divisions on how to go after Big tech”, Brussels: Politico, 14 December 2020.

39 FRANKE, Ulrike., “Europe needs a plan for AI in the military realm”, *The Security Times*, [en línea], February 2020. <https://www.the-security-times.com/europe-needs-plan-ai-military-realm/>. [Consulta: 29 junio 2021].

para desarrollar estrategias nacionales de IA ha desempeñado un rol importante a la hora de alentar a los EM-UE a involucrarse en el tema⁴⁰. La cooperación en las áreas antes descritas pueden ser unas buenas áreas de cooperación a desarrollar dentro del marco de las relaciones transatlánticas y con la OTAN⁴¹.

Por otro lado, el Parlamento Europeo también menciona que la UE puede también reaccionar a la propuesta de la Comisión Nacional de Seguridad de Estados Unidos sobre la creación de una “Asociación de Tecnología de Seguridad Atlántico-Pacífico”⁴² para promover la interoperabilidad en defensa e inteligencia entre Europa y la región Indo-Pacífico.

- *Inteligencia Artificial General*. En este punto la UE remarca que se debería de prestar más atención e invertir en seguridad, así brindar más experiencia técnica a los gobiernos en esta área.

Respecto a este último tema, el *Center for Long-Term Resilience* publicó un informe en junio 2021 titulado “*Future Proof. The Opportunity to transform the UK’s Resilience to extreme risks*” donde identifica que la IAG es una amenaza para la humanidad⁴³.

CONCLUSIÓN

La historia de la humanidad nos muestra que existen períodos de grandes avances tecnológicos que son los que marcan una nueva era dentro de la evolución de los seres humanos. La tecnología ha permitido a los seres humanos facilitar su vida, trabajo y elevar su calidad de vida. Las primeras herramientas de piedra, bronce o acero sirvieron para facilitar el desarrollo de la agricultura, la pesca, defenderse de los animales o realizar las primeras construcciones que le sirvieron como viviendas para protegerse de las inclemencias del tiempo. Es

40 EUROPEAN COMMISSION, *Coordinated Plan on Artificial Intelligence*. COM (2018) 795 final, Brussels: European Commission. [en línea], 2018-12-07 <https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence>. [Consulta: 29 junio 2021].

41 KONAEV, Margarita., CHAHAL, Husanjot., “The path of Least Resistance, Multinational Collaboration on AI for Military Logistics and Sustainment”, Washington: Center for Security and Emerging Technology at Georgetown University, April 2021.

42 US National Security Commission on Artificial Intelligence, Final Report. [en línea], March 2021 <https://reports.nscai.gov/final-report/table-of-contents/>. [Consulta: 29 junio 2021].

43 ORD, Toby., MERCER, Angus., & DANNREUTHER, Sophie., *Future Proof. The opportunity to transform the UK’s Resilience to extreme risks*. TheCentre for Long-Term Resilience. [en línea], June 2021 <https://drive.google.com/file/d/1LHn3nzxF2p68SfhwIPLCb5FMaMLq1dk6/view>. [Consulta: 29 junio 2021].

decir, sirvieron para la vida y subsistencia humana. Pero también esas mismas herramientas sirvieron para combatir y hacer guerras, por consiguiente, para hacer daño a otros seres humanos. La brújula, la imprenta, las máquinas a vapor, la electricidad, la energía nuclear, el internet son claros ejemplos en los cuales las sociedades han dado un salto cualitativo como humanidad.

La IA es sin duda uno de esos hitos que marcan un antes y un después dentro de los períodos de la evolución humana. Ahora bien, la cuestión es que, en la mayoría de las anteriores etapas de evolución humana, éstas cumplían con su definición “evolución”, que significa proceso de transformación, en este caso de la especie humana, a través de cambios producidos en sucesivas generaciones. La diferencia, en el caso de la etapa que nos presenta la IA es que dichos cambios se producen en una misma generación. La era digital que trae consigo una Industria 4.0 y que ha generado una globalización 4.0, demanda una transformación de una sociedad que vaya de 1.0 a 4.0 en una misma generación. Y aquí uno de los desafíos que se tiene como humanidad.

El reconocimiento universal de los derechos humanos ha sido un logro que ha tardado siglos en ser aceptado por el conjunto de la sociedad internacional. Ahora el mundo cuenta con una serie de tratados internacionales y normas que velan por la defensa de dichos derechos. Algunas tecnologías de la IA vulneran muchos de los derechos humanos, por tanto, regular la IA es también un acto de velar por el cumplimiento internacional de los derechos humanos.

Vivimos tiempos convulsos y difíciles, de grandes transformaciones a la vez, nunca el mundo ha tenido que enfrentar el cuestionamiento de varios de sus sistemas al mismo tiempo: político, económico, social y cultural. El resurgimiento y la emergencia de grandes potencias mundiales reconfiguran el sistema internacional que imperaba a lo largo del siglo XX. El siglo XXI agrupa a potencias con gran pasado histórico en los asuntos internacionales y nuevas potencias (China y Estados Unidos) sí a eso le sumamos las grandes empresas del sector de la IA que poseen un poder económico y geopolítico mayor que las grandes potencias y que la gran mayoría de los países, entonces es evidente que vivimos un nuevo orden mundial.

En este nuevo orden mundial producido por el proceso de la digitalización, la UE y sus Estados Miembros se encuentran en el medio de todas las potencias (grandes y medianas). La diplomacia de la UE puede ser el mejor instrumento para recuperar un espacio geopolítico que hoy, tanto la UE como los EM-UE están perdiendo. Con la propuesta de desarrollar una gobernanza ética de la IA, la UE puede lograr ser un mediador global entre las grandes potencias y consolidarse como un actor líder en cuestiones sobre el futuro de la gobernanza 4.0, y ser un agente que gestione el proceso de digitalización 4.0; articulando los distintos intereses entre los países grandes, medianos y pequeños.

En este sentido la propuesta por el gobierno danés de 2017 sobre una *Techplomacy* y un *Tech Ambassador* son respuestas adecuadas a nuestro tiempo. La iniciativa del Ministro de Asuntos Exteriores danés propone desarrollar relaciones con las grandes empresas del sector de la Industria 4.0 (como pueden ser las “BAT” o las “GAFAM”), por varios motivos. Desde entender todos los avances tecnológicos para poder luego regularlos como Estado hasta poder comprender sus implicancias en el conjunto de las relaciones internacionales. Es este tipo de acción exterior que tiene que desarrollar la UE, es decir una política europea en *Techplomacy*, e incluir dentro de ésta las cuestiones concernientes a la gobernanza ética de la IA.

La iniciativa danesa representa un cambio de paradigma y la manera de concebir las relaciones internacionales⁴⁴ acorde a los tiempos en que vivimos, esta acción eleva la tecnología y la digitalización a un prioridad transversal de política exterior y seguridad⁴⁵ de los países. “Hemos sido demasiado ingenuos durante mucho tiempo sobre la revolución tecnológica. Necesitamos asegurarnos de que los gobiernos democráticos establezcan los límites para la industria tecnológica, y no al revés”⁴⁶.

Por otro lado, no se puede querer regular la IA sin regular el ciberespacio. La ciberdiplomacia como aplicación de la diplomacia a los problemas que se crean por causa de las tecnologías que surgen en el ciberespacio⁴⁷.

Estados Unidos y China son las grandes potencias en el ciberespacio, sus empresas son las que controlan gran parte del flujo de información que se produce en

44 TORRES JARRÍN, Mario & RIORDAN, Shaun. “Techplomacy. Hacia la búsqueda de una regulación del ciberespacio y la gobernanza de Internet”, en Beltrame de Moura. *O Direito Internacional Privado Europeu: entre a harmonização e a fragmentação*, Brasil: Emais Editora y Librería Jurídica. Publicación financiada el Programa Erasmus + de la Unión Europea, 2019.

Ver también: TORRES JARRÍN, Mario. & RIORDAN, Shaun. *T20 Policy brief. The cyber diplomacy of constructing norms in cyberspace*. T20/G20 Secretariat. [en línea] 2020-11-24 https://www.g20-insights.org/wp-content/uploads/2020/11/T20_TF5_PB4.pdf [Consulta: 29 junio 2021]

TORRES JARRÍN, Mario & RIORDAN, Shaun. *Global Policy Perspectives Report “Techplomacy and the Tech Ambassador”*. Salamanca-Stockholm: European Institute of International Studies, 2020.

45 MINISTRY OF FOREIGN AFFAIRS OF DENMARK. Foreign and Security Policy Strategy 2017-2018. Copenhagen: Ministry of Foreign Affairs, June 2017.

46 MINISTRY OF FOREIGN AFFAIRS OF DENMARK, Office of the Tech Ambassador of Denmark, [en línea] 2020-11-24 <https://techamb.um.dk/> [Consulta: 29 junio 2021] Ver también: KLYNGE, Casper., EKMAN, MIKAEL., & WAEDEGAARD, Nikolaj Juncher. Diplomacy in the Digital Age: Lessons from Denmark’s Techplomacy initiative, *The Hague Journal of Diplomacy*, 15 (1-2), 185-195, Leiden: BRILL, 2020.

47 RIORDAN, Shaun & TORRES JARRÍN, Mario. *Global Policy Perspectives Report “Cyberdiplomacy”*. Salamanca-Stockholm: European Institute of International Studies, 2020.

el mismo y además las que poseen el poder en lo que se respecta a la administración y gestión de datos. Rusia, India y Unión Europea van en un segundo orden, pero no cabe duda que la diferencia radica entre los países que quieren regular la interacción del individuo con las empresas relacionadas con el sector estratégico de la IA y los que no quieren, o peor aún, los que quieren regular para sus propios beneficios políticos.

Los ciberataques son las nuevas guerras del siglo XXI, no es necesario lanzar misiles atómicos o teledirigidos hacia otro país para causarle daño a un país considerado enemigo, hoy en día basta con realizar un ciberataque a su sistema financiero, sistema de salud, sistema de telecomunicaciones o fuentes de energía para causar grandes daños y perjuicios a un país tanto a nivel económico, como político y social. De ahí que los gobiernos sean recelosos sobre el poder geopolítico de las “Big Tech”, sobre todo las que están vinculadas al sector de las telecomunicaciones.

Muchos países se preparan para pasar sus redes móviles 4G a 5G, buscando un aumento en la velocidad de descargas diez veces superior, hecho que cambiaría radicalmente la forma de transmitir información, lo que repercutiría en la forma de producir, comunicar, investigar, estudiar y trabajar. Empresas europeas como la sueca Ericsson o Nokia de Finlandia están en la vanguardia del sector 5G a nivel de la UE. Representado a las empresas estadounidenses encontramos a Cisco, Juniper Networks o Qualcomm. En el caso de China, Huawei es la empresa líder. El recelo por parte de los países occidentales hacia la empresa china reside en cómo va a manejar los datos y si esta empresa entregará los datos al gobierno chino. Técnicamente, China podría argumentar el mismo punto por parte de las empresas estadounidense y europeas que operen en su territorio. Lo cierto es que existe una desconfianza mutua, y de ahí que la IA para muchos países sea considerada una política de Estado⁴⁸.

La Comisión de Seguridad Nacional sobre Inteligencia Artificial de Estados Unidos publicó un informe en 2021, el cual indica que China está decidida a superar en liderazgo de IA a los Estados Unidos, por tanto, el gobierno estadounidense debería establecer una alianza tecnológica con sus socios habituales para hacer frente a China. Dicha alianza se debe basar en los valores democráticos frente al uso de la IA como herramienta de represión y vigilancia por parte de China. El informe menciona a la UE y a países asiáticos como Japón y Corea del Sur como socios de dicha alianza⁴⁹. Con respecto a la vigilancia, Zuboff nos señala que la

48 MORENO, Luis y PEDREÑO Andrés. *Europa frente a EE.UU y China: prevenir el declive en la era de la inteligencia artificial*, Arizona: University of Arizona Egyptian Exped, 2020.

49 NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE: Final Report. National Security Commission on Artificial Intelligence, 2021.

amenaza que se cierne sobre nosotros, como sociedad, no es ya la de un Estado “Gran Hermano” totalitario, sino la de una arquitectura digital omnipresente: un “Gran Otro” que opera en función de los intereses del capital de la vigilancia⁵⁰.

La creación de normas internacionales en la gobernanza ética de la IA es un imperativo no solo legal sino humano a ser definido y defendido. Las iniciativas emprendidas por parte de la UE respecto a desarrollar una IA centrada en el ser humano, que proteja los principios éticos basados en los valores que propugna la UE es un buen camino a seguir, y representa toda una oportunidad para que la UE lidere dicho proceso de regulación promoviendo sus valores y principios rectores. Los principios éticos propuestos por parte de la UE va más allá de los principios éticos acordados en el marco de la OCDE⁵¹ y del G20⁵².

Mientras se concluye este artículo, la UNESCO⁵³ y la OMS⁵⁴ presentan sus primeras recomendaciones sobre ética en la IA, quedará pendiente hacer seguimiento a estas iniciativas y ver cómo encajan dentro de las propuestas de la UE, y más aún como se pueden articular, a fin de tener una normativa internacional consensuada entre toda la Comunidad Internacional.

Mucha de la tecnología desarrollada desde el área de la IA presenta varias cuestiones éticas sin resolver, las cuales el conjunto de la sociedad no se ha detenido a estudiar, evaluar y considerar sus beneficios o perjuicios, ni mucho menos a investigar su impacto en la humanidad. Hacia dónde queremos ir como sociedad y qué tipo de humanos queremos ser son algunas de las principales cuestiones que nos presenta como reto la IA.

50 ZUBOFF Shoshana, *La era del capitalismo de la vigilancia: la lucha por un futuro humano frente a las nuevas fronteras del poder*. Barcelona: Paidós, 2020.

51 OECD: Recommendation of the Council on Artificial Intelligence. Paris: OECD/Legal Instruments, 2021.

52 ORTEGA, A.: Hacia un régimen europeo de control de la inteligencia artificial. Real Instituto Elcano, ARI 52/2021, [en línea]. http://www.realinstitutoelcano.org/wps/portal/rielcano_es/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/ari52-2021-ortega-hacia-regimen-europeo-de-control-de-inteligencia-artificial [Consulta: 29 junio 2021].

53 UNESCO, *Preliminary Report on the first draft of the Recommendation on the Ethics of Artificial Intelligence*. [en línea] 2021-06-21 <https://unesdoc.unesco.org/ark:/48223/pf0000374266> [Consulta: 29 junio 2021].

54 WORLD HEALTH ORGANIZATION, Ethics and governance of artificial intelligence for health. [en línea] 2021-06-28 <https://www.who.int/publications/i/item/9789240029200> [Consulta: 29 junio 2021].

REFERENCIAS BIBLIOGRÁFICAS

- AMSTRONG, Stuart, *Smarter than us: The rise of machine intelligence*. Machine Intelligence Research Institute, 2014.
- ATLANTIC COUNCIL, *D-10 Strategy Forum*.
- BUGHIN, Jacques; SEONG, Jeongmin; MANYIKA, James y CHUI, Michael. And Joshi, Raoul. *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey Global Institute, 2021.
- BUGHIN, Jacques; SEONG, Jeongmin; MANYIKA, James; HÄMÄLÄINEN, Lari; WINDHAGEN, Eckart y HAZAN, ERIC., MCKinsey Global Institute, February, 2019.
- CAPEK, Karel, *Robots Universales Rossum*. Translate by Clauida Novack-Jones. New York: Penguin Classic, 2004. First edition in Czech as R.U.R. / *Rossumovi Univerzální Roboti* in 1921.
- CONSEJO DE LA UNIÓN EUROPEA, *EU-China Summit: Defending EU interests and values in a complex and vital partnership- Press reléase by President Charles Michel and President Ursula von der Leyen*, 2020.
- CREIGHTON, Jolene, “The “Father of Artificial Intelligence” says singularity is 30 years away”. *Futurism*, February 14, 2018
- DE ASÍS, Rafael, “*Inteligencia artificial y derechos humanos*”. Materiales de filosofía del derecho, N° 2020 / 04. Madrid: Universidad Carlos III de Madrid, 2020.
- DECLARACIÓN CONJUNTA. Cumbre Unión Europea-Estados Unidos “Towards a renewed Transatlantic partnership”. Bruselas: Consejo de la Unión Europea, 15 de junio de 2021.
- EURONEWS, “China rechaza y lamenta las críticas de la cumbre entre Estados Unidos y la UE”, 2021.
- EURONEWS, “Estados Unidos ha vuelto, dice Joe Biden al llegar a Reino Unido para el G7 y una gira europea”, 2021.
- EUROPEAN COMMISSION, *Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. COM (2021) 206 final.
- EUROPEAN COMMISSION, *Coordinated Plan on Artificial Intelligence*. COM (2018) 795 final, Brussels: European Commission.
- EUROPEAN PARLIAMENT, *Artificial Intelligence diplomacy. Artificial Intelligence governance as a new European Union External Policy Tool*. Policy Author: Ulrike Franke Luxembourg: Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies. 2021.
- EUROSTAT, *Artificial Intelligence*, 2021.
- FRANKE, Ulrike, “Artificial Divide. How Europe and America could clash over AI”, London: European Council on Foreign Affairs, January, 2021.
- FRANKE, Ulrike., “Europe needs a plan for AI in the military realm”, *The Security Times*, February 2020.

- GARVIE, Clare; BEDOYA, Álvaro y FRANKLE, Jonathan, *The perpetual Line-Up. Unregulated police face recognition in America*. Washington: Center on Privacy & Technology at the Georgetown University, 2016.
- GOLMAN SACHS EQUITY RESEARCH, *Profiles in Innovation Revisited: AI hardware*, 11 March 2018.
- IMBRIE, Andrew; FEDASIUK, Ryan; AIKEN, Catherine; CHHABRA, Tarun y CHAHAL, Husanjot, *Agile Alliance. How the United States and its allies can deliver a democratic way of AI*. Washington: Center for Security and Emerging Technology at Georgetown University, 2020.
- KONAEV, Margarita y CHAHAL, Husanjot, "The path of Least Resistance, Multinational Collaboration on AI for Military Logistics and Sustainment", Washington: Center for Security and Emerging Technology at Georgetown University, April 2021.
- LARGER, Thibault; SCOTT, Mark y KAYALI, Laura, "Inside the EU's divisions on how to go after Big tech", Brussels: Politico, 14 December 2020.
- LEE, Kai-Fu, *AI superpowers: China, Silicon Valley, and the new world order*. New York: Houghton Mifflin Harcourt Publishing, 2019.
- MCCARTHY, John, "Recursive functions of symbolic expressions and their computation by machine, Part I", Massachusetts: Massachusetts Institute of Technology, 1960.
- MINSKY, Carly, "One former Google exec says there's no hope for Europe's artificial intelligence sector", *Sifted*, [en línea], December, 2018.
- MINSKY, Marvin, "Steps Towards Artificial Intelligence", *Proceedings of the IRE*, Vol. 49. N1, January, 1961.
- MORENO, Luis y PEDREÑO Andrés, *Europa frente a EE.UU y China: prevenir el declive en la era de la inteligencia artificial, Arizona: University of Arizona Egyptian Exped, 2020*.
- OERTEL, Janka, "The new China consensus: How Europe is growing wary of Beijing", London: European Council on Foreign Relation, Policy Brief, September 2020.
- OECD, Recommendation of the Council on Artificial Intelligence. Paris: OECD/Legal Instruments, 2021.
- ORTEGA, A., Hacia un régimen europeo de control de la inteligencia artificial. Real Instituto Elcano, ARI 52/2021.
- ORD, Toby; MERCER, Angus y DANNREUTHER, Sophie, *Future Proof. The opportunity to transform the UK's Resilience to extreme risks*. TheCentre for Long-Term Resilience, 2021.
- OVIDIO, *Metamorfosis*, Madrid: Consejo Superior de Investigaciones Científicas, Vol II, 1994.
- PARLAMENTO EUROPEO, Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas. Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012) (INL)
- PANIER, Alice, "Europe in the Geopolitics of Technology, Connecting the Internal and External Dimensions". Paris: Institut Français des Relations Internationales, April 2021.

- RASO, Filippo; HILLIGOSS, Hannah; KRISHNAMURTHY, Vivek; BAVITZ, Christopher y KIM, Levin., *Artificial Intelligence & Human Rights: Opportunities & Risks*. Cambridge: Berkman Klein Center for Internet & Society at Harvard University, September, 2018.
- REAL ACADEMIA DE CIENCIAS MORALES Y POLÍTICAS, *Anales*. Año LXXI, N 96, Madrid: Real Academia de Ciencias Morales y Políticas, 2019.
- ROSENBLATT, Frank, "The Perceptron: A probabilistic Model for Information storage and organization in the brain" Princeton: *Psychological Review* 65 (6) 1958.
- SCHWAB, Klaus, "The Fourth Industrial Revolution. What it means and how to respond". *Foreign Affairs*, New York: Council on Foreign Relations, December 2015.
- SCHWAB, Klaus, *The Fourth Industrial Revolution*. Geneva: World Economic Forum. 2016.
- TEIGENS, Vasil; SKALFIST, Peter y MIKELSTEN, Daniel. *Inteligencia Artificial: La Cuarta Revolución Industrial*, Cambridge: Cambridge Stanford Books, 2020.
- THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE, 2021.
- TORTOISE, *The Global AI Index*, 2021.
- US NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE, Final Report. March 2021.
- UNESCO, *Preliminary Report on the first draft of the Recommendation on the Ethics of Artificial Intelligence*.
- VON DER LEYEN, Ursula, *A Union that strives for more. My agenda for Europe. Political guidelines for the Next European Commission 2019-2024*. Luxembourg: Publications Office of the European Union, 2019.
- WORLD HEALTH ORGANIZATION, *Ethics and governance of artificial intelligence for health*. Geneva: World Health Organization, 2021.
- ZUBOFF Shoshana, *La era del capitalismo de la vigilancia: la lucha por un futuro humano frente a las nuevas fronteras del poder*. Barcelona: Paidós, 2020.

LA SMARTIFICACIÓN DE LA ADMINISTRACIÓN PÚBLICA: APOSTILLAS CRÍTICAS A PARTIR DEL CASO EUROPEO

*THE SMARTIFICATION OF PUBLIC ADMINISTRATION:
CRITICAL APOSTILLES FROM THE EUROPEAN CASE*

JONATHAN PIEDRA ALEGRÍA

Profesor de Filosofía
Escuela de Filosofía
Universidad Nacional
Heredia/Costa Rica
jonathan.piedra.alegria@una.cr
ORCID: 0000-0003-4532-4415

Recibido: 17/02/2021

Revisado: 21/08/2021

Aceptado: 6/09/2021

Resumen: La Inteligencia Artificial (IA) y las tecnologías derivadas están generando una gran cantidad de cambios en nuestro mundo. Estudios y publicaciones especializados muestran hasta qué punto las tecnologías de Inteligencia Artificial tendrán numerosas implicaciones que cambiarán de manera profunda y disruptiva nuestra sociedad. En el caso concreto de la Unión Europea, en el año 2018, 24 países miembros de la UE (más Noruega) firmaron la *Declaración de Cooperación sobre Inteligencia Artificial* en la cual se comprometían a trabajar juntos con la finalidad de estimular la adopción de la IA en Europa. En este artículo aludimos a algunos ejemplos europeos relacionados con el potencial de la IA en la gestión pública y presentamos argumentos acerca de por qué la IA para uso público es un campo de investigación teórico sin capacidad para ejercer, en la actualidad, el necesario impacto en los futuros diseños de la Administración pública.

Palabras clave: Administración pública, Europa, Inteligencia Artificial, Razonamiento moral.

Abstract: Artificial Intelligence (AI) and derived technologies are generating a great deal of change in our world. Specialized studies and publications show the extent to which Artificial Intelligence technologies will have numerous implications that will profoundly and disruptively change our society. In the specific case of the European Union, in 2018, 24 EU member countries (plus Norway) signed the Declaration of Cooperation on Artificial Intelligence in which they committed to working together to stimulate the adoption of AI in Europe. In this article, we refer to some European examples related to the potential of

AI in public management and present arguments about why AI for public use is a theoretical research field without the capacity to exert, at present, the necessary impact on future designs for public administration

Keywords: Public administration, Europe, Artificial Intelligence, Moral reasoning.

INTRODUCCIÓN

Entre los especialistas de las Ciencias Humanas, Sociales y Jurídicas que investigan en universidades y centros de investigación de Europa y América existe cierto consenso en torno al impacto que ejercerá la Inteligencia Artificial (IA) en el futuro de la burocracia y, en general, en las Administraciones públicas. La IA y otras tecnologías derivadas están generando una gran cantidad de cambios en nuestro mundo, y parece imponerse la previsión según la cual sus implicaciones conducirán a cambios profundos y disruptivos en nuestra sociedad. A pesar del número cada vez mayor de investigaciones relacionadas con la IA, su utilización en el ámbito público es relativamente joven. Tanto así que es común encontrar en la literatura, textos carentes de descripciones adecuadas, así como de sus posibles desafíos. Esta situación se debe, entre otros aspectos, a una ausencia de una definición común u homogénea de la IA y de algunos de sus elementos derivados. Es así, como previamente al desarrollo de los desafíos para su utilización en la Administración pública, es necesario establecer ciertos conceptos sobre la IA. Sobre esto, hay muchas definiciones de IA dependiendo del área o enfoque del cual se trata. Para los ingenieros informáticos, la IA se puede delimitar al desarrollo de programas que muestran o imitan un comportamiento inteligente (Franceschetti, 2018). Esta definición está en consonancia con la propuesta de John McCarthy, un padre fundador de la IA, cuando la define como la ciencia e ingeniería para hacer máquinas y programas inteligentes (McCarthy, 2018). Entendiendo la inteligencia como la capacidad de conseguir un objetivo. Este tipo de definiciones muestran indirectamente que la IA es más un campo de estudio que algo concreto que pueda describirse de manera sencilla. Incluso, es posible dividirla en varios campos como la robótica o el Big Data, por mencionar algunos casos. Verbigracia, la robótica es un campo interdisciplinario de la ingeniería que se ocupa se ocupa de la construcción y operación de los robots físicos (Franceschetti, 2018). A pesar de la falta de acuerdo universalmente aceptado sobre cómo definir un robot, se puede decir *grosso modo* que es una máquina programable con la capacidad de efectuar ciertas acciones de manera semiautónoma o autónoma. En realidad, la IA y la robótica son inicialmente dos campos diferentes, pero se entrelazan cuando ejemplo, se utiliza la técnica de *Machine Learning* para entrenar a

un robot para que ejecute una tarea. La integración de las tecnologías de la IA y la robótica se ha convertido en un tema de creciente interés para ambos campos.

Algo similar ocurre con el Big Data. Este término es utilizado para indicar una gran cantidad de información variada que llega en volúmenes crecientes y con una velocidad cada vez mayor. Estas características hacen que sea imposible para un ser humano realizar un análisis adecuado. No obstante, usando varias técnicas de IA es posible procesar todos esos datos con tal obtener información valiosa. Como se puede observar, existen relaciones significativas entre todos estos campos, aun cuando pueden ser diferenciables. Sin embargo, generalmente, todos coinciden en la capacidad de los robots, máquinas o sistemas para realizar tareas que tradicionalmente han sido ejecutadas por seres humanos, ya que la IA combina diversas habilidades tales como el denominado aprendizaje autónomo, la capacidad de razonar, la resolución de problemas, e incluso la percepción y el reconocimiento (y uso) del lenguaje natural. Su potencial disruptivo se presenta en sectores tan importantes para el diseño de políticas públicas como el energético, el educativo, el relativo a las ciencias de la salud, el transporte e incluso la administración de justicia.

Según Accenture (Mehr, 2017) para el año 2035 las aplicaciones de IA tendrán el potencial de duplicar el crecimiento económico global. En la actualidad, asistimos a una dura competencia para ver quién se convertirá en el líder planetario, es decir, quien será la superpotencia en este tema (Lee, 2020). China pretende invertir cerca de 147.800 millones de dólares americanos para convertirse en 2030 en el mayor innovador global en el campo de la IA (*State Council of the People's Republic of China*, 2017) y los Estados Unidos de América invirtieron en 2016 más de mil millones de dólares americanos en investigación y desarrollo en tecnologías relacionadas con la IA. Europa no se queda atrás, ya que en el año 2019 el gasto relacionado con tecnologías basadas en IA fue de cinco mil millones de dólares americanos, un 49 por ciento más que en el año 2018 (IDC, 2019). La competencia por el liderazgo a escala planetaria tiene un fuerte componente de promesa económica ya que, al cabo, la mayoría de los gobiernos reconocen abiertamente que dicha carrera se basa, sobre todo, en la percepción y el cálculo económico del potencial previamente asignado a la IA. En 2018 los veinticuatro países miembros de la Unión Europea (más Noruega) firmaron la *Declaración de Cooperación sobre Inteligencia Artificial* en la cual estos se comprometían a trabajar juntos con la finalidad de estimular la capacidad y adopción de la IA en Europa. Aunque el documento hace énfasis en los desafíos socioeconómicos y busca garantizar un marco ético y legal para poder ejercer un control democrático sobre la carrera global por el liderazgo en IA, en este artículo enumeraremos algunas de las dificultades y retos sociopolíticos que han destacado investigadores del área de las ciencias humanas, sociales y jurídicas con el objetivo

de dirimir si la sociedad civil europea está en disposición de abordar este debate en los próximos años y cuáles son los factores que incrementarán o, por el contrario, mermará la capacidad reflexiva de la ciudadanía digital.

1. LA AUTOMATIZACIÓN DE LA BUROCRACIA EUROPEA

Los avances tecnológicos basados en IA y el potencial económico asociado están ganando protagonismo en el contexto del sector público. Por ejemplo, Deloitte (2017) estimó que solo en Estados Unidos la automatización de las tareas de los empleados del Gobierno Federal podría llegar a suponer un ahorro de entre 96,7 millones y 1,2 mil millones de horas al año, lo que equivale a 3,3 mil millones y 41,1 mil millones de dólares americanos respectivamente. Si tenemos presente que las organizaciones y órganos del sector público constituyen los empleadores más importantes en la mayoría de los países (Organización Internacional del Trabajo), el ahorro pecuniario para la Administración pública, derivado únicamente de la automatización de algunas funciones (i.e. de carácter repetitivo), sería enorme. Pero el potencial en el sector público-administrativo no podría resumirse únicamente a esto, ya que además permitiría brindar una mejor calidad en los servicios públicos si se siguieran los pasos para dar lugar a lo que se ha denominado la *smartificación de la Administración pública* (Ramió, 2019). Esta expresión suele emplearse para describir el diseño de procesos encaminados a aumentar la inteligencia estratégica de las instituciones en el gobierno de redes públicas y privadas. Pero en los debates en torno a la *smartificación* de las administraciones públicas muchas veces se olvida que uno de los presupuestos objeto de deliberación es el de si, con estos procesos, se conseguirá aportar valor público a las actividades lideradas por administraciones capaces de interactuar con un modelo determinado de ciudadanía digital. Está fuera de toda duda que la implementación de la IA en el sector público ejercerá un impacto significativo sobre nuestras sociedades. Parece evidente que dichas implementaciones influirán sobre el diseño y la consecución de objetivos y políticas públicas. Tanto así, que no es extraño encontrar la introducción de este tipo de tecnología como una forma de garantizar la eficiencia en los procesos gubernamentales, logrando así una prestación efectiva de los servicios públicos en su totalidad. (Doorsamy *et al*, 2020). Sin embargo, todos los pronósticos tienen en común el hecho de que esto se realiza sin precisar qué modelo de ciudadanía digital se propugna y cuáles pueden ser las consecuencias sociopolíticas y económicas a escala global (en este caso europea) derivadas de las innovaciones disruptivas patrocinadas. No menos problemática es la tarea de dirimir si debiéramos calificar como positivos o negativos en relación con los intereses de la futura ciudadanía digital un escenario en el que

en el cual la IA (i.e *Machine learning*) deba manejar una cantidad masiva de datos (*Big Data*) porque o bien no existan suficientes expertos capaces de analizarlos o bien no se confíe en sus decisiones.

En principio, para la mayoría de las personas podría parecer evidente cuán positivo es el hecho de que una tecnología basada en IA pueda automatizar procesos burocráticos rutinarios en una escala de tiempo a la que no tiene acceso el ser humano en la consecución de tareas cognitivas equivalentes. Sea o no favorable la opinión de las personas ante el escenario de la automatización, lo cierto es que el primer fenómeno objeto de atención para el sentido común es la extrema reducción de cargas administrativas de la que quedarían liberados los funcionarios públicos junto a la opinión extendida de que existiría una correlación entre este factor y la obtención de beneficios en la resolución de la asignación de recursos para asumir tareas de elevada complejidad. ¿A qué se debe la tendencia de las personas a asumir que existe una correlación entre la automatización de procesos burocráticos, la liberación de tareas rutinarias entre el personal administrativo y la consiguiente obtención de beneficios derivados de este modelo de gestión? En el desarrollo y aplicación de la IA a las administraciones públicas no se encontrará la respuesta, ya que esta percepción se basa en el impacto ejercido por la cultura política en la que se hallan los ciudadanos, y en la asunción de un modelo de teoría de elección racional que no da cuenta de los efectos generados por la automatización de las tareas, ni de la digitalización de los procesos.

1.1. GESTIÓN 4.0

De hecho, si tenemos presente la velocidad con que se suceden los cambios tecnológicos resulta prácticamente inevitable que la Administración pública innove hasta tal punto en sus métodos tradicionales como para lograr no solo una mayor eficiencia en sus servicios sino, además, una mayor participación ciudadana, una mejor rendición de cuentas y un nivel de interoperabilidad constante entre los distintos niveles de gobernanza. Aun así, existe un país europeo que ha seguido esta vía, Estonia, considerado el país europeo número uno en ciberseguridad (Balví, 2017). El 98 por ciento de los ciudadanos poseen identidades digitales (ID Digital) y el 46,7 por ciento emite el voto *online*. En 1997 el país adoptó un modelo de digitalización que ha tenido como consecuencia que hoy el 99 por ciento de los servicios públicos estén digitalizados lo que haría de Estonia «el país más digital» del mundo. A este fenómeno de digitalización se le asocia el hecho de que en Estonia se ahorre un equivalente al dos por ciento del PIB en salarios. Incluso, en la página web oficial destinada a promocionarse, el país se presenta como «the world's most advanced digital society.»

El alto nivel de digitalización de Estonia fue determinante para que, en mayo del 2019, se iniciara un proceso de implementación de 16 soluciones de inteligencia artificial (o *kratts*) en el sector público (*Report of Estonia's AI Taskforce*, 2019). Para la mitología de Estonia, un Kratt es un ser mágico animado por un alma humana, compuesto de varios objetos (usualmente paja y elementos domésticos). En lo esencial, un Kratt era un sirviente diseñado para ayudar a los seres humanos. Es por esto por lo que la figura del Kratt en Estonia (no es una denominación que se haya internacionalizado) es utilizada como una analogía de la IA, así como de todas sus complejas relaciones con los seres humanos. En este contexto, un kratt es una aplicación práctica basada en IA completamente automática en ocasiones simplemente *software* y en otras instalada en *hardware* tales como vehículos autónomos, drones, robots etc. Estas nuevas tecnologías son parte de la estrategia nacional de IA de Estonia y del desarrollo de lo que se ha denominado E-State motivo por el cual nos encontramos ante un claro caso de una democracia robotizada (Moreno & Jimenez, 2018).

Otros países de la Unión Europea han tomado una posición menos activa, aunque siempre en la línea de la implementación de la IA. Italia, por ejemplo, ha desarrollado una estrategia digital que se enmarca en el contexto de la *Agencia Digital Europea*. El país entiende la transformación digital de la Administración pública como una búsqueda del crecimiento económico, promoción del empleo, calidad de vida y mayor participación democrática. En 2017 se aprobó un plan que buscaba el desarrollo de «cuatro pilares»: (1) ecosistemas digitales o áreas de política (e.g. ciudadanía digital, transporte, turismo, educación.); (2) infraestructuras inmateriales (e.g. manejo de datos de la Administración pública como datos de los bancos, sistema de pagos); (3) infraestructuras físicas (e.g. centros de datos, cloud storage, etc.); y (4) ciberseguridad. Esta búsqueda de la transformación digital del país ha dado lugar a diferentes proyectos de política pública tales como: *Electronic Health Record* (un expediente o registro médico digital), *E-billing* (sistema de facturación electrónica de uso obligatorio para la administración pública), *PagoAP* (plataforma para que ciudadanos y/o empresas pueden pagar a la Administración) y el *Equipo de Respuesta a Emergencias Informáticas de la Administración Pública* (CERT-PA) (una estructura que opera dentro de la *Agencia per l'Italia digitale* y es responsable del procesamiento de los problemas de seguridad en las administraciones públicas). Estas nuevas tecnologías basadas en IA lo que buscan es mejorar la calidad de vida de los ciudadanos italianos, permitiendo un mayor grado de accesibilidad a los servicios públicos, reduciendo el coste de determinados servicios públicos, así como del gasto social, un aspecto, este último que suele relacionarse con la supuesta reasignación de recursos. Estas políticas públicas a menudo se presentan bajo el lema de que la automatización y

gestión basada en IA ofrecerá a la ciudadanía la oportunidad de interactuar con el Estado de manera más rápida y personalizada.

El Gobierno de Suecia y el Nordic Council of Ministers emitieron en 2018 el documento *AI in the Nordic-Baltic región* en el que, en conformidad con el conjunto de países de la Unión Europea, se afirma que la aplicación de la IA en el diseño e implementación de políticas públicas conducirá a la resolución de muchos problemas sociales. Una de las principales diferencias de este texto con respecto a la *Declaración de Cooperación sobre Inteligencia Artificial* estriba en resaltar que los países nórdicos y la región báltica poseen una «alta madurez digital», así como la necesidad de evitar una regulación innecesaria que impediría precisamente las bondades de su desarrollo. Suele insistirse en que esta declaración es una hoja de ruta bastante general, que no pormenoriza ningún tema; no obstante, es importante comentar en detalle algunos de sus puntos. Los países nórdicos se posicionan como líderes digitales («digital front-runners») en el contexto mundial (no únicamente en el europeo), prueba de ello es la transformación digital operada en el sector empresarial nórdico, principalmente en las PYMES. Pero es importante resaltar el hecho de que a la hora de concebir y defender la digitalización de los servicios administrativos se asumió un enfoque comprometido con mejorar la vida de los residentes y no en la reducción de costes (Randall & Belina, 2019). La infraestructura digital ha sido en este país más relevante a nivel regional que estatal con el objetivo de evitar la brecha tecnológica e impedir disparidades entre el ámbito urbano y el rural. Prueba de ello es que se desarrollara una idea de «Smart cities» muy particular, según la cual lo «Smart» de una ciudad no provenía de la tecnología en sí misma, sino de las redes de cooperación y la capacidad de resolución de conflictos que abría la implementación tecnológica. Consiguientemente, la digitalización era percibida más un proceso humano que técnico (Randall & Belina, 2019), que podría mejorar la interacción y la comunicación entre los ciudadanos y la Administración pública.

En Noruega las municipalidades son organismos semiautónomos que se encuentran en el nivel administrativo noruego más elemental. Hay 356 municipalidades cuyo tamaño varía. Van desde las más pequeñas como Utsira que tiene 212 habitantes hasta el caso de Oslo con 647 676 habitantes. Las municipalidades asumen responsabilidades tales como la gestión y mantenimiento de los sistemas educativos o de las vías públicas locales, entre otras prestaciones sociales básicas. Servicios como sensores automatizados de tráfico, tecnologías para el bienestar o intercambios digitales de información, han ido tomando una relevancia mayor en las municipalidades. En el año 2018 se implementaron tres nuevas herramientas digitales y de automatización en los municipios bajo las siguientes denominaciones: *DigiSos*, (plataforma digital para solicitar asistencia social financiera), *DigiHelse* (sistema para la visualización de acuerdos alcanzados

y establecer diálogos con la Administración) y *Minside* (servicio de información sobre temas relativos a la municipalidad). Este tipo de digitalización es lo que ellos mismos llaman «alta madurez digital» algo que, a la postre, permitirá dar a Noruega el salto a las tecnologías de IA en la Administración pública. La existencia de todas estas iniciativas nos permite analizar resultados más concretos sobre la implementación actual de la IA en el servicio público noruego. Según Mikalef, Fjørtoft & Torvatn (2019) las tecnologías que más se han implementado en las municipalidades noruegas están relacionadas con la comunicación y la interacción con los ciudadanos. Un 28,9 por ciento utiliza agentes de interacción inteligente (i.e. *chatbots* y otros tipos de *Intelligent Conversational Agents*), un 21,1 por ciento ha puesto en funcionamiento traducciones en tiempo real (i.e. traducción voz a voz o voz a texto) y un 15,7 por ciento maneja de manera automatizada las solicitudes y el manejo de datos. Si bien no es un número muy alto, la situación global muestra un desarrollo continuando en la incorporación de la IA al servicio público y, en particular, en áreas como la salud, la ciberseguridad, la automoción autónoma, la investigación financiera y, en especial, la detección de fraudes, y, sin duda alguna, la automatización de todos los procesos burocráticos si bien, por encima de cualquiera de estos ámbitos, se mantiene un constante interés por la «interacción inteligente con la ciudadanía.»

Todos estos casos parecen atestiguar el potencial de la IA en la gestión pública, ya sea desde una perspectiva meramente económica —de eficacia en la gestión— como desde la perspectiva de las políticas de bienestar. Sin embargo, no todo es tan simple como parece. Aunque el uso de tecnologías inteligentes puede actuar como facilitador de la innovación, la sostenibilidad, la competitividad y la consulta ciudadana, lo cierto es que estas dimensiones se plantean por lo general asumiendo enfoques paradójicamente indeterminados, en los que no se problematizan los conceptos y concepciones de ciudadanía y estado digitales que se asumen incluso en los casos en los que existen rutas de acción concretas. La IA parece ser más una ideología que un aspecto tecnológico. En medio de este enamoramiento obsesivo por lo tecnología —*technological infatuation*— (Gauld, & Goldfinch. 2006) es a partir de donde se plantea una cultura de renovación administrativa por medio de IA en la gestión pública desde una perspectiva excesivamente optimista, que parece reducir los problemas de gobernanza a problemas técnicos.

En el contexto descrito, observamos cómo se promueve la implementación de la IA en el sector público sin antes mencionar cuál es el modelo de Administración pública subyacente, lo cual, en el mejor de los casos, impide visualizar correctamente las implicaciones del uso de la IA en la esfera de la gobernanza. En los documentos oficiales que hemos citado solo se alude a los aspectos tecnológicos de las aplicaciones de inteligencia artificial, lo que suena bastante vago sin un

modelo de Administración pública sólido que retrate las implicaciones para la gobernanza del estado administrativo. Por ello, algunos autores se preguntan si esta nueva cultura tecno-administrativa implica realmente, acaso un modelo de organización holocrático, es decir, un sistema más horizontal basado en la autogestión (Ramió, 2019). La respuesta a ese dilema no queda clara, precisamente porque se necesitan hacer explícitos y justificar desde un punto de vista ético-jurídico no solo conceptos y concepciones de la robotización de las administraciones públicas, también es preciso hacer explícito cuáles son las aspiraciones de bienestar humano y los compromisos asumidos con la sostenibilidad y el entorno ecosocial. En ausencia de todos estos debates, que exigen un tiempo de madurez no acelerado además del diseño de espacios para la deliberación pública, las políticas emprendidas por la Unión Europea y otros países han sido calificadas por algunos expertos como propias de democracias robotizadas. (Moreno y Jiménez, 2018)

En un marco como este, las vulnerabilidades que pueden crear una falla crítica o un grave problema son temas de gran relevancia. Ya en el 2014 un *Informe independiente sobre el voto electrónico en Estonia* (<https://estoniaevoting.org/>) mostraba serias preocupaciones sobre las garantías democráticas inherentes a esta forma digital de sufragio. Según este estudio, existirían graves brechas en la seguridad operativa y de procedimiento, junto con la arquitectura del sistema que expone el sistema de votación a ataques cibernéticos que podrían alterar el signo de los votos o incluso derivar en una situación de disputa los resultados de las elecciones. Problemas que eran bien conocidos desde hace años, debido a la prohibición del voto electrónico en los Países Bajos y Alemania. En el 2008, la votación electrónica fue suspendida en los Países Bajos. Todo comenzó cuando un grupo activista llamado: «No confiamos en las computadoras de votación» (*Wij vertrouwen stemcomputers niet*) demostró que los sistemas que se utilizaban en este tipo de votación podrían, en determinadas circunstancias, poner en peligro el secreto del voto. El primer problema que señalaron fue la facilidad para reemplazar los chips de programación, lo que permitía que un posible atacante cibernético tuviera acceso a la cuenta de la máquina pudiendo ejecutar cualquier tarea que quisiera. Lo más preocupante es que este tipo de ataques podían pasar desapercibidos debido a la falta de verificabilidad. Esto, generó una alerta adicional, en vista, de que el software utilizado no era de libre acceso (ni tampoco se liberó el código cuando se le solicitó a la empresa) y, por lo tanto, no era posible auditarlo. Para este grupo era claro, que cuando un proceso no puede ser verificado adecuadamente debido a las características de la tecnología utilizada (como la IA) no se debería implementar en áreas sensibles como los sistemas críticos para la seguridad (Yeung *et al*, 2020), o en las elecciones. Otro inconveniente importante que descubrieron es que era posible espiar a la máquina de votación por medio de un complejo ataque conocido como *Tempest* (Jacobs y Pieters, 2009).

Estos aspectos, junto a los problemas de seguridad en los almacenes donde se resguardaban las máquinas de votación, la imposibilidad de realizar un auditaje de manera efectiva porque no hay registro en papel, así como la dependencia del gobierno con los proveedores privados (empresa NEDAP) de las computadoras de votación (Oostveen, 2010), terminaron con 20 años del uso del voto electrónico en los Países Bajos.

Este escenario también tuvo resonancia en Alemania. Un grupo de hackers conocido como «Chaos Computer Club» (que trabajaba en colaboración con los activistas de los Países Bajos) solicitó la prohibición de la utilización de las máquinas de votación alemanas (CCC, 2006), ya que eran de la misma empresa proveedora (NEDAP) de su vecino país y por lo tanto corrían los mismos riesgos que los neerlandeses. Con esta finalidad, presentaron una demanda colectiva en la que solicitaban la suspensión del voto electrónico hasta que el gobierno pudiera demostrar la transparencia y verificabilidad de este tipo de sistemas. En la votación electrónica no es posible determinar cómo se procesan los votos dentro la computadora, por lo que el proceso se vuelve bastante opaco, reduciendo, en el mejor escenario posible, la capacidad de escrutinio y supervisión a la *expertise* de un grupo muy pequeño de técnicos especialistas. Como respuesta a esto, en el año 2009 la Corte Constitucional Federal Alemana (*Bundesverfassungsgericht*) declaró inconstitucional la votación electrónica debido a la falta de transparencia y verificabilidad de los ciudadanos, puesto que los votantes no tenían forma de saber que pasó realmente con su voto (se ocupaba conocimiento especializado) y por el contrario se les pedía una confianza ciega en las tecnologías de votación electrónica. Estas situaciones vulneraban el principio del carácter público de una elección, ya que no era posible garantizar de manera confiable los procesos de verificación en el conteo (o recuento de los votos), por lo que en este caso no se satisfacían las exigencias constitucionales fundamentales en un Estado de Derecho.

A pesar de que existe una amplia literatura sobre todos estos elementos, los entusiastas de este tipo de tecnologías suelen minimizarlos y con frecuencia se presentan estos problemas como si fueran exclusivamente técnicos cuando lo cierto es que afectan a la dimensión procedimental y por tanto a la legitimidad de democrática de los comicios.

Pero la votación es solo un tema más dentro de muchos otros, cuando hablamos sobre los riesgos. Tomemos, por ejemplo, las identidades digitales (ID Digital) en Estonia. Esta identidad digital es según Margit Ratnik (directora de la Oficina de Identidad y Estado de la Junta de Policía y Guardia Fronteriza) la «piedra angular de nuestra sociedad digital». Según datos oficiales (Politsei.ee, 2017) esta ID Digital es la base fundamental para el acceso a 1500 e-servicios públicos (i.e registros médicos o el sistema de votación), así como a 5000 servicios privados. Por números como estos es que los entusiastas de la automatización o digitalización de

la Administración pública consideran como un caso emblemático la situación de este país báltico. Sin embargo, ser «la sociedad digital más avanzada del mundo», también la convierte en la más dependiente tecnológicamente y vulnerable a ciberataques del mundo. En el año 2017 se encontró una grave vulnerabilidad en el cifrado de las Electronic ID Cards que afectó a 760000 personas (BBC, 2017), es decir a más de la mitad de la población del país. Este riesgo potencial permitía que la información pudiera ser robada o que un ciber atacante pudiera hacerse pasar por esa persona suplantando su identidad. El riesgo afectó a los chips utilizados en tarjetas de identificación, permisos de residencia e identificaciones digitales. Frente a esto, el gobierno estonio bloqueó estas ID cards, impendiéndole a miles de ciudadanos el acceso a los servicios gubernamentales. Como medida correctiva, los ciudadanos afectados debían actualizar de forma remota (o en su defecto en las oficinas policiales) el software de sus certificados para seguir utilizando los servicios electrónicos.

Algunas personas pasaron días (BBC, 2017) sin poder realizar la actualización electrónica y otras 18000 personas todo no pudieron realizar este cambio debido a problemas técnicos (News.err, 2017). Si bien algunos días, no parecen ser mucho tiempo, ciertamente lo son cuando este tipo de identificación es «la piedra angular de la sociedad». Durante este tiempo, estas personas no tenían participación real en el *E-State*. Todos sus derechos políticos y sociales se vieron comprometidos por una vulnerabilidad en los procesos de identificación. Aún cuando los elementos básicos de la interacción estatal se mantuvieron (en el mundo analógico) esta situación puso en entredicho la esfera de los derechos de más de la mitad ciudadanos estonios, de una manera inusitada en tiempos modernos. Aspectos como los previos, deberían ser suficientes para analizar con más cautela las supuestas ventajas de este tipo de procesos. En vista de que las características tecnológicas influyen no solo el proyecto mismo de automatización los servicios, sino en el concepto mismo de ciudadanía. Como muestran los casos que hemos citado, el uso de estas tecnologías puede generar una transformación radical en los procesos democráticos o incluso en las nociones tradicionales de participación ciudadana.

2. ¿UNA ADMINISTRACIÓN “INTELIGENTE”?

Si en la literatura en ciencia política es bien conocido el proceso de judicialización de la política, nos encontraríamos ante una amenaza distinta consistente en la nueva suplantación de la política por medio de su robotización. La automatización de nuestras sociedades no es un tema de ciencia ficción, ocurre actualmente y frente a la promesa económica de la IA, es muy probable que transfigure los

espacios políticos tradicionales, con el consiguiente impacto social y cultural, que genera dilemas éticos y jurídicos que van más allá de las descripciones tecnocráticas en torno a la mera implementación de la IA. Tomemos como ejemplo la «sociedad digital más avanzada del mundo» (Estonia) que se presenta a sí misma como una utopía tecnológica (al menos en cuanto a la gestión pública). Cuando nace un bebé este se convierte instantáneamente en un e-ciudadano. En el país en el que —en palabras Marten Kaevats, asesor nacional digital del Gobierno de Estonia— lo «digital es un estilo de vida» todos los ciudadanos se han convertido en «tecnociudadanos» (Echevarría & Almendros, 2019), seres que no forman parte de una nación, sino de un *E-State*. Su identidad no depende de alguna delimitación corporal o perspectiva subjetiva, sino que existe únicamente en virtud de procesos de identificación tecnológicos (ID Digital)¹. Los derechos políticos y sociales se reducen a la existencia de un usuario y a la interacción con un tecno-entorno despersonalizado que se ha convertido en el nuevo espacio público. Ahora el ágora es «Smart» y más eficiente. Si bien las formas constitutivas de interacción estatal siguen existiendo, ahora se trasladan a espacios más «inteligentes».

1 Una alternativa (al menos teóricamente), pero no exenta de problemas, podría ser la utilización de un Gemelo Digital para relacionarnos con la Administración pública. Ya que esto permitiría mantener la identidad física junto con la digital en nuestras interacciones con la Administración, sin llegar a transformarnos en meros usuarios tecno-despersonalizados, como parece que sucede en el caso anteriormente señalado. Un gemelo digital (*Digital Twin*) es una representación dinámica y actualizada de un objeto o sistema físico. Es básicamente un modelo virtual que interactúa con el sistema físico a lo largo de su ciclo de vida. (Grieves and Vickers, 2017). Por lo general, se utiliza en el contexto empresarial para realizar predicciones sobre el futuro de un sistema que se está probando. También se utiliza para valorar la interacción de un objeto o producto real en un ecosistema determinado con tal de optimizar o validar su funcionamiento (Semeraro *et al*, 2021). Sus principales usos se encuentran en la fabricación y manufactura, la gestión del ciclo de vida de un producto y en todo lo que se ha llamado como la industria 4.0. Sin embargo, su utilización para seres humanos es prácticamente inexistente hoy en día, en vista de que somos organismos vivos muy complejos (en los cuales el ciclo de vida completo es muy difícil de analizar), a diferencia de una entidad física usualmente inanimada. A pesar de esto, se ha planteado la posibilidad teórica de un «Human Digital Twin» (Shengli 2021) en ámbitos médicos-asistenciales con tal de resolver ciertas necesidades que los servicios médicos no pueden satisfacer. Este «Gemelo Digital Humano», sería una copia online de una persona real en nuestro mundo físico. Este «Gemelo» es principalmente una base de datos digital en tiempo real de una persona en una computadora o en el ciberespacio. Por lo tanto, cuando la información de esa persona cambia (esto depende de la manera en cómo se registren y sincronicen los datos), también se modificarían los datos registrados. Además, el «Gemelo» analizaría todos los datos utilizando diversas tecnologías (i.e Deep Learning) generando recomendaciones, predicciones o incluso diagnósticos. Sin embargo, a pesar de sus posibles beneficios teóricos, en la práctica su implementación en la Administración pública (o en cualquier otra área) es, como mínimo, una tarea que implicaría una cantidad descomunal de esfuerzo y trabajo especializado (sin contar la capacidad computacional) que presentaría muchos dilemas sobre todo en cuanto al tema de manejo y uso de datos personales, así como aspectos relacionados con la privacidad. Por lo que se hace necesario mayores estudios e investigaciones empíricas con tal concluir su factibilidad, no solo teórica, sino técnico-práctica.

Por consiguiente, se ha producido una transformación tecnológica de la ciudadanía que genera nuevos desafíos sociopolíticos y ético-jurídicos a la Administración pública. Esta situación tiene un impacto sobre la esfera de los derechos humanos la cual habría de obligar a los operadores políticos en general a analizar con más detenimiento las supuestas ventajas inherentes de la digitalización y robotización de los procedimientos democráticos y administrativos pues no es lo mismo garantizar el derecho a acceder a los servicios de la Administración pública por medio de una herramienta digital que tener la obligación de hacerlo únicamente por esa vía. Además, debido a su casi total grado de digitalización ya no existe una verdadera interacción humano-Administración pública. Según Sheryl Brahmam (2015), profesora en la Universidad Estatal de Missouri entre el 10 por ciento y el 50 por ciento de nuestras interacciones con agentes conversacionales (i.e. *chatbots*) son abusivas. ¿Qué quiere decir esto? Básicamente, que nuestra interacción con estos agentes no es precisamente placentera, en cierto modo debido a que, en la comunicación humana, no se pueden eliminar las expectativas y la necesidad de obtener explicaciones complejas (que nos llevan a realizar consultas), un aspecto humano determinante que debería tener presente la Administración en el diseño de su interacción con la ciudadanía. En este sentido, la automatización de respuestas puede generar experiencias frustrantes entre los usuarios de los servicios públicos. Si ya es difícil lidiar con un ser humano cuando solicitamos algo, la frustración rápidamente puede ir en aumento cuando tratamos con agentes que aparentan serlo. Se une a todo ello, el problema del diseño de algoritmos sesgados, que reproducen visiones estereotipadas desde el punto de vista de la etnia, el género, la clase social, etc., un tema que supo poner de manifiesto en una entrevista Brahmam (2015) cuando afirmó que es mucho más común encontrar comentarios sexistas hacia *chatbots* que se presentan como «mujeres» que entre quienes asumen rasgos «masculinos». ¿Debe este fenómeno obligar a la Administración a repensar el diseño de sus *chatbots*? ¿Es la Administración responsable de los efectos que determinados diseños tecnológicos tienen sobre los niveles de agresividad exhibidos por los usuarios? ¿Tienen esos niveles un impacto posterior sobre las relaciones interpersonales? ¿Debería clasificarse y reconocerse conforme a indicadores convalidados a nivel europeo el grado de agresividad generado en la interacción humano-máquina?

De momento, estas cuestiones no parecen ser problemas relevantes en los proyectos de *smartificar* la Administración pública.

CONCLUSIONES

A pesar del potencial de la IA para el sector público y del número creciente de países que apuestan por esta vía para la modernización de la Administración, la IA para uso público es un campo de investigación en el que escasean los diagnósticos exhaustivos acerca de los desafíos ético-jurídicos asociados a la aplicación de estas sofisticadas tecnologías. En estas páginas se sostiene que este déficit prospectivo está estrechamente relacionado con la falta de análisis conceptual robusto y una comprensión política común sobre la IA. Asimismo, el marcado optimismo tecnológico y la poca claridad política e institucional con la que se desarrollan y aplican estas tecnologías, hace más difícil que tenga lugar un acercamiento reflexivo en torno a cuál debería ser el modelo de aplicación de la IA en la Administración pública de los estados europeos. Contrasta con esta situación de indefinición en torno a cuáles habrían de ser los fundamentos ético-jurídicos y sociopolíticos que habrían de guiar a los distintos estados europeos, el hecho de que la aplicación de la IA en la gestión pública parece tener el poder de marcar un rumbo a los problemas de gobernanza, así como los relacionados con la ciudadanía digital.

*Agradezco todos los comentarios y aportes para la realización de este texto, realizados por la profesora María González Navarro de la Universidad de Salamanca. Un especial agradecimiento para los revisores anónimos que con sus comentarios mejoraron la versión definitiva de este documento.

REFERENCIAS BIBLIOGRÁFICAS

- BALVÍ, Muriel, *Los 7 secretos del país más digital del mundo*. Infobae. 27/11/2017. <https://www.infobae.com/tendencias/innovacion/2017/11/25/los-7-secretos-del-pais-mas-digital-del-mundo/>. [Consulta: 20/08/2020].
- BBC. Security flaw forces Estonia ID 'lockdown'. 2017 <https://www.bbc.com/news/technology-41858583> [Consulta : 16/08/2021]
- BRAHNAM, Sheryl, "Episode #121" in *Techemergence*. Nov 1, 2015. Podcast, 0h32m27s, <https://emerj.com/ai-podcast-interviews/calling-siri-names-youre-not-alone-a-closer-look-at-misuse-of-ai-agents/>
- CCC | CHAOS COMPUTER CLUB. Chaos Computer Club fordert Vervot von Wahlcomputern in Deutschland,". 2006. <https://www.ccc.de/updates/2006/wahlcomputer> [Consulta: 01/08/2021]
- DOORSAMY, Wesley y BABU SENA PAUL, Tshilidzi Marwala, *The disruptive fourth industrial revolution : technology, society and beyond*. Cham, Switzerland: Springer, 2020.
- ECHEVARRÍA, Javier y LOLA, S. Almendros, *Tecnopersonas. Cómo las tecnologías nos transforman*. Gijón: Ediciones Trea S.L., 2020.

- EGGERS, William, VIECHNICKI, Peter y SCHATSKY, David, *AI-augmented government*. United States. Deloitte, 2017.
- FRANCESCHETTI, Donald, *Principles of robotics & artificial intelligence*. H. W. Wilson Publishing, 2018.
- GOVERNMENT OFFICE AND MINISTRY OF ECONOMIC AFFAIRS AND COMMUNICATIONS. *Report of Estonia's AI Taskforce*. Estonia, 2019.
- GRIEVES, Michael y VICKERS, John, "Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems". In: *Transdisciplinary Perspectives on Complex Systems*. Springer, 2017: 85–113.
- IDC, *Automation and Customer Experience Needs Will Drive AI Investment to \$5 Billion by 2019 Across European Industries*. IDC. 01/04/2019. <https://www.idc.com/getdoc.jsp?containerId=prEMEA44978619> [Consulta: 20/08/2020].
- JACOBS, Bart y WOLTER, Pieters, "Electronic Voting in the Netherlands: From Early Adoption to Early Abolishment". En: Alessandro Aldini, Gilles Barthe y Roberto Gorrieri (eds.), *Foundations of Security Analysis and Design V*, t. 5705. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science). 2009: 21–144.
- LEE, Kai-Fu, *Superpotencias de la inteligencia artificial*. Barcelona: Editorial Planeta, 2020.
- MCCARTHY, John, "What Is AI? / Basic Questions". 2018. <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>. [Consulta: 01/08/2021]
- MEHR, Hila, *Artificial Intelligence for Citizen Services and Government*. United States: Harvard Ash Center Technology & Democracy Fellow, 2016.
- MIKALEF, Patrick, OLSEN FJØRTOFT, Siw y Y. TORVATN, Hans, "Artificial Intelligence in the Public Sector: A Study of Challenges and Opportunities for Norwegian Municipalities". In *Digital Transformation for a Sustainable Society in the 21st Century*. Lecture Notes in Computer Science, edited By Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., Krogstie, J., & Mäntymäki, M. Norway: Springer, 2019: 267-277.
- MORENO, Luis y JIMÉNEZ, Raúl, *Democracias Robotizadas. Escenarios futuros en Estados Unidos y la Unión Europea*. Madrid: Los libros de la Catarata, 2018.
- NEWS. EE., 35,000 people have renewed ID card certificates. 2017. <https://news.err.ee/640156/35-000-people-have-renewed-id-card-certificates> [Consulta : 16/08/2021]
- NORDIC COUNCIL OF MINISTERS, AI in the Nordic-Baltic region. Sweden. 2018
- OOSTVEEN, Anne-Marie, "Outsourcing Democracy: Losing Control of e-Voting in the Netherlands." En: *Policy & Internet* 2 (4), 2010: 196–215. DOI: 10.2202/1944-2866.1065.
- POLITSEI. EE., *Estonia resolves its ID-card crisis*. 2017. <https://www2.politsei.ee/en/uudised/uudis.dot?id=801245> [Consulta: 16/08/2021]
- RAMIÓ, Carles, *Inteligencia Artificial y Administración Pública. Robots y Humanos compartiendo el servicio público*. Madrid: Los libros de la Catarata, 2019.
- RANDALL, Linda y BERLINA, Anna, *Governing The Digital Transition in Nordic Regions: The Human Element*. Sweden: Nordregio Report, 2019.

- SEMERARO, Concetta, LEZOCHE Mario, PANETTO, Hervé y DASSISTI, Michele, “Digital twin paradigm : A systematic literature review”. *Computers in Industry*, 130, 103469. 2021. <https://doi.org/10.1016/J.COMPIND.2021.103469>
- SHENGLI, Wei, “Is Human Digital Twin possible ?” *Computer Methods and Programs in Biomedicine Update*, 1, 100014. 2021. <https://doi.org/10.1016/J.CMPBUP.2021.100014>.
- YEUNG, Karen, HOWES, Andrew y POGREBNA, Ganna, “AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing”. En Dubber, M., Pasquale, F. & Das, S., *The Oxford handbook of ethics of AI*. New York, NY: Oxford University Press, 2020.

LA ENCRUCIJADA DE LA FILOSOFÍA HUMANÍSTICA DE LA TECNOLOGÍA EN EL SIGLO XXI. ENTRE LA NUEVA ILUSTRACIÓN EUROPEA Y LA FILOSOFÍA POSTEUROPEA CHINA

*THE HUMANISTIC PHILOSOPHY OF TECHNOLOGY IN THE 21ST
CENTURY AT THE CROSSROADS. BETWEEN THE NEW EUROPEAN
ENLIGHTENMENT AND CHINESE POST-EUROPEAN PHILOSOPHY*

MONTSERRAT CRESPIÓN PERALES

Doctora en Filosofía
Licenciada en Derecho y Filosofía
Departamento de Filosofía
Facultad de Filosofía
Universidad de Barcelona
Barcelona/España
m.crespin@ub.edu
ORCID: 0000-0002-0190-9692

Recibido: 19/04/2021

Revisado: 1/06/2021

Aceptado: 6/09/2021

Resumen: Este artículo revisa críticamente una muy reciente incorporación a la “filosofía humanística de la tecnología”: la defendida por el filósofo Yuk Hui alrededor de su “tecnodiversidad”. Se describe cuál es el ambiente actual en el que se discuten asuntos relacionados con la tecnología y con la inquietud que suscita la inteligencia artificial. Esta desazón desemboca, al menos, en dos tendencias. Una, la representada por un anhelo de “nueva ilustración” europea. La otra, propone una filosofía de la tecnología no-europea y no-moderna, por la que aboga Hui. Seguidamente, se discute el abordaje de este autor. Sostiene que, si se consiguieran articular las múltiples historias técnico-culturales, se desmontaría el monismo tecnológico eurocéntrico y se validaría el pluralismo –las diversas “cosmotécnicas”– superándose así la que a sus ojos es la indeseable sincronización tecnológica mundial. Se enuncian los problemas filosóficos e histórico-intelectuales del asunto de fondo que mueve a Hui: la necesidad de una filosofía de la tecnología china por motivos históricos y razones políticas para superar la Modernidad. El estudio concluye evaluando esta filosofía sino-humanística de la tecnología.

Palabras clave: nueva ilustración, filosofía posteuropea, Yuk Hui, tecnodiversidad, cosmotécnica, orientalismo.

Abstract: This article reviews critically one of the most recent additions to the “humanistic philosophy of technology”: the one advocated by the Hong Kong philosopher Yuk Hui around his “techno-diversity”. To do this, firstly, the text describes the actual intellectual environment in which philosophical issues related to technology, and the concern around artificial intelligence, are discussed. Discussions leads to, at least, two trends. One, represented by a longing for a European “new enlightenment. The other, proposing a non-European and non-modern philosophy of technology, which is precisely the one that Hui defends. Secondly, the text analyses Hui’s approach that affirms that a validation of the technological pluralism –the multiple “cosmotechics” will overcome what he sees as an undesirable global technological synchronization. The essay will elucidate the philosophical and intellectual-historical problems around the underlying issue that moves Hui: the need to develop a philosophy of Chinese technology for historical reasons and political reasons to overcome or reset Modernity. The study concludes evaluating the merits and weaknesses of this Sino-humanistic philosophy of technology.

Keywords: New Enlightenment, Post-European Philosophy, Yuk Hui, Technodiversity, Cosmotechics, Orientalism.

INTRODUCCIÓN

Desde el siglo XIX¹, una parte de lo que se presenta como “filosofía de la tecnología” correlaciona con movimientos de reacción a la Modernidad. La “tecnología”, enunciada así, sin mayor especificación, aparece como una decantación crítica a la etapa históricamente movediza que conocemos por “Modernidad”, o, más concretamente, a las “respuestas ideológicas”² que encubren el problema de datarla para, luego, identificar y clasificar sus rasgos, autores o conceptos. Cierta “filosofía de la tecnología” representa una de las capas de la literatura filosófica que se edifica en torno al análisis de la Modernidad. Por eso es importante empezar distinguiendo enfoques en las aproximaciones filosóficas a la tecnología.

Mitcham distinguió convenientemente entre dos tipos de análisis que coinciden en tener como objeto central a la tecnología, pero que divergen en propósito y enfoque filosóficos: una dirección analítica es la impulsada por ingenieros y

1 Normalmente a partir de la publicación en 1877 de *Grundlinien einer Philosophie der Technik* de Ernst Kapp (1808-1896) Véase KAPP, Ernst, *Elements of a Philosophy of Technology. On the Evolutionary History of Culture*. (Ed. Jeffrey West Kirkwood and Leif Weatherby; Trad. Lauren K. Wolfe; Afterword Siegfried Zielinski). Minneapolis, London: University of Minnesota Press, 2018.

2 FORTUNY, Francesc J., “El problema dels orígens de la Modernitat”, *Anuari de la Societat Catalana de Filosofia-IEC*, 2, 1988, 108. <<http://www.raco.cat/index.php/AnuariFilosofia/article/view/211633>> [Consulta: 13 abr. 2021].

tecnólogos cuya pregunta filosófica por la tecnología está regida por su praxis; la otra, concebida por él como “filosofía humanística de la tecnología”, compete a filósofos que tratan de entender, desde fuera de la praxis tecnológica y, habitualmente, también fuera de la científica, el “fenómeno” moderno de la tecnología.³ Por supuesto, entre las dos visiones o puntos de vista se producen cruces, diálogos y confluencias.

Desde el interior de los estudios filosóficos, el esquema predominante es el de la orientación humanística que, en seguida, se conecta con una serie de nombres –Ortega y Gasset, Heidegger, Jonas, Arendt, etc.– en los que predomina un estudio de los aspectos problemáticos de la moderna-sociedad-tecnológica. Esta constelación de filósofos y filósofas dan por sentado el fenómeno mismo de la tecnología porque su objeto de estudio primero y último no es la tecnología, sino las relaciones de esta con la moral, la política, la sociedad o la metafísica⁴.

En este artículo se revisará críticamente una de las más recientes incorporaciones a la literatura de esta “filosofía humanística de la tecnología” que es la defendida por el filósofo hongkonés Yuk Hui (許煜) alrededor de su sistematización de la “tecnodiversidad”. Para ello se expondrá, primero, la tonalidad del ambiente intelectual en el que hoy en día se discuten asuntos relacionados con la tecnología y, particularmente, la inquietud que surge con los avances en el terreno de la inteligencia artificial (en adelante, IA). Esta desazón desemboca, al menos, en dos tendencias. Una, la representada por un anhelo de “nueva ilustración”. La otra, la que, precisamente, quiere desengancharse de los vagones de la idea progresiva de avance tecnológico proponiendo una filosofía de la tecnología no-europea y no-moderna, que es justamente por la que aboga Hui. En segundo lugar, se discutirá el abordaje que propone Hui mediante el cual, según sostiene, si se consiguieran articular las historias técnico-culturales para desmontar el monismo tecnológico eurocéntrico y validar, así, el pluralismo resultante de las diversas “cosmotécnicas”, se superaría la, a sus ojos, indeseable sincronización o convergencia mundial. Esta cuestión requerirá acometer directamente el asunto de fondo que mueve a Hui y que no es otro que la necesidad de imaginar y desarrollar una filosofía de la tecnología china por motivos históricos y razones políticas: una filosofía que supere o resetee la Modernidad europea. El estudio concluirá con

3 MITCHAM, Carl, *Thinking through technology. Path between Engineering and Philosophy*. Chicago, London: The University of Chicago Press, 1994, 13-14.

4 FRANSENN, Maarten, LOKHORST, Gert-Jan, VAN DE POEL, Ibo, “Philosophy of Technology”. En: ZALTA, Edward N., *Stanford Encyclopedia of Philosophy*. [en línea] Stanford (Ca.): Stanford University, 2009 (actualización en 2018). <<https://plato.stanford.edu/archives/fall2018/entries/technology/>>. [Consulta: 13 abr. 2021]. ISSN: 1095-5054.

la evaluación de esta filosofía sino-humanística de la tecnología desvelando, en particular, su arraigado orientalismo y el suelo ideológico que le sirve de apoyo.

1. UNA IA SIN FILOSOFÍA Y EL RETORNO DEFENSIVO DE LA “ILUSTRACIÓN”.

El proyecto sin ánimo de lucro OpenMind del grupo financiero BBVA publica desde el año 2008 una colección de libros cuyo objetivo es analizar y presentar a un público amplio, no especializado, las cuestiones que considera más apremiantes para nuestro tiempo. El volumen publicado en 2019 llevaba por título *¿Hacia una nueva Ilustración? Una década trascendente*. Según se puede leer en su presentación, el objetivo común de las aportaciones publicadas en su interior, fruto de las reflexiones de figuras internacionales relevantes en diferentes campos es, en palabras de Francisco González, revisar los cambios observados en la última década para encarar las decisiones individuales y colectivas en el futuro.⁵ La publicación delinea algunos de los trayectos posibles del futuro, mayoritariamente focalizados en el desarrollo científico y tecnológico, sin faltar, claro está, la cuestión de la IA o la robótica. Y, afirma González, de ello se obtienen dos conclusiones: constatar el impacto de la revolución tecnológica sobre la humanidad y la necesidad de promover “una nueva Ilustración”, que define como “[...] un amplio diálogo para establecer unas nuevas bases filosóficas y éticas que sustenten una economía, una sociedad, una cultura y una regulación adaptadas al nuevo entorno científico-tecnológico, [...]”⁶.

Esta concepción de la “nueva ilustración” recoge, sin duda, el sentido que le diera José Manuel Sánchez Ron, físico, historiador de la ciencia y académico de la Real Academia Española, en su libro *La Nueva Ilustración: Ciencia, tecnología y humanidades en un mundo interdisciplinar* (2011).⁷ La tesis que atraviesa el libro de Sánchez Ron y que permea en la perspectiva de la colección de ensayos de OpenMind/BBVA es la de la interdisciplinariedad, trasuntada como “diálogo”, en las palabras antes citadas de Sánchez, entre ciencia, tecnología y humanidades, para instituir una nueva fundamentación filosófica y ética. Por otro lado, resulta

5 GONZÁLEZ, Francisco, “Prólogo”. En: VV.AA., *¿Hacia una nueva Ilustración? Una década trascendente*. Madrid: OpenMind/BBVA, 2019, 2. <<https://www.bbvaopenmind.com/libros/hacia-una-nueva-ilustracion-una-decada-trascendente/>> [Consulta: 07 abr. 2021]

6 Ibid., 3.

7 SÁNCHEZ RON, José Manuel, *La Nueva Ilustración: Ciencia, tecnología y humanidades en un mundo interdisciplinar*. Oviedo: Ediciones Nobel, 2011.

esclarecedor que en 2019 la posibilidad o viabilidad de la “nueva ilustración” se ponga entre signos de interrogación.

Poco antes, una figura clave para entender plenamente la historia reciente de Estados Unidos y, particularmente, su política internacional, Henry Kissinger (1923-), publicaba en junio de 2018 un artículo en la revista *The Atlantic* titulado “How the Enlightenment Ends”⁸ en el que relata los peligros de la IA. Se puede sintetizar la postura de Kissinger en dos proposiciones. La primera, que “Philosophically, intellectually—in every way—human society is unprepared for the rise of artificial intelligence.”⁹ La segunda, que “The Enlightenment started with essentially philosophical insights spread by a new technology. Our period is moving in the opposite direction. It has generated a potentially dominating technology in search of a guiding philosophy.”¹⁰

La cautela debida no permite encarar aquí la verdad o falsedad de la primera proposición. Sí, en cambio, será fundamental indagar algo en el significado intrínseco de la afirmación de Kissinger cuando habla de la falta de una filosofía rectora de la IA. O, dicho de otro modo: que esta revolución tecnológica avance huérfana de una filosofía que la guíe.

A Kissinger le alarma la ausencia de un “proyecto nacional” estadounidense centrado en explorar la IA, su alcance, implicaciones y su relación con las “tradiciones humanísticas”.¹¹ Teme, en definitiva, que Estados Unidos quede atrás frente a otros países que sí que han apostado por un impulso nacional fuerte en el campo de la IA y, en particular, en decidir cuáles deben ser los principios filosóficos y éticos de esta. Aunque no lo menciona explícitamente, está en su recámara, seguro, la centralidad que la IA tiene en la agenda nacional de China, un país sobre el que ha reflexionado suficientemente el exsecretario de estado.¹² Ahora bien, si se deja a un lado la guerra geopolítica alrededor de la IA que sobrevuela en el escrito del analista,¹³ lo crucial está en ver cómo proyecta la inquietud compartida por muchos al desconocer qué razón filosófica ocupará el espacio vacío. Esto es,

8 KISSINGER, Henry, “How the Enlightenment Ends”, *The Atlantic*, June 2018. <<https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>> [Consulta: 07 abr. 2021]. Yuk Hui le dedica a este texto un extenso comentario. Véase HUI, Yuk, “¿Qué comienza después del fin de la Ilustración?”, en *Fragmentar el futuro. Ensayos sobre tecnodiversidad*. (Trad. Tadeo Lima). Buenos Aires: Caja Negra Editora, 2020, 65-84.

9 KISSINGER, Henry, “How the Enlightenment Ends”, art. cit.

10 *Ibid.*

11 *Ibid.*

12 Véase su KISSINGER, Henry, *China*. Barcelona: Editorial Debate, 2012.

13 Para las cuestiones referidas a la configuración de la nueva revolución tecnológica y, en particular, el debate sobre los principios éticos, esto es, la diatriba, tanto ético-política como material, en torno a la hegemonía o, bien, la dependencia, tecnológica y de la IA en la Unión Europea y China, remito a CRESPIÑ PERALES, Montserrat, «Las máquinas de miembros conscientes y sus vigías: La

qué dimensiones filosóficas –saber, dirección y forma– modificarán el diagnóstico de una “IA sin filosofía” para establecerse como “la” filosofía de la IA. En parte, este desasosiego –el último momento ilustrado que anuncia el dominio de una IA desmesurada– es una secuela más de lo que se conoce como “la narrativa filosófica de la Ilustración”¹⁴ y la pervivencia de la discusión que suscita este período de las luces que arranca en el siglo XVII. Ahora desvela el temor de que el territorio de la “era de la razón”, vapuleado, especialmente durante el siglo XX en manos de sus más acerbados críticos –Foucault, Adorno y Horkheimer, como los más ampliamente conocidos–, y vaciado, sea ocupado plenamente por algo así como una “contrailustración eterna”. Sea, pues, como cesación de la “ilustración” sobrepasada por la IA (Kissinger), sea como afán por renovarla adaptada al presente (Sánchez Ron), la “ilustración” es un fenómeno que, no sin ironía, encajaría con el conocido síntoma bautizado por Freud como “retorno de lo reprimido” que iba de la mano de otro, el de la “defensa secundaria”.¹⁵

Los debates sobre “ilustración” y tecnología son una de las puertas de entrada que, junto a otras, evidencian la indeterminación en torno a qué se entiende por filosofía de la IA. También son discursos que ayudan a comprender de qué manera, también difusa, algunos autores europeos y norteamericanos imaginan el retorno ilustrado como “antídoto”. Y, por contraste, este ejercicio facilita detectar discursos oponentes, como el que aquí se analizará seguidamente que sería una suerte de alternativa: la filosofía posteuropea.

En los siguientes apartados se intentará comprender qué significa y qué propone esta “filosofía posteuropea” en la obra de uno de sus impulsores¹⁶: el ingeniero informático y filósofo, Yuk Hui, discípulo del recientemente fallecido Bernard Stiegler (1952-2020). Para ello, se procederá, primero, con una clarificación conceptual de lo que Hui entiende por “filosofía posteuropea” en conexión con los otros conceptos que van de su mano, “tecnodiversidad” y

ética de la inteligencia artificial en China y Europa», en PELEGRÍN, Àngels (ed.), *Encrucijadas: China desde el presente*. Granada: Editorial Comares, 2021, 113-147.

14 PETERS, Michael A., “The enlightenment and its critics”. *Educational Philosophy and Theory* [En línea], 51, 9, 2019, 892. DOI: 10.1080/00131857.2018.1537832

15 FREUD, Sigmund, “Nuevas puntualizaciones sobre las neuropsicosis de defensa” (1896). En: *Obras completas. Volumen 3 (1893-99). Primeras publicaciones psicoanalíticas*. Segunda reimpresión. Buenos Aires: Amorrortu editores, 1991, 157-184.

16 Ciñéndonos a filósofos japoneses, otros nombres que se han sumado a este proyecto son los de Hidetaka Ishida (石田英敬) (1953-) filósofo y especialista en literatura francesa, particularmente centrado en la obra de Foucault, o Hiroki Azuma (東浩紀) (1971-), crítico literario igualmente marcado por la filosofía postmoderna francesa de autores como Derrida. Hui, Ishida y Azuma compartieron un evento que llevaba por título “Is a Post-European Philosophy of/in Technology Possible?” el pasado 20 de agosto de 2019 que se puede recuperar en el siguiente enlace <<https://vimeo.com/onde-mand/genron20190820>> [Consulta: 14 abr. 2021]

“cosmotécnicas”, tal y como se encuentran en su *The Question Concerning Technology in China. An essay in Cosmotecnics* (2016)¹⁷, y en la colección de artículos publicados por el autor entre 2017 y 2020 traducidos al español bajo el título *Fragmentar el futuro. Ensayos sobre tecnodiversidad*^{18,19}. Tras exponer y comentar la fundamentación conceptual básica sobre la que el autor vertebra su “filosofía posteuropea”, se discutirá su modo de armar esa vía filosófica no-europea y no-moderna recuperando para ello la tópica de “superar la modernidad”. El antecedente de esta premisa, “superar la modernidad”, se encuentra en el eslogan (近代の超克-*kindai no chōkoku*) que sintetiza los contenidos del coloquio de julio de 1942 en el que participaron un amplio y variado grupo de intelectuales japoneses²⁰. Entre ellos, el filósofo Keiji Nishitani (西谷啓治) (1900-1900), en el que Hui centra especial atención porque la obra filosófica del japonés es una suerte de reverberación de la filosofía de Heidegger, el otro pilar sobre el que hace pivotar gran parte de su estrategia argumental.²¹ Así pues, se discutirán los términos, mostrando algunas inconsistencias en las que incurre un autor que, por otro lado, es consciente de que tiene que convencer a unos lectores que presupone, no sin falta de razón, desconocedores de la historia de la filosofía del este asiático, pero fascinados por ella. Polemizar y, especialmente, desvelar la capa (auto)orientalista que recubre los argumentos de Hui, ayudará a calibrar rigurosamente la validez o invalidez de su filosofía posteuropea como ejemplar contemporáneo de filosofía humanística de la tecnología. Esto es, tantear si, como pretende, su modo de responder a nuestra época tecnológica a través de un pasaje hacia un “pensamiento no-europeo y no-moderno”²², o su propuesta de usar China como ejemplo para describir la “otra cara” de la Modernidad y encarar la actual superación de la

17 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotecnics*. Falmouth: Urbanomic, 2016.

18 HUI, Yuk, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit.

19 Yuk Hui es un autor muy prolífico. El lector que quiera conocer otras publicaciones del filósofo puede dirigirse a la información detallada de su bibliografía consignada en su web personal: <<http://www.digitalmilieu.net/yuk/>> [Consulta: 1 jun. 2021]

20 Véase la traducción al inglés del simposio en CALICHMAN, Richard F. (Ed. y Trad.), *Overcoming modernity. Cultural identity in Wartime Japan*. New York: Columbia University Press, 2008.

21 Hui asume los discursos y conceptualizaciones de postguerra en lo referido a la “Escuela de Kioto”, dando por buena la tríada de nombres ideológicamente sesgada –Kitarō Nishida (西田幾多郎) (1870-1945), Hajime Tanabe (田辺元) (1885-1962) y Keiji Nishitani– porque le encajan en la crítica a la modernidad y el nihilismo heideggeriano, como se verá más adelante. Obvia, y es una ausencia grave, al filósofo, relacionado con la Escuela de Kioto, pero crítico agudo de la filosofía idealista del movimiento, que sí dedicó una parte de su producción a reflexionar sobre la tecnología: Jun Tosaka (戸坂潤) (1900-1945).

22 HUI, Yuk, “Prefacio”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 13.

Modernidad o “resetearla”²³, se alejan, como él defiende, de la estación final en la que desembocó la propuesta japonesa en la década de 1940 –la guerra y la devastación ideológicamente enmarcada en el ultranacionalismo– o si es imposible alejarse de ese destino. Es decir, desviarse de la deriva a la que lleva lo que Rieu describe como “síndrome” de crisis, en absoluto privativo de Japón, sino patrón que se encuentra, en sus palabras:

[...] in every modern society and they seem to constitute a step in its own evolution. This syndrome expresses the collective memory of a society at a given moment. It conditions the relationship of this society to its past, the understanding of the present and the capacity of individuals and groups to be actors responding to their present situation²⁴.

En conclusión, algunas de las preguntas básicas que se intentará responder son: La filosofía posteuropea, ¿es la razón filosófica que guiará convenientemente los caminos tecnológicos y de la IA? ¿La “superación de la modernidad” resignifica su sentido originario para dar cabida filosófica hoy al liderazgo –material– de China? Las distintas consideraciones alrededor de estas cuestiones servirán para evaluar el mérito o demérito de esta filosofía de la tecnología no-europea y no-moderna- y del suelo ideológico que la sostiene.

2. FILOSOFÍA POSTEUROPEA, TECNODIVERSIDAD Y COSMOTÉCNICAS

En el prefacio a la traducción al español de los textos de *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, Yuk Hui sintetiza su intención de redespigar la pregunta por la técnica cimentada tras el *diktum* heideggeriano en su “La pregunta por la técnica” (1949, 1953)²⁵. Hui extrae lo que para él es el núcleo de la tesis heideggeriana que expone del siguiente modo:

[...] he [Heidegger] announced that the essence of modern technology is nothing technological, but rather enframing (*Ge-stell*)— a transformation of the relation

23 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechnics*, op. cit., 52.

24 RIEU, Alain-Marc, “The syndrome of ‘overcoming modernity’: Learning from Japan about ultra-nationalism”. *Transtext(els) Transcultures* 跨文本跨文化 [En línea], 9, 2015-08-15, 1-2. <<http://journals.openedition.org/transtexts/552>> DOI : <https://doi.org/10.4000/transtexts.552> [Consulta: 07 abr. 2021].

25 Véase HEIDEGGER, Martin, “La pregunta por la técnica”, en *Conferencias y artículos*. (Trad. Eustaquio Barjau). Barcelona: Ediciones del Serbal, 1994, 9-37.

between man and the world such that every being is reduced to the status of 'standing-reserve' or 'stock' (*Bestand*), something that can be measured, calculated, and exploited. Heidegger's critique of modern technology opened up a new awareness of technological power, [...] this text in particular, portray the shift from *techne* as *poiesis* or *bringing forth* (*Hervorbringen*) to technology as *Gestell*, seen as a necessary consequence of Western metaphysics, and a destiny which demands a new form of thinking: the thinking of the question of the truth of Being²⁶.

La clave del diálogo que Hui plantea y, por tanto, su desobstruir la indagación de Heidegger, es, a un tiempo, una respuesta desde Asia Oriental. Al preguntarle sobre por qué escoger a Heidegger como punto de partida, responde:

Heidegger's reading of modern technology as an accomplishment of Western metaphysics is for me both a pivot and a shortcut to expose his own limits and to reopen the question of technology. To liberate Eastern philosophy is to re-activate it, to give it wings so that it can escape the marginalization by Western technology (metaphysics), and comprehend the latter from a new standpoint. I think that it is only by doing so that we can develop an Eastern "critical theory" or "critical philosophy." To reopen the question of technology is such a liberation and reinvention²⁷.

A su parecer, la pregunta –y la respuesta– por la técnica en la voz de Heidegger está limitada por su dependencia en la “metafísica occidental”. Por tanto, reabrirla es canalizarla para pensar la técnica desde la “filosofía oriental” (*Eastern philosophy*) liberándola de su subordinación occidentalocéntrica. De este modo, a lo que Hui aspira es a defender una “teoría crítica oriental”, necesaria para que la tecnodiversidad se oponga a la “cultura monotecnológica” occidental.²⁸ Así lo expone:

La búsqueda de la tecnodiversidad propone reabrir la cuestión de la técnica: en vez de entender a la tecnología como un universal antropológico, insta a redescubrir una multiplicidad de cosmotécnicas junto con sus respectivas historias y con las posibilidades que ofrecen para hacer frente hoy a la tecnología moderna²⁹.

Entendido su ejercicio filosófico de esta manera, no se está tanto ante un reabrir la pregunta, sino ante una impugnación de la premisa que sostiene que la tecnología sea un universal antropológico. Se sobreentiende, porque el pensador

26 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotecnics*, 3.

27 LOVINK, Geert, HUI, Yuk, "Interview: For a Philosophy of Technology in China". *Parrhesia* [En línea], 27, 2017, 52, <<http://parrhesiajournal.org/parrhesia27/parrhesia27.pdf>> [Consulta: 11 abr. 2021].

28 HUI, Yuk, "Prefacio", *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 12.

29 *Ibid.*, 9.

elude aportar una definición clara y distinta de cómo entiende el adjetivo “universal” y, con ello, en qué exacto aspecto está lo controvertible, que para Hui “universal” equivale a una reducción o asimilación de la variedad cultural de tecnologías –en otro pasaje habla de “tecnologías milenarias de la India y de China, de las civilizaciones maya e inca, de los nahuas, los mapuches o los diferentes pueblos del Amazonas”³⁰ o pone a la medicina china como ejemplo de “cosmotécnica”.³¹ A su juicio, problematizar este universal antropológico-técnico que subsume la variedad en el interior de un monismo “occidental” significa defender el pluralismo tecnológico.

Su defensa parte de reformular “antinómicamente” la pregunta por la técnica. Desgraciadamente no se acaba de entender en qué sentido una antinomia ayuda al propósito de rearticular las diferencias³² de las muchas “técnicas”. Esto porque el autor no explica en qué sentido quiere emplear como método interrogativo lo que no es sino un conflicto de posiciones, de hecho, literalmente, una contradicción de leyes. Bien, pero, entonces, ¿en qué nos dice que consiste la contradicción –la fórmula “antinómica”– en la que se cae al pensar la tecnología? En una tesis y una antítesis:

Tesis: la tecnología es un universal antropológico; puede ser entendida como exteriorización de la memoria y liberación de los órganos, tal como lo han formulado antropólogos y filósofos de la tecnología.

Antítesis: la tecnología no es un universal antropológico; es posibilitada y constreñida por cosmologías particulares que van más allá de la funcionalidad o utilidad. Por consiguiente, no existe una única tecnología, sino múltiples cosmotécnicas³³.

Nótese que más que una antinomia, hay una toma de partido, inicial y final, por la “antítesis”. Frente a la (supuesta) incompatibilidad de ambas proposiciones –su contradicción–, el filósofo no quiere resolver la antinomia como tal. Tal cosa supondría hacer idénticas, normativamente hablando, a las dos leyes en conflicto. No: aborda la disolución de la antinomia sin encararla de frente, optando por una oblicua defensa de la “antítesis”. Se está ante una pura petición de principio. Y aquí, por tanto, la formulación queda en sí misma desmontada, pues no hay tal cosa como un conflicto genuino que pediría, obligadamente, una resolución. Lo que se expone es una conflictividad meramente aparente, puesto que el autor ha

30 Ibid., 10.

31 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechics*, op. cit., 18.

32 HUI, Yuk, “Prefacio”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 11.

33 Ibid., 11.

apostado desde el principio por una de las razones –no existe una única y universal tecnología, sino una multiplicidad de cosmotécnicas–.

Para demostrar la validez de esta proposición parte de la siguiente hipótesis: en China nunca ha existido la técnica tal y como ha sido definida desde la filosofía europea, pues, a su juicio, el eurocentrismo ha nivelado, equivocadamente, las múltiples percepciones culturales sobre la técnica.³⁴ De ahí que, para desanudar esa idea equivocada –que todas las técnicas son iguales o, lo que es lo mismo, que la tecnología sea un universal antropológico–, proponga desocultar las historias singulares de las cosmologías técnico-culturales, esto es, los trayectos de las “cosmotécnicas” que validan la “tecnodiversidad”, es decir, el pluralismo tecnológico.

El autor debe, y así lo intenta, definir “cosmotécnica” de la que ofrece la siguiente explicación:

[...] it means the unification between the cosmic order and the moral order through technical activities (although the term *cosmic order* is itself tautological since the Greek word *kosmos* means order). The concept of cosmotechnics immediately provides us with a conceptual tool with which to overcome the conventional opposition between technics and nature, and to understand the task of philosophy as that of seeking and affirming the organic unity of the two³⁵.

Bien: unificación del orden cósmico y moral. Pero ¿desde qué principio cosmológico? Lo primero que hace es retornar a la raíz del pensamiento chino antiguo que se encuentra, según él lo ve, en la no-oposición de pares filosóficamente bifurcados desde Europa como técnica y naturaleza, por ejemplo. A su entender, lo que singulariza la cosmología china y la hace diferente de la causalista y “opositiva” cosmología europea es la “resonancia”. Es así como el autor traduce la noción china *Ganying* (感應)³⁶ en vez de hablar de “pensamiento correlativo” (o analógico) que Hall y Ames definen como el rasgo que se encuentra en las cosmologías chinas que pivota en la asociación de imágenes o conceptos y no en la causalidad física; en procedimientos analógicos informales que presuponen la asociación y la diferenciación que encuentran en patrones culturales compartidos y no en la necesidad causal³⁷.

34 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechnics*, op. cit., 9.

35 Ibid., 19-20.

36 Ibid., 27.

37 HALL, David L., AMES, Roger T., “Chinese philosophy”. En: GRAIG, Edward, *Routledge encyclopedia of philosophy*. [En línea]. London: Routledge, 1998.<<https://www.rep.routledge.com/articles/overview/chinese-philosophy/v-1>> [Consulta: 13 abr. 2021] DOI: 10.4324/9780415249126-G001-1

Para Hui, el pensamiento correlativo, la resonancia, es la repercusión de los dos órdenes, el humano y el del Cielo (天人合一-Tiān rén hé yī), e implicaría la homogeneidad entre todos los seres y una organicidad en la relación entre parte y parte y entre parte y todo³⁸. El elemento clave de la resonancia está, sin duda, en la noción de Cielo (天-Tiān).

Antes de avanzar, conviene aportar una breve nota aclaratoria sobre el significado de este vocablo esencial en la terminología del pensamiento chino. La idea del “Cielo” no apunta a una dimensión trascendente, más allá de los límites de lo natural, sino a una inmanencia, a un permanecer en los límites cósmico-naturales. Sería un espacio en el que se produce la continua comunicación del Aquí y el Ahora, en palabras de Bauer³⁹, dentro del cual los individuos están rodeados de todo lo que conforma y penetra en la realidad. La noción emerge con la figura de la deidad ancestral Shangdi (el señor de arriba) durante la dinastía Shang (1600-1046 a.C) y es posteriormente ajustada cuando “Cielo” desplaza a esta deidad suprema durante la dinastía Zhou (1046-256 a.C). Pasa a asumir una doble carga significativa: “Cielo” es culto y reverencia, pero también es un principio físico-natural, origen de los fenómenos ambientales. La parte cultural se reflejará en las ceremonias del emperador cuando se convierte en “el hijo del Cielo” (天子-Tiānzǐ) y ostenta el mandato del Cielo (天命-Tiānmìng). En la ceremonia de respeto al Cielo, él hace frente a sus responsabilidades como autoridad de gobierno de la sociedad humana y de respeto a la correlación –analogía– entre su persona, el Cielo y la tierra. De acuerdo con Anne-Hélène Suárez:

Inicialmente, [el mandato del cielo] se refería a la potestad divina puesta en manos del soberano por decisión del cielo; idea que justificaba el que, si un soberano degeneraba, poniendo así en peligro la prosperidad del país, fuera legítimo destruirlo, pues había dejado de ser digno de la confianza celeste. Posteriormente, ya en la época de Confucio, adquirió un doble sentido más amplio. Se trata, por una parte, del orden natural de las cosas. Por otra, del conjunto de principios con los que el cielo dota a los hombres y la misión que éstos tienen en el mundo⁴⁰.

Retomando la cuestión de fondo, Hui subraya la unificación del Cielo y lo humano como punto de partida para su alternativa filosófico-tecnológica que, hoy, a sus ojos, demanda la renovación, “liberación” y “reinención” de la filosofía oriental. A su parecer, la filosofía oriental debe desmarcarse de la metafísica occidental

38 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechnics*, op. cit., 27.

39 BAUER, Wolfgang, *Historia de la filosofía china*. (Trad. Daniel Romero). Barcelona: Herder, 2009, 45.

40 CONFUCIO (Maestro Kong), *Lun Yü. Reflexiones y enseñanzas*. (Trad., Introd. y notas Anne-Hélène Suárez). Barcelona: Kairós, 2002, 164.

para situarse nuevamente en su esfera metafísica propia: en su “cosmología moral” o “metafísica moral”⁴¹ que, como ya se ha dicho, para el pensador depende de la idea de Cielo como fundamento moral⁴².

En resumen, partiendo del caso de China, la cosmotécnica, sendero de la tecnodiversidad, contribuiría idealmente a un “pluralismo de la tecnicidad”⁴³ que, al final, sirviera para “fragmentar el mundo según la diferencia, en vez de universalizar por medio de lo igual; inducir lo igual a través de la diferencia, en vez de deducir la diferencia a partir de lo igual.”⁴⁴ Para Hui esta fragmentación es la que entarima las diversidades filosóficas, estéticas y tecnológicas⁴⁵. Este mosaico impulsará a la filosofía posteuropea cuya misión es desmantelar “[...] la tendencia totalizante de la dialéctica y libera[r] aquello que ha sido eclipsado por la concepción de la historia como progreso”⁴⁶, de tal suerte que retroceder en la historia (factual pero, sobre todo, tecnológica, según la entiende nuestro autor) sea una resistencia a la uniformización europea. Propone, pues, que la filosofía “después de Europa” sea el punto de llegada de esa revisión histórica y abra el camino para un proyecto filosófico-tecnológico que permita una nueva “[...] relación entre lo tradicional y lo moderno, lo local y lo global, Oriente y Occidente”⁴⁷.

3. ¿“SUPERAR LA MODERNIDAD” EN EL SIGLO XXI? EL REVISIONISMO DE UNA DOCTRINA POLÉMICA

Esta tensión intelectual entre las dos “mundovisiones” científico-tecnológicas de China y Occidente, no es novedosa. Cuenta con muchos antecedentes. Una muestra significativa se encuentra en las reflexiones de un filósofo influenciado fuertemente por el pragmatismo de John Dewey (1859-1952), Hu Shi (胡適) (1891-1962), que las volcaría del siguiente modo en su tesis doctoral, *El desarrollo del método lógico en la antigua China*, defendida en el año 1917 en la Facultad de Filosofía de la Universidad de Columbia:

41 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechinics*, op. cit., 29.

42 Ibid., 27.

43 Ibid., 56.

44 HUI, Yuk, “Sobre la conciencia desventurada de los neorreaccionarios”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 40.

45 HUI, Yuk, “Limite y acceso”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 136-137.

46 Ibid., 137.

47 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechinics*, op. cit., 57.

Now that China has come into contact with the other thought-systems of the world, it has seemed to some that the lack of methodology in modern Chinese philosophy can now be supplied by introducing into China the philosophical and scientific methods which have developed in the Western world from the time of Aristotle to this day. This would be sufficient if China were contented to regard the problem of methodology merely as a problem of “mental discipline”; in the schools or even as one of acquiring a working method for the laboratories. But as I look at it, the problem is not really so simple. The problem as I conceive it is only one phase of a still larger and more fundamental problem which New China must face.

This larger problem is: How can we Chinese feel at ease in this new world which at first sight appears to be so much at variance with what we have long regarded as our own civilization?⁴⁸

Los precedentes de la discusión y, de ahí, de la tensión descrita, van formando parte del amplísimo sedimento de la literatura académica alrededor de una pregunta, hasta cierto punto, circular, dentro del campo de la sinología: “¿Por qué la revolución científica no tuvo lugar en China?”⁴⁹ La verbalización que aporta Hu Shi y la más difundida en el contexto académico anglosajón, la formulada por Joseph Needham (1900-1995)⁵⁰ –al que remite Hui en muy distintos momentos en sus obras⁵¹, sirven para poner a contraluz y desmoronar la presunta innovación de nuestro filósofo, presentado rápida y descuidadamente como “superestrella del pensamiento”⁵².

48 HU, Shi, *The development of the logical method in Ancient China*. Shanghai: The Oriental Book Company, 1922, 6.

49 Se recoge la formulación de la pregunta según el catedrático emérito de Cultura China e Historia de la Ciencia de la Universidad de Pennsylvania, Nathan Sivin (1931-). Véase SIVIN, Nathan, “Why the Scientific Revolution Did Not take Place in China –or Didn’t It? *Chinese Science*, 5, 1982, 45-66. <<http://www.eastm.org/index.php/journal/article/viewFile/473/404>>. Véase también la versión revisada por el autor de 2005 en su página personal. <<http://ccat.sas.upenn.edu/~nsivin/scirev.pdf>> [Consulta: 13 abr. 2021]

50 NEEDHAM, Joseph, *The Grand Titration: Science and Society in East and West*. Toronto: University of Toronto Press, 1969, 16: “Why did modern science, the mathematization of hypotheses about Nature, with all its implications for advanced technology, take its meteoric rise only in the West at the time of Galileo?”. Es interesantísimo contemplar cómo responde Sivin a la misma y concluye que la pregunta está equivocada pues impone un patrón científico-histórico a China presuponiendo que las transformaciones científicas no se produjeron en el país de igual modo que se produjeron en Europa. En síntesis: la pregunta se muestra cómo una interrogación falsa o mal formulada. Sivin confía en que se modificará esta perspectiva para salir del círculo vicioso, en el que cae, también, Yuk Hui, al intentar amoldar a China en, justamente, el proceso científico-histórico acaecido en Europa. (Véase SIVIN, art. cit., 2005, 20-21).

51 Por ejemplo en HUI, Yuk, “Prefacio”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op. cit., 11, y en HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotchnics*, op. cit., 165-168.

52 ROBLES, José María, “Yuk Hui, la nueva superestrella del pensamiento: ‘El desarrollo tecnológico en China es tan rápido que da miedo’”, *El Mundo*, 17 de noviembre de 2020. <<https://>

La actualización de la pregunta a la que nos dirige Hui podría sobreescribirse perfectamente sobre la de Hu Shi de 1917: ¿cómo puede China acomodarse al nuevo mundo marcado por la revolución tecnológica del siglo XXI? El filósofo hongkonés inicia el camino hacia la respuesta apuntalando la alteridad o incommensurabilidad de la “disciplina” o “mentalidad filosófica” chinas. Porque, para Hui, esa “mentalidad” otra se explica en el hecho de que la filosofía china

[...] is based on an organic rather than a mechanical form of thinking- [...] It seems that in the Chinese philosophical mentality, the cosmos has a rather different structure and nature than in that of the West; and that the role of the human and its way of knowing are also determined in a different way, in coherence with the cosmos⁵³.

Esta bifurcación orgánico-mecánica o inductivo-deductiva, su sinónimo más habitual, tampoco es nueva. Sí intriga entender por qué un filósofo la recupera ahora y qué se sigue de ello.

Para los académicos familiarizados con la filosofía del este asiático, no deja de sorprender leer en publicaciones tan recientes, comercializadas como “rompedoras”, que en el siglo XXI estas insistan en manejar generalizaciones obsoletas como “mentalidad filosófica china”. O, más grave aún, que sus autores descuiden el mínimo rigor aplicando, sin mayor justificación, una equivalencia engañosa en las historias filosóficas al usar, como hace Yuk Hui, “East” (Oriente) como equivalente a “East Asia” (Asia Oriental) en el que incluye a China, Japón, Corea (sic) y un “etc.” indeterminado (¿India?) porque, a su juicio, todos estos territorios están influenciados por el confucianismo, el budismo y, hasta cierto punto, dice, por el daoísmo⁵⁴.

Esta taxonomía no es que sea anticuada, un producto más del perenne orientalismo, sino que, además, fue sometida hace tiempo a una rigurosa crítica filosófica que, para que no se olvide, conviene recordar. Como explicara perfectamente el académico jesuita Gino K. Piovesana (1917-1996) al polemizar contra las “generalizaciones hechas a la ligera” sobre la filosofía japonesa⁵⁵, cualquier estudioso serio debe ir en contra de esta “tendencia interpretadora” que pretende abrazar “el todo del pensamiento oriental u occidental”.⁵⁶ Es tendencia estrecha

www.elmundo.es/papel/historias/2020/11/17/5fb4000921efa031448b4627.html> [Consulta: 12 abr. 2021]

53 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotronics*, op. cit., 202.

54 Ibid., Nota 2, 3.

55 PIOVESANA, G.K., *Pensamiento japonés contemporáneo*. (Trad. J. M^a Sancho SJ). Madrid: Editorial Razón y Fe, 1967, 209.

56 Ibid., 212.

y, de suyo, falsa. Refiriendo a la superficialidad de nomenclaturas como “mentalidad oriental” o “pensamiento occidental”⁵⁷ escribe Piovesana:

[...] Los técnicos, tanto del Oriente como del Occidente, saben bien que tales términos no tienen más que una connotación geográfica. Nakamura⁵⁸, y con él todos los técnicos en este campo, se opone a una terminología que viniera a incorporar en un solo modo de pensar el pensamiento de los indios, chinos, tibetanos y japoneses. Cada pueblo tiene sus propias características y, por lo tanto, hablar de mentalidad oriental “espiritual” o “místico-intuitiva”, en contraste con “materialista” o “positivista” de Occidente, no es más que ignorar todas las tendencias diferentes y contrastadas que han florecido tanto en el Oriente como en el Occidente⁵⁹.

El manejo, entonces y ahora, de estas tipologías “ideales” sobre tendencias filosófico-culturales, refleja los fallos contemporáneos a la hora de presentar la historia intelectual china, que es la que aquí ahora reviste interés. Fallos que han sido expuestos por diversos autores como Marián Gálík (1933-)⁶⁰ o Benjamin A. Elman (1946-). Sin ir más lejos, Elman señala que la versión simplista de la historia de la filosofía puesta en circulación por los reformistas del estudio del pensamiento chino moderno como Liang Qichao (梁啟超) (1873-1929) o el mencionado Hu Shi⁶¹ es consecuencia de su dejarse llevar por la metodología estadounidense o por el enfoque alemán de la *geistesgeschichte* [historia intelectual]. Elman atribuye esta distorsión al hecho de que estos intelectuales, a caballo entre el siglo XIX y el XX, asumieron los métodos sociológicos foráneos y, particularmente, las tesis weberianas que impulsarían interpretaciones “refractarias”⁶².

La “novedosa” “filosofía posteuropea” de Yuk Hui sigue anclada en el mismo esquema, replicando una proyección sociológica deudora de la mirada alemana que esconde algunos meandros interesantes que pueden pasar desapercibidos. Así

57 Ibid., 212.

58 El autor hace referencia a la figura de Hajime Nakamura (中村元) (1912-1999) conocido por su notable obra *Maneras de pensar de los pueblos orientales: India, China, Tíbet, Japón*. Véase NAKAMURA, Hajime, *Ways of Thinking of Eastern Peoples. India, China, Tibet, Japan*. Honolulu: University of Hawai'i Press, 1964.

59 PIOVESANA, *Pensamiento japonés contemporáneo*, op. cit., 213.

60 GÁLÍK, Marián, “*Quo Vadis Philosophia?: Musings about the Necessity of Sino-Western Dialogue*”, *Journal of Chinese Philosophy*, 31, 2004, 65-81. DOI: <https://doi-org.sire.ub.edu/10.1111/j.1540-6253.2004.00143.x> [Consulta: 11 abr. 2021]

61 Para una introducción a la relevancia intelectual de estos pensadores chinos, véase CRESPIÓN PERALES, Montserrat, “*Pensament xinès contemporani*”, en PRADO-FONTS, Carles (Coord.), *Pensament i religió a Àsia Oriental*. Barcelona: FUOC, 2011, 50-68. <<http://hdl.handle.net/10609/49701>> [Consulta: 1 jun. 2021]

62 ELMAN, Benjamin A., “*The Failures of contemporary Chinese Intellectual History*”, *Eighth-Century Studies*, 43, 3, 2010, 376. <<https://muse.jhu.edu/article/388529>> DOI: 10.1353/ecs.0.0191 [Consulta: 11 abr. 2021]

por ejemplo, cuando Hui hace uso de una conferencia de Heidegger de 1936 que llevaba por título *Europa und die philosophie*, no afina lo suficiente al presentar qué significa “Asia” para el filósofo alemán. No está pensando, como interpreta Hui, en una categoría nítida de los “no-europeos”⁶³, zanjando así la cuestión. La lectura heideggeriana se basa en, primero, concebir que solamente hay filosofía en Alemania, esto es, en aquel pensar determinado por las características étnicas y por el más vulgar nacionalismo cultural esencialista. Y, segundo, en aplicar la misma operación, pero al enrasar “Asia” que, en su ideación, representa a la alteridad de su idea de Europa = Alemania. La literalidad del texto aclara por sí sola esta cuestión:

Something shall here be said, for the moment, about German philosophy and thereby about philosophy in general. Our historical Dasein experiences with increasing distress and clarity that its future is equivalent to the naked either/or of saving Europe or its destruction. The possibility of saving, however, demands something double: (1) the protection [*Bewahrung*] of the European *Völker* from the Asiatic, (2) the overcoming of their own uprootedness and splintering⁶⁴.

Karl Löwith (1897-1973), al que, sorprendentemente, no cita ni una sola vez en su *The Question Concerning Technology in China. An essay in Cosmotecnics*, ausencia clamorosa teniendo en cuenta que la segunda parte del libro Hui discute la conexión histórico-filosófica entre modernidad y nihilismo, compuso un ensayo titulado “El nihilismo europeo” (1940) mientras residía en Japón, país en el que encontró asilo como judío.⁶⁵ Löwith explica allí de qué manera la identidad europea depende de su opuesta esencial: Asia. Describe, pues, el mecanismo de nivelación Europa/Asia como sigue:

Europa es un concepto que no proviene de sí mismo, sino de su oposición esencial con Asia. (...) “ereb” (la tierra de la oscuridad o del sol poniente) y “asu” (la tierra

63 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotecnics*, op. cit., 287.

64 HEIDEGGER, Martin, “Europe and German Philosophy”, (Trad. Andrew Haas). *The New Yearbook for Phenomenology and Phenomenological Philosophy* [En línea], VI, 2006, 331. <https://www.pdcnet.org/nyppp/content/nyppp_2006_0006_0331_0340> [Consulta: 12 abr. 2021].

65 El ascenso de Hitler al poder en 1933 como canciller de Alemania supuso para Löwith que se truncara su trayectoria académica en su país natal y comenzara su cadena de exilios que le llevarían a Italia primero (1934-1936) y a Japón después (1936-1941) como profesor en la Universidad de Sendai, gracias a la mediación de un pensador japonés de relevancia, Shūzō Kuki (九鬼周造) (1888-1941). Véase LÖWITH, Karl, *Mi vida en Alemania antes y después de 1933. Un testimonio* (1940). (Trad. Ruth Zauner). Madrid: Visor, 1992.

del sol naciente). Europa es en su origen y mientras sea fiel a sí misma una potencia contra-asiática, en lo político y en lo espiritual⁶⁶.

Europa y Asia, como conceptos o tipos “ideales”, no como realidades tangibles e históricas, son fundamentales para entender lo que está entrelíneas en la denominación rasa que aplica Heidegger. Pero, además, sin ni siquiera tener que ir, otra vez, a Heidegger, es claro que el linaje de su pensamiento forma parte de la estructura de la filosofía de la historia hegeliana que hace pender el concepto de “Europa” de aquel con el que se comunica inversamente, Asia y, en particular, los “otros” que necesita para justificar el discurrir astuto de la historia: China e India.⁶⁷ Por otro lado, en 1936 [Heidegger] y en 1940 [Löwith], Asia es la región de Asia Oriental pero, también, la Unión Soviética. Y es justamente la asiática-Rusia aquella que se conjuga con mayor precisión como el doble otro de Europa. Leyendo nuevamente a Löwith:

La Europa del siglo XIX ya no vivía con la fe en una verdadera misión; se limitaba a difundir por doquier sus mercancías y su civilización científico-técnica. Contra esta Europa de las empresas capitalistas e imperialistas cuya ética estaba falta de leyes y cuyo espíritu no tenía un objetivo se alzó entonces por primera vez en el siglo pasado una voz crítica proveniente de pueblos extraeuropeos que atacaba los fundamentos más propios de Europa. Occidente se vio desde entonces en el espejo crítico de Rusia (Kireievski, Dostoievski y Tolstoi) y de China (Ku-hung-ming)⁶⁸ (...) ⁶⁹

Por consiguiente, se echa en falta que Hui indague más en la distorsión heideggeriana y en el juego de espejos identitarios que subyace en unos textos cuyo contexto es el nazismo que, además, pervive en las conferencias sobre la técnica escritas una vez acabada la Segunda Guerra Mundial. Y, paralelamente, defrauda que Hui persevere en la imperecedera insistencia circular sobre la desgastada dicotomía Oriente-Occidente, un concepto filosófico, siempre retóricamente eficaz, que, como define Pattberg, se sostiene en:

[...] claiming that the two cultural hemispheres, East and West, developed diametrically opposed, one from the particular to the universal and the other from the

66 LÖWITH, Karl, “El nihilismo europeo” (1940), en *El hombre en el centro de la historia. Balance filosófico del siglo XX*. (Trad. Adan Kovacsics). Barcelona: Herder, 1998, 57.

67 Véase HEGEL, G.W.F., *Introducciones a la Filosofía de la Historia Universal*. (Trad. Román Cuartango). Madrid: Istmo, 2005.

68 Gu Hongming (辜鴻銘) (1857-1928), escritor, contrario al movimiento reformista chino de finales del siglo XIX.

69 LÖWITH, Karl, “El nihilismo europeo” (1940), *El hombre en el centro de la historia. Balance filosófico del siglo XX*, op. cit., 59.

universal to the particular; the East is more inductive while the West is more deductive. Together they form an equilibrium...⁷⁰

Aunque la complementariedad que proclama Pattberg es distinta de la fragmentación que propugna Yuk Hui, ambos parten del mismo esquema divisorio de dos geografías imaginadas y proyectadas en bloque, homogéneas, alejadas de cualquier pluralidad o diversidad en su interior o fuera de sí. Por consiguiente, hay una veta que conecta a la “filosofía posteuropea” con el (auto)orientalismo y el nacionalismo identitario. Pero, quizá, es filosóficamente más controvertido y esencial que al lado de esta dicotomía, el autor rescite el proyecto de “superar la modernidad”, atravesado por el espíritu del nihilismo moderno. Hui emplea el espectro de esa modernidad abisal para poner a China frente al peligro que le acecha actualmente y que debería evitar: caer en la, siguiendo a Stiegler, “inconsciencia tecnológica”⁷¹, motor de una modernidad europea caracterizada por el olvido de los límites de los instrumentos técnicos⁷².

De nuevo, la sombra heideggeriana se yergue sobre el método que emplea nuestro autor para proponer una nueva “superación de la modernidad”. Escoge algunas de las obras del filósofo japonés Keiji Nishitani quien, como ajustadamente lo presentara González Valles, es un pensador sincretista, preocupado, por un lado, en estudiar y profundizar en su conocimiento del pensamiento europeo, fundamentalmente alemán, y, por otro, en el pensamiento búdico⁷³. No obstante, es un sincretismo desnivelado: los modismos y las preocupaciones filosóficas de Nishitani –también las budistas– están profundamente surcadas por la obra de Heidegger con el que mantuvo diálogo directo⁷⁴. En cierto modo, Hui sigue este mismo sendero y transita desde Heidegger a Nishitani y de Nishitani a Heidegger

70 PATTBERG, Thorsten, *The East-West Dichotomy*. New York: LoD Press, 2009, i. En el libro, Pattberg mira hacia habla de una nueva ola de “iluminación oriental” dirigida a Occidente (79-80) que se mezcla con una “[...] institutionalized overstatement and euphemism in sensation-seeking media or some individuals’ fancies may bring, even the most frivolous ambition to remedy the past failures of the Asian for the glory of her future cannot cover the fact, [...] that China in particular had not developed or not sufficiently developed anything in the way of science and technology that could compete with the –rather lucky than good– Western Imperialist’s model, [...]”.

71 Hui remite a STIEGLER, Bernard, *Technics and Time 3: Cinematic Time and the Question of Malaise*. Stanford: Stanford University Press, 2010.

72 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotronics*, op. cit., 42.

73 GONZÁLEZ VALLES, Jesús, *Historia de la Filosofía Japonesa*. Reimpresión. Madrid: Tecnos, 2002, 501.

74 El japonés sería becado por el gobierno para ampliar sus estudios en el extranjero, algo que le llevaría en 1936 a la Universidad de Friburgo donde pasaría dos años que le permitirían conocer a Heidegger, imbuyéndose especialmente de la interpretación que este hiciera de la filosofía de Nietzsche al que en aquel entonces dedicaba sus lecciones. Las lecciones están recogidas en dos volúmenes en HEIDEGGER, Martin, *Nietzsche*. (Trad. Juan Luis Vermal). Barcelona: Ediciones Destino, 2000.

para proponer a China el ejercicio necesario para dotarla hoy de una verdadera “conciencia tecnológica”. De hecho, para liderar y ocupar el espacio vacío de una filosofía humanística de la tecnología que, se dijo, al parecer todavía no tiene un rostro definido.

Ya se explicó con Rieu que el síndrome de “superar la modernidad” no es un fenómeno únicamente japonés, aunque sea mayoritariamente conocido y estudiado entre los especialistas de la filosofía japonesa del siglo XX y diseccionado como uno de los elementos, entre otros, útiles para comprender el marco ideológico que amparó el ultranacionalismo en Japón. Ilumina mucho más que un mero episodio local. En palabras de Rieu, proyecta la experiencia psicosocial de una sociedad que se percibe a sí misma como trastocada o distorsionada, atravesada por una experiencia disruptiva –llámese aquí “Modernidad”:

It is the subjective experience of an estranged society shared by a whole community, of past or recent evolutions, which have disrupted or distorted an established order or continuous path. This experience is a sense of loss felt intimately by individual subjects, with intense anxiety about the present and future of their community. In these moments, society has lost the imagined ground, the frame of its history, and the grand narratives at the core of its identity⁷⁵.

Esa sensación de desarraigo atraviesa la historia moderna de Japón. De hecho, como se encuentra en un documento que sigue siendo clarividente, la tesis doctoral de un casi desconocido filósofo, Tsunezo Kishinami (岸波修根蔵) (1877-?) titulada *The Development of Philosophy in Japan* (1915), es reacción y fruto del período ilustrado que viviría Japón a partir del inicio de la era Meiji (1868-1912).⁷⁶ Kishinami se muestra especialmente lúcido identificando que la concienciación nacional, pintada de conservadurismo y reacción, germinó gracias a la integración de nociones e ideas ilustradas y, en particular, después de que se combinaran con el subjetivismo, el historicismo y el nacionalismo de cuño alemán:

Secondly, the reactionary spirit arose and was the chief factor in bringing the movement to its end. In fact, the introduction of Western culture and science into

75 RIEU, art. cit., 3-4.

76 Para más detalles sobre la era Meiji y, en particular, el período de ilustración en la filosofía moderna de Japón, en el que se detalla el proceso de paso del liberalismo ilustrado inicial hacia el conservadurismo de la intelectualidad que surgiría de, por ejemplo, los autores [Arinori Mori (森有礼) (1847-1889), Amane Nishi (西周) (1829-1894), Hiroyuki Katō (加藤弘之) (1836-1916), o Yukichi Fukuzawa (福澤諭吉) (1834-1901)] que formarían parte de la sociedad intelectual *Meirokeisha* (明六社), activa a partir de 1873 y cuyo lema era “Civilización e ilustración” (文明開化-*bunmei kaika*-), véase CRESPIÓN PERALES, Montserrat, “Filosofía y pensamiento contemporáneo: Sincretismo japonés”, en FERNÁNDEZ, Julián (ed.), *Japón, el archipiélago de la cultura, Volumen 2: Tierra de Wa - Religión*. Barcelona: Mediatres, 2020, 135-207.

Japan moved hand in hand with the development of political affairs. [...] But when these [political] changes had been made and the new modes of government were firmly established, the people naturally began to reflect, and came to be aware of their native conditions as these had historically developed. In other words, the Japanese national consciousness was awakened. This turn of current thought, strange to say, was hastened by the Enlightenment movement itself. For in contributing to the establishment of constitutional government, it aided in supplying the conditions under which a greater degree of self-consciousness could be attained. When Japan became truly aware of itself, its gaze inevitably returned upon its past and the Enlightenment, in its cosmopolitan character, was at an end⁷⁷.

En consecuencia, ese malestar que aparece en su forma desquiciada en la década de los 1940 tiene su origen en la vivencia de los años vertiginosos de modernización e industrialización del país. Forma parte de un espíritu finisecular y va encaramándose en un haz de temores entre los que destaca el miedo a perder la soberanía en el período de los imperialismos predatorios europeos. A juicio de Swale, la élite gobernante e intelectual compartían ansiedades: miedo al colapso social y a caer subyugados por los poderes colonizadores occidentales ya presentes en otros territorios de Asia⁷⁸.

Esta atmósfera de angustia y agitación sociopolíticas cristalizan en el famoso (o maldito)⁷⁹ simposio “La superación de la modernidad”, auspiciado por tres críticos literarios, Tetsutarō Kawakami (河上徹太郎) (1902-1980), Hideo Kobayashi (小林秀雄) (1902-1983) y Katsuichirō Kamei (亀井勝一郎) (1907-1966), que juntó en 1942 en la ciudad de Tokio a trece intelectuales, entre ellos, Nishitani. Si a finales del siglo XIX se hablaba en Japón de renovación moral (矯風-*kyōfū*) y cultural, en el período de entreguerras y bélico, la renovación significará “sanar” de la infección de europeísmo y americanismo que había carcomido la “verdadera” naturaleza e identidad japonesas, en palabras de Hideo Tsumura (津村秀夫) (1907-1985).⁸⁰ Para Nishitani, el origen del desarraigo también es claro –Europa y Modernidad–:

In general, what is called “modern” means European. [...] Modernity in Japan is also grounded upon European elements, as introduced following the Meiji

77 KISHINAMI, Tsunezo, *The Development of Philosophy in Japan*. Princeton: Princeton University Press, 1915, 11.

78 SWALE, Alistair D., *The Meiji Restoration. Monarchism, Mass Communication and Conservative Revolution*. London: Palgrave MacMillan, 2009, 11.

79 Así adjetivado en ARAKI, Tōru, “2. Tokyo, 1942; Le colloque maudit [Dépassement de la modernité]”, *Ebisu*, 6, 1994, 75-95. DOI: <https://doi.org/10.3406/ebisu.1994.1586>.

80 Así lo expresaba este crítico de cine en su intervención en el simposio: “Following the Taisho period, however, Japan imported a way of life to accompany these modern conveniences, which gradually came to infect the country’s life and customs.” en CALICHMAN, Richard F., *op. cit.*, 119.

Restoration. [...] If we compare this with the introduction of Chinese culture following the age of Prince Shōtoku, for example, a marked difference seems to appear. Chinese culture was introduced in such a way that all of the fields were interrelated around Buddhism and Confucianism. This difference may perhaps be attributed to the fact that Western culture was already divided into various specialized fields, which made it so-called advance culture. [...] ⁸¹

Así pues, para hacer frente a esa pérdida de identidad se trataba de recuperar la “energía moral”⁸² connatural, intrínseca, a la esencia ética del estado japonés. Japón, moralmente restituido, estaba llamado a liderar a toda la región de Asia Oriental –la conocida “Esfera de co-prosperidad asiático-oriental” o “Gran Asia Oriental”–⁸³ y carearse con Europa, intelectual, económica, social, militar y, por supuesto, “culturalmente”, para dar nacimiento a una “nueva esfera cultural asiático-oriental”⁸⁴.

Habiendo descrito y comentado el perfil ideológico ultranacionalista del coloquio, tiznado de jerga nacionalsocialista, ya se puede conectar aquel pasado con nuestro presente. ¿La resignificación de la “superación de la modernidad”, convenientemente enmarcada en el tecno-humanismo, justifica el afán o anhelo de liderazgo –material e intelectual– de China hoy? Curiosamente, Hui sostiene que el proceso de la “Modernidad” europea no tuvo lugar en China, aunque sí la “modernización” que equivaldría a equipararse con las grandes potencias:

China has modernised itself over the past century in order to “catch up with the UK and outstrip the US” (超英趕美, a slogan proposed by Mao Zedong in 1957); now it seems to be at a turning point, its modernisation having reached a level that allows China to situate itself among the great powers⁸⁵.

La sutilidad de la diferenciación es digna de comentario. En cierto modo es análoga a la distinción entre “modernizar” y “occidentalizar” que Howland⁸⁶ maneja para el caso japonés a partir del período Meiji: “occidentalizar” equivaldría al movimiento de la 文明開化-*bunmeikaika* (civilización e ilustración), dirigido a edificar un Japón fuerte siguiendo la estela del patrón europeo en relación con el progreso industrial, de conocimiento, educativo, etc. Así, “occidentalizar” sería

81 NISHITANI, Keiji, en *Ibid.*, 51.

82 Nishitani emplea el concepto alemán “moralische Energie” empleado por Leopold Von Ranke (1795-1886).

83 NISHITANI, Keiji, en CALICHMAN, *op. cit.*, 59-63.

84 TSURUMA, Hideo, en CALICHMAN, *op. cit.*, 117.

85 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotronics*, *op. cit.*, 7.

86 HOWLAND, Douglas R., *Translating the West. Language and Political Reason in Nineteenth Century Japan*. Honolulu: University of Hawai'i Press, 2002, 5.

anivelarse al progreso material, también a la conceptualización de civilización que se maneja ya en el siglo XIX. Modernizar, por el contrario, comportaría asumir el esquema de valores europeo. La distinción entre Modernidad y modernización de Hui reedita la yuxtaposición moral/técnica del eslogan de Sakuma Shōzan (佐久間象山) (1811-1864), “Moralidad oriental, técnicas occidentales” (東洋の道德 西洋の芸術-Tōyō no dōtoku, seiyō no geijutsu)⁸⁷.

Esto último clarifica la reflexión del autor y, a la vez, la pregunta que aquí se le formulaba a su tesis. Si Hui distingue entre Modernidad y “modernización”, y China no pasó por el proceso de la “Modernidad” ni tampoco, por tanto, por el problema de haber de superar la “inconsciencia tecnológica”, fruto enfermo de la “voluntad de poder” europea⁸⁸, ¿qué se supone, pues, que debe superarse? No sería la “Modernidad” como esquema moral, sino, en todo caso, la distorsión que se produzca por parte de China al liderar el poder tecnológico presente y futuro.

Aplicando el argumento del hongkonés, si China ostenta ya y se ve empujada a través de la “voluntad de poder”, en su particular encrucijada, el país puede ser el vector acelerador de lo que el pensador prevé como futuro apocalíptico al que desembocaría la modernización y la uniforme Modernidad –la destrucción cultural, medioambiental, social y política–⁸⁹. O podría separarse de esa fatalidad. Siguiendo este camino deductivo, esto es, que China es ahora la voluntad de poder tecnológica del mundo, se aclara mejor la interrogación que formula el pensador: “But is the China of today ready to take up the question of technology and to give it sufficient reflection from the perspective of its own culture and tradition?”.⁹⁰ Evidentemente, por cultura y tradición se refiere a la antes explicada “resonancia”, connatural a la susodicha “mentalidad filosófica china”, y diametralmente opuesta al pensamiento europeo cuyo pilar es el principio de no-contradicción. ¿Satisface esta salida tradicionalista⁹¹, este retorno a la cosmología moral antigua –salvaguar-

87 Véase sobre el sentido del eslogan junto a una exposición de la historia del desarrollo de las concepciones sobre la tecnología en Japón, INUTSUKA, Yū, “The historical development of the concept of technology in Japan”. *MECHANE. International Journal of Philosophy and Anthropology of Technology*, 2020, 235-245. <<https://mimesisjournals.com/ojs/index.php/mechane/article/view/338/253>> [Consulta: 16 abr. 2021].

88 HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotronics*, op. cit., 241-242.

89 Ibid., 7.

90 Ibid., 242.

91 A pesar del autor que en innumerables ocasiones ofrece a sus lectores excusas no pedidas que son acusaciones manifiestas, como se sabe. En una entrevista reciente dijo: “Cosmotécnica no es sinónimo de nacionalismo. No es fascismo, no es una identidad política”, en FONTEVECCHIA, Jorge, “Yuk Hui: ‘La cosmotécnica no es nacionalismo, no es fascismo, no es una identidad política’”, *El perfil*, 17 de octubre de 2020. <<https://www.perfil.com/noticias/periodismopuro/yuk-hui-la-cosmotecnica-no-es-nacionalismo-no-es-fascismo-no-es-una-identidad-politica.phtml>> [Consulta: 12 abr. 2021]. O en HUI, Yuk, “Máquina y ecología”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, op.

dar aquello que la recuerda “ceremonia del té, caligrafía, etc.”⁹² para esquivar tamaño futuro de destrucción y extinción que da por seguro nuestro autor? Según sostiene, nos dirigimos hacia el Apocalipsis:

[...] Después de la Guerra Fría, el incremento en la competencia ha resultado en una cultura monotecnológica que ya no busca el equilibrar progreso económico y progreso tecnológico, sino que los asimila al tiempo que avanza hacia un final apocalíptico. La competencia basada en la monotecnología está devastando los recursos naturales de la Tierra en aras de la maximización de ganancias e impide a los actores adoptar caminos o direcciones diferentes, es decir, bloquea la “tecnodiversidad”⁹³.

NOTA FINAL: DESACOPLAR EL FUTURO

Concluyendo, ¿es filosóficamente satisfactoria esta filosofía sino-humanística de la tecnología, contrapunto de la otra vía, la de “nueva ilustración”?

La conocida distinción entre apocalípticos e integrados de Umberto Eco puede resultar de utilidad para cerrar este estudio y enmarcar la encrucijada de la filosofía humanística de la tecnología en el siglo XXI que bascula entre la nueva ilustración europea, de la que queda pendiente un estudio pormenorizado, y la filosofía posteuropea china de Hui, extensamente anotada.

Eco distinguía entre las dos actitudes o reacciones del siguiente modo:

El Apocalipsis es una obsesión del *dissenter*, la integración es la realidad concreta de aquellos que *no disienten*. La imagen del Apocalipsis surge de la lectura de textos *sobre* la cultura de masas; la imagen de la integración emerge de la lectura de textos *de* la cultura de masas⁹⁴.

Hui imagina una caída irrecuperable⁹⁵ hacia el fin del mundo si se sigue la senda de la cultura monotecnológica –representa el perfil del disidente que describiera Eco. Y, como se ha visto, para evitar esa caída, esa destrucción, esa pérdida

cit., 131: “Lo local no tiene por qué ser sinónimo de etnocentrismo o nacionalismo. [...] Lo local no significa aquí política identitaria, sino la capacidad de reflexionar sobre el devenir tecnológico de lo local y de evitar replegarnos a alguna forma de tradicionalismo, para que múltiples localidades puedan estar en condiciones de inventar su propio pensamiento y futuro tecnológicos”.

92 HUI, *The Question Concerning Technology in China. An essay in Cosmotronics*, 223.

93 HUI, “Cien años de crisis. La cultura monotecnológica ante el brote epidémico”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, 95.

94 ECO, Umberto, *Apocalípticos e integrados*. Barcelona: Editorial Lumen, 1984, 13.

95 Ibid., 12.

de cosmos, en su lenguaje⁹⁶, defiende con vehemencia su antítesis de la “tecnodiversidad” que no resulta en una teorización filosóficamente robusta.

Tras este comentario filosófico interno de las tesis del autor, una parte del cual ha desvelado su “moldura” filosófica (auto)orientalista y, además, la falta de matices al presentar al lector euro-norteamericano recortes de la historia filosófica del este asiático, resulta y resalta, primero, su especulativismo ecléctico. Su apoyatura y dependencia de la tesis, también apocalíptica, de Heidegger y su epígono Nishitani es, en extremo, delicada. El calco de la pregunta heideggeriana por la técnica trasladado a China emborrona su propuesta filosófica, dejando sin atender cuestiones relevantes. Podando la hojarasca mística que rodea gran parte de la exposición de Hui, se podría discutir el indubitable papel que juega la historia sociocultural múltiple del mundo enfrentada a los retos de la revolución tecnológica evitando los riesgos de una caída en el relativismo moral. Como señala convenientemente Coeckelberg:

[...] dada la diversidad y pluralidad de perspectivas sobre este asunto [la IA] dentro de las sociedades y las diferencias culturales entre sociedades, las cuestiones sobre una vida buena y con sentido coexistiendo con la tecnología pueden recibir diferentes respuestas en distintos sitios y contextos, y en la práctica estarán sujetas a todo tipo de procesos políticos que no tienen por qué generar consensos. Reconocer esta diversidad y pluralidad puede llevarnos a un enfoque pluralista. Este también puede tomar la forma del relativismo⁹⁷.

Asimismo, esta orientación sino-humanística en la que predomina la desazón metafísica alrededor de la tecnología, y no el tratamiento concreto y específico de asuntos tecnológicos per se, rima con otros fenómenos. La “fragmentación” o “tecnodiversidad” puede ser el nombre filosófico para lo que, política y económicamente, ya hace años que se conoce como “desacoplamiento”. El discurso del “desacople” reaparece cada vez que se producen grandes crisis mundiales, económicas o de otra índole y apunta a la separación entre China y otros países emergentes de las economías más industrializadas y, en particular, de EEUU⁹⁸. Parece una interpretación viable, o, al menos, con sentido, que “fragmentar el

96 Es significativo que Hui ponga como pórtico de su *The Question Concerning Technology in China. An essay in Cosmotronics* (v), la siguiente cita de la obra *Apocalipsis* (1930) de D.H. Lawrence (1885-1930): “When I hear modern people complain of being lonely then I know what has happened. They have lost the cosmos”.

97 COECKELBERG, Mark, *Ética de la inteligencia artificial*. (Trad. Lucas Álvarez Canga). Madrid: Cátedra, 2021, 120.

98 RUIZ, Àlex, “¿El retorno del decoupling? Mito y realidad en el desacoplamiento de las economías emergentes.”, *Documentos de economía “la Caixa”*, 16, Diciembre 2009, 5. < https://www.caixabankresearch.com/sites/default/files/content/file/2016/09/de16_esp.pdf > [Consulta: 16 abr. 2021].

futuro” pueda leerse también como “desacoplar el futuro”. Hui no emplea ni la noción ni el verbo, pero sí otro tótem conceptual que ha inundado la literatura académica filosófica: descolonizar.⁹⁹ Fragmentar, desacoplar, descolonizar son componentes en el centro de la “filosofía posteuropea” de Hui y encajan perfectamente en lo que Pluckrose y Lindsay llaman “mentalidad postmoderna aplicada”:

The West has constructed the idea that rationality and science are good in order to perpetuate its own power and marginalize nonrational, nonscientific forms of knowledge production from elsewhere. *Therefore, we must now devalue white, Western ways of knowing for belonging to white Westerners and promote Eastern ones (in order to equalize the power imbalance)*¹⁰⁰.

Se puede afirmar que estas voces asiáticas de la diplomacia tecnológica parece que están sincronizadamente enredadas en el mismo nudo filosófico que padecemos en la academia europea tras adoptar acríticamente las lentes del constructivismo. Detrás de la “filosofía posteuropea”, además de las lagunas y problemas extensamente expuestos, se amaga, según lo describiera Boghossian, el más rudo constructivismo y, en particular, una conocida doctrina, “la de la validez igual”, elevada por amplios sectores a dogma de fe: “Existen muchas formas radicalmente distintas, pero “igualmente válidas”, de conocer el mundo, de las cuales la ciencia es sólo una”.¹⁰¹ Este dogma no solamente es contraintuitivo y se invalida de suyo cuando se buscan bases objetivas, evidenciadas y justificadas epistemológicamente. También representa una línea más de la deriva filosófica hacia cogitaciones simbólicas alejadas de lo material, lo real y lo verificable.

Concluyendo con una licencia metafórica: si lo que se quiere es rasgar el velo de seda de los retos que plantea la realidad tecnológica hace falta algo más que plumas especulativas.

99 HUI, “Prefacio”, *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, 12: “[...] Es en este contexto social y político que resulta urgente reabrir la cuestión de la tecnología y emprender la búsqueda de múltiples cosmotécnicas. Esta búsqueda es fundamentalmente un proyecto de descolonización, [...]”.

100 PLUCKROSE, Helen, LINDSAY, James, *Cynical theories. How Activist Scholarship Made Everything about Races, Gender, and Identity -and Why This Harms Everybody*. Durham: Pitchstone Publishing, 2020, 76. Cursivas en el original.

101 BOGHOSSIAN, Paul, *El miedo al conocimiento. Contra el relativismo y el constructivismo*. (Trad. Fabio Morales). Madrid: Alianza Editorial, 2012, 17.

REFERENCIAS BIBLIOGRÁFICAS

- ARAKI, Tōru, “2. Tokyo, 1942; Le colloque maudit [Dépassement de la modernité]”, *Ebisu*, 6, 1994, 75-95. DOI: <https://doi.org/10.3406/ebisu.1994.1586>.
- BAUER, Wolfgang, *Historia de la filosofía china*. (Trad. Daniel Romero). Barcelona: Herder, 2009.
- BOGHOSSIAN, Paul, *El miedo al conocimiento. Contra el relativismo y el constructivismo*. (Trad. Fabio Morales). Madrid: Alianza Editorial, 2012.
- CALICHMAN, Richard F. (Ed. y Trad.), *Overcoming modernity. Cultural identity in Wartime Japan*. New York: Columbia University Press, 2008.
- COECKELBERG, Mark, *Ética de la inteligencia artificial*. (Trad. Lucas Álvarez Canga). Madrid: Cátedra, 2021.
- CONFUCIO (Maestro Kong), *Lun Yu. Reflexiones y enseñanzas*. (Trad., Introd. y notas Anne-Hélène Suárez). Barcelona: Kairós, 2002.
- CRESPÍN PERALES, Montserrat, “Pensament xinès contemporani”, en PRADO-FONTS, Carles (Coord.), *Pensament i religió a Àsia Oriental*. Barcelona: FUOC, 2011, 50-68. <<http://hdl.handle.net/10609/49701>> [Consulta: 1 jun. 2021]
- , “Filosofía y pensamiento contemporáneo: Sincretismo japonés”, en FERNÁNDEZ, Julián (ed.), *Japón, el archipiélago de la cultura, Volumen 2: Tierra de Wa - Religión*. Barcelona: Mediatres, 2020, 135-207.
- , “Las máquinas de miembros conscientes y sus vigías: La ética de la inteligencia artificial en China y en Europa”, en PELEGRÍN, Àngels (ed.), *Encrucijadas: China desde el presente*. Granada: Editorial Comares, 2021, 113-147.
- ECO, Umberto, *Apocalípticos e integrados*. Barcelona: Editorial Lumen, 1984.
- ELMAN, Benjamin A., “The Failures of contemporary Chinese Intellectual History”, *Eighteenth-Century Studies*, 43, 3, 2010, 371-391. <<https://muse.jhu.edu/article/388529>> DOI: 10.1353/ecs.0.0191 [Consulta: 11 abr. 2021]
- FONTEVECCHIA, Jorge, “Yuk Hui: ‘La cosmotécnica no es nacionalismo, no es fascismo, no es una identidad política’”, *El perfil*, 17 de octubre de 2020. <<https://www.perfil.com/noticias/periodismopuro/yuk-hui-la-cosmotecnica-no-es-nacionalismo-no-es-fascismo-no-es-una-identidad-politica.phtml>> [Consulta: 12 abr. 2021].
- FORTUNY, Francesc J., “El problema dels orígens de la Modernitat”, *Anuari de la Societat Catalana de Filosofia-IEC*, 2, 1988, 108-111. <<http://www.raco.cat/index.php/AnuariFilosofia/article/view/211633>> [Consulta: 13 abr. 2021].
- FRANSENN, Maarten, LOKHORST, Gert-Jan, VAN DE POEL, Ibo, “Philosophy of Technology”. En: ZALTA, Edward N., *Stanford Encyclopedia of Philosophy*. [en línea] Stanford (Ca.): Stanford University, 2009 (actualización en 2018). <<https://plato.stanford.edu/archives/fall2018/entries/technology/>>. [Consulta: 13 abr. 2021]. ISSN: 1095-5054.
- FREUD, Sigmund, *Obras completas. Volumen 3 (1893-99). Primeras publicaciones psicoanalíticas*. Segunda reimpresión. Buenos Aires: Amorrortu editores, 1991.
- GÁLIK, Marián, “Quo Vadis Philosophia?: Musings about the Necessity of Sino-Western Dialogue”, *Journal of Chinese Philosophy*, 31, 2004, 65-81. DOI: <https://doi-org.sire.ub.edu/10.1111/j.1540-6253.2004.00143.x> [Consulta: 11 abr. 2021]

- GONZÁLEZ VALLES, Jesús, *Historia de la Filosofía Japonesa*. Reimpresión. Madrid: Tecnos, 2002.
- HALL, David L., AMES, Roger T., "Chinese philosophy". En: GRAIG, Edward, *Routledge encyclopedia of philosophy*. [En línea]. London: Routledge, 1998. <<https://www.rep.routledge.com/articles/overview/chinese-philosophy/v-1>> [Consulta: 13 abr. 2021]. DOI: 10.4324/9780415249126-G001-1
- HEGEL, G.W.F., *Introducciones a la Filosofía de la Historia Universal*. (Trad. Román Cuartango). Madrid: Istmo, 2005.
- HEIDEGGER, Martin, *Conferencias y artículos*. (Trad. Eustaquio Barjau). Barcelona: Ediciones del Serbal, 1994.
- , *Nietzsche*. (Trad. Juan Luis Verma). Barcelona: Ediciones Destino, 2000.
- , "Europe and German Philosophy", (Trad. Andrew Haas). *The New Yearbook for Phenomenology and Phenomenological Philosophy* [En línea], VI, 2006, 331-340. <https://www.pdcnet.org/nyppp/content/nyppp_2006_0006_0331_0340> [Consulta: 12 abr. 2021].
- HOWLAND, Douglas R., *Translating the West. Language and Political Reason in Nineteenth Century Japan*. Honolulu: University of Hawai'i Press, 2002.
- HU, Shi, *The development of the logical method in Ancient China*. Shanghai: The Oriental Book Company, 1922.
- HUI, Yuk, *The Question Concerning Technology in China. An essay in Cosmotechnics*. Falmouth: Urbanomic, 2016.
- , *Fragmentar el futuro. Ensayos sobre tecnodiversidad*. (Trad. Tadeo Lima). Buenos Aires: Caja Negra Editora, 2020.
- INUTSUKA, Yū, "The historical development of the concept of technology in Japan". *MECHANE. International Journal of Philosophy and Anthropology of Technology*, 2020, 235-245. <<https://mimesisjournals.com/ojs/index.php/mechane/article/view/338/253>> [Consulta: 16 abr. 2021].
- KAPP, Ernst, *Elements of a Philosophy of Technology. On the Evolutionary History of Culture*. (Ed. Jeffrey West Kirkwood and Leif Weatherby; Trad. Lauren K. Wolfe; Afterword Siegfried Zielinski). Minneapolis, London: University of Minnesota Press, 2018.
- KISHINAMI, Tsunezo, *The Development of Philosophy in Japan*. Princeton: Princeton University Press, 1915.
- KISSINGER, Henry, *China*. Barcelona: Editorial Debate, 2012.
- , "How the Enlightenment Ends", *The Atlantic*, June 2018. <<https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>> [Consulta: 07 abr. 2021].
- LOVINK, Geert, HUI, Yuk, "Interview: For a Philosophy of Technology in China". *Parrhesia* [En línea], 27, 2017, 48-63, <<http://parrhesiajournal.org/parrhesia27/parrhesia27.pdf>> [Consulta: 11 abr. 2021].
- LÖWITH, Karl, *Mi vida en Alemania antes y después de 1933. Un testimonio* (1940). (Trad. Ruth Zauner). Madrid: Visor, 1992.
- , *El hombre en el centro de la historia. Balance filosófico del siglo XX*. (Trad. Adan Kovacsics). Barcelona: Herder, 1998.

- MITCHAM, Carl, *Thinking through technology. Path between Engineering and Philosophy*. Chicago, London: The University of Chicago Press, 1994.
- NAKAMURA, Hajime, *Ways of Thinking of Eastern Peoples. India, China, Tibet, Japan*. Honolulu: University of Hawai'i Press, 1964.
- NEEDHAM, Joseph, *The Grand Titration: Science and Society in East and West*. Toronto: University of Toronto Press, 1969.
- PATTBERG, Thorsten, *The East-West Dichotomy*. New York: LoD Press, 2009.
- PETERS, Michael A., "The enlightenment and its critics". *Educational Philosophy and Theory* [En línea], 51, 9, 2019, 886-894. DOI: 10.1080/00131857.2018.1537832
- PIOVESANA, G.K., *Pensamiento japonés contemporáneo (1963)*. (Trad. J. M^a Sancho SJ). Madrid: Editorial Razón y Fe, 1967.
- PLUCKROSE, Helen, LINDSAY, James, *Cynical theories. How Activist Scholarship Made Everything about Races, Gender, and Identity -and Why This Harms Everybody*. Durham: Pitchstone Publishing, 2020.
- RIEU, Alain-Marc, "The syndrome of 'overcoming modernity': Learning from Japan about ultra-nationalism". *Transtext(e)s Transcultures* 跨文本跨文化 [En línea], 9, 2015-08-15, 1-2. <<http://journals.openedition.org/transtexts/552>> DOI : <https://doi.org/10.4000/transtexts.552> [Consulta: 07 abr. 2021].
- ROBLES, José María, "Yuk Hui, la nueva superestrella del pensamiento: 'El desarrollo tecnológico en China es tan rápido que da miedo'", *El Mundo*, 17 de noviembre de 2020. <<https://www.elmundo.es/papel/historias/2020/11/17/5fb4000921efa031448b4627.html>> [Consulta: 12 abr. 2021]
- RUIZ, Àlex, "¿El retorno del decoupling? Mito y realidad en el desacoplamiento de las economías emergentes.", *Documentos de economía "la Caixa"*, 16, Diciembre 2009, 5. <https://www.caixabankresearch.com/sites/default/files/content/file/2016/09/de16_esp.pdf> [Consulta: 16 abr. 2021].
- SÁNCHEZ RON, José Manuel, *La Nueva Ilustración: Ciencia, tecnología y humanidades en un mundo interdisciplinar*. Oviedo: Ediciones Nobel, 2011.
- SIVIN, Nathan, "Why the Scientific Revolution Did Not take Place in China -or Didn't It? *Chinese Science*, 5, 1982, 45-66. <<http://www.eastm.org/index.php/journal/article/viewFile/473/404>>. [Consulta: 16 abr. 2021].
- SWALE, Alistair D., *The Meiji Restoration. Monarchy, Mass Communication and Conservative Revolution*. London: Palgrave MacMillan, 2009.
- VV. AA., *¿Hacia una nueva Ilustración? Una década trascendente*. Madrid: OpenMind/BBVA, 2019. <<https://www.bbvaopenmind.com/libros/hacia-una-nueva-ilustracion-una-decada-trascendente/>> [Consulta: 07 abr. 2021]

MISCELÁNEA

TRANSHUMANISMO, EL ÚLTIMO SUSPIRO DE DIOS: ANÁLISIS DEL PROYECTO TRANSHUMANISTA DESDE LA FILOSOFÍA POSTHUMANISTA POSMODERNA

*TRANSHUMANISM, THE LAST BREATH OF GOD:
ANALYSIS OF TRANSHUMANISM PROJECT FROM
POSTHUMANISM POSTMODERNIST PHILOSOPHY.*

LUIS LAVINA FAUSTMANN

Doctorando de Filosofía
Facultad de Filosofía
Universidad de Barcelona
Barcelona/España
luislavina@hotmail.com
ORCID: 0000-0002-1347-9845

Recibido: 2/12/2021
Revisado: 16/06/2021
Aceptado: 6/09/2021

Resumen: El presente artículo analiza desde una postura crítica la ideología transhumanista y su defensa de la necesidad inminente de la mejora técnica hacia la sociedad posthumana. Para articular la crítica, se aborda el concepto de posthumanismo. Esta noción de posthumanismo es comprendida de manera diametralmente opuesta por los autores transhumanistas y los posthumanistas posmodernos. El análisis de la diferente forma de comprender la superación de lo humano de ambas corrientes permitirá finalmente contraponerlas para denunciar cómo el proyecto transhumanista es un nuevo humanismo. En ese sentido, y en tanto que filosofía heredera de la Ilustración, supone una aspiración peligrosa, dado que mantiene una confianza arrogante en la razón y la tecnología, está profundamente despoliticizada, se contrapone necesariamente a aquellos considerados menos humanos, y mantiene en su núcleo un liberalismo individualista y visión utópica/escatológica del progreso.

Palabras clave: Bioética, Transhumanismo, Técnica, posthumanismo, posmodernidad.

Abstract: This article proposes a critical view of the transhumanist ideology and its support to the supposed imminent and necessary technical enhancement towards a posthuman society. The concept of posthumanism is analyzed in order to articulate the critique. Posthumanism is understood in opposed ways by transhumanist authors and postmodern posthumanists. The analysis of both ways of comprehending the overcoming of the human

is useful because the contraposition shows that the transhumanist project is a new humanism. In this sense, and as a philosophy derived from the Enlightenment, transhumanism is a dangerous aspiration that: maintains an arrogant confidence in reason and technology, is deeply depoliticized, is necessarily based on the opposition to those considered less human, and maintains at its core an individualistic liberalism and a utopian/eschatological vision of progress.

Keywords: Bioethics, Transhumanism, Technique, posthumanism, postmodernity.

1. EL PROYECTO DE MODIFICACIÓN HUMANA: LA PREGUNTA DE NUESTRA ERA

En los últimos dos siglos la técnica se ha desarrollado a pasos agigantados, aumentando las posibilidades y los límites de la acción humana. Con el continuo avance de estas libertades técnicas surgen nuevas responsabilidades éticas y nuevas cuestiones filosóficas. Sin embargo, la pregunta de hasta dónde debemos desarrollar las posibilidades que la nueva tecnología propone cobra ahora un sentido más radical que nunca. Por vez primera no es el medio natural, lo externo, sino la propia naturaleza humana lo que se muestra como técnicamente disponible. Los avances en ingeniería genética, que han dado los primeros resultados efectivos en clonación y modificaciones del ADN; la informática, con el desarrollo de inteligencias artificiales y las primeras síntesis de seres orgánicos y cibernéticos (cyborgs); o la farmacología, con medicamentos altamente eficaces en la regulación del comportamiento y el ánimo, han encontrado un cada vez más numeroso grupo de autores que ven de manera optimista estas posibilidades como el camino hacia la “necesaria superación” del ser humano: los transhumanistas.

Esta filosofía transhumanista es definida como “a way of thinking about the future that is based on the premise that the human species in its current form does not represent the end of our development but rather a comparatively early phase” [una forma de pensar sobre el futuro que se basa en la premisa de que la especie humana en su forma actual no representa el final de nuestro desarrollo, sino una fase relativamente temprana]¹. Así, el transhumanismo propone que el ser humano abandone su posición pasiva con respecto su propia evolución, sustituyendo “la selección natural por una selección consciente”². A pesar de que el término transhumanista englobe una serie de autores con posturas diversas, como

1 Bostrom, N. *The Transhumanist FAQ* [en línea]. Version 2.1. Oxford: World Transhumanist Association, 2003. <<https://www.nickbostrom.com/views/transhumanist.pdf>> [Consulta 15/02/2021] p. 4.

2 Harris, J. *Enhancing Evolution: The Ethical Case for Making People Better*. Nueva York: Princeton University Press, 2007.

inmediatamente se verá, lo esencial de su posición es una defensa optimista de la mejora de la condición humana por medios técnicos.

El presente ensayo trata de formular un enfoque desde el que articular una crítica al proyecto transhumanista. Para ello, se propone una mirada global que atienda a las raíces filosóficas y las aspiraciones de dicha ideología. En ese sentido se hace un análisis del manido término de posthumanismo. El recorrido por este concepto y las distintas maneras de entenderlo resulta de entrada relevante porque hay una gran confusión en la bibliografía académica en cuanto a su significado. El posthumanismo es un término en disputa, y tanto los transhumanistas como los posthumanistas críticos (de vertiente posmoderna) lo utilizan para referirse a un proyecto en el cual el concepto de humano sea superado. Centrar este ensayo en torno al término de posthumanismo puede parecer que nos aleja de la cuestión principal –esta es, si debe el ser humano mejorarse–. Sin embargo, entender las diferencias fundamentales que existen entre la forma de pensar *lo posthumano* de los transhumanistas y de la vertiente posmoderna nos servirá para el objetivo final de este trabajo. El trabajo plantea un criterio para distinguir el término posthumanismo propio de la posmodernidad, de la búsqueda de una sociedad posthumana fruto del transhumanismo. Mientras que el posthumanismo realiza una crítica que trata de deconstruir el concepto de humano y señalar cómo el proyecto humanista es una herramienta de dominación; el transhumanismo en última instancia necesita redefinir y remarcar las fronteras de lo humano, además de ser una corriente heredera de las filosofías humanistas. La tesis principal de este ensayo es que esta diferencia es crucial, dado que sitúa a ambas corrientes en sentidos opuestos; y que en última instancia dicha distinción sirve para desentrañar los peligros del transhumanismo.

Como segundo pilar desde el que articular la crítica, se apuntan algunas nociones acerca de la crítica posmoderna de las ideas ilustradas de progreso y utopía. Estas ideas son relevantes en primer lugar para contextualizar la corriente posthumanista posmoderna. Pero, además, permiten dismantelar de nuevo lo problemático de la utopía técnica transhumanista. La confianza transhumanista en el progreso técnico supone una secularización de las nociones teológicas de providencia y salvación.

En definitiva, el objetivo principal del trabajo es demostrar cómo, desde los pilares interrelacionados de posthumanismo y posmodernidad, se puede formular una crítica certera que sirva para desentrañar los peligros de la ideología transhumanista. El transhumanismo es un nuevo humanismo, que se nutre de ideas propias de la Ilustración pero que además está preocupantemente despolitizado. En tanto que proyecto humanista, se debe asentar sobre la distinción de aquello considerado *menos humano*. El transhumanismo mantiene en su seno un ideal de emancipación muy exclusivo y que se articula en referencia al Otro, al que

debe dominar. Y, en tanto que nueva utopía, demuestra una confianza arrogante en la razón y en las ilusiones de emancipación, progreso y utopía, que lo hacen instrumento idóneo para dominar bajo una supuesta legitimidad.

2. POSIBILIDADES TÉCNICAS- LA PREGUNTA NECESARIA

El proyecto transhumanista busca la mejora del ser humano a través de la técnica. Los medios por los que se puede alcanzar dicha superación de las limitaciones del ser humano son diversos, pero se pueden establecer dos modalidades distintas aunque complementarias: las mejoras informáticas y el biomejoramiento humano³. En la primera vertiente encontramos prácticas tan diversas como el desarrollo de nuevas facultades por medio de la fusión con la máquina⁴, el control de las decisiones e impulsos violentos por medio de electrodos⁵, o el uploading de la mente a una interfaz informática, entre otros. Por el lado del biomejoramiento las prácticas más importantes son: la modificación del ADN de la descendencia, la terapia génica en uno mismo, y el control de emociones y el ánimo a través de la farmacología⁶.

No es del interés de este ensayo detenerse entre las diferencias y razones de las distintas corrientes políticas y filosóficas transhumanistas (transhumanistas democráticos, libertarios, etc), puesto que en su mayoría estos proyectos comparten la misma vocación de mejorar técnicamente al ser humano y las diferencias

3 Esta distinción ha sido tomada de Diéguez, Antonio. *Transhumanismo, la búsqueda tecnológica del mejoramiento humano*. Barcelona: Herder, 2017.

4 Por ejemplo, Neil Harbisson se considera a sí mismo como el primer cyborg. Para profundizar, Harbisson, N. I listen to color [vídeo en línea] Ted Talks, 2012 <https://www.ted.com/talks/neil_harbisson_i_listen_to_color?> [Consulta: 1/02/2021].

5 José Manuel Rodríguez Delgado, neurólogo de origen español, fue pionero durante el pasado siglo en sus estudios sobre el control mental por medio de impulsos eléctricos. Apostaba porque dichos controles permitiesen evitar conductas violentas y otros problemas conductuales alcanzando una sociedad con mayor grado de civilización. Para profundizar, ver, Rodríguez Delgado, J. M. *Control físico de la mente: hacia una sociedad psicocivilizada*. Madrid: Espasa-Calpe, 1983

6 Una primer intuición parece indicar que la segunda opción, la modificación de nuestra propia estructura orgánica es la que produce un mayor vértigo moral. Esta distinción se debe en parte a que la mejora genética suele entenderse como eugenesia, esto es, mejora de la descendencia. En cambio, la robotización del ser humano parece verse como un proceso que afecta al propio individuo, y que siempre puede revertirse voluntariamente. Sin embargo, cuando se analizan de manera crítica, ambas posturas tratan de pre-diseñar al ser humano, de hacerlo "máquina" en función de unos objetivos o finalidades concretas. Además, parece probable considerar que serán las TIC las que tengan un efecto de transformación más profundo a corto plazo en las sociedades actuales. De hecho, cabe preguntarse en Occidente, donde una gran mayoría de la población tiene la posibilidad de acceso instantáneo a un dispositivo informático que lleva siempre consigo, como es el móvil, si no somos ya una sociedad cyborg.

que presentan, cuando se analizan críticamente dichas posturas, son sutiles⁷. La esencia del proyecto transhumano anhela llevar esta mejora hasta la superación, hasta la trascendencia de la propia especie humana para alcanzar una condición posthumana. La argumentación bioética que utilizan busca centrarse en el *qué* mejorar y en el *cómo* distribuir los medios para hacerlo, olvidando la pregunta de *por qué* mejorar. Para evaluar éticamente el transhumanismo, más que acudir a posturas esencialistas o a hipotetizar consecuencias bioéticas, se debe analizar cuál es su proyecto posthumano, en que valores y perspectivas se apoya. En palabras de Le Dévédec: “More than a bioethical issue to be managed and regulated, human enhancement is a fundamentally political issue – in the philosophical and basic sense of the word – “[más que una cuestión bioética que regular y gestionar, la mejora humana es un asunto fundamentalmente político –en el sentido filosófico y elemental de la palabra–]”⁸.

En las próximas secciones se analizarán las raíces filosóficas del proyecto transhumano. Para ello, primero abordaré el polémico concepto de *posthumanismo* en dos vertientes diferenciadas. La primera, se refiere a la consecución final del proyecto transhumano, una sociedad posthumana, *más* que humana. La segunda, que está íntimamente ligada a la filosofía posmoderna, se basa en una defensa de una sociedad que haya superado la categoría de “humano”, puesto que esta encierra en su seno una ambición de dominar a colectivos racializados, a la mujer, a los seres vivos, a la misma naturaleza... Posteriormente, la presente reflexión tratará de demostrar cómo ambos posthumanismos son antagónicos. En gran medida, el posthumanismo trata de borrar los límites de la categoría de humano, de deconstruir dicha noción puesto que considera que este tipo de definiciones esenciales han sido vehículo histórico para legitimar la dominación. Al contrario, el proyecto transhumanista precisa de apoyarse en una doble definición

7 Dentro del transhumanismo un sector defiende que las mejoras tecnológicas del ser humano deben de hacerse con criterios de igualdad social, en lo que ha sido definido como transhumanismo democrático (ver sobre todo la obra de James Hughes, o la postura de Peter Singer). Sin embargo, este sector es minoritario, dado que realmente hay un escaso debate político en torno al problema de la desigualdad en las mejoras entre las filas transhumanistas. De hecho, muchos de los autores más importantes sostienen, desde un enfoque radicalmente liberal, que deben ser los individuos quienes de manera autónoma decidan qué quieren modificarse (de este último sector ver sobre todo la obra de Max More y su Extropian Institute, o a Nozick defendiendo el “supermercado genético” en NOZICK, R. *Anarquía, Estado y Utopía*. México: FCE, 1988, p. 302). Sin embargo, si bien entre ambas posturas hay un conflicto en torno a cómo repartir las mejoras, todas las corrientes transhumanistas asumen que es posible, alcanzable y deseable mejorar al ser humano. El grueso de la crítica filosófica aquí efectuada va hacia las raíces de dicho proyecto, a los problemas de la aspiración de mejorar y trascender la humanidad, por lo que atañe a los dos bandos del debate político transhumano.

8 Le Dévédec, N. *Unfit for the future? The depoliticization of human perfectibility, from the Enlightenment to transhumanism*. *European Journal of Social Theory*, 21, 4, 2018, 488-507; p. 500.

cerrada sobre lo humano. Como se verá, los transhumanistas definen la naturaleza humana con una serie de características concretas que tienen unas implicaciones filosóficas y políticas importantes: la naturaleza humana es deficiente, el ser humano es un ser esencialmente técnico, es posible y deseable trascender la propia humanidad... A la vez, en tanto que asume que es abordable el reto de definir hacia dónde debemos mejorar, el transhumanismo necesita redefinir una noción cerrada de aquello que debemos ser: unas características cerradas de lo que es un transhumano.

Así, finalmente se aprovechará la crítica que hace el posthumanismo de vertiente posmoderna a la categoría de “humano” y al proyecto “humanista” para dismantelar las aspiraciones del proyecto posthumanista de los transhumanos, desenmascarando finalmente las incongruencias y equivocaciones intrínsecas de esta última utopía. En definitiva, el transhumanismo es el último hijo de la modernidad ilustrada, la última gran narrativa, en la que un liberalismo individualista y una visión utópica/escatológica del progreso técnico constituyen su núcleo fundamental.

3. POSTHUMANISMO(S)

El posthumanismo es un “concepto paraguas”, tal y como lo define la académica Francesca Ferrando⁹. Con la idea de concepto paraguas Ferrando se refiere al hecho de que el término posthumano es actualmente un concepto en disputa, que se piensa y se quiere resignificar desde diversas perspectivas en ocasiones muy divergentes: anti-humanismo, New Materialism, transhumanismo, posthumanismo cultural y posthumanismo crítico, etc. Todas ellas, sin embargo, concuerdan en la idea fundamental de que el concepto de humano no sirve ya para hablar del ser humano del siglo XXI y que, por lo tanto, es necesario redefinir nuestra concepción de lo humano a través del término “posthumano”. En el presente trabajo nos ocuparemos de dos de estas vertientes que se disputan el término: por un lado, el posthumanismo como proyecto de mejora transhumanista; y por otro, los posthumanismos cultural, crítico y filosófico, de clara vertiente posmoderna.

En aras de la concreción, resulta necesario resaltar que los límites entre ambas posturas son difusos, y que una hay una gran confusión en la bibliografía académica a la hora de delimitar correctamente las diferencias existentes entre

9 Ferrando, F. “Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations” *Existenz*, 8, 2, 2013, 26-32.

cada postura.¹⁰ En primer lugar, esta confusión se debe a que ambos posthumanismos comparten una crítica a entendernos como humanos, puesto que esta categoría encierra unos prejuicios que nos impiden repensarnos y avanzar hacia una nueva sociedad (aunque los fines del transhumanismo y del posthumanismo posmoderno sean distintos). De hecho, algunos autores transhumanistas, como Max More, se han apropiado de manera confusa o tergiversada de la filosofía posthumana posmoderna y su crítica al concepto de humano. En segundo lugar, muchos de los autores posthumanistas posmodernos entienden los cambios técnicos en el cuerpo humano como una estrategia óptima para escapar de las categorías fijas de “lo humano”.¹¹ Del mismo modo, autores que defienden un proyecto transhumanista, como Sloterdijk, comparten una visión similar a la de los autores posmodernos acerca de los límites y errores del proyecto humanista, aunque extraen conclusiones opuestas.¹² En definitiva, hay numerosos ejemplos en los que confluyen de un modo u otro una defensa de la modificación humana por medio de la técnica, junto con una conciencia crítica de raíz posmoderna sobre la necesidad de deconstruir la noción de humano. Pese a esta aclaración,

10 Por ejemplo, Antonio Diéguez, en *Transhumanismo*, *op. cit.*, habla de transhumanismo cultural para referirse a la vertiente posthumanista posmoderna. Y son otros muchos los que utilizan el término posthumanismo refiriéndose exclusivamente al transhumanismo: Pérez Escudero, A. “Darwin y el posthumanismo.” *Eikasia, Revista filosofía*, 32, 2010, 170-201; o, Fukuyama, F. *Our posthuman future: Consequences of the biotechnology revolution*. New York: Farrar, Straus and Giroux, 2003. Pepperell, por su parte, se refiere al transhumanismo como una fase previa de transición hacia la sociedad posthumana, Pepperell, R. *The Posthuman Condition. Consciousness beyond the brain*. Portland: OR: Intellect Books, 2003.

11 Por ejemplo, Dona Haraway concluye su *Manifiesto Cyborg* con la afirmación: “prefiero ser un cyborg que una diosa” (p. 38). Con esta sentencia Haraway recalca que la categoría de Cyborg nos permite escapar de un concepto cerrado de humano, siempre pensado en términos de género binarios. La autora Shulamith Firestone también defiende que la tecnología es el medio por el que superar la diferencia sexual que permita crear una sociedad postgénero. Ambos son buenos ejemplos de teorías posthumanistas que, en este caso desde el feminismo, consideran la tecnología como la herramienta fundamental para deconstruir la categoría de humano (por lo que transhumanismo y posthumanismo convergen). Haraway, D. *Manifiesto Cyborg, El sueño irónico de un lenguaje común para las mujeres en el circuito integrado* (Trad. Manuel Tanel). University of California, 1984.

12 Sloterdijk, en el polémico discurso de Elmau, realiza una dura crítica al proyecto humanista. El filósofo entiende que este se fundamenta en la distinción del humanizado y el bárbaro, y que esta diferencia establece una distancia insalvable entre criadores (los que establecen su “proyecto de mejora” o educación concreto), y los criados. Sin embargo, sus conclusiones parecen indicar que las contradicciones del proyecto humanista son inevitables, y que, por lo tanto, lo mejor que se puede hacer es aceptar las distinciones anteriores y desarrollar una antropotecnia, o técnica del mejoramiento humano, que no enmascare dichas diferencias entre criadores y criados (como sí hacía el proyecto humanista), sino que las haga explícitas. Sloterdijk es un ejemplo de un autor que, desde una visión crítica con el humanismo, acaba llegando sin embargo a una postura transhumanista. Sloterdijk, P. “Reglas para el Parque Humano. Una respuesta a la Carta sobre el Humanismo” Conferencia pronunciada en el castillo de Elmau, Baviera, 1999 (Trad: Adolfo Váquez Rocca). Consultado online en *Revista Observaciones Filosóficas*.

diferenciar las dos corrientes sigue siendo crucial, dado que el interés último de ambos proyectos es diferente y, en gran medida, contradictorio.

Anteriormente mencionamos que todo posthumanismo se basa en la tesis de que el término de humano no es adecuado para referirse a la (post)humanidad de la actualidad y del futuro. Sin embargo, el posthumanismo posmoderno –de ahora en adelante posthumanismo– y el posthumanismo transhumanista –de ahora en adelante transhumanismo, salvo que se especifique lo contrario– entienden de manera opuesta, por un lado, en qué sentido el término de humano ya no es válido, y, en segundo lugar, cómo se debe superar este término. Esto es, ¿qué significa ser posthumano?

3.1. EL TRANSHUMANISMO Y EL HUMANO COMO PREJUICIO

Los transhumanistas tienden a explicar que la técnica hace tiempo que ha ido alterando nuestro mundo natural y nuestra propia naturaleza, por lo que la mejora del ser humano no es más que una continuación de la historia natural del ser humano. En la misma línea, Savulescu explica que no hay una diferencia moral relevante entre tomar hierbas curativas para paliar el dolor de una enfermedad, e introducir información genética de otras especies para permitir la regeneración continua de células humanas cuya degeneración causa la vejez¹³. Según su punto de vista, la vida tiene valor “instrumental”. Así, una enfermedad supone un daño moral porque impide realizar ciertas actividades agradables, como disfrutar de la familia. Del mismo modo, alargar la vida es éticamente bueno porque permite realizar esas mismas actividades. Aceptando esta perspectiva, la conclusión es que no hay una diferencia radical entre mejorar y curar, y que el ser humano siempre ha utilizado la técnica como instrumento para satisfacer sus deseos y necesidades. La sociedad posthumana sólo representa entonces un paso natural de la humanidad: el desarrollo histórico de la técnica nos ha ido permitido progresivamente cada vez más medios de cura y mejora¹⁴.

Continuando con esta perspectiva, los transhumanistas opinan que lo que se conoce como ser humano es el resultado de un proceso natural de evolución, y que el siguiente paso, el posthumanismo, es por lo tanto igual de natural. En este punto nos detendremos posteriormente, puesto que la lectura transhumanista del

13 Savulescu, J. *¿Decisiones peligrosas? Hacia una bioética desafiante*. Madrid: Grupo Anaya, 2012

14 Este argumento de la continuidad, compartido por la mayoría de los transhumanistas, ha sido extensamente criticado en López Frías, F. J. “Continuidad de las innovaciones tecnológicas: el reto de las intervenciones biomédicas de mejora humana” *ISEGORÍA* 48, 2013, 213-228.

materialismo darwiniano conduce de manera paradójica a un existencialismo absoluto y a entender al hombre como posibilidad infinita. Lo que nos ocupa por ahora es entender que para los transhumanistas el concepto humano debe ser desechado porque se basa en un prejuicio –en el sentido de Francis Bacon, esto es, en una creencia no racional, no científica– que impide y frena el avance científico y de la especie humana. Los transhumanistas se oponen a entender al ser humano como ente con una esencia determinada o inamovible. Max More lo resume así: “los transhumanistas no ven la naturaleza humana como un fin en sí mismo, ni como perfecta, ni como poseedora de ningún derecho a nuestra lealtad”¹⁵. En definitiva, la crítica del transhumanismo al concepto de humano podría resumirse en la idea de que la existencia de una esencia humana fija es una creencia falsa, un prejuicio con altas connotaciones religiosas que frena el avance científico y la libertad individual de auto transformarse. Además, estos autores señalan que, de existir algo así como la esencia humana, no hay ningún argumento moral de peso que nos impida alterarla. Los transhumanistas rechazan el concepto de humano puesto que esta es una *categoría prejuiciosa* que nos ancla en el pasado, es una categoría conservadora.

3.2. EL POSTHUMANISMO Y EL HUMANO COMO MECANISMO DE PODER.

El posthumanismo de influencia posmoderna vierte toda su atención al problema y las premisas del concepto de humano. Es esta una crítica mucho más profunda y radical. El posthumanismo trata de desenmascarar las mentiras de *lo humano* no únicamente como prejuicio del pasado, si no como mecanismo de poder. Los autores de esta vertiente evidencian que históricamente el concepto de lo humano ha servido como una herramienta de dominación¹⁶. Por ejemplo, el iusnaturalismo ilustrado, derivaba de una supuesta naturaleza humana una serie de privilegios o derechos que reservaba únicamente a aquellos humanos que fuesen hombres, blancos, heterosexuales y propietarios. Francesca Ferrando expone que el concepto de humano, si se analiza históricamente, ha sido siempre exclusivo: “no todos los seres humanos eran considerados como tal”¹⁷. De hecho, la noción

15 Citado en A. Diéguez, *op. cit.*, p.33. Original de More, Max, “The Philosophy of Transhumanism”. En More, M./Vita-More, N. (eds.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology and Philosophy of the Human Future*. Número monográfico, Chichester, John Wiley & Sons, 2013, pp. 3-17; p. 4.

16 Ahluwalia, P. “Two Senses of the Post in Posthumanism”. En: Banerji, D. & Paranjape, M. R. (Eds), *Critical Posthumanism and Planetary Futures*. New Delhi: Springer, 2018, 131-141, p. 134.

17 Ferrando, “Transhumanism, posthumanism...”(*op. cit.*), p. 28.

de humano no es peligrosa por el mero hecho de excluir por olvido a determinados colectivos de dicha categoría, sino porque necesita de esa misma exclusión para formularse. El humano siempre se enfrenta a lo Otro. Esta manera dicotómica de pensar es, según la filósofa italiana, consustancial a la filosofía occidental:

Western hegemonic accounts had mostly been dualistic, based on opposites which can be exemplified in the classical pairs: nature/culture, female/male, black/white, gay/hetero etc. Such a dualistic approach reflected the historical praxis of war («us» against the «enemy»). [La hegemonía cultural occidental ha sido mayoritariamente dualística, basada en oposiciones que pueden ser ejemplificadas en los pares clásicos: naturaleza/cultura, mujer/hombre, blanco/negro, gay/hetero etc. Este enfoque dualista reflejaba la praxis histórica de la guerra (“nosotros” contra el “enemigo”)]¹⁸.

La autora italiana expresa aquí algunas de las categorías o parejas más evidentes, sin embargo, el concepto moderno de humano está cargado y relacionado con diversas dualidades antagónicas más, que se interrelacionan y connotan mutuamente. Por ejemplo, los dualismos cartesianos e ilustrados mente/cuerpo, razón/emoción, verdad objetiva/ mentira subjetiva, se traducen a una lectura binaria de los sexos en la cual la mujer queda representada como ente guiado por la pasión y el instinto, por lo sensible, cuya esencia es la de ser un cuerpo de deseo o de reproducción. La mujer queda reflejada en ese espejo como menos humana: es animal, es natural (vs civilizada), es elemento de engaño y mentira, corpórea, sangrante y sensible.

Resulta crucial entender cómo las categorías dicotómicas de la modernidad se superponen, se connotan unas a otras, creando una determinada idea de lo humano que se fundamenta en la diferencia frente a lo menos humano (mujer o bárbaro) o a lo no humano (animal o naturaleza). En definitiva, los posthumanistas denuncian que la noción de humano esconde y significa un afán de dominio, de control, de poder¹⁹. Este dominio se manifiesta de dos maneras distintas. Por un lado, lo humano como superior está en derecho de disponer de *aquello que no es humano*: la contrapartida a los derechos humanos es, para los posthumanistas, la destrucción del medio ambiente (antropoceno) y el dominio del resto de seres vivos (especismo). En segundo lugar, encontramos que el término humano históricamente ha sido definido con las características de aquellos que detentan el poder. Así, lo humano es entendido como lo masculino, occidental, normativo, etc; contraponiéndose a *lo que es menos humano*. Es decir, hay humanos que lo son más que otros. Los posthumanistas han evidenciado cómo el concepto de humano va

18 Ferrando, F. “Posthumanism”. *Tidsskrift for kønnsforskning* nr, 2, 2014b,168-172; p. 169.

19 Ahluwalia, P. “Two Senses of the Post in Posthumanism”, *op. cit.*, p. 134.

ligado al concepto de humanismo, como proyecto de guía “civilizadora”, y cómo este es en realidad un instrumento de dominación.

El mito del Estado de Naturaleza en la filosofía política moderna es ejemplo que evidencia como, debido a que se formula a través de las categorías dicotómicas que enuncia Ferrando, la noción de humano es el mecanismo por el que se articula y legitima la opresión²⁰. Este Estado de Naturaleza es caracterizado en su versión más radical como una etapa de conflicto continuo e instintivo por unos recursos limitados, por la propia seguridad, etc. No es casual que Hobbes comparase en aquella su cita al humano con los lobos: en ese Estado Natural los humanos son bestias, salvajes, aún dominados por los instintos y la violencia²¹. Sin embargo, el contractualismo se basa en la afirmación de que es posible trascender a ese conflicto continuo gracias al pacto racional. Así, la razón permite el salto hacia una existencia política y ciudadana, verdaderamente humana.

La trampa aparece entonces en un segundo momento, cuando aquellos que han adquirido su renovada posición de ciudadanos, los más humanos en tanto que racionales, se declaran legítimos guías del proceso de humanización. Así, los europeos en la Ilustración, gracias a su desarrollo y civilización, se conciben como superadores del Estado de Naturaleza. Ellos aparecen representados como libres, científicos, racionales y políticos. Frente a ellos, y desde su perspectiva, un gran número de sociedades del mundo se mantenían sin civilizar, sin progreso²². Siguiendo con el análisis de las dicotomías interrelacionadas, estas sociedades sin política ni ciencia estaban más cerca del estado de naturaleza, eran más salvajes, menos libres, más animales...

Los europeos modernos, de este modo, consideraban que tenían el derecho, o más bien, la obligación, de guiar hacia esa libertad –civilizar– a las otras culturas y países²³. John Stuart Mill, un autor relativamente progresista para su época que

20 En Ferrando, F. “Posthumanism” (op.cit).

21 En las lecturas del Estado Natural más benevolentes, como en Rousseau, el humano se sigue significando desde lo animal, aunque se haga desde una naturaleza pacífica e ingenua. La referencia directa a lo natural en ese Estado de Naturaleza no deja lugar a dudas: lo que define lo verdaderamente humano es su diferencia con lo salvaje y bestial. Y esa diferencia está en lo racional, en la política, en el control humanista de sus instintos, etc.

22 Para profundizar en esta concepción dicotómica de la naturaleza y sus implicaciones políticas se recomienda, Latour, B. *Nunca fuimos modernos. Ensayo de antropología simétrica*. Madrid: Siglo veintiuno editores, 2007.

23 Aquí se ha optado por describir este proceso desde la crítica a la colonización. Lo que se ha tratado de señalar es como no sólo convivían, sino que encajaban de manera eficaz, los discursos políticos que defendían una humanización de sus conciudadanos por medio del autogobierno, con aquellos otros discursos contemporáneos que justificaban el dominio colonial. Pero, del mismo modo, se pueden encontrar ejemplos similares que legitiman otro tipo de opresiones. Por ejemplo, en Locke, J, en *Segundo ensayo sobre el gobierno civil: un ensayo sobre el verdadero origen, alcance y finalidad del gobierno civil*. (C. Piña, Trad.) Buenos Aires: Editorial Losada, 2002; en las secciones que tratan

defendía la extensión del sufragio en Reino Unido, escribía a la vez: “India nos necesita, estos son territorios y gente que suplica por nuestra dominación, [...] Sin Inglaterra la India caería en la ruina”²⁴. El hombre blanco moderno, libre, culto, instruido, racional y científico –el humano–, se contraponen y debe dominar al negro primitivo, esclavo, idiota, instintivo, y supersticioso. El proceso mediante el cual se ejerce este dominio es el proyecto “humanista”. Sartre resume en el prefacio al libro *Los condenados de la tierra* los resultados últimos de este proyecto en Argelia:

Primero hay que afrontar un espectáculo inesperado: el striptease de nuestro humanismo. Helo aquí desnudo y nada hermoso: no era sino una ideología mentirosa, la exquisita justificación del pillaje; sus ternuras y su preciosismo justificaban nuestras agresiones²⁵.

El posthumanismo cultural y crítico, se ve claro ahora, mantiene una herencia con la teoría decolonial y la denuncia de autores como Fanon o Said. Algunos autores han simplificado, en mi opinión equivocadamente, este posthumanismo crítico apuntando a que este es poscolonialismo con una moral antiespecista. Por ejemplo, Monirul Islam afirma: “Posthumanism plays against the revered term ‘human’ and valorises the nonhuman other. Postcolonialism counters racism, and posthumanism counters speciesism, because speciesism for the posthumanist is comparable to sexism and racism” [El posthumanismo se enfrenta al venerado término de humano y trata de reconocer al otro no humano. El poscolonialismo se opone al racismo, y el posthumanismo se opone al especismo, porque el especismo es para los posthumanistas comparable al sexismo o el racismo] ²⁶.

la justificación del derecho de propiedad, el filósofo defiende que algunos con mayor capacidad para planificar son los que acumulan propiedad privada. Esto, lejos de ser un agravio, es positivo para el conjunto de la sociedad. Además, en Locke, puesto que aquellos propietarios son los que planifican de manera coherente, y los que tienen interés en que se proteja la seguridad y la propiedad privada, son los únicos con derechos de participación política en un modelo de sufragio censitario muy restrictivo. De este modo, Locke justicia a la vez la desigualdad social y política haciendo referencia a las diferentes capacidades de razonar: los más racionales son los que crean las instituciones civilizadoras, como el dinero y el Estado; el resto, sin embargo, están ligados a los instintos cortoplacistas e irracionales del Estado Natural. La razón, que es lo que nos permite superar la naturaleza más animal del Estado de Naturaleza, está presente en mayor grado en aquellos que son propietarios y que deben, por lo tanto, dominar políticamente. Para profundizar en la antropología desigual de Locke, consultar, Macpherson, C. *La teoría política del individualismo posesivo: De Hobbes a Locke*. Madrid: Trotta, 2005.

24 Mill, J. S, citado en, Said, E. *Culture and imperialism*. Nueva York: Vintage, 1994, p. 66.

25 Sartre, J. P. “Prefacio”. En Fanon, F. *Los condenados de la tierra*. México: Fondo cultura económica, 1961, p. 21.

26 Islam, M. “Posthumanism: Through the postcolonial lens”. En Banerji, D. & Paranjape M. R. (Eds), *Critical Posthumanism and Planetary Futures*. New Dehli: Springer India, 2016, p. 120.

Aunque una conciencia ecológica sea consustancial al pensamiento posthumanista (sobre todo en la vertiente crítica), el posthumanismo va mucho más lejos que simple anti-especismo anticolonial. La reflexión en torno a las categorías dicotómicas realizada por autoras del posthumanismo como Ferrando, merece una especial atención. Parece más acertado leer el posthumanismo como una reflexión y síntesis cuidadosa, que viene del poscolonialismo, pero también de otras teorías críticas como el feminismo o la teoría queer, y que trata de proponer nuevas maneras de considerar lo humano que no caigan en esencialismos ni dicotomías. En suma, el posthumanismo debe entenderse como una crítica exhaustiva a la modernidad, que supera en ocasiones los límites del posmodernismo convencional.²⁷

3.3. POSTHUMANISMO CRÍTICO Y POSMODERNIDAD

Puesto que el posthumanismo crítico se puede entender como una vertiente de la filosofía posmoderna, y debido a que esta última nos servirá a la hora de desentrañar el proyecto transhumanista, conviene detenerse brevemente en entender la crítica del posmodernismo a la modernidad. La Ilustración europea es resumida por Kant como el proyecto emancipatorio por el cual el “hombre”, a través de la razón, sale de su “autoculpable minoría de edad”²⁸. El hombre moderno, técnico, toma las riendas del progreso histórico y lo conduce hacia el ideal de sociedad humanista, libre y emancipada. La crítica posmoderna ha desmantelado una a una

27 Por ejemplo, Foucault sentencia uno de sus más famosos libros con la muerte del hombre. Foucault, M. *Las palabras y las cosas: una arqueología de las ciencias humanas*. Madrid: Siglo veintiuno editores, 1968, p. 378. Haciendo una clara analogía con la muerte de dios de Nietzsche, el historiador y filósofo francés critica el sujeto moderno y sus contradicciones internas. El posthumanismo, a pesar de que podría entenderse como la consecuencia lógica de esta idea, y ha sido definido como una corriente que propone “el fin del universo centrado en el hombre” (en Pepperell, R. *The Posthuman Condition*, 2003, p. 171), realiza una crítica a esta idea de Foucault: “posthumanism does not rely on any symbolic death: such an assumption would be based on the dualism dead/ alive, while any strict form of dualism has been already challenged by posthumanism, in its post-dualistic process-ontological perspective” [el posthumanismo no se basa en ninguna muerte simbólica: tal suposición se basaría en el dualismo vivo/muerto, mientras que precisamente el posthumanismo, por su ontología de los procesos, desafía cualquier forma estricta de dualismo] Ferrando, F. *Transhumanism, posthumanism.. (op.cit)*, p. 31.

28 Kant, I. *¿Qué es la ilustración? Y otros escritos de ética, política y filosofía de la historia*. (R. R. Aramayo, Trad.) Madrid: Alianza Editorial, 2013, p. 87. Resulta sumamente representativo Max More, autor al que nos hemos referido, defina de igual modo el proyecto transhumanista: “we have reached our childhood’s end” [hemos alcanzado el fin de nuestra infancia] More, Max. “A letter to Mother Nature”. *The transhumanist reader: classical and contemporary essays on the science, technology, and philosophy of the human future*, 2013, pp. 449-450.

las grandes narrativas que este ideal ilustrado proponía: la confianza en el progreso, en el racionalismo, el afán de dominio, la emancipación y la utopía ilustrada...

No es objeto de este ensayo detenerse en demasía en el grueso volumen de esta crítica. A pesar de ello, creo que sí resulta interesante entender la modernidad a la luz del concepto de secularización. La secularización debe entenderse en este ensayo como la idea que sostiene que la modernidad “laica” ha recibido una serie de transferencias del cristianismo. De este modo, los espacios e ideas antes dominados por la religión, en la modernidad han heredado elementos teológicos. Por ejemplo, Karl Lowith explica que las nociones cristianas de Providencia y Salvación se ven sustituidas en la modernidad por el Progreso y la emancipación utópica²⁹. Así, la Verdad revelada se ve sustituida por la Verdad alcanzada por medio del método racional/científico. El análisis de la secularización, como herramienta para comprender de las creencias heredadas del cristianismo, es una perspectiva filosófica fundamental para desmontar la falsedad escondida tras los ideales emancipatorios de la modernidad. Nietzsche anticipó que, tras la muerte de dios, aún quedarían muchos ídolos a los que destruir a martillazos.

De manera paralela, autores como Adorno y Horkheimer han formulado una crítica directa al proyecto de la Ilustración. La confianza en que la razón permitiría finalmente alcanzar los ideales ilustrados se ve truncada de manera radical con la barbarie de las guerras mundiales³⁰. Resulta crucial comprender que, según los autores de la teoría crítica, es precisamente en esa confianza arrogante en el progreso y la razón donde radica el error fatal del mito ilustrado, que posibilitó la barbarie posterior. La raíz fundamental del problema está en que esa apariencia de verdad universal, además de ser falsa, es el medio a través del cual se articula la diferencia de poder y la dominación de colectivos discriminados, tal y como se ha visto con el concepto de humano.

Detenerse brevemente en la crítica posmoderna a la idea de progreso nos resulta interesante porque este es el segundo pilar desde el que formular una oposición al proyecto transhumanista. Si este aboga por *la mejora técnica del ser humano*, dicha aspiración se torna compleja y peligrosa si atendemos a las críticas filosóficas que se han vertido hacia las ideas de progreso y mejora (por el lado de la posmodernidad), y a las nociones de humano y humanismo (desde el posthumanismo crítico). El transhumanismo mantiene en su seno la confianza en la razón y la técnica como medios fiables para alcanzar una sociedad posthumana

29 Lowith, K. *Historia del mundo y salvación: los presupuestos teológicos de la filosofía de la historia*. Buenos Aires: Katz Editores, 2007. Para profundizar sobre el funcionamiento de las transferencias de la secularización: Monod, J. C. *La querrela de la secularización. De Hegel a Blumenberg*. Madrid: Amorrortu, 2015.

30 Adorno, T. W., & Horkheimer, M. *La dialéctica de la ilustración*. Madrid: Trotta, 1994.

con características utópicas. Además, precisa de apoyarse en una definición concreta, esencial y cerrada de lo que es el *ser humano* y de lo que debe ser el *ser transhumano* que lo sitúa como una corriente antagónica al posthumanismo que se ha visto en el último apartado.

Es de hecho esta diferencia en torno al concepto de humano la que hace que la distinción aquí propuesta entre el transhumanismo y el posthumanismo tenga relevancia y sea útil a la hora de clarificar la problemática a analizar. Caben algunas palabras más acerca de esta distinción entre el transhumanismo y el posthumanismo, dado que algunos autores consideran que, teniendo ambos diferentes bases filosóficas, confluyen en una postura tecnófila: “it is ironical that critical posthumanism depends on the liberating discourse of transhumanism that promises to make the disembodied or borderless existence (that posthumanism aims at) posible” [es irónico que el posthumanismo crítico dependa del discurso transhumanista que promete traspasar las fronteras de la existencia corpórea]³¹.

Andy Miah afirma, en línea con la sentencia anterior: “I suggest that posthumanism encounters an ongoing undecidability over the value of transgressing boundaries and that posthumanism is characterized by undecidability in relation to biological change” [sugiero que el posthumanismo se encuentra en una incapacidad de situarse sobre el valor de transgredir los límites, y que está caracterizado por la imposibilidad de posicionarse frente a la modificación biológica]³².

La idea subyacente en este planteamiento es que el posthumanismo, puesto que precisa de la tecnología para la consecución de sus ideales, no tiene forma de articular una crítica contra el transhumanismo. Sin embargo, parece claro que el interés de este posthumanismo posmoderno hacia la tecnología es tangencial. Por ejemplo, algunos autores han señalado que prácticas como los tatuajes, las escarificaciones o los piercings son también herramientas que permiten alterar la comprensión del cuerpo humano³³. Del mismo modo, la artista Orlan utiliza la cirugía no para mejorarse, si no para deformarse plásticamente y realizar una denuncia de los cánones de belleza convencionales. Incluso el cyborg-feminismo de Donna Haraway debe ser entendido como una denuncia al orden patriarcal actual y no como una utopía tecno-futurista. La propia autora ha manifestado en numerosas ocasiones su desacuerdo con el transhumanismo y sus aspiraciones tecnológicas.

31 Islam, M. “Posthumanism: Through the postcolonial lens”, *op. cit.*, p. 117.

32 Miah, A. “Posthumanism: A critical history”. En Gordijn, Bert & Chadwick, Ruth (Eds.). *Medical Enhancements & Posthumanity*. Nueva York: Routledge, 2007, p. 89.

33 Ferrando, F. “The body”. En Ranisch, R. & Lorenz Sorgner, S (Eds) *Post-and transhumanism. An Introduction*. Fráncfort de Meno: Peter Lang, 2014, 213-226

Recalco que no se trata de afirmar que el posthumanismo y el transhumanismo no tengan conexiones. Anteriormente he mencionado las características que comparten y cómo diversos autores se encuentran a caballo entre ambos. Lo que sí que resulta evidente es que el objetivo del posthumanismo posmoderno es pensarnos y actuar desde un punto de vista crítico con el concepto de humano, pero no es necesario para ello ninguna transformación. Ferrando, que también rebate el transhumanismo tecnófilo, explica que una diferencia fundamental radica en que, desde el posthumanismo, se puede afirmar sin modificación alguna de nuestro cuerpo que ya somos posthumanistas, puesto que esta es una perspectiva y una praxis, no un “ser más que humanos”³⁴. El interés de esta vertiente del posthumanismo en la técnica radica en que ser distintos gracias a la modificación puede ayudar a entendernos de manera distinta, pero esa técnica nunca es un fin en sí mismo para el posthumanismo posmoderno³⁵. Del mismo modo, el interés del posthumanismo en determinadas teorías científicas, especialmente en la teoría de la evolución darwiniana y el desarrollo de inteligencias artificiales, radica en que, siendo ambas fruto de la ciencia moderna, muestran las propias contradicciones internas de la filosofía ilustrada y su intento de sostener al humano como excepcional sujeto racional/trascendental.

Al contrario de lo que proponen Andy Miah y otros muchos –a saber, que el posthumanismo posmoderno no puede establecer un discurso que rechace el posthumanismo transhumanista– considero que la raíz misma de la denuncia a la modernidad formulada por el posthumanismo posmoderno es una crítica eficaz que desenmascara las contradicciones del transhumanismo. En gran medida, el posthumanismo quiere deconstruir la noción de humano, hacer más difusos los límites de dicha categoría y establecer una comprensión de la humanidad que no se formule de manera dicotómica frente a lo animal. Si atendemos a este posicionamiento en pro de diluir las fronteras³⁶; el proyecto transhumanista es diametralmente opuesto, ya que su anhelo de traspasar los límites de lo humano sólo puede hacerse una vez estos han sido remarcados. La superación del concepto de humano desde el posthumanismo es un ejercicio de *deconstrucción*; el

34 Ferrando, F. *Posthumanism.*, *op. cit.*, p. 170.

35 De manera análoga a como, para autores queer como Paul Preciado, determinados avances tecnológicos relativos a métodos de reproducción, tratamientos hormonales y las cirugías de reasignación de sexo contribuyen a desarticular las nociones clásicas de lo que es el género; para muchos posthumanistas las nuevas tecnologías son una herramienta útil para repensarnos en tanto que humanos de manera distinta. Preciado, Paul B. *Testo yonqui*. Anagrama, 2020. Es evidente, sin embargo, que ni en uno ni en otro caso el fin último es la transformación del sexo biológico o de la propia naturaleza por medios técnicos; sino dinamitar la concepción de lo masculino y lo femenino, o las categorías culturales de lo humano.

36 Varela, L., & Alvarado, J. T. “Posthumanismo e hibridación”. *Pensamiento*, 75, 283, 2019, 307-319.

proyecto transhumanista, en cambio, sólo puede pasar por reafirmar lo humano y *reconstruir* una nueva humanidad transhumana.

En otras palabras, el transhumanismo es una ideología ilustrada, que mantiene una confianza plena en la razón como medio para alcanzar “el verdadero bien hacia el que mejorarnos”; un fuerte individualismo; una esencia humana fija, basada en la imperfección; una confianza escatológica en el progreso ligada a un determinismo tecnológico; y, por último, un optimismo desenfrenado en los resultados de la técnica, que permite sostener la idea de que el transhumanismo sea la religión del siglo veintiuno.

4. TRANSHUMANISMO ILUSTRADO: EL NUEVO HUMANO DE LOS TRANSHUMANOS

Algunos transhumanistas tienden a hacer explícitos sus antecedentes modernos. Según Nick Bostrom, “el transhumanismo hunde sus raíces en el humanismo racionalista” y se basa en la “la superación de las limitaciones humanas a través de la razón, la ciencia y la tecnología”³⁷. Hughes también defiende que el transhumanismo “forma parte de la familia de las filosofías de la Ilustración”³⁸. Resulta lógico, pues, que muchas de las contradicciones que fueron analizadas en las secciones anteriores se mantengan en el proyecto transhumanista. Hay otros autores que quieren alejarse de la tradición ilustrada, como Max More. Sin embargo, considero que es claro, como intentaré justificar en adelante, que los ideales de la modernidad europea subyacen al conjunto del pensamiento transhumanista. En esta sección se analiza cómo el “posthumano” de los transhumanistas, precisamente porque proviene del “humano” moderno, se fundamenta en la oposición frente al Otro en la doble conjugación que ya hemos señalado: al Otro no humano –naturaleza y seres vivos– y a aquellos Otros humanos que se consideran más alejados del ideal transhumano.

37 Bostrom, N. “Una historia del pensamiento transhumanista”. *Argumentos de la Razón Técnica*, 14, 2011, p. 162.

38 Hughes, J. “Contradictions from the Enlightenment Roots of Transhumanism”. *Journal of Medicine and Philosophy*, 35, 2010, p. 622.

4.1. LA PARADOJA TRANSHUMANISTA: MATERIALISMO EXISTENCIALISTA

Como ya se ha señalado, el proyecto transhumanista no puede superar de manera crítica la noción de humano, sino que tiene que redefinirla de un modo que apoye su proyecto de superación. Así, una lectura de la esencia humana, muy extendida entre los transhumanistas, es lo que puede ser considerado un materialismo existencialista. Muchos autores favorables al transhumanismo (Pérez Escudero, Bostrom) se basan en la teoría de la evolución darwiniana para explicar que el ser humano es un animal, y defienden que sólo desde filosofías de la trascendencia o religiosas se puede argumentar sosteniendo posturas esencialistas o bioconservadoras (como las de Sandel, Habermas o Fukuyama). Sin embargo, esta crítica al antropocentrismo típico de las religiones, pero también de la modernidad, acaba conduciendo de manera paradójica a un antropocentrismo transhumanista aún más radical. El comienzo de la argumentación se basa en que somos animales naturales, por lo que categorías como instinto/libertad, natural/artificial, naturaleza/cultura, o incluso técnica/instinto, no son más que intentos falaces de situar una diferencia ontológica que nos separe del animal. De este modo, toda actuación del humano es tan natural como cualquier acto animal.

Considerarnos como seres que tienen una responsabilidad para con la naturaleza, parecen decir algunos transhumanistas, es una manera de situarnos por encima de ella. Pero, la postura a la que se acaban adscribiendo es, en definitiva, liberarse de toda responsabilidad para con esa misma naturaleza o nuestro cuerpo natural: si el ser humano es un animal, si todos sus actos son naturales, y si su naturaleza es en esencia evolutiva, soy absolutamente libre para modificarla. Este contradicción produce que en el mismo texto se citen a Pico de la Mirándola, a Nietzsche, a Ortega, o a Heidegger, al lado de Darwin o autores de la escuela positivista de Viena³⁹. Un autor español critica esta contradicción exponiendo que:

En última instancia, si la naturaleza humana es plenamente natural, entonces es técnicamente disponible, y si la naturaleza humana simplemente no existe, entonces tenemos la tarea de inventarla técnicamente. La antropotecnia, está indicada en ambos casos⁴⁰.

En definitiva, los transhumanistas parecen concluir que, puesto que el ser humano no existe por ningún motivo o fin superior, debe buscar trascenderse:

39 Este hecho ocurre a menudo en los textos de Nick Bostrom, por ejemplo, en Bostrom, N. "Una historia del pensamiento transhumanista" (op. cit.). También en otros autores como Pérez Escudero, "Darwin y el posthumansimo" (op.cit.).

40 Marcos, A. "Nuevas perspectivas en el debate sobre la naturaleza humana". *Pensamiento* 71, 269, 2015, p. 1242.

puesto que *no hay nada por encima* del humano, *podemos ser algo* más que humanos. Así su visión antropológica, que en un primer momento se contrapone a cualquier postura filosófica o teológica que sitúe algún ente sobre el ser humano o una concepción trascendental del mismo, acaba asumiendo que el ser humano es por lo tanto lo más elevado que hay, ser absolutamente libre. El proyecto, finalmente, se acaba traduciendo en un dominio de la naturaleza que no difiere de aquel propio de la modernidad en el cual se asumía un salto ontológico entre el ser humano y el resto de seres vivos.

4.2. EL HUMANO COMO SER DEFICIENTE

Una de las herencias más claras que el transhumanismo retoma de la tradición ilustrada/humanista es una visión antropológica del ser humano como ser limitado o insuficiente. En la modernidad esa desafección se enfocaba hacia las características más animales o instintivas de lo humano, y la figura paradigmática era el miembro del Estado de Naturaleza o el salvaje de las colonias; pero también el pobre con poca capacidad de planificación o la mujer (históricamente asociada a la naturaleza, lo irracional, etc). Desde el transhumanismo se ha resignificado ese mismo humano que debe ser superado, pero se hace abandonando las connotaciones animales y adquiriendo matices sanitarios y biológicos: el ser humano es deficiente, inadaptado, etc. En definitiva, el transhumanismo comparte una visión general de la naturaleza humana como enferma, que debe ser modificada. Esta visión –en nada parecida a la del posthumanismo– justifica la mejora del ser humano. La mejora, ¿hacia dónde? El proyecto transhumano, según su particular razón, pasa por mejorar individuos. Como ha notado Diéguez, suele pasarse por alto el hecho de que la relación entre individuos mejorados y una sociedad mejor no es tan clara como el transhumanismo parece asumir⁴¹.

Con respecto a la concepción antropológica pesimista que mantienen los transhumanistas, Le Dévédec ha formulado una brillante crítica que merece ser examinada⁴². El autor primero analiza una serie de textos, como *La carta a la Madre Naturaleza* de Max More, o el libro *Human engineering and climate change*, de Sandberg Liao y R. Roache. En este último se proponen una serie de mejoras transhumanistas para solventar los problemas del cambio climático. Algunas de las propuestas son realmente extravagantes, como reducir el tamaño

41 Diéguez, A. *Transhumanismo, la búsqueda tecnológica del mejoramiento humano*. Barcelona: Herder, 2017.

42 Le Dévédec, N. "Unfit for the future? The depoliticization of human perfectibility, from the Enlightenment to transhumanism". *European Journal of Social Theory*, 21, 4, 2018, 488-507.

de los seres humanos para limitar el consumo, o mejorarnos moralmente para hacerlos más ecológicamente responsables. Le Dédevec concluye afirmando que esta concepción antropológica conduce a afirmar que: “it is never our social and political environment that is maladapted or problematic; rather, it is the human beings in flesh and blood” [nunca es la situación política la que resulta inadaptada o problemática, sino el mismo ser humano en carne y hueso]⁴³. Así, se produce una naturalización de los problemas políticos, a la vez que una despolitización de las cuestiones sociales que deben resolverse ahora por la técnica biológica:

It is always the human being that is unfit for the future; a future which is always naturalized and so taken for granted. Never does transhumanism, whatever its variation, seriously consider that it is this very future – the capitalist future of indefinite growth and technical progress – that is fundamentally maladapted to establishing a truly human society. [Siempre es el ser humano el que está inadaptado para el futuro; un futuro que siempre está naturalizado y que se da por sentado. El transhumanismo, sea cual sea su variante, nunca considera seriamente que es este mismo futuro –el futuro capitalista de crecimiento indefinido y progreso técnico– el que está fundamentalmente inadaptado para establecer una sociedad verdaderamente humana]⁴⁴.

4.3. LA ESENCIA TÉCNICA Y EL OTRO

Otra contradicción interesante se produce por el hecho de que los transhumanistas critiquen el concepto cerrado de “humano” desde una posición en ocasiones similar a la de los posthumanistas, pero luego no duden en proponer un nuevo humanismo. Esta paradoja se traduce en la afirmación simplificada de que, puesto que no tenemos una esencia humana fija, lo humanamente *bueno* es que nos modifiquemos nuestra *moral* para dejar de ser humanos y llegar a ser posthumanos⁴⁵. Esta sentencia hace patente que, lejos de lo que algunos transhumanistas puedan afirmar, esta filosofía sí que se basa en una determinada esencia de lo que es y lo que debe ser el humano –*no hay naturaleza humana, pero lo natural es mejorarnos*, como hemos hecho siempre–.

43 Íbid, 58.

44 Íbid, 62.

45 Bostrom, en un texto realmente curioso en el que simula una carta de un transhumano futuro a la actualidad, escribe: “Puedes decir que soy feliz, que me siento bien. Puedes decir que siento dicha interminable. Pero estas palabras se utilizan para describir experiencias humanas. Lo que siento está mucho más allá de los sentimientos humanos ordinarios, tal y como mis pensamientos se encuentran más allá de los pensamientos humanos.” (p. 4). Pero, a la vez, toda la argumentación se basa en explicarnos por qué esa supuesta utopía debería parecernos, a nosotros, humanos, moralmente deseable. Bostrom, N. “Letter from Utopia”. *Studies in Ethics, Law, and Technology*, 2,1, 2008, 1-7.

Los transhumanistas intentan, como hemos mencionado con anterioridad, demostrar que hay una continuidad entre la técnica de los primeros homínidos y la actual bioingeniería. Según esta perspectiva, la esencia del ser humano radica en utilizar unos conocimientos técnicos cada vez mayores (cuantitativamente), para hacer su existencia más cómoda y satisfacer sus deseos ⁴⁶. De este modo, no habría diferencia alguna entre encender un fuego y modificar nuestro cuerpo para permitirle soportar temperaturas menores. Esta tesis cae en la incongruencia de afirmar que se debe mejorar y alterar nuestra humanidad (es decir, sostener que la naturaleza humana es fallida) porque va en nuestra naturaleza técnica, y porque es lo que continuamente hemos hecho a lo largo de la historia (a la vez que supone una ruptura con la continuidad, dado que se busca trascender lo humano). Además, lo que esconde esta posición es que la técnica, como lo esencialmente humano, es lo que nos contrapone de nuevo al Otro. Un autor transhumanista explica que “expanding our potential is precisely the primary distinction of our species” [expandir nuestro potencial es precisamente la principal distinción de nuestra especie] ⁴⁷. Del mismo modo, entender al humano como esencialmente técnico enmascara que esta misma técnica científica tiende a asociarse al occidente moderno:

Indeed, there is a technological cultural complex that celebrates progress, religion, whiteness and modernity over what are deemed static, primitive or even terrorist societies. Critically, it was machine versus human or animal power; science versus superstition and myth; synthetic versus organic; progressive versus stagnant, that defined the relationship between the colonizer and the colonized. [De hecho, existe un complejo cultural tecnológico que celebra el progreso, la religión, la blancura y la modernidad frente a lo que se considera sociedades estáticas, primitivas o incluso terroristas. Críticamente, fue la máquina frente al poder humano o animal; la ciencia frente a la superstición y el mito; lo sintético frente a lo orgánico; lo progresivo frente a lo estancado, lo que definió la relación entre el colonizador y el colonizado]⁴⁸.

Así, nuestra esencia técnica por un lado, nos separa del no humano: el animal, que no tiene progreso técnico. Por otra parte, sitúa a algunos, los occidentales y su ciencia moderna, por encima de los menos posthumanos. El transhumano futuro es igual de excluyente que el humano clásico.

46 Julian Savulescu es uno de los autores que mejor representan esta línea del transhumanismo. De hecho, el mismo autor tiende a utilizar el término de biomejoramiento, en vez de transhumanismo, en un esfuerzo de presentar sus tesis como menos radicales, como la continuación del mejoramiento progresivo que el ser humano siempre ha buscado. Savulescu, J. *¿Decisiones peligrosas?* (op. cit.).

47 Kurzweil, R. “Human Body Version 2.0” [texto en línea], 2003. <<http://www.kurzweilai.net/human-body-version-20>> [Consultado: 5/02/2021].

48 Ahlawulia, “Two Senses of the Post in Posthumanism”, *op. cit.*, p. 135.

A propósito del humanismo, Ahluwalia señala: “In a sense, all forms of humanism to date have been involved in the project of empire” [De algún modo, todas las formas de humanismo hasta la actualidad han estado asociadas al proyecto imperialista]⁴⁹. Vemos como al transhumanismo, que es definido por sus propios partidarios como un nuevo humanismo, se le puede aplicar también esta sentencia. Algunos autores han ido aún más lejos y han señalado cómo la ratio industrial moderna y burguesa domina primero las especies de animales y plantas; luego, la propia sociedad a través de la biopolítica (controles de la natalidad y otras regulaciones sobre la vida); y, por último, han apuntado a la nueva industria genética e informática como el subsecuente paso para un asalto hacia la dominación de nuestros propios cuerpos⁵⁰.

4.4. EL DEBER SER TRANSHUMANOS: UN TRANSHUMANO ME HABLA DESDE EL FUTURO

Por último, se podría argumentar que el transhumanismo, puesto que busca mejorar al humano y es un nuevo humanismo, debe mantener una determinada concepción de lo que este (post)humano debe ser. Nick Bostrom en *Letter from Utopía*, pone en boca de un supuesto posthumano: “Soy uno de tus posibles futuros. Si todo va bien, un día te convertirás en mí”⁵¹. El posthumano sustituye al humano moderno como modelo al que se debe aspirar. Así, todo lo que se acerque más a esa determinada concepción de cyborg-biomejorado, será mejor valorado, del mismo modo que la humanidad moderna se “medía” en tanto que mayor o menor semejanza con el ideal de hombre, blanco, racional, culto, propietario... Ferrando advierte que: “the posthuman overcoming of human primacy, though, is not to be replaced with other types of primacies (such as the one of the machines)” [la superación posthumana de la centralidad de lo humano, sin embargo, no debe remplazarse con nuevas primacías (como la de las máquinas)]⁵².

En esta sección se ha visto cómo, lejos de deconstruir el término de humano, el transhumanismo lo reconstruye de una manera dual. Por un lado, el humano, en tanto que ser plenamente natural, puede modificarse a su antojo. Además, para los transhumanistas muchos problemas sociales se pueden solucionar mejorando o corrigiendo la naturaleza humana. Así, en un movimiento contradictorio, los problemas políticos se naturalizan a la vez que el cuerpo natural se vuelve

49 Íbid, 134.

50 Martorell Campos, F. “Al infierno los cuerpos: El transhumanismo y el giro posmoderno de la utopía”. *Thémata. Revista de Filosofía*. 2012, 489-496.

51 Bostrom, N. “Letter from utopía”, *op. cit.*, p. 1.

52 Ferrando, F. “Posthumanism, Transhumanism,..”, *op. cit.*, 29.

artificial. Con esta visión el cuerpo humano se ve como deficiente o enfermo. Desde su particular tecnofilia, el transhumanismo defiende a la vez que estos errores pueden corregirse a través de la técnica, que aparece como una de las características esenciales de la naturaleza humana. Resulta paradójico que se trate de naturalizar de este modo un proyecto que busca en definitiva trascender la esa misma naturaleza.

Por otra parte, parece claro que, en tanto que proyecto de mejora, si la naturaleza humana es insuficiente, cabe definir cuáles son las características que la naturaleza transhumana debe poseer. De este modo, se puede sospechar que el posthumano transhumanista no rompe con las dicotomías y el exclusivismo del humano moderno, si no que las reconfigura en torno a términos nuevos. Simplemente sustituye el sujeto tras-cendental kantiano por un sujeto trans-humano, y su imperativo categórico por un imperativo tecnológico que parece ser más peligroso que el antiguo humanismo. Este nuevo transhumano adquiere primacía, y se sobrepone al simple humano, que es visto ahora como deficiente o inadaptado. Contrariamente a esa postura, el posthumanismo crítico propone un anti-exclusivismo, que rechace cualquier “dualismo o antítesis, desmitificando cualquier polarización ontológica a través de la práctica posmoderna de la deconstrucción”⁵³. Resulta claro cómo la crítica del posthumanismo posmoderno al concepto de humano moderno se puede trasladar para evaluar, y desmontar, la filosofía transhumanista y su ideal posthumano.

5. TRANSHUMANISMO ILUSTRADO- EL ÚLTIMO SUPIRO DE DIOS

Anteriormente se ha visto cómo la teoría de la secularización puede ser una herramienta útil para la posmodernidad y su labor de “desmitificar”, tal y como propone Ferrando. La filosofía transhumanista mantiene una fuerte visión escatológica de la historia, basada en las nociones de progreso y utopía que, según expusimos anteriormente, constituyen la traducción a la modernidad de los conceptos cristianos de providencia y salvación. Estas aspiraciones no son rechazables simplemente por ilusorias, sino que, históricamente, la construcción de supuestas utopías, la posesión de una verdad trascendental, y la guía en el proceso humanista o civilizador han sido los instrumentos legitimadores de la dominación por excelencia.

53 *Ibid.*, 30.

5.1. PROGRESO Y PROVIDENCIA

De acuerdo con la visión ilustrada, los transhumanistas confían que ha habido una serie de progresos, que ellos centran casi con exclusividad en lo tecnológico, y que permiten una continua superación de los límites naturales del ser humano. Esta postura, recordemos, se complementa con el argumento de la continuidad que ya hemos mencionado con anterioridad. El avance tecnológico parece ser, para algunos autores, imparable:

Por muy convincentes que puedan ser los argumentos éticos, es difícil ver cómo pueden detener durante mucho tiempo el siguiente paso, en especial si lo que está en juego es la posibilidad de prolongar indefinidamente la vida humana, vencer enfermedades incurables y mejorar nuestras capacidades cognitivas y mentales⁵⁴.

Las sociedades humanas se desarrollan así siguiendo un determinismo técnico, que sólo puede ser acelerado o retrasado, pero nunca detenido. Esta perspectiva, además, se utiliza para presionar a las sociedades occidentales bajo el argumento de que debemos prepararnos para una suerte de carrera armamentística, reinventada ahora en carrera biogenética, que podemos ganar o perder, pero de ninguna manera no participar: las regulaciones y prohibiciones en occidente, conducirían a “incrementar más que disminuir el riesgo si el desarrollo se llevase a cabo soterradamente o en regiones del mundo menos dubitativas”⁵⁵. Este argumento se basa en una reinterpretación del “choque de civilizaciones”, que nos enfrenta ahora principalmente a China, como supuesto adalid terrorífico y sin escrúpulos de la bioingeniería y la eugenesia social⁵⁶. Esta visión del mundo, de herencia spengleriana, parece indicarnos que, si no nos mejoramos más rápido, otras regiones acabarán por dominarnos. Lo mejor pues, es adelantarnos para seguir dominando nosotros, los verdaderos (post)humanos.

5.2. UTOPIA Y SALVACION

Por otra parte, el ideal de progreso transhumanista no estaría tan impregnado de un olor a religión del siglo XXI si no fuese por lo exageradamente optimista

54 Harari, Y. *Homo Deus: breve historia del mañana*. Barcelona: Debate, 2014, 28.

55 Bostrom, N. “Una historia del pensamiento transhumanista” (op. cit.) 2011, 25.

56 Moravec, refiriéndose en este caso al desarrollo de la inteligencia artificial, afirmaba: “Si los EEUU detuvieran unilateralmente su desarrollo tecnológico, pronto sucumbirían al poder militar de los soviéticos o a los éxitos económicos de sus socios comerciales”. Moravec, H. “Los vagabundos”. En Minsky, M (ed.). *Robótica*. Barcelona: Planeta, 1986, 113.

de sus propuestas. Desde la vertiente de mejora informática una serie de científicos, como Moravec, esperan lo que han denominado Singularidad Técnica. Con este nombre se refieren al esperado acontecimiento en el que el ser humano cree una inteligencia artificial capaz de diseñar una máquina más inteligente que ella misma. Para estos autores se produciría así una especie de explosión exponencial de conocimiento que daría lugar a una “supercivilización” sin precedentes. De hecho, esta nueva sociedad de super-inteligencias, según el autor, probablemente nos haga prescindibles, con lo que se asume la extinción de la humanidad como un paso necesario para la posthumanidad hacia la que Moravec avanza optimista. Lo extravagante de las propuestas podría pasarse por alto si no fuese por el hecho de que no es el desvarío de un único autor, si no que esa misma explosión de conocimiento, la Singularidad Técnica, ya cuenta con su propia universidad en la que Google y otras muchas empresas del entorno de Silicon Valley y la criticada “ideología californiana” han destinado muchas de sus esperanzas y de sus millonarias inversiones⁵⁷.

Más allá de estas visiones que pueden parecer demasiado estrambóticas para ser realmente atractivas, el núcleo de la filosofía transhumanista parece proponer un “paraíso terrenal”: la salvación aquí en la tierra, última mundanización del proceso secularizador. Así, los transhumanistas afirman que se pueden eliminar determinados aspectos negativos de la condición humana: el sufrimiento, como en la utopía hedonista de Pearce en la que propone acabar por medio de drogas e ingeniería genética con el dolor, y guiar así nuestras acciones por una gradación de placeres (*The hedonistic imperative*); la muerte, que se halla presente en el horizonte de la mayoría de los transhumanistas; el control de nuestros propios recuerdos por medio de fármacos o chips informáticos; la criminalidad, como en la sociedad psicocivilizada de Rodríguez Delgado...

La inmortalidad, sueño anhelado de toda religión, es producida en la sociedad del siglo XXI por medios técnicos. Nick Bostrom, en la ya citada *Letter from Utopia*, escribe así acerca de la “jaula de la muerte”:

No aceptes que es bueno para ti que tus amigos enfermen y mueran en una jaula. No asumas que es una bendición el estar confinado para siempre tras la verja de la estupidez. No creas que no hay nada que merezca ser experimentado más allá de tus limitaciones psíquicas actuales. Desde que una criatura peluda cogió las dos primeras piedras de pedernal y empezó a golpearlas para hacer una herramienta, tus ancestros han estado golpeando esos barrotes, y se están aflojando a cada momento. El día de la ruptura se acerca⁵⁸.

57 Sobre la ideología californiana, Barbrook, R., & Cameron, A., “The californian ideology”. *Science as Culture*, 6, 1, 1996, 44-72.

58 Bostrom, “Letter from utopía”, *op. cit.*, p. 9.

En este curioso pero representativo texto Bostrom propone romper cuatro jaulas (vejez, malestar emocional, limitaciones cognitivas y muerte), para así “liberar el espíritu humano”⁵⁹.

Otra de las variantes de la inmortalidad que merece la pena analizar es el “uploading” de la mente a un dispositivo informático (Minsky, Kurzweil, Moravec, etc). Muchos autores han señalado que esta último anhelo se relaciona con en el proyecto del humanismo clásico, desde tiempos griegos, centrado en el (auto) dominio del cuerpo, ligado a las pasiones, a lo carnal, lo sensible y al dolor. Lo humano dominando a la parte animal del hombre. Descartes, en la modernidad, afianzó esta oposición con la dualidad *res cogitans* y *res extensa*. Martorell critica de manera eficaz esta idea que está en la base de la modificación transhumanista:

La transbiomorfosis evidencia el pelaje gnóstico y dualista del transhumanismo. A ojos suyos, el cuerpo corresponde al orden del tener, no al orden del ser. Y como es “algo” que se tiene, puede reformarse, sustituirse o abandonarse sin menoscabar la supervivencia del yo⁶⁰.

Francesca Ferrando, desde el posthumanismo, se ha mostrado crítica con dicha concepción dicotómica de cuerpo/mente y plantea un interrogante crucial:

My question to this seemingly “neutral” body being redesigned is: how are the histories and herstories of the historical human body going to affect our posthuman future? The body, as a biological and figurative locus of socio-political interactions, is hardly neutral. [Mi pregunta a este cuerpo aparentemente “neutro” que se está rediseñando es: ¿cómo van a afectar al futuro posthumano la historia y la *herstory* –historiografía con perspectiva de género– del cuerpo humano? El cuerpo, como lugar biológico y simbólico de las interacciones sociopolíticas, no es neutral]⁶¹.

CONCLUSIONES

Más allá de las críticas que se puedan hacer a la simplicidad y lo caduco de las dicotomías que se han presentado, lo que se hace patente es que el transhumanismo es la utopía o la religión de nuestra era, marcada esta por un gran culto a lo tecnológico. Sólo una religión y una salvación técnicas son posibles en una sociedad donde la crítica posmoderna ha resquebrajado las utopías emancipadoras de carácter político. Aplicando el esquema dual de secularización/crítica

59 Íbid, p. 9.

60 Martorell Campos, F. “Al infierno los cuerpos”, *op. cit.*, p. 494.

61 Ferrando, F. “Posthumanism, Transhumanism...”, *op. cit.*, p. 10.

posmoderna que aquí hemos bordeado es necesario traspasar el grueso de la crítica a desmontar, o desmitificar, los ideales de la nueva técnica del siglo XXI.

Es en este sentido que se puede sostener la idea de que el transhumanismo es el último suspiro de dios... Una última esperanza utópica de aquellos que aún creen en la Verdad como medio para llegar al paraíso, esta vez, en la tierra. Esta utopía, repetimos, no es perjudicial únicamente por ser una ilusión, sino porque, a lo largo de la historia, cada vez que se ha enunciado la Verdad ha sido para ejercer un dominio oculto sobre el Otro, camuflado de supuesta salvación (sacerdote), civilización o humanismo (hombre, burgués e ilustrado) y, ahora, trans-humanización (técnico). Además, la nueva utopía no sólo acoge en su seno todas las contradicciones que tenía el viejo humanismo, sino que cuenta intrínsecamente con una potencialidad técnica sin precedentes y está peligrosamente despolitizada. Algunos de los críticos han señalado el preocupante hecho de que el transhumanismo es “exageradamente renuente, al contrario que las tecno-utopías anteriores, a insuflar política igualitarista a su propuesta y despejar las serias dudas que suscita”⁶².

La filosofía debe analizar críticamente este pensamiento que sin duda alguna cobrará cada vez más peso en el debate público de las décadas venideras, y que tratará de imponer con su Verdad técnica y su ratio industrial un dominio sobre el cuerpo. En la época posmoderna en la que nos encontramos, parece difícil encontrar las certezas morales desde las que formular una crítica rigurosa y filosóficamente sólida al proyecto transhumanista. Sin embargo, la distinción que aquí se ha propuesto, en torno a cómo el posthumanismo y el transhumanismo tienen objetivos fundamentalmente opuestos, puede ayudar a clarificar el debate y a señalar las raíces del nuevo humanismo transhumano. Así, por un lado, con la crítica posthumanista al término de humano como mecanismo de poder; y por otro, con la crítica posmoderna a la confianza arrogante en la razón y el progreso como teología secularizada, es posible desmontar las aspiraciones y peligros del proyecto transhumanista.

Este proyecto se constituirá, según el análisis aquí expuesto, en base a dos antagonismos: el menos posthumano, y el no humano. Las transformaciones técnicas serán sin duda alguna reservadas para unos pocos, principalmente para aquellos que son ya los *más humanos* (ricos, blancos, hombres, etc). Algunos transhumanistas ya han bordeado la idea de que los posthumanos no tienen por qué tener los mismos derechos que los seres humanos, puesto que tendrán mayores capacidades cognitivas y morales ⁶³. Una vez más, igual que con los primeros derechos del “hombre” de la Ilustración, el proyecto emancipatorio

62 Martorell Campos, F. “Al infierno los cuerpos”, *op. cit.*, p. 496.

63 Lauritzen, P. “Stem Cells, Biotechnology, and Human Rights: Implications for a Posthuman Future”. *The Hastings Center Report*, 35, 2, 2005, 25-33.

transhumanista privilegia de manera exclusiva a unos pocos. Por otra parte, el posthumano está íntimamente ligado a un incremento continuo de la explotación de la naturaleza, incluso algunos transhumanistas proponen ya colonizar nuevos planetas ⁶⁴, precisamente para paliar los problemas ecológicos que el mismo crecimiento económico y técnico ha causado. El posthumano puede y debe, para trascenderse, enfrentarse y vencer al Otro, en este caso, la naturaleza y los seres vivos. El transhumanismo es una filosofía con muchos oscuros, y sólo analizando desde el posthumanismo crítico el verdadero proyecto que sustenta se pueden analizar las contradicciones, desigualdades y discriminaciones que corren el riesgo de producirse.

REFERENCIAS BIBLIOGRÁFICAS

- ADORNO, T. W. y HORKHEIMER, M., *La dialéctica de la ilustración*. Madrid: Trotta, 1994.
- AHLUWALIA, P., "Two Senses of the Post in Posthumanism". En BANERJI, D. & PARANJAPPE, M. R. (eds.), *Critical Posthumanism and Planetary Futures*. New Delhi: Springer, 2018, 131-141.
- BARBROOK, R. y CAMERON, A., "The californian ideology". *Science as Culture*, 6, 1, 1996 44-72.
- BOSTROM, N., "The Transhumanist FAQ" [en línea]. Version 2.1. Oxford: World Transhumanist Assotiation, 2003. <<https://www.nickbostrom.com/views/transhumanist.pdf>> [Consulta: 15/02/2021]
- , "Letter from Utopia". *Studies in Ethics, Law, and Technology*, 2,1, 2008, 1-7.
- , "Una historia del pensamiento transhumanista". *Argumentos de la Razón Técnica*, 14, 2011, 157-191.
- DIÉGUEZ, A. *Transhumanismo, la búsqueda tecnológica del mejoramiento humano*. Barcelona: Herder, 2017.
- FERRANDO, F., "Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations". *Existenz* 8,2, 2013, 26-32
- , "The body". En RANISCH, R. & LORENZ SORGNER, S. (eds.), *Post-and transhumanism. An Introduction*. Fráncfort de Meno: Peter Lang, 2014, 213-226.
- , "Posthumanism". *Tidsskrift for kjønnsforskning nr. 2*, 2014b, 168-172.
- FOUCAULT, M., *Las palabras y las cosas: una arqueología de las ciencias humanas*. Madrid: Siglo veintiuno editores, 1968.
- FUKUYAMA, F., *Our posthuman future: Consequences of the biotechnology revolution*. New York: Farrar, Straus and Giroux, 2003.

⁶⁴ Liao, S. Matthew; Sandberg, Anders; Roache, Rebecca. "Human engineering and climate change". *Ethics, Policy & Environment*, 15, 2, 2012.206-221.

- HARARI, Y., *Homo Deus: breve historia del mañana*. Barcelona: Debate, 2014.
- HARAWAY, D., *Manifiesto Cyborg, El sueño irónico de un lenguaje común para las mujeres en el circuito integrado* (Trad. Manuel Tanel). University of California, 1984.
- HARBISSON, N., "I listen to color" [vídeo en línea] Ted Talks, 2012 <https://www.ted.com/talks/neil_harbisson_i_listen_to_color?> [Consulta: 1/02/2021].
- HARRIS, J., *Enhancing Evolution: The Ethical Case for Making People Better*. Nueva York: Princeton University Press, 2007.
- HUGHES, J., "Contradictions from the Enlightenment Roots of Transhumanism". *Journal of Medicine and Philosophy*, 35, 2010, 622-640.
- ISLAM, M., "Posthumanism: Through the postcolonial lens". En BANERJI, D. & PARANJAPE M. R.(Eds), *Critical Posthumanism and Planetary Futures*. New Dehli: Springer India, 2016, 115-130.
- KANT, I. *¿Qué es la ilustración? Y otros escritos de ética, política y filosofía de la historia*. (R. R. Aramayo, Trad.) Madrid: Alianza Editorial, 2013.
- KURZWEIL, R., "Human Body Version 2.0" [texto en línea], 2003. <<http://www.kurzweilai.net/human-body-version-20>> [Consultado: 5/02/2021].
- LATOUR, B., *Nunca fuimos modernos. Ensayo de antropología simétrica*. Madrid: Siglo veintiuno editores, 2007.
- LAURITZEN, P., "Stem Cells, Biotechnology, and Human Rights: Implications for a Post-human Future". *The Hastings Center Report*, 35, 2, 2005, 25-33.
- LE DÉVÉDEC, N., "Unfit for the future? The depoliticization of human perfectibility, from the Enlightenment to transhumanism". *European Journal of Social Theory*, 21, 4, 2018, 488-507.
- LIAO, S. Matthew; SANDBERG, Anders y ROACHE, Rebecca, "Human engineering and climate change". *Ethics, Policy & Environment*, 15, 2, 2012, 206-221.
- LOCKE, J., *Segundo ensayo sobre el gobierno civil: un ensayo sobre el verdadero origen, alcance y finalidad del gobierno civil*. (C. Piña, Trad.) Buenos Aires: Editorial Losada, 2002.
- LÓPEZ FRÍAS, F. J., "Continuidad de las innovaciones tecnológicas: el reto de las intervenciones biomédicas de mejora humana". *ISEGORÍA* 48, 2013, 213-228.
- LOWITH, K., *Historia del mundo y salvación: Los Presupuestos Teológicos De La Filosofía De La Historia*. Buenos Aires: Katz Editores, 2007.
- MACPHERSON, C., *La teoría política del individualismo posesivo: De Hobbes a Locke*. Madrid: Trotta, 2005.
- MARCOS, A., "Nuevas perspectivas en el debate sobre la naturaleza humana". *Pensamiento* 71, 269, 2015 1239-1248.
- MARTORELL CAMPOS, F., "Al infierno los cuerpos: El transhumanismo y el giro posmoderno de la utopía". *Thémata. Revista de Filosofía.*, 2012, 489-496.
- MIAH, A., "Posthumanism: A critical history". En GORDIJN, Bert & CHADWICK, Ruth (eds.). *Medical Enhancements & Posthumanity*. Nueva York: Routledge, 2007, 71-94.

- MONOD, J. C., *La querrela de la secularización. De Hegel a Blumenberg.* . Madrid: Amorrortu, 2015.
- MORAVEC, H., “Los vagabundos”. En MINSKY, M (ed.). *Robótica*. Barcelona: Planeta, 1986, 99-119.
- MORE, Max, “A letter to Mother Nature”. *The transhumanist reader: classical and contemporary essays on the science, technology, and philosophy of the human future*, 2013, 449-450.
- , “Transhumanist declaration” [en línea] En *Humanity+*, 2017: <<https://humanityplus.org/transhumanism/transhumanist-declaration/>> [Consulta: 9/11/2020]
- NOZICK, R., *Anarquía, Estado y Utopía*. México: FCE, 1988.
- PEPPERELL, R., *The Posthuman Condition. Consciousness beyond the brain*. Portland: OR: Intellect Books, 2003.
- PÉREZ ESCUDERO, A., “Darwin y el posthumanismo”. *Eikasia, Revista filosofía*, 32, 2010
- PRECIADO, Paul B., *Testo yonqui*. Anagrama, 2020.
- RODRÍGUEZ DELGADO, J., *Control físico de la mente: hacia una sociedad psicocivilizada*. Madrid: Espasa-Calpe, 1983.
- SAID, E., *Culture and imperialism*. Nueva York: Vintage, 1994.
- SANDEL, M., *Contra la perfección*. Barcelona: Marbot ediciones, 2007.
- SARTRE, J. P., “Prefacio”. En FANON, F. *Los condenados de la tierra*. México: Fondo cultura económica, 1961, 5-16.
- SAVULESCU, J., *¿Decisiones peligrosas? Hacia una bioética desafiante*. Madrid: Grupo Anaya, 2012.
- SLOTERDIJK, P., “Reglas para el Parque Humano. Una respuesta a la Carta sobre el Humanismo” Conferencia pronunciada en el castillo de Elmau, Baviera, 1999 (Trad: Adolfo Váques Rocca). Consultado en *Revista Observaciones Filosóficas*.
- VARELA, L. y ALVARADO, J. T., “Posthumanismo e hibridación”. *Pensamiento*, 75, 283, 2019, 307-319.

ELOGIO TRÁGICO DEL REFLEJO. SOBRE LA NOSTALGIA ANTROPOLÓGICA DE LA IDENTIDAD

*TRAGIC PRAISE OF THE REFLECTION. ON THE
ANTHROPOLOGICAL NOSTALGY OF IDENTITY*

JOAQUÍN ESTEBAN ORTEGA

Doctor en Filosofía y en Educación
Profesor Director
Departamento de Humanidades
Facultad de Ciencias Sociales.
Universidad Europea Miguel de Cervantes
Valladolid / España
jesteban@uemc.es
ORCID: 0000-0001-5414-1878

Recibido: 19/04/2021

Revisado: 18/05/2021

Aceptado: 6/09/2021

Resumen: El reflejo de uno mismo siempre ha generado inquietud y sospecha. Con casi seguridad vernos y reconocernos nos ha puesto desde el primer momento en relación directa con nuestro verdadero alcance y nuestra caducidad. El tiempo líquido en el que la iconicidad de nuestra naturaleza se proyecta constantemente sobre espacios imposibles de virtualidad parece requerir el contrapunto de un elogio trágico del reflejo. El presente trabajo pretende rastrear algunos argumentos que puedan servir de recordatorio de la importancia crítica que tiene el reflejo para la construcción de la identidad humana. Lo haremos desde una perspectiva hermenéutico-narrativa dialogando con algunos relatos emblemáticos en los que no se ha pretendido nunca la invisibilización de la muerte. Intentaremos volver a poner de manifiesto esa tensión trágica implícita en el carácter contradictorio de la vida y, de ese modo, aquellos extremos esenciales que parecen hoy ya olvidados en el éxtasis epocal de la autofoto hueca.

Palabras clave: Doble; reflejo; mortalidad; inmortalidad; alteridad; identidad; autorretrato; invisibilidad.

Abstract: One's reflection has always generated uneasiness and suspicion. Seeing and recognizing ourselves places us, from the very beginning, in a direct relationship to our true scope and our caducity. Liquid time, in which the iconicity of our nature is constantly projected onto impossible spaces of virtuality, seems to require the counterpoint of a tragic

praise of reflection. This paper aims to trace some arguments that may serve as a reminder of the critical importance of reflection for the construction of human identity. We will do it from a hermeneutic-narrative perspective, in dialogue with some emblematic stories in which making death invisible has never been pretended. We will try to bring to light again the tragic tension involved in the contradictory nature of life and to vindicate those essential extremes that seem to have been forgotten today in the epochal ecstasy of the hollow selfie.

Keywords: Double; reflection; mortality; immortality; otherness; identity; self-portrait; invisibility.

...somos, no sé cómo, dobles de nosotros mismos...”

M. de Montaigne, *Ensayos*.

No podemos tener claro qué fue primero: caer en la cuenta de que la opacidad de nuestro cuerpo, de nosotros mismos, generaba una proyección de oscuridad en el suelo o en las rocas, o que al acercarnos al agua a beber o a refrescarnos se generaba una imagen inconsistente pero constante vinculada a nuestra presencia. ¿Cuál era el estatuto de realidad de esas extrañas entidades? ¿Qué relación podrían tener con aquellas otras que se producían de manera algo descontrolada en nuestro interior mientras cerrábamos los ojos para descansar y dormir? Sombras, sueños, reflejos, la silueta de las propias manos representadas en una pared al apoyarlas ensangrentadas tras una intensa sesión alimenticia de supervivencia: en todo ello acontece el ineludible acompañamiento originario del desdoblamiento antropológico. A fin de cuentas, en tanto que el propio silencio de lo real habita en nosotros mismos especialmente, no hay posibilidad alguna de que nos podamos ver, por mucho que lo pretendamos, si no es a través de la repetición o de la mediación de algún reflejo o alguna suerte de duplicación de nosotros mismos. El reto que se nos plantea es si esa inevitabilidad del desdoblamiento nos devuelve la verdad de nosotros mismos o si nos falsea irremediablemente, si está sometida a la manipulación de terceros o somos dueños de nuestra propia alteridad, si terminamos vaciándonos en el espectro que nos inventamos o si lo convertimos, sin temor, en auténtico interlocutor de nuestras contradicciones. El tiempo líquido en el que la iconicidad de nuestra naturaleza se proyecta constantemente sobre espacios imposibles de virtualidad parece requerir el contrapunto de un elogio trágico del reflejo. En lo que sigue intentaremos rastrear algunos argumentos que puedan servir de recordatorio de la importancia crítica que tiene el reflejo para la construcción de la identidad humana. Lo haremos desde una perspectiva hermenéutico-narrativa dialogando con algunos relatos emblemáticos para rescatar de ellos aquellos extremos esenciales que parecen hoy ya olvidados en el éxtasis epocal de la autofoto hueca.

1. LA ENERGÍA TRÁGICA DEL DESDOBLAMIENTO

El ser humano desde muy pronto, como señala con acierto Román Gubern¹, lo que comprendió fue que su identidad escidente y escindida era muy vulnerable ya que se iba desparramando por ahí en forma de huellas, sombras y reflejos. Lo peor fue que constató que “aquellas derivaciones aparienciales” podrían ser manipuladas en contra de uno con magias y otras prácticas rituales. No obstante, cabe establecer algún matiz entre estas prolongaciones, ya que, si la sombra y el reflejo son lumínicas, la huella, por su parte, es más bien corpórea y volumétrica. Además, esta última tiene la peculiaridad de no requerir la presencia sincrónica del referente; cosa que sí ocurre en las primeras. Con estas entidades del desdoblamiento la especie tomó también conciencia de la multiplicidad, de la fragmentación y de su caducidad y contingencia; asunto este que parece haber sido olvidado por completo en la actualidad. Para Otto Rank, en su clásico libro sobre el asunto², la imagen propia reflejada en el espejo, y la reticencia e inquietud que genera tanto su presencia como su ausencia, tiene una relación directa con el miedo a la muerte. Insiste en la figura del *Doppelgänger*, ese ser doble y errante que nos acompaña de manera invisible siempre. Según las tradiciones nórdica y germana, es mejor no verlo para que, de ese modo, no quede anunciada ya nuestra muerte. Esa conexión del doble con la muerte es lo que nos posibilita hablar de inmortalidad del alma, ya que esta es el doble originario del cuerpo.

El rito funerario, desde sus primeras manifestaciones más arcaicas y prehistóricas, ha tenido siempre como objeto el deseo y la aspiración de que el muerto no desaparezca del todo y pueda permanecer entre nosotros de alguna manera. Sin duda, en primer lugar, por los vínculos y la afectividad establecidos en vida, pero especialmente como esperanza y proyección de supervivencia en lo referido a nuestro propio destino. Esta presencia invisible y latente del muerto y de la muerte en la memoria y en la imaginación recibió el nombre de espíritu (*spiritus*, lat.; *psykhe*, grieg.), cuyo significado es “aliento”. En las lenguas germánicas, como recuerda Dastur³, remite con más precisión, e incluyendo el campo semántico de

1 GUBERN, Román, “Del rostro al retrato”. *Anàlisi*, 27, 2001, 37.

2 RANK, Otto, *El doble*. Floreal Mazía. Buenos Aires: Orión, 1982 (Or. 1925). Además de este libro, planteado desde presupuestos psicoanalíticos y antropológicos, nos permitimos destacar algunas referencias importantes. Desde una perspectiva más histórico-comparatista el libro de TYMMS, Ralph, *Doubles in Literary Psychology*. Cambridge: Bowes & Bowes, 1949; también el sustantivo y sugerente análisis de COTARELO, Ramón, *La fábula del otro yo*. Valencia: UNED-Centro Francisco Tomás y Valiente, 2005; y, por último, la específica pero amplia investigación de BALLESTEROS, Antonio, *Narciso y el doble en la literatura fantástica victoriana*. Cuenca: Universidad de Castilla-La Mancha, 1998.

3 DASTUR, Françoise, *La muerte. Ensayo sobre la finitud*. María Pons. Barcelona, Herder, 2008, p. 34.

la inquietud y el horror ante el espíritu-otro, a la presencia-ausencia del fantasma (*ghost* o *Geist*) de aquel desaparecido que continúa estando presente entre los vivos.

Sin aliento, de Edgar Allan Poe⁴, es un frenético y perturbador relato en el que un hombre, en el momento en el que se encontraba chillando a su esposa en una discusión, se queda sorprendentemente sin aliento. La fantasía de Poe hace que estar sin aliento no implique estar sin vida; sin embargo, tal circunstancia redundante en un encadenamiento de acontecimientos violentamente sorprendentes en los que se ve golpeado, vejado, rajado, ahorcado..., sin terminar de morir, arrojado a una fosa común en la que reconoce a su vecino Airedesobra, que de la misma forma sorprendente sobrevive mal con la doble respiración y el doble aliento expirado por el protagonista cuando se lo sustrajo. En este caso no es la sombra lo que se pierde, sino el aliento. El hálito de la vida siempre se ha vinculado con el alma como decíamos. Lo asombroso aquí es que la expiración no sea mortal explícitamente, aunque lo sea de manera implícita. Hay en el relato de Poe una evocación expresa sobre la constante presencia de la muerte tal y como anuncia la misma pérdida fisiológica de respiración. Se trata de una suerte de amenaza que da la vida.

El aliento del espíritu es entendido, por tanto, como una proyección doble de lo que somos después de nuestra muerte, pero que ya habita en nosotros mientras estamos vivos. Este doble, en este sentido, es un acompañamiento constante e inquietante de lo que no somos. El doble y la muerte, de esta manera, se identificarían en la resistencia antropológica de la vida y en la aspiración a la supervivencia.

Edgar Morin, en su libro de referencia sobre la muerte, sustentó buena parte de sus argumentos sobre esta antropología del doble. En su opinión, esta presencia de lo otro condiciona nuestra propia concepción imaginaria de la vida: “El doble es el núcleo de toda representación arcaica concerniente a los muertos. Pero este doble no es tanto la reproducción, la copia conforme y *post mortem* del individuo fallecido, sino que acompaña al vivo durante toda su existencia, lo dobla, y este último lo siente, lo conoce, lo oye, lo ve, según una constante experiencia diurna y nocturna, en sus sueños, en su sombra, en su imagen reflejada, en su eco, en su aliento, en su pene e incluso en sus gases intestinales”⁵. Lo interesante, por tanto, de esta ineludible insistencia antropológica en el desdoblamiento es que, al ponerse en relación directa con la alteridad de la muerte, deja de ser la mera fabulación sobre la que sustenta el sueño de la vida. Recuperar la muerte en

4 POE, Edgar Allan, *Cuentos completos*. Carlos del Pozo. Barcelona: Penguin, 2015, pp. 73 y ss.

5 MORIN, Edgar, *El hombre y la muerte*. 2ª ed. Barcelona: Kairós, 1994, p. 142.

nuestras “sociedades evanescentes”⁶ pasa por la convivencia y la continuidad de la vida y la muerte en recíproco desdoblamiento. Ello no evita la exigencia trágica de tener que habitar una existencia atravesada por la contradicción y por la enfermedad de la ambivalencia. Y lo peor: la tendencia es a no soportar en absoluto esta enfermedad esencial de la vida descargando de contingencia al doble que somos para convertirlo en una nube de humo sin consistencia.

2. LA PATOLOGÍA ANTROPOLÓGICA DE LA DUPLICACIÓN

Una poderosa figura sobre otras nos invita a centrar nuestra reflexión sobre la desesperada enfermedad del desdoblamiento. En *el Fausto* de Goethe habita un paradigma. Este esquema arquetípico de un individuo solo, retado por lo real, por el destino y por su propia subjetividad, por el amor, por la ambición o por el deseo de conocimiento, situado ante algún tipo de situación irresoluble o límite, precisamente por los límites a los que le somete su propia individualidad, que se encuentra con alguna figura mediadora, fantástica, onírica y diabólica, real o imaginaria, que representa precisamente el contrapunto de contradicción tanto ontológica como antropológica y moral y que supone la posibilidad de superar la aporía vital al liberar la energía de la alteridad implícita en cada uno de nosotros. Este esquema, como decimos, se hace recurrente en buena parte de las tramas en las que se produce un desdoblamiento o duplicación de las almas, las sombras, los reflejos, las máscaras, o la impostura. La lucha de Fausto, el poder semántico del personaje, tiene que ver con el potencial arquetípico que refleja la propia condición humana desde múltiples puntos de vista. Buena muestra de ello se encuentra justamente en el inicio del relato cuando Fausto se nos presenta torturado por su deseo de conocer, su aspiración de ir algo más allá, pero a la vez atravesado por el reconocimiento de su imposibilidad: “... no podemos saber nada. Esto casi me quema el corazón. Claro que soy más sabio que estos necios, teólogos, doctores y escritores; no me afligen escrúpulos ni dudas, ni me dan miedo inferno ni demonio... Pero he perdido toda la alegría; no creo saber nada con sentido, ni supongo poder enseñar nada, ni a nadie mejorar ni convertir”⁷. Este escepticismo trágico que anima al personaje le sitúa en una posición límite y extrema de lucha y de contención moral, pero también de desesperación, al mismo tiempo que quizás sea la causa de que se haya convertido en el favorito de Dios. Fausto lo intenta

6 ROCHE, Juan Antonio, *La sociedad evanescente*. Barcelona: Anthropos, 2009.

7 GOETHE, Johann Wolfgang von, *Fausto*. José María Valverde. Barcelona: Planeta, 1980, p. 15.

también con la magia, pero todo es inútil hasta el extremo, como sabemos, de pretender quitarse la vida. Lo real es silencioso. La única alternativa es el doble.

Percibimos en la obra, en cuya complejidad no nos detendremos, una progresiva relajación del personaje, tras el pacto de sangre, ante las trampas de seducción que propone Mefistófeles regresándole a la pérdida juventud y a posiciones de poder y privilegio social. El empeño del diabólico por ganar la famosa apuesta hecha a Dios al principio del relato sobre conducir al mal al incorruptible se convierte en la motivación propia del desdoblamiento y en el implícito recordatorio constante de mortalidad que este conlleva: “Dos almas, ay, habitan en mi pecho nos dice Fausto y quieren una de otra separarse; una, con recio afán de amor, se aferra al mundo, con sus miembros abrazados; otra, fuerte, se eleva desde el polvo a los campos de los nobles abuelos”⁸.

Ese misterioso lugar inconcebible de un alma doble en el que trascender, al que remite la desaparición de la conciencia de uno mismo, requiere ser tematizado y materializado para acallar el rumor de la implícita desesperación que conlleva. Se hará todo lo necesario; incluso alimentar la posibilidad de que el ser humano se convierta en el generador de una nueva especie de vida que pueda proyectarse en el tiempo como una duplicidad por él creada. Si nos situamos reductivamente mucho más acá de toda la inflación imaginaria de la iconografía que sustenta el mito de Frankenstein nos damos cuenta de que la fuerza semántica y la potente intuición que alberga el relato de Mary Shelley sintetiza esta pasión por la continuidad, además, por supuesto, de la pasión filosófica por el origen de la vida, el control de la naturaleza y el misterio de la muerte. Ya se nos anunció en su momento que “el sueño de la razón produce monstruos”.

Es algo muy conocido ya que los seres humanos, los grupos sociales, pueden definirse especialmente bien a partir de lo que excluyen. Esta idea tiene que ver con la concepción que se tenga de lo monstruoso, que es todo aquello que se quiere ocultar de nosotros mismos; pero que también tiene que ver con que lo excluido es lo que no se da, lo inaprehensible. La exclusión de lo que está ahí pero no se da es una reacción humana a su silencio. Además, lo normal, para poder serlo, necesita de la anormalidad. La identidad social queda delimitada en el imaginario de lo extraño, de los extraños, los parias, las criaturas abominables, los monstruos⁹. El monstruo, en su radical alteridad, nos permite identificarnos, nos dice quiénes somos y nos exige aceptar los negros susurros desplazados de nuestra alma. Víctor Frankenstein, desde el primer momento en el que se da cuenta hasta dónde le ha llevado el ímpetu de su desmesura, abomina de la criatura de manera enfermiza;

8 *Ibid.*, p. 34.

9 PLANELLA, Jordi, *Los monstruos*. Barcelona: UOC, 2007.

abomina del terror al que le somete haberse asomado al otro lado, a la sombra de sí mismo materializada en una carne-otra, para responder al hueco permanente que le generaba la existencia. El monstruo es obviamente el doble de Víctor que, liberada la sujeción de la censura moral por medio del obsesivo narcótico por la manipulación del secreto de la vida a través de la ciencia y la técnica, hace que fluyan las más íntimas y oscuras pulsiones agresivas de autodestrucción hacia los suyos, hacia sus seres más queridos. En este sentido nos podríamos preguntar en qué medida no era la propia Mary Shelley la que creó este eco inconsciente de su propia monstruosidad haciendo aflorar el posible complejo de culpa por haber sido, con su nacimiento, la causa de muerte de su admirada madre, la escritora y feminista Mary Wollstonecraft. Sea como fuere la figura literaria de Víctor Frankenstein se convierte en un claro paradigma de nuestra reflexión, pero de manera especial, la criatura.

Como siempre ocurre en todas las leyendas parece que la peripecia de aquella noche de lluvia en la que se gestó el relato en Villa Diodati, la casa que Lord Byron alquiló en el entorno de los Alpes junto al lago, aunque sustentada por la amistad, estaba atravesada por una cierta tensión emocional en la convivencia de aquellos amigos¹⁰. El poeta Percy B. Shelley se había tenido que marchar de Inglaterra huyendo de las deudas con su mujer Mary y con la hermanastra de ésta, Claire, que esperaba un hijo de Lord Byron. Por su parte Lord Byron mantenía una tensa relación con su acompañante el médico y escritor John William Polidori. Además, a este cruce de convivencias se le unía el fuerte vínculo intelectual que unía a Shelley con Lord Byron. Por tanto, en este contexto, la famosa noche del 17 de junio de 1816, todos estos personajes, acompañados también de la condesa Potocka y de otro amigo de Lord Byron llamado Matthew G. Lewis, como consecuencia de leer juntos relatos de fantasmas se propusieron escribir una historia de este tipo cada uno. El reto se mantuvo varios días en los que se formularon diversas propuestas, pero la que trascendió, como sabemos, fue la de Mary.

Aunque no parece difícil adivinar en el soporte de la historia la presencia de aquellas más que probables conversaciones entre poetas y estudiosos sobre los principios de la vida y de la muerte, sobre los límites del instrumento corporal humano y sobre los nuevos usos de la energía a través de la electricidad para el dominio de la naturaleza, lo cierto es que en *Frankenstein, o el moderno Prometeo (1817-1831)* nos encontramos con una magnífica novela en la que se vislumbran los primeros impulsos de la ciencia ficción en el marco de la incipiente revolución industrial. Además, aunque pudiéramos encontrar precedentes en las

¹⁰ Nos servimos para este asunto de los datos aportados por FERNÁNDEZ VALENTÍ, T. y NAVARRO, Antonio José en su libro titulado *Frankenstein. El mito de la vida artificial*. Madrid: Nuer, 2000, pp. 32 y ss.

figuras de Prometeo, el Golem o Fausto, lo cierto es que el mito de la criatura se origina en este relato ya que no se encuentra ninguna correspondencia exacta en algún elemento del folclore¹¹.

Como Fausto, Victor Frankenstein, termina siendo víctima de la desmesura y del veneno del conocimiento¹². El diablo rebelde que quiere ser autónomo se encuentra presente. Habiendo leído *El paraíso perdido* de Milton¹³ en su sorprendente periodo de autoformación, la criatura se identifica con la reacción de Satán al ver la violencia con la que se le hace explícita la falta de respuesta sobre su origen y su destino. No se lo explica hasta que toma conciencia efectiva de ser “lo otro” en el momento de observar su reflejo en las aguas del estanque. A Narciso, el brillo saturado del reflejo que produce su imagen le nubla la vida encerrándole en su belleza; lo veremos más adelante. Al monstruo de Frankenstein, su reflejo termina ofreciéndole la posibilidad de ir más allá del bien y del mal, p mejor, quedarse más acá del bien y del mal, porque en la zona hueca y silenciosa del otro lado al que pertenece no hay destino. Lo que sí que hay es el sufrimiento de la soledad y de sentir el horror de ser únicamente un doble: “... mi cuerpo es una repulsiva imitación del tuyo, más horrorosa cuanto mayor es su semejanza”¹⁴. De hecho, la exigencia que le propone a su creador es precisamente la de dejar de ser “el otro” a partir de la creación de “un otro” para él en el que descargar su oquedad. En el doble nos descargamos de vacío. Un doble que le permitiera liberar su angustia y que le permitiera alguna clave de autocomprensión. Nuevamente la referencia a Milton parece clara aquí en tanto que Eva se convierte en el contrapunto de la soledad. “¿Me estás diciendo que a cada hombre le está destinada una esposa y a cada animal su pareja, pero que yo tendré que estar siempre solo?”¹⁵. Creador y criatura han de ser considerados como espejos de sí mismos. Tras desencadenarse la culminación de la amenaza de la criatura con la muerte de Elisabeth y con la reacción de deseo de venganza y persecución hasta el último confín de la tierra para acabar con ese terrible destino provocado por él mismo, percibimos bien cómo en el fondo Víctor lo que hace es perseguirse a sí mismo; enfrentarse con la patológica y trágica terapia de su reflejo.

11 BALLESTEROS, Antonio. *Narciso y el doble*, p. 96.

12 SHELLEY, M., *Frankenstein, o el moderno Prometeo*. Silvia Alemany. Barcelona: Penguin, 2015, p. 103.

13 *Ibid.*, pp. 218-219.

14 *Ibid.*, p. 219.

15 *Ibid.*, p. 265.

3. EL REFLEJO COMO “PATOLOGÍA TERAPÉUTICA” DE LA IDENTIDAD

Es posible que la enfermedad de la sombra exija más que el reflejo ya que en ella no hay cesura aparente¹⁶. La sombra terminal, como reduplicación constante, nos acompaña en continuidad con nuestra entidad y nuestra opacidad. Sin embargo, entre el reflejo y nosotros existe una importante distancia espacial a partir de la cual se producen otros procesos patológicos que tienen que ver también con la identidad. Es imposible comenzar a reflexionar sobre estos asuntos sin referirnos a Narciso.

En el mito de Narciso se produce una doble circunstancia especulativa: la de la voz y la de la imagen. Como sabemos bien, a la peripecia del hermoso joven le acompaña el trágico destino al que se ve sometida Eco, la enamorada y rechazada ninfa. Es en la gran síntesis de las *Metamorfosis* de Ovidio¹⁷ donde encontramos la versión mejor identificada por todos. Ovidio nos narra cómo se cruza abruptamente el destino de Eco y Narciso. Nos relata cómo la ninfa fue castigada por Juno por haberse dado al juego de entretenerla, hablando y contándole historias, propuesto por Júpiter para disfrutar con las otras Ninfas. El castigo por su casi obligada traición fue el de perder la iniciativa de la voz, conservando únicamente la posibilidad de expresarse repitiendo las últimas palabras de su interlocutor. Su voz, de esta manera, se convirtió en un mero espejo confuso de las palabras ajenas; un espejo de palabras en el que el otro jamás podría reconocerse. Quizás esa agresión implícita en la falta significativa de réplica es lo que, entre otras cosas, predispusiera tan negativamente a Narciso cuando se encontró con Eco por primera vez. La ninfa, absolutamente enamorada del bello adolescente desde el momento en el que le vio asomar en los campos, como con sus repeticiones se veía incapaz de transmitirle lo que sentía, se abalanzó sobre él generando el rechazo inmediato del joven. Tras el rechazo, avergonzada, se retiró del mundo escondiéndose en las cuevas, sometiéndose a una suerte de anorexia mítico-espiritual, y desapareciendo progresivamente (“...todo el jugo de su cuerpo se va hacia los aires”..., vv. 399.)¹⁸ hasta quedar únicamente sus huesos, que se convirtieron en piedras, y su voz. En ella únicamente vive el sonido de todo aquel que quiera hablarle a la naturaleza.

16 Hemos tratado la implicación de la sombra como duplicación de lo real en nuestro trabajo titulado *El escorzo melancólico de lo real. Ensayo sobre el decrecimiento hermenéutico y los límites de la desnudez*. Granada: Comares, 2020, pp. 89 y ss.

17 OVIDIO, *Metamorfosis*. Consuelo Álvarez y Rosa María Iglesias. Madrid: Cátedra, 2016, Libro III, vv. 339-510, pp. 293-300.

18 *Ibid.*, p. 295.

Por su parte, el acuático destino de Narciso bien pudiera ser causa de su origen, ya que nació del estupro perpetrado por el río Cefiso sobre la ninfa Liríope. El anciano Tiresias, cegado por Juno, pero vidente sobre el presente y el futuro gracias a Júpiter, vaticinó sobre Narciso que podría seguir vivo si conseguía no conocerse a sí mismo. Este contradictorio designio particular con respecto al celeberrimo del oráculo de Delfos podría servirnos como imagen para expresar, a pesar que en nuestros días pueda parecer lo contrario, la reticencia implícita que existe ante los reflejos de uno mismo por vincularse tal reflejo con los otros mundos del más allá. Recordaremos cómo algunos personajes como Drácula, en sí mismos entendidos como “lo otro” no reflejan su imagen en este mundo. Pues bien, a pesar de tales designios del anciano, después de una jornada de caza, el hermoso muchacho sediento se acercó a una charca a beber pasando por alto las prohibiciones al respecto hechas por su madre. Al verse quedó absorto, y relajando la altanería que la que había rechazado a todos sus enamorados pretendientes, entre los que estaba Eco, se desdibujó de deseo y de amor por aquel joven que le miraba con aquellos hermosos ojos y adornado de aquellos cabellos dignos de Baco y de Apolo. Tras diversos intentos infructuosos por acariciarlo y besarlo, y ver que al hablar los labios del otro se movían sin emitir sonido, llega la anagnórisis, es decir, cayó en la cuenta de que el objeto de su deseo era él mismo: “¡Ese soy yo! Me doy cuenta; y no me engaña mi imagen: me abraso de amor por mí, y muevo y sufro las llamas. ¿Qué puedo hacer? ¿Debo ser rogado o rogar? ¿Qué voy a rogar entonces? Lo que deseo está conmigo: mi riqueza me ha hecho pobre.” (vv. 464-467, p. 298). Incapacitado para separarse de sí mismo, igual que su fiel Eco que le acompañaba en su destino final, comenzó a abandonarse, a no comer, a no dormir, hasta el último suspiro en el que reconoció haber amado en vano a aquel reflejo de sí mismo. Igual que Eco, Narciso experimentó también una transformación convirtiéndose tras su muerte en una “flor azafranada que rodeaba el centro con blancas hojas”.

Probablemente este padecimiento patológico del vernos reflejados a nosotros mismos sea síntoma de la necesidad ineludible de construirnos una identidad personal donde habitar ante una experiencia de lo real carente de eco. Como sabemos, el psicoanálisis de Lacan se ocupó de ello. Por nuestra parte, al margen de las dudas y la discusión técnica sobre la originalidad de la teoría lacaniana sobre el “estadio del espejo”¹⁹, lo que nos interesa recordar muy brevemente es que Lacan concibe esta fase especular como “una identificación, es decir, una transformación (una duplicación, añadiríamos desde nuestro propio discurso) que

19 GONZÁLEZ REQUENA, Jesús, “El Estadio del Espejo ‘de Jacques Lacan’. Crónica de una mascarada (1)”. *Trama y Fondo. Lectura y teoría del texto*, nº 37 (2014).

se produce en el sujeto cuando asume una imagen”²⁰. En origen, el *infans*, tiene una concepción fragmentada de su cuerpo que análogamente la podríamos interpretar como la falta de vínculo con la totalidad propia de los seres humanos. En nuestro caso ya hemos hablado del componente trágico de la escisión y lo que ello implica. Parece, pues, necesario una terapéutica imaginaria que dé continuidad ficcional a una concepción del yo y de la identidad. Esta es la que, según Lacan, se obra en el reflejo de uno mismo, que en el fondo es, propiamente hablando, un proceso de metamorfosis y de alienación. Es el orden simbólico el constituyente para el sujeto. El sujeto se constituye a partir de la ilusión del desdoblamiento que él ha formado en su conciencia a partir de su reflejo. Pongamos un ejemplo de la dificultad de este proceso.

“¡Lo he hecho! ¡Lo he hecho! ¡He retirado la sabana de mi armario y me he mirado en el espejo! exclamaba el protagonista de *Diario de un cuerpo*, de Daniel Pennac, cuanto tenía trece años . He decidido que se había terminado. He hecho caer la sábana, he apretado los puños, he respirado muy hondo, he abierto los ojos y me he mirado. ¡ME HE MIRADO! Era como si me viera por primera vez. Me he quedado mucho rato delante del espejo. Realmente no era yo por dentro. Era mi cuerpo, pero no era yo. Ni siquiera era un compañero. Me repetía: ¿Tú eres yo? ¿Eres tú, yo? ¿Yo soy tu? ¿Somos nosotros? No estoy loco, sé perfectamente que jugaba con la *impresión* de que no era yo, sino un muchacho abandonado en el fondo del espejo. Me preguntaba desde hacía cuánto tiempo estaba allí”²¹. El jovencito que se propone redactar un diario sobre su cuerpo tras una experiencia traumática al ser abandonado en un bosque y atado a un árbol por una broma excesiva de sus compañeros scouts ha tenido tapado su espejo del armario desde siempre con una sábana para no verse, para no ver su cuerpo. Decide retirar la sábana, pero más que identidad lo que encuentra allí es a un extraño. En ese extraño que es la imagen reflejada del propio cuerpo se condensa la tensión de lo que falta. Las palabras, la escritura del relato de su cuerpo tendría como objetivo la expiación de esa extrañeza, la reducción de la distancia de ese desconocimiento, aún a sabiendas de que jamás podrá rellenarse esa grieta. “Algo en mi imagen sigue siéndome ajeno”²², declara nuestro protagonista en sus reflexiones de la treintena. Quizás por ello se inventara a ese amigo imaginario al que alertar de las presiones de mamá y que regresaría justamente en el momento de la despedida. Dodo, se llamaba. La última entrada del diario, cuando ya no

20 LACAN, J. “El estadio del espejo como formador de la función del yo [*je*] tal como se nos revela en la experiencia psicoanalítica”. En *Escritos I*. México: Siglo XXI, 2009, p. 100.

21 PENNAC, Daniel, *Diario de un cuerpo*. Manuel Serrat. Barcelona: Mondadori, 2012, p. 28.

22 *Ibid.*, p. 132.

quedan casi fuerzas, es para sosegar a ese otro irreconocible del espejo en el momento de la despedida: “Y ahora, mi pequeño Dodo, habrá que morir. No tengas miedo, te enseñaré”²³.

Más allá de las motivaciones psicológicas que expliquen la construcción de la personalidad, lo cierto es que la proyección en el reflejo para identificarnos también cursa una carga importante de catarsis y de reparación del dolor. Por eso nuestro oxímoron patológico-terapéutico sobre el reflejo incluye el alivio; si bien, un alivio ambivalente. El retrato, como veremos expresamente con Dorian Gray, puede ser ese ámbito especular en el que depositar toda la angustia, el terror, el deseo y la inquietud que provoca la falta. Pero quizás sea el autorretrato consciente y con contenido el lugar en el que mejor se acomoda esta pasión catártica y especular del desdoblamiento.

Existe un atávico miedo al retrato en las culturas tradicionales; pensemos, por ejemplo, cómo a los pieles rojas no les gusta que les fotografíen porque temen que el alma les abandone al pasar a la imagen retratada o pintada. Esta reticencia podría ser universal; sin embargo, parece que los condicionantes culturales la nublan e incluso la purgan mediante la exposición, o la sobreexposición, iconográfica de uno mismo; y no solo en tiempos aparentemente antidualistas como los nuestros.

Como decíamos, aunque el retrato tiene también una fuerte carga antropológica de duplicación y de interacción con lo otro, lo que creemos que nos ayuda más ahora a avanzar en nuestro discurso es efectivamente el fenómeno del autorretrato como paradigma de la necesidad de poder ver en lo otro, no en mí, lo que no se soporta: el dolor²⁴. No nos podremos servir en este momento de la deriva fugaz y vacía de la imagen autorepresentativa que supone el fenómeno actual del *selfie*. Se escapan a las intenciones de nuestro trabajo analizar los múltiples aspectos culturales contenidos en dicho fenómeno. Lo que nos interesa más es esa concepción tradicional de la autoimagen como inserción de la vida en el tiempo y en el espacio, en la caducidad y en las referencias. En esta concepción tradicional el desdoblamiento no solo puede ser terapéutico, sino que también aspira a ser una fuente de conocimiento y de permanencia; la posibilidad de rescatar al alma de la muerte y el olvido gracias a la repetición, en la que se expía el potencial devastador del tiempo.

A pesar de que la práctica del autorretrato está presente desde los orígenes en los pueblos antiguos (Egipto) con finalidades rituales, será en la modernidad, en el Renacimiento, especialmente con Durero cuando los artistas dejen de asomar

23 *Ibid.*, p. 323.

24 MOLINA RUIZ, I., *El autorretrato como canalizador del dolor*. Granada: Universidad de Granada, 2015.

en sus obras, camuflados tímidamente, para autopresentarse abiertamente como motivo en sí mismo. Parece que cuando la modernidad tuvo que responder al silencio absolutista del Dios nominalista²⁵ con el proceso de autonomía de los sujetos, se superaron las reticencias sobre que fueran los individuos los que protagonizaran los motivos artísticos. Además, liberados en alguna medida de la indiferencia divina, parecía adecuado el momento de recordar que el hombre era una suerte especial de autorretrato de Dios. Sin duda, estos elementos de renovación del imaginario dieron pie a la legitimación de la representación individual en retratos y en autorretratos. Ahora bien, quizás sea algún siglo más tarde, en el XVII, cuando con Rembrandt, y sus más de ochenta autorretratos, podamos concretar un trabajo constante de especulación y desdoblamiento en el que se intentaba encontrar alguna respuesta a la enfermedad (Rembrandt padecía hidropesía) y al sentido íntimo del yo. Prácticamente hasta el último y dramático de sus autoreflejos pictóricos de 1669, en el que su mirada nos delata lo muy consciente que era ya de su final, podemos corroborar cómo todos esos reflejos fueron una aspiración de autocomprensión y de terapia contra el paso del tiempo.

Otros emblemas pictóricos muy conocidos, por distintas circunstancias, de esta analgésica alienación en la autoimagen proyectada en el lienzo serían Van Gogh, Egon Schiele, Edward Munch, James Ensor, Frida Kahlo, Francis Bacon, entre otros. En todos, además de una búsqueda permanente de sí mismos, se produce esa constante ontológica mediante la cual los seres humanos liberan el dolor y la angustia que genera el cruel e indiferente silencio de lo real.

A Van Gogh aquella circunstancia biográfica de haber nacido exactamente un año después, el mismo día (30 de marzo), que su hermano muerto, de que sus padres le pusieran exactamente el mismo nombre y de que le inscribieran en el registro de nacimientos con el mismo número, le convirtió sin querer en un doble patológico al que de partida se le encomendaba la tarea de sosia; dejando al margen sus perturbaciones psíquicas endógenas. No es de extrañar que se presentara también como otro de los pintores que más autorretratos ha realizado, probablemente como síntoma de su íntima soledad, pero sin duda también como respuesta paliativa a una identidad sustitutoria y nunca propia. Las conocidas cartas a su otro hermano Theo ratifican, en este sentido, el ansia enfermiza de un eco en el cual reflejarse y en el que sustentarse.

Cada uno de los artistas mentados nos daría para alguna disquisición particular sobre su terapéutica específica. No nos detendremos. Únicamente nos gustaría destacar, entre tantos posibles, aquel sobrecogedor autorretrato del pintor

25 BLUMENBERG, Hans, *La legitimación de la Edad Moderna*. Pedro Madrigal. Valencia: Pre-Textos, 2008.

simbolista suizo Arnold Böcklin titulado *Autorretrato con la muerte tocando el violín*, de 1872, en el que, como el título indica, el artista, y con él nosotros también, se encuentra sorprendido en el espacio abisal que se abre en el reflejo de su pintura por aquello que no se da nunca pero que está siempre presente condicionando nuestra vida de manera latente e invisible. Böcklin consigue durante un solo instante sorprenderse bailando la danza que propone la muerte al compás de la melodía que resuena en el eco de su violín. El espejo del retrato de sí le ha permitido, sin esperarlo, verse en su totalidad.

4. LA AMBIVALENCIA SINIESTRA DEL REFLEJO

Con este cuadro de Böcklin constatamos cómo resulta imposible prescindir del reflejo. Se trata de una forma de generar el alivio del eco donde no lo hay y, a la vez, un recordatorio valioso de la caducidad con el que podemos enriquecer la vida sin engaños de virtualidad vacía. La peripecia y consecuencias que pudiera tener una pérdida fantástica e imaginaria de nuestro reflejo la encontramos, entre otras, en *La historia del reflejo perdido* (1815), de E.T.A. Hoffmann²⁶. En este relato el joven Erasmus Spikher deja en su casa a su mujer y a su hijo para emprender un viaje por los países del sur. Tras diversas peripecias ante la posibilidad de perder a su nueva y enigmática amante Giulietta, y tras un accidentado asesinato en el que se ve envuelto, accede a la diabólica petición de esta de cederle su reflejo depositándolo en un espejo. Tal circunstancia, que inicialmente parecía permitir un efímero consuelo, se convierte progresivamente en una amenaza constante ante el rechazo y la desconfianza de todos; especialmente de su propia mujer y de su hijo a su regreso a casa.

Del mismo modo, el caso de las dos estupendas películas de principios de siglo XX con el mismo título, *El estudiante de Praga*, además de insistir expresamente en el idéntico motivo que en el relato de Hoffmann, es paradigmático de una duplicación que habla sobre la duplicación misma: las dos películas se convierten en dobles de sí mismas. Tanto la primera de 1913, de Paul Wegener, como la segunda, de 1926, de Henrik Galeen, cuentan estructuralmente la misma historia. Únicamente advertimos algunas diferencias de carácter formal que tienen que ver con un mayor estatismo en la edición de la primera con respecto a la segunda y con que el anunciado expresionismo implícito de la primera se terminará

26 HOFFMANN, E.T.A. "La historia del reflejo perdido". En MOLINA FOIX, Juan Antonio (Ed.), *Álter ego. Cuentos de dobles (Una antología)*. Aurora Nolla. Madrid: Siruela, 2007, pp. 39-62.

de ratificar en la segunda. La trama se concreta de la siguiente manera. El joven Balduin, enamorado de una rica heredera inaccesible por la brecha social y económica, accede a la petición del diabólico personaje de Scapinelli (más diabólico en la de Wegener que en la de Galeen) a cambiarle aquello que desee de lo que se encuentre en su habitación por una incontable cantidad de piezas de oro. La cierta posición de ventaja en el intercambio que suponía Balduino al ser consciente de la precariedad absoluta de su estancia se torna en terrible angustia cuando lo que Scapinelli le reclama sorprendentemente es el reflejo de su imagen en el espejo. La tensión iconográfica de la secuencia en las dos películas es grande en el preciso momento en el que la imagen sale del espejo y se anuncian, de manera fantasmal e implícita, los trágicos acontecimientos. Como consecuencia del trato Balduino se convierte en un joven acaudalado y benefactor de sus compañeros estudiantes; sin embargo, ha de disimular constantemente ante los demás cuando se sitúa ante un espejo. En el transcurso de un baile se promete amor con la joven a pesar de su ya adquirido compromiso con otro noble del agrado del padre. De esta situación de tensión surge un duelo de honor entre los dos pretendientes. Balduino no podrá evitar la muerte del caballero, a pesar de haber prometido por su honor al padre de la joven que no utilizaría su gran pericia reconocida con la espada para salvarle la vida. La causa será que su otro yo, su reflejo perdido, se le adelanta, le suplanta en el duelo y cumple con la dramática victoria prevista. Desde ese momento Balduino no es recibido por la nobleza, y consecuentemente tampoco por su prometida, entrando en un bucle destructivo de juego, mala vida, desesperación y rechazo, tanto de la universidad como de sus propios colegas al enterarse que rompió su palabra de honor en el duelo. En un último y desesperado intento consigue entrar en el cuarto de su amada, pero la situación empeora al sincerarse y al darse cuenta esta de la ausencia de su reflejo en el espejo. En esta secuencia final entra en juego el doble haciendo inevitable un oscuro e imposible duelo de ultratumba al querer destruir la inexistente insustantividad de su otro yo. Balduino se ve perseguido por su reflejo en una secuencia magistral que termina precipitándose en la propia autodestrucción en el momento en que se dispara a sí mismo a través del otro. En la versión primera de Wegener, esta secuencia final se inicia con ambas alteridades jugando una siniestra partida de cartas como expresión manifiesta de la presencia inminente de la muerte²⁷. Se juega contra el tiempo, por si cupiera algún estertóreo aplazamiento. El reflejo, en su extraña independencia, es la propia muerte.

27 Esta constante simbólica y arquetípica de la muerte jugando una última partida con quien ya ha sido avisado la recordamos bien a través del modo en que lo dio forma Ingmar Bergman en *El séptimo sello* (1956).

Jugar con la muerte que está al otro lado, el reflejo y el original como analogía del bien y del mal, la imposibilidad de seguir vivo ante la pérdida del reflejo: todos estos aspectos que hemos podido recabar tanto en el relato de Hoffmann como en las películas de Galeen y Wegener, se encuentran también sintetizados en la historia de *Drácula* (1897) del escritor irlandés Bram Stoker. Drácula podría ser *alter ego* de cualquiera que quiera dominar el tiempo y ser inmortal, de cualquiera que quiera permanecer rejuveneciendo. Sin duda, su reconocido atractivo, centrado en su componente erótico y transgresor, se une a la posibilidad de vivir eternamente en el espacio misterioso de la muerte para los vivos. Doblarnos en una alteridad existencial de ese tipo es un íntimo deseo que habita en lo más profundo de cada uno de nosotros. La cultura y el orden establecido ha establecido las adecuadas vías rituales y místicas para sobrellevar esa carencia; sin embargo, el atractivo de la propuesta de Stoker se encuentra en que habitar ese otro mundo creado no implicaría una idealización utópica de este, sino una vida de no muertos dando rienda suelta a los deseos y pulsiones en eterna repetición. Algo así de transgresor no se puede permitir; y menos en aquella mentalidad victoriana de la época de Stoker. Los miedos y la presencia de la muerte tienen que estar regulados para facilitar la tarea a las estructuras del orden y del poder. La hipótesis de que la “infección” de Mina pudiera ser definitiva, como la de Lucy, no se puede dar bajo ningún punto de vista porque significaría una concesión a la liberación pulsional de las mujeres en muchos aspectos. Todas las grandes instituciones (la religión, la ciencia-medicina, la educación) concentradas en la figura del profesor Van Helsing, se movilizan para impedir que Drácula se convierta en “el padre o guía de una nueva raza de hombres y mujeres que seguirán su camino en la muerte y no en la vida”²⁸.

No cabe ahora, por nuestra parte, insistir en el alcance del mito. Lo que nos interesa es centrarnos exclusivamente en lo sustancial de esta novela para la reflexión de este apartado; y por ello deberíamos recordar bien que una de las características que definen a Drácula es que no tiene reflejo: el vampiro es un no-vivo que, por mucho que lo parezca, no tiene imagen en este mundo. Mientras se afeitaba, ya inquieto y desconfiado por el comportamiento de su anfitrión y por el clima general del castillo, Jonathan Harker sintió de manera imprevista la visita del conde llevándose la sorpresa de no poderlo ver reflejado en su pequeño espejo a pesar de tener seguridad plena de que se encontraba detrás de él. Nos relata la perplejidad en la que quedó sumido de esta manera: “Cuando hube contestado al saludo del conde, me puse a mirar otra vez por el espejo, tratando de comprender cómo había podido engañarme. No había el menor error: sabía que el conde se

28 STOKER, B., *Drácula*. Mario Montalbán. Penguin: Barcelona, 2015, p. 430.

hallaba detrás de mí, casi a mi lado, y solo tenía que volver la cabeza para verle. Pues bien, ¡el espejo no reflejaba su imagen! El espejo reproducía todo cuanto había a mi espalda, pero ahí no había el menor signo de ser humano... aparte de mí”²⁹. Sin reflejo, por tanto, no hay vida porque la vida no es concebible sin su resolución: sin la muerte. La tensión entre el reflejo y la muerte es similar a la del intento de sujeción del paso del tiempo en la multiplicidad de imágenes que los seres humanos reproducen y representan de sí mismos. Dorian Gray puede que sea el paradigma más conocido de desdoblamiento, no tanto en el reflejo, como en aparente distancia de la propia representación. Sin duda Stoker tuvo presente el libro de Oscar Wilde, intentando sustraerse de las consecuencias biográficas que le estaban deparando al dandi, pero muy cercano a esa sensibilidad transgresora.

La vida y la obra de Óscar Wilde fueron también un fiel reflejo y representación de sí mismo. Verdaderamente se puede decir que fueron un dramático recuerdo de que quien queda atravesado por la autonomía y el silencio de la belleza deberá estar preparado para aceptar un destino trágico y de sufrimiento. “Todos tendremos que sufrir, y sufrir terriblemente, por lo que los dioses nos han dado”³⁰, le dice el artista pintor Basil Hallward a su amigo Lord Henry Wotton al comienzo del relato sobre Dorian Gray. La aspiración de Wilde de mantener la distinción espiritual reivindicando el poder de la mentira del arte y de la belleza frente a su decadencia³¹ en un tiempo de adocenamiento técnico y moral es la que nos hace, desde un principio, intuir la fatalidad del desdoblamiento, entendido en este caso en un sentido muy amplio. “En toda distinción física e intelectual hay una fatalidad”³², señala Basil. Se trata de la fatalidad de la escisión y, sobre todo, la fatalidad de aquellos espíritus que son capaces de percibir su silencio y sus consecuencias. La estética de Wilde no renuncia a una aristocracia espiritual sobrevenida, ni a la exigencia de responder al adormecimiento generalizado. Es en lo oculto, es decir, en aquello que está en el otro lado, donde se encuentra el destino de quien habita la fractura de la belleza. Basil, atravesado de arriba abajo por Dorian y su íntimo vínculo vital con la belleza, da muestra del poder que tiene la ocultación de lo otro en lo que uno se refleja. “Cuando una persona me gusta muchísimo nunca digo su nombre a nadie. Sería como ceder una parte de ella. Cada vez amo más el secreto. Parece ser lo único que puede volver misteriosa o maravillosa la vida moderna para nosotros”³³. Late, en todas estas reflexiones

29 *Ibid.*, p. 85.

30 WILDE, Oscar, *El retrato de Dorian Gray*. Mauro Armíño. Barcelona: Austral (Espasa), 2010, p. 32.

31 Cf. *Id.*, *La decadencia de la mentira. Un comentario*. Javier Fernández. Barcelona, Acanthilado, 2014.

32 *Id.*, *El retrato de Dorian Gray*, p. 32.

33 *Ibid.*, p. 32.

del capítulo inicial, la tensión que se establece entre lo explícito y lo implícito, lo aparente y lo real, lo presente y lo oculto. Se trata de una tensión que configura el escenario de la alienación posible. Se anuncia la conocida tragedia del relato en diversos momentos; como, por ejemplo, cuando Basil refiere a su amigo Lord Henry su reticencia a exponer el magnífico retrato de Dorian. “La razón que me impide exponer ese cuadro dice es mi miedo a haber mostrado en él el secreto de mi alma”³⁴. En el cuadro se encuentra el artista. Se ha alienado. Pero tiene conciencia de que es imposible sujetar la belleza en el tiempo. Su fulgor es perecedero, en analogía directa con la repetición constante de aquello que nunca de manifiesta. La educación y la información se convierten en un artillero moderno de conservación para perdurar, y la mente de un hombre perfectamente informado, como apunta Lord Henry, es algo horrible³⁵. El destello del tiempo de la belleza, la caducidad de cada repetición de ese murmullo que apunta constantemente al otro lado, es la temática y la motivación de este extraordinario relato de Wilde en el que sentimos cómo, a pesar de todo, la vida se nos escapa de las manos.

En 1891, año de publicación de la última versión de *El retrato de Dorian Gray*, Wilde ya se encontraba enfrentado a su propio desdoblamiento personal y al debate contradictorio de la representación de su propia vida. ¿Qué era lo que reflejaba el espejo de Oscar Wilde? ¿Cómo sería el retrato que permanecería en una hipotética representación de su imagen? Probablemente la dicotomía existencial de Wilde era la causa de que su esteticismo estuviera demasiado condicionado por una ególatra vanidad entendida como armazón de supervivencia. En este sentido el narcisismo propio y de sus personajes podría corresponder a esa especie de dandi decadente y cínico que, paradójicamente, se abastece de la propia decadencia y que no puede prescindir de la alteridad social. El dandismo de Baudelaire, por el contrario, correspondería más bien a un autoextrañamiento y marginación voluntaria del yo, sin hacer tantas concesiones sociales como Wilde pudiera haber hecho en el controvertido y fatal caso de la relación con Lord Alfred Douglas³⁶; o del mismo modo, tan opuesto al del joven aristócrata Des Esseintes, en la novela de Joris-Karl Huysmans, *A contrapelo* (1884)³⁷, cuyo cinismo, pesimismo y misantropía se convierten en expresión clara, sin concesiones, ante la hipocresía de la sociedad moderna. No debemos olvidar, por lo mismo, que es precisamente esta obra la que se supone que Lord Henry Wotton entrega a Dorian Gray y le conduce a un camino de horror y de deprivación, probablemente a

34 *Ibid.*, p. 34.

35 Cf. *Ibid.*, p. 41.

36 Cfr. BALLESTEROS, Antonio, *Narciso y el doble*, pp. 303-304.

37 HUYSMANS, J.-K., *A contrapelo*. 8ª ed. Juan Herrero. Madrid: Cátedra, 2016.

partir de una aplicación extrema de los planteamientos de Des Esseintes. En todo caso, este destino de Dorian-Wilde ratificaría el desfase explícito entre una estética vital sustentada en la belleza y la imposibilidad trágica de llevarla a cabo debido a la inexorable contingencia del paso del tiempo.

Aunque nos detendremos brevemente en la duplicidad más evidente, la novela está atravesada en su estructura por los desdoblamientos: en primer lugar, como acabamos de sugerir, el del propio Wilde con sus tres alteridades: Basil, Dorian y Lord Herry; en segundo lugar, la de Basil como artista con su obra, en la cual no solo se proyecta, sino que se desintegra depositando en ella su alma, lo cual a la postre le costará la vida; en tercer lugar, la de Lord Henry, quien, aparentemente sujeto en su cinismo, no puede más que volcarse como mentor sobre su iniciado pupilo en el que aspira también permanecer con su influencia (“Proyectar el alma propia en una forma refinada, y dejarla recrearse allí por un momento” (...) “No había nada que no se pudiese hacer con él. Se le podía convertir en un titán o en un juguete”³⁸); y, evidentemente, en último lugar, el propio Dorian.

Al mirarse en el retrato, a través de los ojos de Lord Henry, Dorian consigue reconocerse y tomar conciencia de sí en su totalidad. La imprecación de que sea la representación la que sufra el paso del tiempo, las consecuencias del mal que habita en el alma y la perversión de la vida a la que lleva este hecho del auto-reconocimiento, parece desencadenar esa siniestra y fáustica concesión trágica cuya única resolución posible será, desde un principio, la inexorable y presentida muerte. De alguna manera el que Dorian, como en su momento también hiciera con la sombra de su alma el joven pescador del otro relato de Wilde al que hicimos referencia en el apartado anterior, se quisiera independizar de su doble para dejar en evidencia y en carne viva la escisión de la vida con la muerte o de los seres humanos con el indiferente mutismo de lo real, explícita y anuncia, frente a los antídotos del olvido, el irremisible destino de la extinción.

Conforme avanza el tiempo y nuestro protagonista se ve irremisiblemente envuelto en esta fatal enajenación percibimos cómo este desdoblamiento se convierte en especial al requerir simultáneamente la representación del cuadro, en creciente estado de degeneración, y el reflejo en el espejo de la imagen fría y radiante de una belleza descarnada. “... de pie, con un espejo, frente al retrato que Basil Hallward le había hecho, miraba alternativamente la malvada y envejecida cara del lienzo y el hermoso y joven rostro que le sonreía desde el cristal pulido. La intensidad misma del contraste solía agudizar su sensación de placer. Fue enamorándose cada vez más de su propia belleza e interesándose más cada vez en la

38 WILDE, Oscar, *El retrato de Dorian Gray*, pp. 66-67.

corrupción de su propia alma”³⁹. Dorian estaba ya más que enfermo de un trágico narcisismo. Siempre lo estuvo desde que comenzó a reconocerse en situación de privilegio habitando el otro lado. La reduplicación del mito es más que evidente desde todos los puntos de vista. Ahogarse en el abismo de su propia alma. Cuanta más vida tenía más vida deseaba. Instalado en una ambigüedad absoluta, al sentirse acosado por la amenaza del hermano de Sybil Vane, se mostraba “enfermo de un frenético miedo a morir y, sin embargo, indiferente a la vida”⁴⁰. Llegó un momento en que ya no soportaba su propia belleza, ni verse reflejado en el espejo que durante todo aquel tiempo había certificado su perdurabilidad. Lo rompió. Lo pisoteó. Y con la intención de expiar el dolor de la conciencia que se concentraba en el cuadro decidió destruirlo, destruyéndose. Nuevamente la constante de duplicación muestra su trágica paradoja al aliviar eventualmente la carga contingente de la vida, al mismo tiempo que reclama de manera inexorable la urgente resolución del autoexterminio exigido como vanidoso peaje de esta *hybris* que nos provoca el temor y la ignorancia.

5. CONTRA LA INVISIBILIZACIÓN DE LA MUERTE

Deberíamos en este momento, como contraste y liberación del carácter siniestro del doble de Gray, hacer brevemente mención de esa lectura aparentemente amable y compleja que implica el desdoblamiento, no de uno mismo solo sino de la totalidad del mundo en el espejo, en *Alicia a través del espejo* (1871) de Lewis Carroll. Lo que nos gustaría recordar, en todo caso, son tres aspectos: en primer lugar, que la casa del otro lado del espejo genera dudas y desconfianza, hasta el punto de que Alicia lo utiliza como posible castigo en el que recluir al gatito negro por no haberla hecho caso en sus juegos. Probablemente es ella misma la reticente en la tarea de profundizar en su propia identidad y de esa manera convertirse en una mujer adulta; en segundo lugar, que Alicia tiene muy claro que se trata de un lugar de transgresión, por tanto un objeto de deseo, especialmente porque todo está invertido e implica una realidad doblada (“... aquí no habrá nadie que me regañe por acercarme al fuego”⁴¹); y, en tercer lugar, que el mundo invertido es un sueño, una ilusión, en el que, a pesar de todo, la acción queda regulada por las propias reglas de la partida de ajedrez. Se trata de un sueño generado por el desafío de sentirnos interpelados por un mundo que late ante nosotros en silencio,

39 *Ibid.*, p. 172.

40 *Ibid.*, p. 256.

41 CARROLL, Lewis, *Alicia a través del espejo*. 3ª ed. Jaime de Ojeda. Madrid: Alianza, 2016, p. 173.

aunque anhelamos con vehemencia su complemento. En el sueño, el espejo se envuelve en una bruma plateada y brillante disolviéndose en las manos de Alicia. De este modo, cuando se convierte en liminaridad pura, es decir, en el mismo momento en que se produce la fusión de la niña con su reflejo, se hace invisible en el ámbito de lo real. Lo sabemos bien: por mucho que despertara Alicia de su sueño, para nosotros siempre formará parte de lo invisible, de lo desdoblado. Pero existe, en este mismo marco en el que nos desenvolvemos, otro caso más específico de invisibilidad del doble.

Sabemos que algunas de las obras literarias que forman parte del acervo cultural y de sus proyecciones más significativas han trascendido no especialmente por su gran valor literario sino por la potencia de esa intuición central que encierran sustentada y dirigida a los intereses profundos de los seres humanos como tales o como integrantes de una época específica. El caso de *El hombre invisible* (1898), de G.H. Wells es de este tipo. A nuestro entender se trata de una obra no especialmente destacable desde un punto de vista literario, pero igual de sugerente que la mayor parte de todas las obras de Wells en su empeño por vislumbrar hacia dónde se orientaban los retos del hombre en la época de la eclosión de la ciencia y la técnica. No nos detendremos ahora en este asunto concreto de cuáles son los límites y las consecuencias del uso técnico y tecnológico de la ciencia. Lo que más nos interesa de la novela es la cuestión, simbólicamente expresada, de la identidad, y de la posibilidad, o no, del reflejo de esta. Se podría decir que la estructura antropológica profunda del relato se haya en la ambivalente tensión que vive la especie humana, por una parte, por querer hacerse invisible para resguardarse y protegerse de los depredadores y los peligros, es decir, para sobrevivir y alejar a la muerte, y, por otra, por la necesidad imperiosa de que los otros nos reconozcan a través de la proyección de nuestra imagen, es decir, que nos vean, para poder reconocernos a nosotros mismos.

Pocas veces los diferentes pensadores que, de una manera u otra, se han ocupado del tema central de la identidad se han planteado el asunto desde la no presencia de un algo, de alguna suerte de imagen, que se pueda identificar. ¿Cómo cabría hablar de una identidad que excluya su visibilidad, sea esta implícita o explícita? Parecería, por tanto, que nos vemos impelidos a identificar de manera inmediata existencia con visibilidad; hasta el punto de afirmar que lo invisible es inexistente, con todo lo que ello implica. Los seres humanos desde siempre le han tenido pavor a perder la identidad, independientemente de que supieran lo que era esto. En este caso que evaluamos ahora es la literatura, de manera metafórica, la que nos anuncia la invisibilidad como una analogía de la carencia; de aquella sabiduría trágica que nos recuerda que el yo siempre resulta ser posterior a la oquedad y que se construye sustancial o trascendentalmente desde ese vacío que exige siempre un complemento.

El hombre invisible funciona como fábula de la desmaterialización de la imagen del hombre. Además nos damos cuenta de cómo la despechada soberbia de científico del protagonista, de Griffin, que, como Frankenstein o Jekyll, quieren ir más allá de los límites, le hace enfermar de una vanidad narcisista que le impone hacerse famoso de la noche a la mañana y que le permitirá dominar el misterio, el poder y la libertad⁴². Hasta el punto de, para continuar la investigación que le permitiría tener éxito en su proyecto de invisibilización, robarle a su padre dinero que no era suyo y provocarle, por ello, la muerte. Griffin, al obtener el éxito buscado, consiguió que el hombre invisible se convirtiera en un doble perverso de sí mismo. En el momento mismo de invisibilizarse y hacerse otro, el doble es también una especie del desdoblamiento ya que obviamente se ha convertido en dos: en un hombre y en su invisibilidad. Griffin carecerá también de cualquier tipo de reflejo ante el espejo porque desdoblándose hacia la desaparición ha perdido cualquier tipo de imagen y la posibilidad de quedar reflejado. La consecuencia de esto, como ya sabemos, es la transformación en monstruo. Toda esta peculiaridad se convierte en un arma de desdicha y de poder con la que zafarse, hacer el mal, robar e intentar dominar a los demás ante el rechazo de su extrañeza en el comportamiento y en el aspecto siempre tapado hasta el último recoveco de su cuerpo. De manera obvia el destino del *Hombre Invisible* tendrá que ver en una forzada concepción negativa del narcisismo cientificista, lo que finalmente le convertirá en un iluminado que gracias a su descubrimiento se veía capaz de sembrar el terror e instaurar un nuevo orden bajo su dominio. Lo que constatamos cuando consiguen darle caza y matarle tras la última pelea con Kemp es que nuevamente la fantasía literaria corrobora como lo que se invisibilizaba no era tanto la vida como la muerte, ya que Wells se encarga de describir al final el modo en que, muerto Griffin, el cuerpo, las arterias, los huesos, la piel, etc., recuperaba la forma materia haciéndose presente. La muerte nos vuelve a hacer visibles en la época invisible de la hipervisibilidad. Se acaba con la peligrosa invisibilidad voluntaria y libre de quien reclama el poder de manera individual y activa. La invisibilidad debe seguir teniendo que ver con el anonimato instituido. No debemos olvidar, en este sentido, que el doble manifiesta aquí un trasunto social. Al ser entendido como algo invisible, podría ser fiel reflejo de la propia invisibilidad del individuo alienado en el anonimato de la masa social que daba cuerpo al espíritu de la Revolución Industrial, época a la que, como sabemos, pertenece el relato y la preocupación literaria de H.G. Wells en su conjunto.

Nuestra época parece que se rige por un predominio de la clonicidad hueca generalizada en la hipervisibilidad, como decimos. Hoy podríamos corroborar que

42 WELLS, H.G., *El hombre invisible*. Miguel Ángel Pérez. Madrid: Alianza, 2015, p.142.

se está materializando esta duplicación yendo más allá de la tematización moderna de las propuestas fantásticas, animistas, diabólicas, mecanicistas, etc., propias del final del siglo XVIII, principios del XIX y sus proyecciones. La transparencia y el exceso líquido de visibilidad saturan los reflejos y ahuecan la potencia semántica de la duplicidad. Con el talante de una hermenéutica narrativa, hemos querido reflexionar, junto a algunos relatos esenciales sobre los que se proyecta el imaginario ambivalente del reflejo, con objeto de insistir en la necesidad de recuperar la radicalidad trágica del doble. Se trata de rescatar la muerte y su tremenda alteridad para la vida. El doble, desde esta perspectiva entendido, es recordatorio, y no olvido, de la indiferencia sagrada de lo real. El hueco vacío nos devuelve la vida robada porque hace que lo otro no sea alienante. Justamente lo contrario a lo que les ocurría a aquellos extraños personajes de *La invención de Morel*⁴³, de Adolfo Bioy Casares, cuya naturaleza era totalmente desconocida tanto para el narrador como para los lectores. Es cierto que al final de esta novela descubrimos que se trata de imágenes producidas por la máquina de Morel, pero su extraño desdoblamiento las hacía inalcanzables. Únicamente estaban vivos de manera virtual, en tanto que imágenes y aire. Fuera de la isla habían perdido la vida, estaban muertos. Eran seres que se habían perdido para siempre en su desdoblamiento ficticio, a pesar de haber pretendido todo lo contrario.

REFERENCIAS BIBLIOGRÁFICAS

- BALLESTEROS, Antonio, *Narciso y el doble en la literatura fantástica victoriana*. Cuenca: Universidad de Castilla-La Mancha, 1998.
- BLUMENBERG, Hans, *La legitimación de la Edad Moderna*. Pedro Madrigal. Valencia: Pre-Textos, 2008.
- CARROLL, Lewis, *Alicia a través del espejo*. 3ª ed. Madrid: Alianza, 2016.
- CASARES, Bioy, *La invención de Morel*. 2ª ed. Barcelona, Austral (Espasa), 2016.
- COTARELO, Ramón, *La fábula del otro yo*. Valencia: UNED-Centro Francisco Tomás y Valiente, 2005.
- DASTUR, Françoise, *La muerte. Ensayo sobre la finitud*. Barcelona, Herder, 2008.
- ESTEBAN ORTEGA, Joaquín, *El escorzo melancólico de lo real. Ensayo sobre el decrecimiento hermenéutico y los límites de la desnudez*. Granada: Comares, 2020.
- FERNÁNDEZ VALENTÍ, T. y NAVARRO, Antonio José, *Frankenstein. El mito de la vida artificial*. Madrid: Nuer, 2000.
- GOETHE, Johann Wolfgang von, *Fausto*. Barcelona: Planeta, 1980.

43 CASARES, Bioy, *La invención de Morel*. 2ª ed. Sara Josefina Demaría. Barcelona, Austral (Espasa), 2016.

- GONZÁLEZ REQUENA, Jesús, “El Estadio del Espejo ‘de Jacques Lacan’. Crónica de una mascarada (1)”. *Trama y Fondo. Lectura y teoría del texto*, nº 37 (2014).
- GUBERN, Román, “Del rostro al retrato”. *Anàlisi*, 27, 2001.
- HOFFMANN, E.T.A. “La historia del reflejo perdido”. En MOLINA FOIX, Juan Antonio (Ed.), *Álter ego. Cuentos de dobles (Una antología)*. Aurora Nolla. Madrid: Siruela, 2007.
- HUYSMANS, J.-K., *A contrapelo*. 8ª ed. Madrid: Cátedra, 2016.
- LACAN, J. “El estadio del espejo como formador de la función del yo [je] tal como se nos revela en la experiencia psicoanalítica”. En *Escritos I*. México: Siglo XXI, 2009.
- MOLINA RUIZ, I., *El autorretrato como canalizador del dolor*. Granada: Universidad de Granada, 2015.
- MORIN, Edgar, *El hombre y la muerte*. 2ª ed. Barcelona: Kairós, 1994.
- OVIDIO, *Metamorfosis*. Madrid: Cátedra, 2016.
- PENNAC, Daniel, *Diario de un cuerpo*. Manuel Serrat. Barcelona: Mondadori, 2012.
- PLANELLA, Jordi, *Los monstruos*. Barcelona: UOC, 2007.
- POE, Edgar Allan, *Cuentos completos*. Barcelona: Penguin, 2015.
- RANK, Otto, *El doble*. Buenos Aires: Orión, 1982 (Or. 1925).
- ROCHE, Juan Antonio, *La sociedad evanescente*. Barcelona: Anthropos, 2009.
- SHELLEY, M., *Frankenstein, o el moderno Prometeo*. Barcelona: Penguin, 2015.
- STOKER, B., *Drácula*. Mario Montalbán. Penguin: Barcelona, 2015.
- TYMMS, Ralph, *Doubles in Literary Psychology*. Cambridge: Bowes & Bowes, 1949.
- WELLS, H.G., *El hombre invisible*. Madrid: Alianza, 2015.
- WILDE, Oscar, *El retrato de Dorian Gray*. Mauro Armíño. Barcelona: Austral (Espasa), 2010.
- , *La decadencia de la mentira. Un comentario*. Javier Fernández. Barcelona, Acan-tilado, 2014.

CIENCIA Y RELATO VEROSÍMIL: ¿ES POSIBLE HACER FILOSOFÍA DE LA NATURALEZA?

SCIENCE AND PLAUSIBLE STORIES: ¿IS IT POS-
SIBLE TO DO PHILOSOPHY OF NATURE?

MANUEL C. ORTIZ DE LANDÁZURI
Profesor Contratado. Doctor en Filosofía
Departamento de Filosofía
Universidad de Navarra
Pamplona/España
mcruz@unav.es
ORCID: 0000-0002-1043-4879

Recibido: 30/04/2021
Revisado: 29/07/2021
Aceptado: 21/09/2021

Resumen: frente a los grandes descubrimientos de la ciencia, parece que la filosofía ya no tiene un papel relevante acerca de qué es la vida y cuál es el lugar del ser humano en el cosmos. La ciencia sería la que, mediante un conocimiento objetivo y basado en la experiencia empírica, revela los secretos de la vida y el universo. En este artículo pretendo mostrar que las ciencias son un tipo de relato con un grado de verosimilitud restringido a un campo de aplicación muy concreto. Más allá del nivel de relato descriptivo-mecánico, el ser humano necesita un relato más profundo acerca de la vida y su papel en el cosmos: aquí es donde entra en juego la filosofía, como un tipo de relato diferente, con unos criterios de verosimilitud distintos. Como paradigma del problema de la superposición de relatos, trataré de mostrar algunos problemas que ha suscitado la teoría sintética de la evolución, como una postura científica, pero que a la vez se ha vuelto filosófica.

Palabras clave: Foucault, neodarwinismo, Popper, razón, verosimilitud.

Abstract: In the face of the great discoveries of science, it seems that philosophy no longer has a relevant role to play in the question of what life is and what is the place of human beings in the cosmos. Science would be the one that, through objective knowledge based on empirical experience, reveals the secrets of life and the universe. In this article I try to show that science is a kind of story with a degree of plausibility restricted to a very specific field of application. Beyond the level of the descriptive-mechanical account, human beings need a deeper account of life and its role in the cosmos: this is where philosophy comes into play, as a different kind of account, with different criteria of plausibility. As a

paradigm of the problem of overlapping narratives, I will try to show some of the problems raised by the synthetic theory of evolution, as a scientific stance, but one that has become philosophical at the same time.

Keywords: Foucault, neodarwinismo, Popper, reason, verisimilitude.

1. INTRODUCCIÓN

¿Puede el ser humano tener conocimiento profundo del mundo físico, “más allá” de las descripciones de las ciencias naturales? Evidentemente, la cuestión del conocimiento de la realidad es un problema que difícilmente se podría abordar en unas pocas páginas; sin embargo, en este artículo quisiera tratar esta cuestión desde la perspectiva de verosimilitud del relato. De algún modo, podemos entender que toda explicación es en realidad un relato que proporciona una imagen, una historia, en la cual hay protagonistas y acciones. Basta tomar un libro de bioquímica para ver que allí se habla de “células”, “moléculas”, “átomos de hidrógeno”, etc.: elementos que interactúan y se engarzan en procesos (historias). Nuestra manera de entender algo es a través de un relato pero, por otro lado, nos damos cuenta que no todos los relatos son igualmente objetivos ni igualmente verosímiles. Si uno dijera que la tierra gira porque hay unos duendes bajo tierra que están haciendo girar unos mecanismos, diríamos que no es verosímil, porque no tenemos noticia alguna de esos duendes. El relato no es creíble porque no hay datos empíricos que lo avalen, ni tampoco una razón argumentada que lo sustente.

Entendemos mediante relatos, y a la vez necesitamos relatos que vayan más allá de lo meramente descriptivo y material y proporcionen respuestas profundas a los problemas de la existencia humana. En este sentido, siempre habrá un relato metafísico acerca de la vida y la naturaleza. La pregunta entonces, objeto de este artículo, es si es posible elaborar un relato filosófico verosímil que vaya más allá del relato de la ciencia. Para analizar esta cuestión, trataré en primer lugar de afrontar algunos problemas que podría presentar una propuesta semejante. Por un lado, es preciso cuestionar la visión positivista del Círculo de Viena, que tanta influencia ha tenido a la hora de considerar la filosofía de la naturaleza como un mero instrumento de análisis lógico de las proposiciones científicas. Por otro lado, se hace necesario mostrar algunas deficiencias de la visión del saber-poder de Michel Foucault, que podrían derivar en la relatividad absoluta de cualquier relato. Trataré a continuación de mostrar la superposición de relatos que se ha dado en las últimas décadas, especialmente a raíz de la teoría sintética de la evolución. Por último, indicaré algunos criterios de verosimilitud de un relato filosófico, en la línea de Karl Popper.

Sin embargo, antes de seguir adelante, conviene definir qué se entiende por relato. En términos generales un relato es la narración de una acción, es decir, una perspectiva comprensiva acerca de algo que sucede. Si tomamos un suceso, un dato X, ese suceso por sí mismo no explica nada: es preciso integrar el dato, el suceso, en un hilo argumental, enfocarlo desde una perspectiva que permita comprenderlo a la luz de una causa. Los relatos permiten así comprender lo que sucede desde una diversa perspectiva. Ahora bien, lo que está en juego es si cualquier relato es igualmente válido para comprender los sucesos y si todos los relatos sirven para explicar del mismo modo. En este artículo defenderé precisamente que hay criterios de verosimilitud para tomar un relato como válido, y que existen diversos niveles de relato, cada uno con su propio criterio de verosimilitud.

2. CIENCIA Y FILOSOFÍA EN EL CÍRCULO DE VIENA

Hace ya más de cien años que la filosofía del Círculo de Viena, bajo la influencia de Mach, marcó un abismo infranqueable entre la objetividad lógica de la ciencia natural y la subjetividad de la filosofía especulativa. Es quizás Reichenbach quien nos ha dejado textos verdaderamente claros al respecto:

Se debe considerar esta clase de filosofía [metafísica] como una superestructura racional basada en experiencias emocionales. Estos filósofos desarrollan una teoría del conocimiento, una concepción del ser, a veces una especie de cosmología incluso, todo lo cual está construido, evidentemente, sin consideración al análisis lógico y epistemológico. Su única finalidad es la de disfrazar actitudes emocionales mediante un sistema conceptual pseudocientífico. Puede ser que esta forma de poesía conceptual satisfaga emocionalmente a sus admiradores, del mismo modo que las creaciones teológicas y mitológicas han satisfecho a los creyentes durante miles de años¹.

Frente a un relato objetivo, basado en hechos empíricos observables y en el análisis lógico de las proposiciones, la filosofía especulativa supondría un tipo de relato poético, sin ningún tipo de verosimilitud. La moderna filosofía de la naturaleza –piensa Reichenbach– debe ocuparse del análisis lógico de las proposiciones de la ciencia, y nada más (es decir, debe convertirse en filosofía de la ciencia). Así es como lo expresa también Schlick:

La meta de la filosofía de la naturaleza es interpretar el significado de las proposiciones de la ciencia natural y, por lo tanto, la filosofía de la naturaleza no es en

1 REICHENBACH, H. “Fines y métodos de la filosofía de la naturaleza moderna”. En: *Moderna filosofía de la ciencia*. Madrid: Tecnos, 1965, pp. 130-131.

sí misma una ciencia sino una actividad dirigida a la reflexión sobre el significado de las leyes de la naturaleza².

En el fondo, tanto Schlick como Reichenbach piensan que el único relato necesario es el relato descriptivo, basado en hechos observables y leyes causales deterministas, y cuya verosimilitud vendría dada por la experiencia empírica. El segundo tipo de relato posible sería el de la filosofía de la naturaleza, entendida como un criterio de clarificación del primer relato, es decir, como un relato puramente lógico acerca del relato descriptivo. Dados los hechos, la filosofía de la naturaleza establece las condiciones para que el relato sea considerado científico: las leyes de la lógica. La filosofía de la naturaleza sería, en definitiva, filosofía de la ciencia.

Por tanto, para estos filósofos más allá de estos relatos no sería necesario otro tipo de explicación. Basta con las proposiciones científicas para tener conocimiento del mundo. Sin embargo, nos damos cuenta de que hay varios niveles de relato necesarios además de este. Por un lado, una ciencia necesita un nivel de explicación que permita entender los hechos de experiencia. Los puros “hechos” no explican nada. No es posible entender el papel de una molécula de dióxido de carbono al margen de una serie de procesos que tienen lugar en algo que llamamos célula, constituida por diferentes procesos y partes, como la mitocondria, el aparato de Golgi, etc. Incluso la ciencia necesita un relato que no es un puro “dato”, es una elaboración conceptual bastante real. Por otro lado, el ser humano necesita una explicación metafísica del mundo. El querer negar la posibilidad de verosimilitud de cualquier relato más allá del científico termina por otorgar al relato científico valor de verdad absoluta. Se elabora así un relato filosófico que trasciende el relato científico. Lo veremos con más claridad al tratar la cuestión de la teoría sintética de la evolución.

3. POSTMODERNIDAD Y RELATO

Una vez visto que el relato científicista no resulta suficiente, ya que incluso ese relato se fundamenta en un relato filosófico más profundo, con una concepción del conocimiento basada en la experiencia empírica, podríamos pensar que entonces cualquier relato acerca del mundo es válido, siempre y cuando nos resulte convincente. Tal es la visión postmoderna acerca de la ciencia y de la explicación filosófica del mundo: ya que el ser humano tiene necesidad de relato,

2 SCHLICK, M., *Filosofía de la naturaleza*. Madrid: Encuentro, 2002, p. 15.

necesariamente elaboramos explicaciones, pero detrás de estas no hay algo así como una verdad, sino simplemente un relato persuasivo, convincente, al servicio en último término de la estructura de poder imperante.

El ejemplo más claro de esta postura lo encontramos en Michel Foucault, que vincula el saber y el poder en un dinamismo que se retroalimenta. El saber de una época –piensa Foucault– proporciona una visión del mundo, una *episteme*, basada en unos procedimientos y prácticas establecidas por alguna forma de poder (la autoridad científica, por ejemplo). En el fondo, detrás de esa *episteme* lo que hay es una forma de poder que construye un saber con el que además se legitima. Todo saber es un relato elaborado para persuadir y dominar. El poder engendra el saber y, a la vez, el saber proporciona el poder. La única alternativa ante esa situación es mostrar la genealogía del saber a lo largo de la historia, de modo que seamos conscientes de que todo relato es en realidad relativo a un poder determinado. Así pues, no hay lugar para pensar que una teoría es verdadera. Lo único que tenemos son datos, que agrupamos en teorías según procedimientos de un poder establecido. Foucault es precisamente muy cuidadoso en afirmar que él no pretende hacer una teoría, ni siquiera sobre el poder:

Esto no podría ser a lo sumo, ni querría ser, más que el bosquejo de una teoría, no de lo que es el poder sino del poder, con la condición de admitir que éste no es justamente una sustancia, un fluido, algo que mana de esto o aquello, sino un conjunto de mecanismos y procedimientos cuyo papel o función y tema, aun cuando no lo logren, consisten precisamente en asegurar el poder³.

Si Foucault pretendiera hacer una teoría sobre el poder, él mismo estaría cayendo en aquello que quiere evitar: ser presa de unas prácticas y procedimientos científicos que, en realidad, son formas de poder. La única manera de escapar del poder es renunciar a establecer teorías, es decir, relatos definidos, como si fueran verdaderos:

Es la relación verdad/poder, conocimiento/poder lo que me preocupa. Así que, esta capa de objetos, esta capa de relaciones más bien, es difícil de atrapar; y ya que no tenemos una teoría general para entenderlos, soy, si se quiere, un empirista ciego⁴.

Ahora se entiende por qué no hay posibilidad de una teoría verdadera, ni siquiera verosímil. Lo único que se puede hacer es tomar datos, ser un *empirista ciego*, e interpretar como uno pueda.

3 FOUCAULT, M., *Seguridad, territorio, población*. Madrid: Akal, 2008, p. 14.

4 FOUCAULT, M., *Dits et écrits*, Vol. III. Paris: Gallimard, 1994, p. 216.

Mi discurso, lejos de determinar el lugar de donde habla, esquivo el suelo en el que podría apoyarse. Es un discurso sobre unos discursos; pero no pretende encontrar en ellos una ley oculta, un origen recubierto que solo habría que liberar; no pretende tampoco establecer por sí mismo y a partir de sí mismo la teoría general de la cual estos discursos serían los modelos concretos. [...] Se trata de operar un descentramiento que no deja privilegio a ningún centro⁵.

Ahora bien, si esto es así, entonces no hay criterio alguno para establecer la verosimilitud de un relato, más allá de su fuerza persuasiva. Foucault acierta a ver el nexo de unión entre el poder y el saber pero, si de hecho se han dado cambios de paradigma en la ciencia, muchas veces estos no iban en consonancia con las creencias científicas de la época. Basta pensar en la teoría de la relatividad, que tardó en ser aceptada, o en la propia física cuántica. Son los hechos y nuestra mejor explicación de ellos los que permiten modificar el relato para que sea más verosímil. De igual modo, es la argumentación racional la que permite decir si una explicación filosófica sobre el mundo es verosímil o no. Resulta claro que el ser humano tiene necesidad de estos relatos, y a la vez parece que tenemos criterios objetivos para establecer cuándo un relato es verosímil o no.

4. RELATOS ACERCA DE LA EVOLUCIÓN

Si de algo podemos estar seguros es de que el ser humano necesita un relato verosímil sobre el mundo para poder comprenderse a sí mismo. La propuesta de Foucault radica en situar esta necesidad de *episteme* como un elemento configurador del poder. El poder engendra un relato, y a la vez el relato genera un poder. Pero, como hemos tratado de mostrar, no nos basta cualquier relato y, en este sentido, quisiera defender que los relatos admiten diversos grados de verosimilitud, atendiendo a su objeto propio.

En concreto, acerca del mundo natural la cuestión más importante es entender qué es la vida. Que el espacio sea curvo o que haya un principio de indeterminación en las partículas elementales son cuestiones importantes para tener una imagen del universo, pero no supone un relato especialmente relevante para los problemas del ser humano: ¿quiénes somos? ¿qué papel jugamos en el cosmos? En cambio, a la hora de enfrentarnos a la vida, parece que las cuestiones se vuelven más interesantes: ¿es la vida fruto del azar? ¿hay un orden en las especies?

La ciencia obtiene un conocimiento descriptivo de los procesos naturales, pero hay cuestiones que exigen un nivel de relato más profundo. Incluso desde un

5 FOUCAULT, M., *La arqueología del saber*. México: Siglo XXI, 1970, p. 345.

punto de vista casi exclusivamente científico, parece imposible hacer una ciencia sin un nivel de relato que vaya más allá de la pura descripción de procesos. Es necesario interpretar, relacionar esos datos dentro de un modelo, dentro de una teoría. Claro está que la verosimilitud de este relato viene dada por su funcionalidad en la experiencia. Por ejemplo, el modelo de ADN de Watson y Crick resulta verosímil porque permite realizar numerosísimos experimentos exitosos. Pero, incluso más allá de este nivel de experiencia, la ciencia necesita de un relato interpretativo acerca de los modelos: en primer lugar, grandes teorías generales; en segundo lugar, visiones del mundo cargadas de cierta metafísica. No es por eso extraño que los científicos propongan interpretaciones que exceden el campo puro de la ciencia, o que sitúen un relato como el más “científico” y por eso mismo el más verdadero.

Un ejemplo paradigmático de esta situación lo encontramos en la “teoría sintética de la evolución”, que es el relato imperante acerca de cómo se debe entender la evolución de la vida. Se trata de un modelo científico que se ha convertido también en relato filosófico, sin que muchas veces sea sencillo establecer los límites entre uno y otro⁶. La razón es clara: necesariamente hace falta un modelo, un relato, desde el cual comprender los hechos, y así la ciencia da lugar a una especulación de tipo casi metafísico. Ahora bien, conviene precisamente distinguir la teoría sintética como modelo científico y su posterior evolución hacia un relato filosófico.

Como modelo científico (es decir, un relato acerca de la interpretación de los hechos), la teoría sintética simplemente propone que los cambios evolutivos se deben fundamentalmente a modificaciones al azar en el ADN y posterior selección natural. ¿Se trata de un relato verosímil? Sí, dependiendo de cómo utilicemos los términos “azar” y el matiz que se le quiera dar a la palabra “fundamentalmente”. Si por azar entendemos cierta imprevisibilidad que no podemos determinar debido a ignorancia, diríamos que es verosímil que se produzcan mutaciones al azar y que estas puedan ser la causa de la evolución. Decir “azar” aquí es decir “de modo que no es posible determinar con precisión por nuestra incapacidad de hacerlo”, igual que una bola en una ruleta cae en un número “por azar”: la bola está completamente determinada por las leyes de la física a caer donde ha caído, pero nos es muy difícil predecirlo. Si por azar entendemos la realización de probabilidades, entraríamos en un problema matemático que hace difícilmente

6 Ejemplos de esta actitud se encuentran en científicos y filósofos como Dawkins, Monod, Denett, Ruse y Ayala. Cfr. DAWKINS, R., *El relojero ciego*. Barcelona: Labor, 1988, MONOD, J., *El azar y la necesidad*. Barcelona: Orbis, 1985; DENETT, D. C., *La peligrosa idea de Darwin*. Barcelona: Galaxia Gutenberg, 1999; RUSE, M., *Darwin and Design*. Cambridge: Harvard University Press, 2003; AYALA, F. J., *Darwin y el diseño inteligente*. Madrid: Alianza, 2007.

verosímil la evolución. En este caso, se entiende que hay una posibilidad muy remota de que se produzca una variación favorable. Ahora bien, es cuestión de tiempo para que esa posibilidad se realice, ya que existe una ínfima probabilidad de que ocurra. ¿Existe una posibilidad de que tire varias maderas por la ventana y al caer formen una mesa? La posibilidad (como no contradicción lógica) existe, y se podría señalar con una probabilidad matemática. Ahora bien, tengo la certeza de que jamás se formará una mesa aunque tire la madera por la ventana miles de millones de veces.

En el caso de la palabra “fundamentalmente”, la cuestión es si la causa principal de la evolución son las mutaciones azarosas o si por el contrario hay otras causas igual o incluso más importantes para el cambio en las especies. En este sentido, hay voces que sugieren que habría que elaborar una nueva teoría de la evolución, más amplia que la sintética, que recoja aspectos ambientales, epigenéticos, la importancia del desarrollo embrionario⁷ e incluso mecanismos casi lamarckianos⁸. La teoría sintética como relato científico tiene cierta verosimilitud, pero con algunas limitaciones a las que tiene todavía que hacer frente, y no sería extraño que el relato necesite una reelaboración profunda⁹.

Ahora bien, está claro que este modelo ha originado un relato filosófico de corte naturalista (materialista) acerca del origen de la vida y los mecanismos evolutivos que sobrepasa, sin duda alguna, el ámbito de la ciencia. La teoría de la evolución como relato científico no implica el evolucionismo como relato filosófico¹⁰. De hecho, no hay un único relato interpretativo acerca de la evolución, sino que existen alternativas al neodarwinismo de Dawkins o Ruse, tal como se pueden encontrar en Gould¹¹ o Stuart Kauffman¹². Sin embargo, resulta evidente que el neodarwinismo radical de Dawkins y otros ha ejercido una influencia notable a la hora de configurar un relato materialista acerca de la vida humana y, precisamente por ello, merece una especial consideración. Según este relato, la ciencia habría logrado explicar la evolución como una concatenación de modificaciones fortuitas y determinadas por las leyes bioquímicas, que serían la causa del cambio y la diversificación de las especies. En otras palabras: la ciencia habría demostrado que todo orden en la naturaleza es en realidad aparente, es un equilibrio logrado a lo

7 Cfr. LAUBLICHER, M., y MAIENSCHIN, J. (eds.), *Form and Function in Developmental Evolution*. Cambridge: Cambridge University Press, 2009.

8 Cfr. JABLONKA, E.,-LAMB, M. J., *Evolution in Four Dimensions*. Cambridge: MIT Press, 2005, pp. 140-145.

9 Ver WHITFIELD, J., “Postmodern Evolution?”. *Nature* 455 (2018), pp. 281-284.

10 Cfr. ARTIGAS, M., *La inteligibilidad de la naturaleza*. Pamplona: Eunsa, 1992, p. 300.

11 Cfr. GOULD, S., *Ontogeny and Phylogeny*. Cambridge: Harvard University Press, 1977.

12 Cfr. KAUFFMAN, S., *At Home in the Universe : The Search for Laws of Self-Organization and Complexity*. New York : Viking, 1995.

largo del tiempo, pero que no hay nada más que materia y leyes físico-químicas en la naturaleza. La clave de este relato, y por lo que es tan persuasivo y, hasta cierto punto verosímil, es que se presenta una postura filosófica como una explicación científica, basada en los hechos. El mero estudio de los procesos bioquímicos habría revelado que el orden de los seres vivos es, en realidad, un equilibrio complejo pero fortuito, que ha incorporado cambios a lo largo del tiempo. Un relato filosófico se ha elaborado, quizás de modo inconsciente, a partir de un relato científico. El científico podría decir: “yo me limito a describir procesos”, pero aquí hemos dado un paso más: los procesos dibujan un relato mecanístico-determinista por el cual todo orden es, en realidad, equilibrio fortuito. Pero al dar este paso, hemos elaborado un relato explicativo que rebasa la ciencia, pues estamos diciendo que la realidad es materia determinada, y *nada más que eso*. Este último matiz es importante. La ciencia necesariamente considera la naturaleza como materia, y nada más que eso, pero de ahí no demuestra que solo sea materia, sino que considera solo el aspecto material. La cuestión en juego es si podemos explicar todo el orden solo desde ese aspecto material y, frente a este problema, la respuesta es que podemos describir los procesos que han generado el orden, pero no saber por qué hay orden.

5. VEROSIMILITUD DEL RELATO EVOLUCIONISTA NEODARWINISTA

Se podría argumentar que son los resultados de la ciencia los que apuntan a la verosimilitud del relato evolucionista radical, según el cual desde la propia bioquímica se puede explicar la vida. En realidad, la ciencia simplemente llega a la descripción de ciertos procesos bioquímicos, y frente a ciertos problemas se encuentra lejos de haber hallado una respuesta. Por otro lado, en el relato neodarwinista lo que subyace de fondo con un poco de análisis es una petición de principio. Tratemos de verlo. Si estudiamos la realidad desde un punto de vista material-descriptivo, mi comprensión de esa realidad será siempre material y descriptiva. Dado el punto de vista, dado el resultado. Puesto que estudio la naturaleza ateniéndome a lo material, mi resultado siempre será una explicación material, pero de ahí no puedo inferir que entonces la naturaleza se explique solo por lo material: se explicará de modo material en las ciencias naturales, que han optado por ese punto de vista y elaboran un relato en ese nivel. Evidentemente, si parto de que todo es material, la conclusión es que todo es material. La frase “todo ser material puede ser explicado por sus propiedades físico-químicas” resulta una tautología, ya que no estoy afirmando nada más que ser materia=propiedades físico químicas. Esto es algo bastante obvio. Mientras sepamos en qué nivel de relato nos movemos (descriptivo-material), podemos delimitar los criterios de

verosimilitud en tal tipo de relato (resultados experimentales). Ahora bien, si nos movemos hacia otro nivel de relato (explicación filosófica), entonces el grado de verosimilitud no viene dado únicamente por la experiencia empírica.

Ahora bien, ¿es el relato evolucionista materialista un relato verosímil? ¿Podemos explicar la vida por los procesos bioquímicos? Podemos describirla, ciertamente, pero la cuestión del orden queda latente. Sin duda alguna en las últimas décadas han surgido científicos y filósofos que han defendido que mediante el azar (entendido como determinación en la naturaleza) y la selección natural se puede explicar todo el orden y complejidad de los seres vivos:

Entonces, ¿cómo se puede pasar de una bacteria a un camello? Gran parte de la respuesta es: ¡lentamente! El Neodarwinismo explica un proceso de este tipo descomponiendo un gran número de pequeños pasos. [...] Los organismos tienen su estructura y comportamiento afectados por las mutaciones: generalmente para mal, pero ocasionalmente para bien. Los organismos viven, se reproducen y mueren, y los que llevan genes nuevos se reproducen más o menos que otros organismos de la población. Si hay más “información específica compleja” en un camello que en una bacteria, entonces el proceso natural descrito anteriormente es capaz de crear esta información. No hay ningún misterio metafísico en tal creación¹³.

Godfrey-Smith sitúa el relato científico neodarwinista como la última explicación del orden en la naturaleza, como un relato exitoso frente al “misterio metafísico” de otras concepciones. Pero, en realidad, este relato es tan metafísico como otros, ya que supone que la materia por sí misma, desde su propia determinación, es la única fuente de orden. Sin embargo, si esto es así, habría que preguntarse cómo las leyes que rigen la materia permiten una complejidad tan específica en el caso de los seres vivos. La causa del orden sigue siendo un agujero negro que desconocemos, y es preciso comenzar a hacer metafísica: o bien decir que la materia por sí misma se basta a sí misma, o bien admitir otras fuentes de orden, bien sea una fuerza vital presente en la naturaleza (como sostiene el vitalismo), bien sea una causa extrínseca a la naturaleza.

Para Dawkins el orden en la naturaleza no es un gran problema para el relato neodarwinista, ya que desde su perspectiva el azar terminará produciendo alguna modificación favorable. Precisamente por eso el núcleo de la propuesta de Dawkins, desde el punto de vista filosófico, se encuentra en una concepción de la materia según la cual esta realizará en el transcurso del tiempo todas las probabilidades. De este modo, su explicación de la evolución hacia especies más complejas se debe a la necesidad de la posibilidad lógica. Dice Dawkins:

13 GODFREY-SMITH, P., “Information and the Argument from Design”. En: PENNOCK, R. T. (Ed.), *Intelligent Design and Its Critics*. Cambridge: MIT Press, 2001, pp. 587-588.

Mi idea es que, siempre que la diferencia entre los intermediarios vecinos de nuestra serie que conduce al ojo sea suficientemente pequeña, las mutaciones precisas están casi obligadas a aparecer¹⁴.

Dawkins intenta explicar cómo el ojo humano actual puede surgir por pequeñas modificaciones de algo previo. Si es posible imaginar una serie de pasos intermedios, entonces es posible que se haya dado el proceso a lo largo de generaciones. Como actualmente tenemos ojos complejos y es posible rastrear ojos más sencillos e incluso algo parecido al ojo pero sin llegar todavía a serlo (células fotosensibles), entonces es que la evolución se ha dado por pequeñas modificaciones azarosas en el ADN. Si es posible que surjan mutaciones precisas, entonces es que necesariamente se han dado por azar. Sin embargo, aquí estamos ante un argumento que pasa de la posibilidad lógica a la posibilidad real. Yo me puedo imaginar el proceso, pero eso no significa que se haya dado (cuestión distinta es que puedo encontrar evidencias, como versiones más simples del ojo en otros seres vivos, que sugieran que se haya dado) y, todavía más grave, puedo imaginar que se han producido las mutaciones precisas, pero no lo puedo demostrar o, en todo caso, no puedo asegurar que se deban al azar (al relojero ciego). Quizás quede más clara esta idea con el ejemplo que el utiliza al principio de *El relojero ciego* sobre las golondrinas:

Las golondrinas vuelan. Como hemos visto, no es fácil construir una máquina voladora. Si se observan las células de una golondrina, y se colocan juntas al azar, la probabilidad de que el objeto resultante vuele no es, para nuestros propósitos, distinta de cero. No todos los seres vivos vuelan, pero hacen otras cosas que son igualmente poco probables¹⁵.

Dawkins dice: la posibilidad (por azar y selección natural) de que surgiera un ser vivo que volara era ínfima, pero no era cero. Como existen seres vivos que vuelan, significa que esa posibilidad remota se ha dado. Al igual que piensa Monod, nos ha tocado el número en la ruleta de Montecarlo¹⁶. Ahora bien, esto supone una petición de principio: dado que todo se rige por azar y selección natural, los seres vivos se han desarrollado así por azar y selección natural, aunque las probabilidades fueran mínimas. Sin embargo, el relato es poco verosímil porque presupone que la materia realiza todas las probabilidades posibles si se le da el tiempo oportuno. Esto, evidentemente, no es creíble. Desde un punto de vista lógico, ¿hay alguna probabilidad de que si arrojo unas maderas por la ventana se forme una mesa al caer? Desde luego que la probabilidad, aunque ínfima, existe.

14 DAWKINS, R., *El relojero ciego*. Barcelona: Labor, 1988, p. 60.

15 DAWKINS, R., *El relojero ciego*. Barcelona: Labor, Barcelona, 1988, p. 6.

16 Cfr. MONOD, J., *El azar y la necesidad*. Barcelona: Orbis, 1985.

Pero el hecho de que exista la probabilidad no significa que eso vaya a suceder¹⁷. Es más, eso no va a suceder nunca, al menos no mientras no haya algo que prepare las cosas para que las maderas caigan en el lugar conveniente. Solo habría dos modos de admitir que vaya a suceder esa posibilidad: la primera, suponiendo de antemano que la materia se comporta con un determinismo total, y que dado el tiempo oportuno, se realizará la probabilidad establecida (es decir, que todas las probabilidades en algún momento dado se realizarán). En realidad, un relato así no es verosímil ni falsable, porque siempre puedo interpretar los hechos de modo que encajen perfectamente en el relato, que está dado de antemano (si las golondrinas vuelan es porque existía una posibilidad ínfima de que sucediera, y por tanto ha sucedido). La segunda posibilidad sería pensar que existe una posibilidad basada en una tendencia natural a agruparse los elementos de modo que, conforme pase el tiempo, se apilen como una mesa. Lógicamente, esta tendencia no se ha demostrado (por mucho que tiremos las maderas, no hemos comprobado que se forme una mesa). En el caso de la evolución, quizás se podría decir que los elementos básicos tienen una tendencia debido a las cargas electrónicas de los átomos a agruparse en ciertas estructuras moleculares básicas, y estas a su vez en ciertos procesos vitales, etc. Si esto realmente es así (quizás la ciencia pueda algún día demostrarlo) habría que hablar de una tendencia natural en los elementos a la constitución de formas de vida (las golondrinas vuelan porque había una tendencia de los organismos a desarrollarse en esa dirección, dados sus constituyentes básicos). Sin embargo, precisamente el relato filosófico neodarwinista quiere prescindir de cualquier explicación finalista, en clave de “tendencia natural” para poder dar cuenta de la evolución.

Así pues, conviene distinguir el neodarwinismo científico del neodarwinismo filosófico. El relato científico neodarwinista puede ser válido para la explicación de los hechos empíricos, pero no anula consideraciones filosóficas y metafísicas de otro tipo. La ciencia observa la naturaleza desde un paradigma material (y debe hacerlo así), pero se puede observar la naturaleza desde otras perspectivas, y es posible indagar en fundamentos que no son estrictamente materiales. Este es un punto que no podrá ser admitido por quien busque un relato científico mecanicista. Por ejemplo, el filósofo Michael Ruse rechaza el vitalismo porque cree que si una idea (como el *élan vital* de Bergson) no tiene valor predictivo, entonces no tiene valor explicativo alguno. Explicar para Ruse es poder predecir, dar cuenta del mecanismo. Sin embargo, ¿explicar no es también comprender, basándose en razones? ¿No estará Ruse atrapado en el paradigma descriptivo de la ciencia mecanicista? Por otro lado, el propio Ruse cae en una severa contradicción al

17 Cfr. PAPP, D., *Filosofía de las leyes naturales*. Madrid: Espasa-Calpe, 1945, pp. 168-169.

asignar al neodarwinismo basado en mutaciones aleatorias un poder predictivo que no posee:

Gracias a teóricos como Fisher y practicantes como Dobzhanski, los neo-darwinistas tuvieron una mejor comprensión de cómo las mutaciones puramente aleatorias pueden, sin embargo, ser la base de características altamente compleja¹⁸.

En realidad, si el neodarwinismo tiene valor como relato es porque ofrece una explicación comprensiva de los organismos, y no por su valor predictivo (de hecho, no puede predecir cómo van a evolucionar los organismos).

6. CONCLUSIONES: VEROSIMILITUD Y RELATO

Si algo he tratado de mostrar a lo largo de este artículo, es que el ser humano siempre elaborará un relato filosófico acerca del universo. Tal necesidad de explicación viene dada fundamentalmente por la inquietud de saber qué lugar ocupamos en el cosmos: ¿somos productos del azar? ¿hay orden en la naturaleza? Ante estas preguntas, la ciencia puede aportar un relato descriptivo muy verosímil, ya que parte de los hechos de experiencia y únicamente remite a ellos. La ciencia puede relatar bastante bien los acontecimientos a nivel material y empírico. Ahora bien, resulta claro que solo con este nivel de relato no nos basta: necesitamos interpretar los hechos en un relato más amplio, y es aquí donde la filosofía entra en juego. Este nivel de relato tiene también sus propios criterios de verosimilitud, basados en la argumentación lógica y en la intuición intelectual.

Por ejemplo, se ha discutido mucho sobre la posibilidad de universos paralelos, la conocida teoría de los universos múltiples. La base para admitir los “multiversos” es que en el mundo cuántico hay una indeterminación en el comportamiento de las partículas elementales, de modo que no es posible predecir con absoluta certeza (como sí ocurría en la física anterior) cómo se comportará una partícula: mediante las ecuaciones de Schrödinger tenemos un conocimiento probable acerca del mundo cuántico, y es solo cuando interviene el observador que hace colapsar la función de onda, de modo que la partícula elemental se determina por una de las posibilidades que establece la función. Ahora bien, si esto es así, parece que en el universo existe la indeterminación, que se realizan diversas probabilidades sin una causa determinada. Para evitar este problema, una solución es elaborar un relato alternativo: cada vez que se da la posibilidad de varias opciones para la partícula elemental, en realidad se actualizan todas las

18 RUSE, M., *Darwin and Design*. Cambridge: Harvard University Press, 2003, p. 255.

posibilidades a la vez, solo que en universos paralelos. De este modo, surge una explicación filosófica, basada en una metafísica determinista, de muy dudosa verosimilitud. ¿Es razonable pensar que todas las posibilidades se realizan? ¿Por qué deberíamos pensar una cosa así? Un relato filosófico deja de ser verosímil cuando no existe una razón bien argumentada que lo sostenga. Evidentemente, a veces no tenemos una razón del todo definida frente a un problema, pero, cuando existen diversos relatos, es preciso someterlos a crítica racional para ver cuál es el más verosímil.

¿Significa esto que no podemos alcanzar una verdad acerca de la naturaleza, y que todo relato tiene una validez relativa? Ciertamente, sobre la naturaleza quizás nunca podremos tener un relato cerrado, pero decir que un relato es verosímil no quiere decir que sea falso, sino que es una aproximación parcial a la verdad. Acerca de la naturaleza, como afirma Platón en el *Timeo*, no podemos posiblemente tener una verdad absoluta, sino una explicación verosímil, basada en razones: “es conveniente que admitamos acerca de estas cosas un relato verosímil y no buscar más que eso” (*Timeo* 29d). La ciencia obtiene su verosimilitud a partir de su conformidad con los datos empíricos, la filosofía la obtiene a partir de la argumentación lógica y las evidencias intelectuales.

En este sentido, la propuesta que mantengo en este artículo acerca de los relatos filosóficos de la naturaleza se encuentra próxima a la idea de verosimilitud de Karl Popper. Como es bien sabido, Popper piensa que lo que hace que una teoría sea tomada en consideración como científica es su falsabilidad: es decir, que pueda ser contrastada por la experiencia y ser derribada por una teoría mejor. De este modo se alejaba de la propuesta del positivismo lógico:

Los miembros del primer grupo –los verificacionistas o justificacionistas– sostienen, en términos muy generales, que todo lo que puede ser apoyado por razones positivas es indigno de ser creído, y hasta de ser tomado seriamente en consideración. Por otro lado, los miembros del segundo grupo –los refutacionistas o falibilistas– dicen, en términos muy generales, que lo que (en el presente) no es posible, en principio, derribar por la crítica es (en el presente) indigno de ser considerado seriamente; mientras que lo que es posible, en principio, de refutar y, sin embargo, resiste todos nuestros esfuerzos críticos por hacerlo, quizás sea falso, pero de todos modos no es indigno de ser considerado seriamente y hasta de ser creído, aunque solo sea tentativamente¹⁹.

Es decir, Popper propone que un relato resulta verosímil mientras pueda ser considerado críticamente y susceptible de ser refutado. Si resiste la crítica, entonces es verosímil: quizás no del todo verdadero, porque podrá ser superado

19 POPPER, K., *Conjeturas y refutaciones*. Barcelona: Paidós, 1983, p. 279.

o mejorado por un relato mejor, pero sí digno de ser creído actualmente. Ahora bien, ¿supone esta idea de verosimilitud que no podamos tener un relato verdadero? Más bien, lo que supone es que cualquier relato que elaboremos será una aproximación parcial a la verdad, susceptible de mejora²⁰. Nuestros relatos son intentos de atrapar una verdad que se nos escapa, pero de la que podemos tener cierta perspectiva:

La idea de verosimilitud es más importante en los casos de los cuales sabemos que debemos trabajar con teorías que son, en el mejor de los casos, aproximaciones, es decir, teorías de las cuales sabemos que no pueden ser verdades (esto ocurre a menudo en las ciencias sociales). Aun en estos casos, podemos seguir hablando de mejores o peores aproximaciones a la verdad²¹.

Precisamente por esto, es una tarea importante en la educación del ser humano el ser capaz de distinguir los tipos de relatos acerca de la naturaleza. El discurso científico, cuya verosimilitud viene dada por la experiencia y la argumentación lógica; el discurso filosófico, que indaga en los fundamentos últimos acerca del mundo y cuya verosimilitud se basa en la argumentación y la intuición intelectual. Evidentemente, no es posible establecer un mismo criterio de verosimilitud para cada tipo de relato. Para la descripción de los fenómenos, que es el objeto del relato científico, la verosimilitud viene dada por la adecuación de las teorías a los hechos, por su falsabilidad. Para la comprensión de la vida, lógicamente son importantes los datos de la ciencia, pero el relato será necesariamente filosófico, basado en la argumentación y en instancias cognoscitivas distintas al análisis científico de datos. Posiblemente de este modo, tal como sostiene Popper, nunca encontraremos un relato del todo cerrado, sino una aproximación parcial en constante mejora pero que, por lo menos, nos llevará a rechazar como simplista toda forma de reduccionismo. Solo de este modo puede el ser humano llegar a elaborar un relato verosímil sobre sí mismo. Solo mediante la actitud socrática que reconoce la propia ignorancia, podemos buscar una verdad a la que accedemos mediante relatos verosímiles, examinados con juicio crítico.

20 ARANA, J., *El caos del conocimiento: del árbol de las ciencias a la maraña del saber*. Pamplona: Eunsa, 2004, p. 84.

21 POPPER, K., *Conjeturas y refutaciones*. Barcelona: Paidós, 1983, p. 287.

REFERENCIAS BIBLIOGRÁFICAS

- ARANA, J., *El caos del conocimiento: del árbol de las ciencias a la maraña del saber*. Pamplona: Eunsa, 2004.
- ARTIGAS, M., *La inteligibilidad de la naturaleza*. Pamplona: Eunsa, 1992.
- AYALA, F. J., *Darwin y el diseño inteligente*. Madrid: Alianza, 2007.
- FOUCAULT, M., *Seguridad, territorio, población*. Madrid: Akal, 2008.
- FOUCAULT, M., *Dits et écrits*, Vol. III. Paris: Gallimard, 1994.
- FOUCAULT, M., *La arqueología del saber*. México: Siglo XXI, 1970.
- DAWKINS, R., *El relojero ciego*. Barcelona: Labor, 1988.
- DENETT, D. C., *La peligrosa idea de Darwin*. Barcelona: Galaxia Gutenberg, 1999.
- GODFREY-SMITH, P., "Information and the Argument from Design". En: PENNOCK, R. T. (Ed.), *Intelligent Design and Its Critics*. Cambridge: MIT Press, 2001, pp. 575-596.
- JABLONKA, E.,-LAMB, M. J., *Evolution in Four Dimensions*. Cambridge: MIT Press, 2005.
- LAUBLICHER, M., y MAIENSCHIN, J. (eds.), *Form and Function in Developmental Evolution*. Cambridge: Cambridge University Press, 2009.
- MONOD, J., *El azar y la necesidad*. Barcelona: Orbis, 1985.
- PAPP, D., *Filosofía de las leyes naturales*. Madrid: Espasa-Calpe, 1945.
- POPPER, K., *Conjeturas y refutaciones*. Barcelona: Paidós, 1983.
- REICHENBACH, H., *Moderna filosofía de la ciencia*. Madrid: Tecnos, 1965.
- RUSE, M., *Darwin and Design*. Cambridge: Harvard University Press, 2003.
- SCHLICK, M., *Filosofía de la naturaleza*. Madrid: Encuentro, 2002.
- WHITFIELD, J., "Postmodern Evolution?". *Nature* 455 (2018), pp. 281-284.

**CONCEPTUAL STUDIES OF THE HISTORICAL AND
MODERN-DAY INFLUENCE OF CLIMATE ON FORMS
OF GOVERNMENT, POLITICAL AND SOCIO-ECONOMIC
PROCESSES: A VIEW FROM RUSSIA**

*ESTUDIOS CONCEPTUALES SOBRE LA INFLUENCIA HISTÓRICA Y
ACTUAL DEL CLIMA EN LAS FORMAS DE GOBIERNO, PROCESOS
POLÍTICOS Y SOCIOECONÓMICOS: UNA VISIÓN DESDE RUSIA*

ALEKSEY ANISIMOV

Doctor of Juridical Sciences

Volgograd Institute of Management - branch of the Russian Academy of

National Economy and Public Administration under the President of the Russian Federation

Volgograd/Russia

e-mail: anisimovap@mail.ru

ORCID: 0000-0003-3988-2066

Recibido: 28/04/2020

Recibido: 28/04/2020

Aceptado: 6/09/2021

MARINA LEONIDOVNA DAVIDOVA

Doctor of Juridical Sciences

The Head of the Department of Constitutional and Municipal Law

Volgograd State University

Volgograd/Russia

davidovaml@volsu.ru

ORCID: 0000-0001-8392-9592

ERVIN ALIBEKOVICH AHVERDIEV

Master of Laws

Department of Constitutional and Municipal Law

Volgograd State University

Volgograd/Russia

kmp@volsu.ru

ORCID: 0000-0003-4264-2821

Abstract: This paper argues that climate has a significant impact, not only on the development of an economy, the state of the natural objects, morality and culture, but also on systems of government. On the basis of the concept of climate possibilism, the authors of this paper demonstrate patterns relating to the establishment of monocratic government in countries with hot climates and polycratic governments in countries with harsher, colder climates. Therefore, current climate change, i.e., the rise in the average annual temperature, will inevitably lead to the gradual expansion of territories with hot climates. As a result, the polycracy that is common in the modern world will eventually transform into monarchy, which, in turn, will lead to changes in the legal, moral, religious, cultural and other spheres of life. For this reason it is necessary to increase international cooperation in terms of implementing the provisions of the Paris Agreement (2015), and to adopt and implement a set of appropriate measures at the national level that will enable us to slow down the process of losing the cultural identity of countries with a polycratic (republican) form of government.

Keywords: Climate; determinism; possibilism; philosophy; law.

Resumen: Este artículo sostiene que el clima tiene un impacto significativo, no solo en el desarrollo de una economía, en el estado de los objetos naturales, la moral y la cultura, sino también en los sistemas de gobierno. Sobre la base del concepto de posibilismo climático, los autores de este artículo demuestran que existen patrones relacionados con el establecimiento de un gobierno monocrático en países con climas cálidos y gobiernos policráticos en países con climas más duros y fríos. Por tanto, el cambio climático actual, es decir, el aumento de la temperatura media anual, conducirá inevitablemente a la expansión paulatina de territorios con climas cálidos. Como resultado, la policracia que es común en el mundo moderno eventualmente se transformará en monocracia, lo que, a su vez, conducirá a cambios en las esferas legal, moral, religiosa, cultural de la vida. Por ello es necesario incrementar la cooperación internacional en términos de implementación de las disposiciones del Acuerdo de París (2015), y adoptar e implementar un conjunto de medidas adecuadas a nivel nacional que nos permitan frenar el proceso de pérdida de la identidad cultural de países con una forma de gobierno policrático (republicano).

Palabras clave: Clima; determinismo; posibilismo; filosofía; derecho

INTRODUCTION

The global climate change problem has multiple aspects to it and is studied by academics of different social, natural and other sciences. With regard to the law, this problem is usually viewed in terms of the development of international cooperation (implementation of the Paris Agreement provisions (2015) at the national level), a reduction in harmful CO₂ emissions, compensation for the harm caused by the adverse impact of climate change on nature, and people's lives and health, implementation of strategies for adaptation to, and mitigation of, climate change, defining climate-related human rights, implementation and protection of refugees' environment-related rights, etc. A couple of years ago, criminologist revealed that there is a direct connection between the crime rate and the nature of crime, on the one hand, and the change in the air temperature, on the other. Furthermore,

there have been a number of statistical studies that proved the fact that a rise in temperature is accompanied by a rise in the level of hostility, as well as the fact that the most brutal crimes are committed during heat waves¹.

The discussions that are actively held by the representatives of different national scientific law schools draw our attention to extremely important aspects of global climate change, however, they all automatically presume that the problem of global climate change and its unarguably adverse effect (depletion of the ozone layer, the increase in harmful ultraviolet radiation, sea levels rising, the flooding of island states, etc.) only emerged in the 20th century. However, not many of us realise that climate impacts not only the current environmental situation but also the long-term consequences for countries that are facing the risk of an increase in annual average temperature. Recent legal, historical and philosophical research shows that, in some areas, climate determines not only the crop yields or the state of bodies of water, but also the system of government, the mentality of the people, the level of legal culture and the content of legal regulations. The latter means that, historically, climate affected the predominance of a particular form of a country's government. At the same time, it is necessary to understand that, in human history, climate changed for various reasons, besides human activity. On the basis of all that is stated above we present the hypothesis of this research, namely the idea that the global climate changes that we witness today will eventually lead to changes in the political and legal systems of different countries where the climate will get hotter, and influence those countries' economic relations, legal systems (including forms government), morality, etc. That means that the development of appropriate measures to prevent global climate change still conducted at the international level after the Paris Agreement had been adopted, has a greater value than people tend to think, which is why all the efforts of modern states in this respect should be increased in order to avoid any changes in the political and legal order of any country in which democracy seems strong and under no threat at all. History shows us that changes can be slow to materialise and even unnoticeable at first, but inevitable nonetheless.

Thus, in the first part of this article we are going to examine the impact that climate had on forms of government in human history based on modern historical and philosophical studies (climate possibilism); in the second part we are going to examine how the concept of climate possibilism works in particular monocratic and polycratic states; in the third part of the article we will examine the impact of climate on the socio-economic and political-legal determinants of forms of government, as well as the mentality of the population.

1 BARON, R.A. & RICHARDSON D. R., *Human aggression* (Saint-Petersburg: St. Pete, 2001).

1. THE CLIMATE'S HISTORICAL IMPACT ON FORMS OF GOVERNMENT

From ancient times, people have studied the impact that climate has had on social processes. In the 5th century BC, Herodotus wrote that “Egyptians, aside from the Libyans, are the healthiest people in the world because of their unchanging climate”.² Hippocrates, in his turn, discovered the patterns of climate influencing the development of diseases known at that time³. In his work “Politics”, Aristotle wrote that the tribes in the European part of the county that lived in a colder climate were physically strong but not very intelligent and thus not very advanced in the trades. On the other hand, he believed that people in hot countries, for example, in Asian states, are very intelligent, but do not possess enough physical strength and live as slaves⁴. Later, in the 3rd and 4th centuries BC, Polybius devoted a whole chapter in his work titled “The Histories” to the way geographical location and climate influence people’s customs and traditions. He also noted that their harsh manner could be explained by their cold climate⁵. Later still, the influence of climate on social development was studied by Plato, Strabo and Arabian philosopher Ibn Khaldun.

Thanks to the works of Niccolò Machiavelli, Montesquieu, Jean Bodin and others, the Modern Age is marked by the new idea of environmental (climate) determinism. Environmental determinism is a concept in terms of which the social and economic development of nations, countries and societies in general is explained by the geographic factor, namely by geographic location, climate, landscape, water, soil, plant and mineral resources. Montesquieu is usually considered to be the father of this concept. At the same time, the foundations of this concept can be observed in the works of Jean Bodin, who revised human history, while taking into consideration the climate factor. Bodin divided history into three separate periods, the period of the soul, the period of the mind and the period of brute force. According to Bodin, the main role in the first period is played by the people of the south-east who worked on creating religion, developing sciences and learning wisdom. In the second period, the people of the Mediterranean establish states

2 HERODOTUS, *History*. 9 Volumes / translated by G.A. Stratanovsky (Moscow: «AST Publishers», «Ladimir», 2001).

3 JOUANNA J., *Hippocrates* / translated by M. B. DeBevoise (John Hopkins University Press, 2001).

4 ARISTOTLE, *Works*. In 4 Volumes. Vol. 4 / translated from ancient Greek; general ed. A.I. Dovaturova (Moscow: Myisl, 1983).

5 BUZESKUL, V. P., *Introduction to the history of Greece. Resources review and an article on the development of Greek history in the 19th century and in the early 20th century* / Introductory article and general ed. prof. E.D. Frolova (Saint-Petersburg: Publishing house «Kolo», 2005).

and something resembling legislative systems. During the third period, the people from the north discover new crafts, inventions as well as arts. He states that each period of history lasts for around 2000 years.

According to Jean Bodin, the key things that are influenced by climate are human physiology, systems of legislation and authority, and mentality. In his works, the French philosopher divides the planet into three climate zones: the hot equatorial zone, the cold polar zone, and middle mild zone. In terms of mental capabilities, Jean Bodin believed that people from the south are more intelligent than the people from the north, while the people from the north are physically stronger and sturdier than those who live in the south. This idea was justified by the fact that the defeat of the people living in the southern part of the continent did not last long and that the conquerors from the north were not able to sustain their status as the victors.

Therefore, the physical strength of the northern people combined with their limited mental capabilities shaped the mentality of this primitive and tough people. At the same time, it pushed them to systemise their daily life and made them more prudent. J. Bodin considered the southern people to be even more aggressive and hostile than the northern due to the brutality of their methods of fighting and violence towards their enemies. This kind of mentality demonstrated their autocratic tendencies. At the same time, the southern people had their own virtue that was based on the fact that they had an abundance of blessings and subconsciously tried to neutralise their vices⁶.

The French scientist also saw a correlation between climate and the state of the legislative system and governmental authority of a country⁷. For example, he noted that the rules of conduct adopted in the northern part of the continent were based on a sense of community and mutual respect, while the rules of the southern people (who had a more individualised mentality) were based on strict subordination⁸.

The concept of the environmental determinism was finalised in the work of Montesquieu titled "The Spirit of the Laws"⁹. The French philosopher claimed that climate predetermines the physiological diversity of people, their mentality, and their attitude towards slavery and affects the state and its legislative system. In the second chapter of the "Spirit of the Laws", Montesquieu explains how the

6 BODIN, J., *Method for the easy knowledge of history* (Moscow: Nauka Publisher, 2000).

7 GRININ, L., *Lecture: Nature factor in terms of the theory of history*, 2 Philosophy and society, 195-198 (2011).

8 BODIN, J., *Method for the easy knowledge of history* (Moscow: Nauka Publisher, 2000).

9 MONTESQUIEU, Ch., *The Spirit of the Laws* / Ed. A. Mateshuk; Transl. A. Mateshuk; Comment. A. Mateshuk (Moscow: Myisl, 1999).

climate affects people. He believed that low temperatures strengthen the muscles, while high temperatures relax them. The philosopher also underlined the fact that “the inhabitants of warm countries are, like old men, timorous; the people in cold countries are, like young men, brave”. This idea becomes the basis for his conclusion that the inhabitants of warm countries are more inclined to tyranny and slavery. According to Montesquieu, a tyrannical form of government can be explained not only by psycho-physiological peculiarities, but also by social specificity. As an example, he showed that the dependent status of women, which serves to maintain strict and centralised government at the lower level, also influenced the governmental system in general. In contrast, families in the northern countries develop principles of equality and mutual respect that were reflected in their legislative systems.

Further development of the ideas of environmental determinism can be found in the works of British scholars such as Henry Buckle, Francis Bacon, Halford Mackinder and Arnold J. Toynbee, French scientists such as Élisée Reclus, Antoine Barnave, Bernard Le Bovier de Fontenelle and Jean-Baptiste Dubos, German scholars such as Immanuel Kant, Johann Gottfried Herder, Alexander von Humboldt, Friedrich Ratzel and Carl Ritter, American researchers such as A. Guyau, Ellsworth Huntington, Ellen Churchill Semple and George Perkins Marsh, as well as Russian scientists such as Lev Gumilyov, Lev Mechnikov, Georgi Plekhanov and others.

In the British law school we see a new understanding of some aspects of environmental determinism. The most notable representative of this school, Henry Thomas Buckle, wrote that in spite of the fact that the influence of soil and climate is considered indirect, this influence results in some significant consequences¹⁰. Early in social development, an environment helps with the accumulation of goods and leads to the development of people, who later begin conquering nature as a result of their activity. The British scholar wrote that the people who are at the early stage of their development accumulate goods with the help of their environment (for instance, climate, soil, resources), while societies that are already at the next stage of their development accumulate goods with the help of their good environmental practices that are closely connected with the accumulation of knowledge. These processes have different limits, either based on the amount of natural resources, or based on the intellectual and business development of the society. The approach that he described marks the beginning of a new direction

10 BUCKLE, H., *History of Civilization in England* / Translated by K. Bestuzheva-Rumina and N. Tiblena. Vol. 1-2 (St-Petersburg, Moscow: M.O. Volf, 1873).

in terms of environmental determinism, namely climate (geographical) possibilism¹¹. The main distinction of this new school (in comparison with that of the French philosophers) is that the environment is seen as one of the determinants of the social process, starting with creating conditions for human development and finishing with rational activity aimed at the accumulation of wealth¹². Climate determinism was also studied by the representatives of the German school of thought. Carl Ritter wrote that the development of social mentality cannot be predetermined because it is just a result of adaptation to the environmental conditions by developing the traits that are essential for human life and activity. "Each person is a representative of the nature that brought him into this world as well as brought him up. People are the reflection of their homeland. The influence of the landscape on people's personality, build, shape of skull, temper, language and spiritual development is undeniable...The whole existence of humans is connected with their land by thousands of strong roots that cannot be plucked out"¹³. Friedrich Ratzel, one of the founders of the concept of Geopolitics, also considered climate to be the main incentive. After analysing the climate's impact on a person, Friedrich Ratzel came to the conclusion that the moist air of the tropics has a soporific effect that in general leads to social phenomena such as slavery. In contrast, dry air stimulates the nervous system that, in turn, stimulates the rule of the people and an active civil society. Such differences can be traced even in terms of a single state¹⁴.

In Russian legal literature, the concept of environmental (climate) determinism was studied by Lev Mechnikov and Vladimir Klimenko.

Lev Mechnikov believed that, in bad weather conditions, one should cooperate with other members of the society. This makes the structure of social relations more complicated due to the division of labour and further social differentiation. Under favourable conditions, one can get any goods necessary for life and activity without any additional help. Another key geographical factor for the Russian scientist is the aquatic environment, with the help of which we can divide human development into periods in terms of a river, sea and ocean spectrum¹⁵. The river

11 GRININ, L., *Lecture: Nature factor in terms of the theory of history*, 2 Philosophy and society, 195-198 (2011).

12 TRIFONOVA, T. A., LUBISHEVA, A. V. & REPKIN R. V., *Geography. Ecological and geographical and social and economic space: textbook* (Vladimir: Vladimir State University Press, 2013).

13 SAMAROVA, L. R., *Geographical determinism in modern religious studies*, 79 Culture of the people from the Black Sea region, 134-138 (2006).

14 RATZEL, F., *Anthropogeographie*. Vol. 1 (Stuttg., 1921); Ratzel F. *Politische Geographic* (Munch., 1923).

15 GRININ, L., *Lecture: Nature factor in terms of the theory of history*, 2 Philosophy and society, 195-198 (2011).

period is characterised by tyranny, the sea (Mediterranean) period is characterised by oligarchy, while the ocean period is characterised by the building of liberal democracy. At the same time, Lev Mechnikov paid special attention to climate conditions. He noted that, in a hot climate, in spite of its abundant wildlife, “there is still no strong civilisation that could fill a page in the chronicles of human history. The reason for that is the excessive development of organic life in all of its forms.

This abundance of life stymies people’s energy and intellectual development; the inhabitants of the warm climate do not have the key incentive to work, to explore the world around them and to cooperate since everything is given in abundance by nature”¹⁶. That is why “cooperative solidarity” predetermines more personal liberties in terms of rights, as well as in terms of actions that minimise the risk of sustaining tyranny and leads society to democratic and republican forms of government.

Vladimir Klimenko has also devoted many of his works to climate conditions in the course of human history. The author writes: “...It doesn’t seem to be a coincidence that all the periods of massive cultural shift were also periods of global cooling to a greater or lesser degree”¹⁷. Consociation of tribes and unification of people, establishment of new states and the rise in intelligence, as well as cultural and technological breakthroughs should all be connected with the historical periods of climate cooling. In their turn, the periods of climate warming lead to intellectual and spiritual regression. For example, the great period of Chinese philosophy and culture occurred after the year 330 BC, a period of climate cooling. According to Vladimir Klimenko, local climate changes (for instance, in one country or one part of the country) do not always match global climate changes; and notions such as climate “deterioration” (e.g., air cooling, a decrease in the amount of precipitation) or climate “amelioration” (e.g., warming, an increase in the amount of precipitation) are quite subjective¹⁸.

Thus, as time passes the initial understanding of geographical determinism changes into a more low-key rhetoric and eventually turns into the concept of geographical (climate) possibilism. Furthermore, studies such as Geopolitics, for instance, started emerging on the basis of this concept. If we take into consideration climate and geographical location when studying the processes of legislation and state establishment, we realise the necessity of taking a fresh look at the

16 MECHNIKOV, L. I., *Civilization and great historical rivers* (Moscow, 1995).

17 KLIMENKO, V. V., *Climate and the history of early High cultures (3500-500 BC.)*, 4 Vostok Publishers, 5-24 (1998).

18 KARPYUK, S. G., *Climate and Geography in terms of the human dimension (archaic and classical Greece)* (Moscow: Institute of World History (Russian Academy of Sciences), 2010).

established legal categories. A complete disregard for climate conditions as objective determinants can be considered a mistake. We should presume that while climate has a direct influence on people's lifestyle, habits, mentality, and the levels of economic development, as well as the way countries collaborate with each other, it also has an indirect influence on law and governance.

2. CLIMATE POSSIBILISM IN MONOCRATIC AND POLYCRATIC STATES

In our opinion, the connection between the climate and human development cannot be denied. The influence of climate on different spheres of social life, including the development of a state, is well demonstrated in the works of European, American and Russian academics. In terms of historical analysis of the climate's influence on the establishment of types of governance, it is possible to determine (besides a direct influence during the early periods of the states' development) its indirect influence through different stages of establishing social behaviour.

At the early stages, a nurturing or adverse environment shapes a particular kind of mentality in people living in a specific area by means of natural selection. It directly affected the processes of socialisation that later took place with the help of religion. We should underline that the religion itself is a result of the behavioural patterns that were also shaped by environmental conditions. If we look at the works of Montesquieu, Friedrich Ratzel and Lev Mechnikov, as they give examples of relations in families, gender roles in society, people's attitudes to strong beverages and other habits, it is difficult to deny the transformation of some moral standards into religious dogma. If we continue tracing the connection between climate and social development, we should underline that established methods of socialisation, as well as people's dispositions, become an integral part of their culture (mental matrix). In essence, culture also determines the development of the economy and the legal sphere. This last stage is the one that eventually pre-determines the method of governance. Of course, there can be some exceptions in terms of this chain but even in cases of exceptions, the environmental conditions have an impact on humans. In this case, we talk about the development of strong personalities and particular ideological and worldview theories that can suppress the established forms and principles of cooperation between society and the state with the help of these tools.

Before presenting the discovered patterns of climate influence on the type of state, it is necessary to mention the cognitive level at which it was studied. Modern science recognises at least two of these levels, geographical (climate) determinism and possibilism¹⁹.

The difference between these two levels is a casual connection between environmental conditions and the studied phenomenon. The second claims that there is a possible connection between two objects, but only in terms of adaptation, which eliminates direct dependence. Taking into consideration various historical examples and multiple aspects of the development of social life, it only makes sense to talk about climate possibilism²⁰.

The ideas that are stated above allow us to detect the patterns of the influence of a particular climate, in the broad sense, on the form of government. After the further development of these ideas, we can conclude that monarchy is usually established in countries with beneficial climate conditions, while polycracy is usually established in countries with less beneficial conditions. Considering climate as an incentive phenomenon that determines certain behaviours and human thinking characterises it as a particular area of motivation. One can clearly see this in the historical processes that were studied in the works of Vladimir Klimenko. He considers an adverse climate (low temperatures, an abnormal amount of precipitation or a lack thereof, a complex landscape) to be a catalyst for human development²¹.

According to his table of the sequence of climate events and the history of the first high cultures in the 6th–1st millennia BC; a comparison table of the periods of Rome (Byzantium) and Iran starting with the 1st century BC and until the middle of the 7th century AD²²; and a comparison table of the sequence of climate and historical events during the rise of world religions (the 5th century BC to the 7th century AD) we can clearly see a certain pattern. At the same time, in this article we are focusing only on the changes that concern statehood and law, i.e., the establishment, development and dissipation of states. The idea of presenting an adverse climate as an incentive for the development of public affairs also encom-

19 ZAMYATIN, D. N., *Geographical factor in the genesis and dynamic of civilization*. In: *Civilisations in a globalising world. Preliminary results of the interdisciplinary project: According to the materials presented at the scientific conference* / Ed. V.G. Horos. Series «The Library of the Institute of World Economy and International relations». (Moscow: Institute of World Economy and International Relations of the Russian Academy of Sciences, 2009).

20 FEKADU, K., *The paradox in environmental determinism and possibilism: a literature review*, 7 *Journal of Geography and Regional Planning*, 132-139 (2014).

21 KLIMENKO, V. V., *Climate and the history of early High cultures (3500-500 BC.)*, 4 Vostok Publishers, 5-24 (1998).

22 KLIMENKO, V. V., *Climate and the history from Confucius to Muhammad*, 1 Vostok Publishers, 5-31 (2000).

passes the processes of state establishment. Based on the data given by Vladimir Klimenko, we can see that the air temperature during the rise of Ancient Egypt was getting cooler²³. However, it is necessary to note that the location of the first states is characterised by an acceptably warm climate and close proximity to water bodies (in our example, it is the river Nile). This is also proved by the fact that, at that period of time, northern peoples did not even have a state. There is another fact worth mentioning: nowadays the climate in most parts of Italy is mild, but during the early periods of the Ancient Rome (753 BC) it was quite severe. According to the ancient books that reached the present age, the period from 850 BC to 480 BC was marked by a dramatic cooling.

For example, Lucius Junius Moderatus Columella, a Roman author, wrote, based on the well-known book written by the two Sasernae (father and son), that in the early 1st century BC grape and olive growing was difficult in the central part of Italy due to the long, severe winters that sometimes could even prevent “grape or olive shoots from growing”²⁴.

According to the Roman writer Lucius Livius Andronicus, the winters in the middle of the 1st century were so severe that the Tiber River would freeze²⁵. The described situation is related to the thesis that polycracy (in this example it is a republic state) develops only in adverse climates. For instance, similar climate conditions predetermined the establishment of democracy in Athens (as a clear sign of polycracy). According to the works of Herodotus and other historians that have reached our age, the development of democratic institutions in Ancient Greece took place in adverse climate conditions that manifested in the freezing of the Kerch Strait (Cimmerian Bosphorus) and the Bosphorus, as well as in the inability to grow grapes and olives²⁶. According to our thesis, the opposite process takes place when the climate starts getting warmer. In the works of different European scholars, we find that it is very common to witness the development of an individual, centralised system of governance, usually a monarchy, in these warmer periods.

For instance, if we take the Eastern climate as a warm climate then Hegel's statement seems quite representative: The East knew and still knows that only one is free (a monarch, a tsar, a shah or an emperor); “the German world knows that

23 KLIMENKO, V. V., *Climate and the history of early High cultures (3500-500 BC.)*, 4 Vostok Publishers, 5-24 (1998).

24 LAMB, H. H., *Climate, History and the Modern World* (L. and N.Y., 1995).

25 LAMB, H. H., *Climate, History and the Modern World* (L. and N.Y., 1995); *Climate of the Ancient Rome*. In: *Historicus. Social and political journal* / <http://historicus.ru/1650> (date of access: 11.12.2020).

26 VINOGRADOV, Y. A. & MARCHENKO, K. K., *Greeks and Scythians in the northwestern part of the Black Sea region in the 5th century BC*, 1 *Journal of Ancient History*, 80-85 (1995).

everyone is free”²⁷. However, one can argue that in the Middle Ages there were a number of countries in Europe that had a monarch as their head. Today we also have monarchic form of government in countries such as the UK and Japan.

An answer to this can be the following abstract from the book of an American jurist H.J. Berman “Law and Revolution”: “Royal jurists such as Bracton said that the kings have a duty to obey the law, that the king is “under God and the law,” that it is not the king that makes law but the law that makes the king... Eike of Repgau, Bracton, Beaumanoir and others claimed that the king himself was bound by law and that the king’s subjects might even have, in some circumstances, the right to disobey his command if it was unlawful. The belief was rooted, first, in the theological conviction that the universe was subject to law and, second, in the duality of secular and spiritual authorities, which placed both practical and theoretical limitations upon the power of each...third, in the pluralism of secular authorities within each kingdom and, fourth, the belief in the supremacy of law was closely linked to the mutuality of obligations between superiors and inferiors...”²⁸.

Thus, even though we see that the external conditions presume a monocratic form of government similar to the one in hot eastern countries, the internal processes are marked by progressive and applicable democratic principles, which are woven into all the elements of the supreme authority. At the same time, it is necessary to remember the provisions that are stated above and again underline that the form of government should be defined not only by formally established norms, but also by taking into consideration the actual conditions in which the government functions. We should also mention that the impact that climate has on the form of government can be traced not only in ancient times, but in the Modern Age as well. The French Revolution can be a good example of this. In 1787–1788 the country was in crisis due to the bad harvest, and the severe winter of 1788–1789 completed this process. More and more often, “hordes of ragged vagabonds go out on the streets and one can hear their angry cries of hunger”²⁹. Edicts are issued against speculative activity which drives people to despair. All of these lead to the irreversible consequences – Parisians begin the Revolution.

Thus, “the hunger and the extremely high prices of bread provoked such a strong hatred towards the king and his court that it could not have been provoked by the combined works of François-Marie Arouet Voltaire, Jean-Jacques Rousseau and Denis Diderot”³⁰.

27 HEGEL, G., *Works on politics* (Moscow, 1978).

28 BERMA, H. J., *Law and Revolution* (L., 1983).

29 CARLYLE, T., *The History of the French Revolution* / Trans. from English Yu.V. Dubrovina and E.A. Melnikova (Part I) (Moscow; Thought, 1991).

30 GRIGORIEV, R., *Should we fight fire with fire?* 6 Knowledge is power, 39-42 (2007).

Similarly, the Tsardom of Muscovy in the summer of 1601 was marked by a period of long cold showers, and July, in particular, was marked by the first snow. In late August, the Dnieper was completely frozen and people were sledging on its icy surface as if it was wintertime. There was no harvest that year. In May of 1602 the air temperature was again below zero, which led to the loss of harvest. The following famine left almost one-third of the population dead, which led to massive riots (the Time of Troubles) that stopped in 1613 with the first and only election of a Tsar. He was elected by a national assembly (Zemsky Sobor) that included a huge portion of people and was not the only candidate.

Thus, the idea of the climate possibilism that climate has an impact on the form of government is proved by both scientific studies and examples from world history. These examples give us an opportunity to define the patterns of the climate's influence on social development, namely the fact that a mildly warm climate allows the government to be centralised and separate from the people, which eventually leads to monarchy.

In its turn, an adverse (cold) climate motivates people to be self-organised, to take care of their businesses and determine their future based on personal freedom and cooperation, which leads to polycracy and the development of democratic institutions. We should also underline the connection between social and economic, as well as political and legal, determinants and climate conditions. As a result, we can get a glimpse of the actual effect that environmental conditions have on the type of government.

3. INFLUENCE OF CLIMATE ON SOCIO-ECONOMIC AND POLITICAL-LEGAL DETERMINANTS THAT DETERMINE THE FORM OF GOVERNMENT, AS WELL AS THE MENTALITY OF THE POPULATION

3.1. INFLUENCE OF CLIMATE ON ECONOMIC PROCESSES AND THE MENTALITY OF THE POPULATION

It is obvious that climate is important for the development of an economy. A well-known Arabic philosopher and historian, Ibn Khaldun, wrote that "...if the soil is fertile then there is only one ruler in the state, if the soil is barren than it leads to the rule of many"³¹. It is true that a cold climate was not initially an

31 IBN KHALDUN, *Prolegomena (Muqaddimah)* / Ed. Translated from Arabic and comment. A.V. Smirnova. In: *Historical and philosophical annual journal 2007* (Moscow: Nauka Publishers, 2008).

advantage for people since it resulted in a harsh environment and scarcity. However, as states started developing and the economy started becoming more and more global; the northern countries started developing exponentially. The necessity to adapt and tame nature with the help of various technologies predetermined the economic prosperity of these countries. As Douglass North writes, peoples have different attitudes to goods and resources. Thus, for example, the principle of complete accessibility to economic resources is implemented differently in cold (Western) countries and hot (Eastern) countries. In the first case, complete accessibility is there, but not free³².

While in the second case it is free, but it is not there for the general public. This phenomenon is closely studied by another American economist, Lawrence E. Harrison, who believes that society can take wealth either for granted or as a result of hard work and effort.

It is quite obvious that the second approach to wealth helps people respect private property, obey the principles of business practice and develop an innovative way of thinking. A good example of this can be found the works of E. Harrison, where he describes the economies in Latin America. "A mild climate is perfect for developing good habits (lack of excessiveness, thriftiness and forethought) that enhance economic development, but most of the countries in Latin America belong to the tropical climate"³³. The same ideas can be found in the works of French representatives of the concept of the geographical (climate) determinism.

People in good climates tend to be lazy and do not seem to be very creative or have an innovative way of thinking since there is always enough of everything to meet their needs and they do not need to work hard to have it. In fact, if we look at the world today, we will notice that economy in mild and cold climates is very developed. This idea is proved in the works of American economist Jeffrey Sachs³⁴. Climate can directly influence people's habits and lifestyle, which, in turn, will eventually shape their mentality. As we have already mentioned, there were many scientists who were working on distinguishing and defining behavioural patterns related to climate.

32 NORTH, D., WALLIS, J. J. and WEINGAST B., *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History* (Cambridge and New York: Cambridge University Press, 2009).

33 HARRISON, L. E., *Who prospers? How cultural values shape economic and political success* (New York: Basic Books, Cop. 1992).

34 SACHS, J., *Notes on a New Sociology of Economic Development*, In: Culture Matters. How values shape human progress / Ed. by Lawrence E. Harrison and Samuel Huntington (Moscow: Moscow school of political studies, 2002); Rangel C. *The Latin Americans - Their Love-Hate Relationship with the United States* (New York: Harcourt Brace Jovanovich, 1977).

These were social and culture scientists, legal experts, economists and political scientists. For example, Lawrence E. Harrison discovered that a reasonably mild climate together with a complex landscape and close proximity to the ocean had made Korean and Japanese people hardworking, prudent and hungry for knowledge³⁵. In fact, if we project these conclusions on the people who live in adverse climate conditions, we can discover that their efforts are primarily aimed at preserving and developing existent energy (resources, etc.). This kind of attitude to nature is only possible if people possess qualities such as prudence, accuracy and meticulousness. Another popular concept that is worth our attention is the one that explains the differentiation of Western and Eastern cultures on the basis of the development of brain hemispheres. This approach was supported by such scholars as Vadim Rotenberg and Viktor Arshavskiy. These authors distinguish Eastern and Western ways of thinking based on hemispheric asymmetry. Each hemisphere is characterised by its specific method of processing information³⁶. However, climate conditions also play their role in this system. Adverse climate conditions boost the development of the left hemisphere, while beneficial climate conditions boost the development of the right hemisphere. This is proved by the biological characteristics of each hemisphere³⁷.

3.2. THE IMPACT OF CLIMATE ON THE POLITICAL AND LEGAL DEVELOPMENT OF SOCIETY

Climate conditions have an impact on the development of the following aspects of the state and the law.

It is necessary to note that the level of legal awareness among people depends on their level of intelligence and worldview. The need to survive in harsh climate conditions facilitates the development of logical thinking in people of certain countries. New technologies are invented as an answer to the demand of rational consumption of resources such as raw materials, and solar, air and water energy. The same demand is an incentive for the further development of these technologies. In this case, the most representative examples are those of Germany, the USA and the UK, as confirmed by statistics. According to the data presented in the Human Development Report of the UN issued in 2004, the average adult

35 HARRISON, L. E., *Who prospers? How cultural values shape economic and political success* (New York: Basic Books, Cop. 1992).

36 KESSIDI, F., *Globalisation, functional asymmetry of the brain and the issue of cultural integration*, 1 *Philosophy and society*, 45-52 (2004).

37 ROTENBERG, V. S. and ARSHAVSKY V. V., *Right and Left Hemispheres Activation in the Representatives of Two Different Cultures*, 38-2 *Homeostasis*, 49-57 (1997).

literacy rate in Arabic countries was 77% for men and 57% for women. More than a half of women in Iraq, Egypt and Morocco were illiterate. Only 29% of women and 70% of men in Yemen turned out to be literate. An exceptionally low level of literacy among women is most noticeable in two areas when it comes to legal awareness: first, women obey men, which leads to gender inequality and the inability of women to participate in any governance processes; second, illiterate women channel their orthodox views onto their children, which results in a general decrease in literacy³⁸. But the perception of the world expressed through one's attitude to society is also influenced by climate. Thus, if we take into account the idea of Andrey Pelipenko regarding the struggle for survival and the formation of a "new naturality", we should underline that an adverse climate has a negative impact on the demographic rate. Of course, this cannot have a positive influence on the work of the northern people that is aimed at developing their state or finding new means and resources for survival. However, it does explain why the inhabitants of a cold climate have to think not only of themselves but of others as well. In the legal consciousness and legal systems of such states, this climate-determined phenomenon is reflected in the "principle of humanity" that is in fact a mentality based on love for others and the overcoming of anything unsustainable³⁹.

Thus, the impartial social relations, the necessity to protect human rights, as well as the support of equality and justice in the countries with an adverse climate justify the need for establishing and developing a legal system without any assistance. This is exactly the reason why all the members of the society live by the principle of the rule of law, since it is based upon a consensus. One of the key traits of democracy is the guarantee that this principle would be followed. And the fact that the law is shaped by the will of people lets us talk about a genuine democracy as an integral part of the republican form of government.

Political and legal doctrines are a reaction to existent critical situations, including the crises in the governmental sphere. Violations of human rights and *coups d'état* are phenomena that give rise to an almost instantaneous negative reaction from peoples in states with cold climates. The mentality of people living in warm climates, as well as their lack of practical skills mean they either do not react to

38 HARRISON, L. E., *Jews, Confucians, and Protestants: Cultural Capital and the End of Multiculturalism* / Translated from English by Y. Kuznetsova (Moscow: Myisl Publishers, 2014); United Nations International Children's Emergency Fund, «Afghanistan: Statistics». In: www.unicef.org/infobycountry/afghanistan_statistics.html (date of access: 11.12.2020).

39 PELIPENKO, A. A., *Sensogenetic concept of the civilisation history*. In: *Civilisations in a globalising world. Preliminary results of the interdisciplinary project: According to the materials presented at the scientific conference* / Ed. V.G. Horos. Series «The Library of the Institute of World Economy and International relations» (Moscow: Institute of World Economy and International Relations of the Russian Academy of Sciences, 2009).

events in the state or turn to philosophical, abstract or religious dogma to find inner peace. This difference is clearly shown the revolutionary spirit of countries. One of the reasons for establishing and developing political and legal doctrines is the people's ability to self-critique. Lawrence Harrison noted that "individual and group self-criticism, self-help and the desire to fight against corruption are considered to be necessary traits and valuable advantages"⁴⁰.

These traits become stronger as we move north. This phenomenon was closely studied in the works of C. Rangel and Jean-François Revel, who pointed out the inability of Latin Americans to be self-critical, which was the reason for their low level of culture and poor quality of life⁴¹. The development of different philosophies and doctrines was also studied in the works of Vladimir Klimenko. The Russian scientist traces the connection between the temperature cooling and the rise of new ideologies. Thus, we can conclude that the climate's influence on the social and economic, as well as political and legal determinants, that define the form of government (type of state) can be either direct or indirect. Today, the influence of climate conditions is becoming more and more indirect. However, the idea of monocratic states being established in beneficial climate conditions and polycratic states being established in adverse climate conditions was proved in the present analysis.

CONCLUSION

In terms of the present research, we have described the main categories with which one can define the patterns by which climate conditions affect the form of government. According to these categories, the forms of government were distinguished dualistically based on the historical period. For example, today we distinguish between monarchy and polycracy. The correlation between types of government and climate conditions together with the works of different political scientists, economists, philosophers, cultural scientists, and history itself led to the following conclusion: beneficial climate conditions result in monarchy, while adverse climate conditions result in polycracy. In terms of proving this conclusion it would be fair to note that there are some exceptions. However, even in the cases

40 HARRISON, L. E., *Jews, Confucians, and Protestants: Cultural Capital and the End of Multiculturalism* / Translated from English by Y. Kuznetsova (Moscow: Myisl Publishers, 2014).

41 Culture Matters. How values shape human progress / Ed. by Lawrence E. Harrison and Samuel Huntington (Moscow: Moscow school of political studies, 2002); Rangel C. *The Latin Americans - Their Love-Hate Relationship with the United States* (New York: Harcourt Brace Jovanovich, 1977).

of these exceptions the importance and the impact of climate conditions cannot be denied. Most of the elements of the political and legal, as well as social and economic determinants, are more or less influenced by the climate.

The present article shows that climate conditions have a direct effect on the established form of government, as well as an indirect effect through social and economic, as well as political and legal, determinants. All in all, it can determine the most suitable ways of governance based on the prevailing determinants. We should also underline that supporting concepts introduced herein does not eliminate the consideration of other theories. In different periods of time these determinants may have a different level of influence. In view of the above, we prefer the original concepts of climate possibilism that are based on the idea that human development is possibly and partially determined by climate.

This means that, today, the efforts of states aimed at preventing global climate change are important not only in terms of preserving the established forms of economic activity (agriculture) or natural features (forests, water places), but also in terms of preserving the European/American cultural identity, which could be lost as a result of the current changes in climate.

The increase in average temperature due to the global climate change will inevitably lead to the expansion of the territory with a hot climate and certain countries will gradually gravitate from polycracy to monarchy, which, in turn, will result in economic, political, legal, moral and other changes, reducing the number of adherents of Christianity in favour of Islam and in the establishment of new customs and traditions. This is yet another reason for international cooperation in implementing the terms of the 2015 Paris Agreement.

BIBLIOGRAPHIC REFERENCES

- ARISTOTLE, *Works*. In 4 Volumes. Vol. 4 / translated from ancient Greek; general ed. A. I. Dovaturova. Moscow: Myisl, 1983. 830 p.
- BARON, R.A. & RICHARDSON D. R., *Human aggression*. Saint-Petersburg: St. Pete, 2001, 352 p.
- BERMAN, H. J., *Law and Revolution*. L., 1983, 293, 537.
- BODIN, J., *Method for the easy knowledge of history*. Moscow: Nauka Publisher, 2000. 412 p.
- BUCKLE, H., *History of Civilization in England* / Translated by K. Bestuzheva-Rumina and N. Tiblena. Vol. 1-2. St-Petersburg, Moscow: M.O. Volf, 1873. 575 p.
- BUZESKUL, V. P., *Introduction to the history of Greece. Resources review and an article on the development of Greek history in the 19th century and in the early 20th*

- century / Introductory article and general ed. prof. E.D. Frolova. Saint-Petersburg: Publishing house «Kolo», 2005. 672 p.
- CARLYLE, T., *The History of the French Revolution* / Trans. from English Yu.V. Dubrovina and E.A. Melnikova (Part I). Moscow: Thought, 1991. 575 p.
- , «Climate of the Ancient Rome». In: *Historicus. Social and political journal* / <http://historicus.ru/1650> (date of access: 11.12.2020).
- , «Culture Matters. How values shape human progress» / Ed. by Lawrence E. Harrison and Samuel Huntington. Moscow: Moscow school of political studies, 2002. 280 p.
- FEKADU, K., «The paradox in environmental determinism and possibilism: a literature review». *Journal of Geography and Regional Planning*, 7, 2014, 132-139.
- GRIGORIEV, R., «Should we fight fire with fire?». *Knowledge is power*, 6, 2007, 39-42.
- GRININ, L., «Lecture: Nature factor in terms of the theory of history». *Philosophy and society*, 2, 2011, 195-198.
- HARRISON, L. E., *Jews, Confucians, and Protestants: Cultural Capital and the End of Multiculturalism* / Translated from English by Y. Kuznetsova. Moscow: Myisl Publishers, 2014. 285 p.
- HARRISON, L. E., *Who prospers? How cultural values shape economic and political success*. New York: Basic Books, Cop. 1992. 280 p.
- HEGEL, G., *Works on politics*. Moscow, 1978. 439 p.
- HERODOTUS, *History*. 9 Volumes / translated by G. A. Stratanovsky. Moscow: «AST Publishers», «Ladimir», 2001. 752 p.
- Ibn KHALDUN. «Prolegomena (Muqaddimah)» / Ed. Translated from Arabic and comment. A.V. Smirnova. In: *Historical and philosophical annual journal 2007*. Moscow: Nauka Publishers, 2008, 187-217.
- JOUANNA, J., *Hippocrates* / translated by M.B. DeBevoise. Baltimore and London: John Hopkins University Press, 2001, 520 p.
- KARPYUK, S. G., *Climate and Geography in terms of the human dimension (archaic and classical Greece)*. Moscow: Institute of World History (Russian Academy of Sciences), 2010. 224 p.
- KESSIDI, F., «Globalisation, functional asymmetry of the brain and the issue of cultural integration». *Philosophy and society*, 1, 2004, 45-52.
- KLIMENKO, V. V., «Climate and the history of early High cultures (3500-500 BC.)». *Vostok Publishers*, 4, 1998, 5-24.
- KLIMENKO, V. V., «Climate and the history from Confucius to Muhammad». *Vostok Publishers*, 1, 2000, 5-31.
- LAMB, H. H., *Climate, History and the Modern World*. L. and N.Y., 1995. 464 p.
- MECHNIKOV, L. I., *Civilization and great historical rivers*. Moscow: Progress, Pangea, 1995. 461 p.
- MONTESQUIEU, Ch., *The Spirit of the Laws* / Ed. A. Mateshuk; Transl. A. Mateshuk; Comment. A. Mateshuk. Moscow: Myisl, 1999. 672 p.
- NORTH, D.; WALLIS, J. J. and WEINGAST, B., *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge and New York: Cambridge University Press, 2009. 328 p.

- PELIPENKO, A. A., «Sensogenetic concept of the civilisation history». In: *Civilisations in a globalising world. Preliminary results of the interdisciplinary project: According to the materials presented at the scientific conference / Ed. V.G. Horos. Series «The Library of the Institute of World Economy and International relations»*. Moscow: Institute of World Economy and International Relations of the Russian Academy of Sciences, 2009. 162 p.
- RANGEL, C., *The Latin Americans - Their Love-Hate Relationship with the United States*. New York: Harcourt Brace Jovanovich, 1977. 302 p.
- RATZEL, F., *Anthropogeographie*. Vol. 1. Stuttgart: Verlag von J. Engelhorn, 1921. 178 p.
- RATZEL, F., *Politische Geographic*. Munchen-Berlin, 1923. 626 p.
- ROTENBERG, V. S. and ARSHAVSKY, V. V., «Right and Left Hemispheres Activation in the Representatives of Two Different Cultures». *Homeostasis*, 38-2, 1997, 49-57.
- SACHS, J., «Notes on a New Sociology of Economic Development». In: *Culture Matters. How values shape human progress / Ed. by Lawrence E. Harrison and Samuel Huntington*. Moscow: Moscow school of political studies, 2002. 320 p.
- SAMAROVA, L. R., «Geographical determinism in modern religious studies». *Culture of the people from the Black Sea region*, 79, 2006, 134-138.
- TRIFONOVA, T. A.; LUBISHEVA, A. V. and REPKIN, R. V., *Geography. Ecological and geographical and social and economic space: textbook*. Vladimir: Vladimir State University Press, 2013. 165 p.
- UNITED NATIONS INTERNATIONAL CHILDREN'S EMERGENCY FUND, «Afghanistan: Statistics». In: www.unicef.org/infobycountry/afghanistan_statistics.html (date of access: 11.12.2020).
- VINOGRADOV, Y. A. and MARCHENKO, K. K., «Greeks and Scythians in the north-western part of the Black Sea region in the 5th century BC». *Journal of Ancient History*, 1, 1995, 80-85.
- ZAMYATIN, D. N., «Geographical factor in the genesis and dynamic of civilization». In: *Civilisations in a globalising world. Preliminary results of the interdisciplinary project: According to the materials presented at the scientific conference / Ed. V.G. Horos. Series «The Library of the Institute of World Economy and International relations»*. Moscow: Institute of World Economy and International Relations of the Russian Academy of Sciences, 2009. 162 p.

HACIA LA HETERONOMÍA DE LA EXPERIENCIA ESTÉTICA: UNA APROXIMACIÓN ALTERNATIVA A LA ESTÉTICA DE DAVID HUME

*TOWARDS THE HETERONOMY OF THE AESTHETIC EXPERIENCE:
AN ALTERNATIVE APPROACH TO DAVID HUME'S AESTHETICS*

JORGE LÓPEZ LLORET

Doctor en Filosofía

Profesor Titular de Universidad

Escuela Politécnica Superior. Departamento de Estética e Historia de la Filosofía

Universidad de Sevilla

Sevilla/España

lopezlloret@us.es

ORCID: 0000-0002-6738-5895

Recibido: 21/04/2021

Revisado: 1/07/2021

Aceptado: 6/09/2021

Resumen: En este artículo se ensaya una aproximación alternativa a la estética de David Hume, partiendo de la tesis fundamental de que este se esforzó por neutralizar la indecidibilidad del juicio estético que resultaba de la centralidad concedida a la experiencia subjetiva. Aprovechando los precedentes de la fisiología de George Cheyne y el ambientalismo de Jean-Baptiste Du Bos, ambos conocidos por Hume, se plantea que Hume redujo a fisiología al sujeto cartesiano, imposibilitando así la viabilidad de la autonomía estética. Para articular esta tesis, además de a sus textos estéticos, se recurre a las conexiones de Hume con la anatomía, a su crítica a la identidad personal y a sus reflexiones sobre la materialidad del pensamiento. Todo esto confluye en la interpretación de la norma del gusto como algo cuya factibilidad muestra la estadística del proceso histórico en la figura del crítico, que representa el estado de salud orgánica del que cualquiera podría disfrutar, todo lo cual hace que la de Hume sea un referente esencial para la nueva estética de la heteronomía.

Palabras clave: Arte, Estadística, Estética, Fisiología, Heteronomía, Historia, Hume.

Abstract: In this paper is tested an alternative approach to David Hume's aesthetics from the underlying thesis that he strove to neutralize the undecidability of aesthetics judgment resulting from the centrality granted to subjective experience. Using the precedents of George Cheyne's physiology and Jean-Baptiste Du Bos' environmentalism, both know to Hume, it is proposed that Hume reduced the Cartesian subject to physiology, thus making

non-viable the aesthetic autonomy. To articulate this thesis, in addition to his aesthetic texts, are used Hume's connections with anatomy, his critique of personal identity, and his reflections on the materiality of thought. All this converges in the interpretation of the standard of taste as something whose feasibility is evinced by the statistics of the historical process in the figure of the critic, who stand for the state of organic health that anyone could enjoy, all of which makes Hume's an essential reference for the new aesthetics of heteronomy.

Keywords: Art, Statistics, Aesthetics, Physiology, Heteronomy, History, Hume.

INTRODUCCIÓN

El empirismo británico dio importancia a la idea de que la belleza no estaba en el objeto experimentado sino en el sujeto que lo experimentaba¹, tesis que David Hume presentó en algunas de sus obras. Con esto se asomó a lo que Kant denominó la antinomia del gusto², la contradicción entre el deseo de alcanzar un criterio común y la diversidad del gusto, abordando su indecidibilidad con una teoría fisiológica, estadística y mecanicista de la evaluación estética³.

Hume resolvió la antinomia erosionando la consistencia del sujeto cartesiano a partir de su admiración hacia la anatomía⁴, la cual le hizo afrontar el problema del gusto con cuatro premisas: 1) tenemos el mismo cuerpo; 2) compartimos el mismo entorno; 3) nuestros gustos divergen por algún defecto en nuestra respuesta corporal al entorno; 4) superados tales defectos los gustos convergerán. Además de a los textos de Hume, para desarrollar esta hipótesis vamos a recurrir a escritos de George Cheyne y Jean-Baptiste Du Bos, conocidos por él antes de comenzar la redacción del *Tratado de la naturaleza humana* (1739-1740). Cheyne le proporcionó un modelo fisiológico y Du Bos otro complementario de determinación ambientalista.

Estructuraremos nuestro argumento siguiendo el orden cronológico de influencias y desarrollo del pensamiento humeano, identificando los instrumentos de los que se sirvió y mostrado el uso que hizo de ellos: en el primer apartado nos

1 TATARKIEWICZ, Wladyslaw, *Historia de seis ideas*. Madrid, Tecnos, 1990, 247-248.

2 KIVY, Peter, *The Seventh Sense. Francis Hutcheson and Eighteenth-Century Aesthetics*. Oxford, Oxford University Press, 2003, 293-294; COSTELLOE, Timothy M., *Aesthetics and Morals in the Philosophy of David Hume*. New York, Routledge, 2009, 37-52; COSTELLOE, Timothy M., *The British Aesthetic Tradition*. Cambridge, Cambridge University Press, 2013, 54.

3 SHELLEY, James, "Hume and the Value of the Beautiful". *The British Journal of Aesthetics*, 51, 2, 2011, 213-222; STROUD, Barry, *Hume*. London, Routledge, 1977, 9 y 13.

4 Disentiremos de RIBEIRO, Brian, "Hume's Standard of Taste and *De Gustibus Sceptic*", *British Journal of Aesthetics*, 47, 1, 2007, 16-28, para quien Hume nunca superó dicha indecidibilidad. Recomendamos ARRIETA URTIZBEREA, Agustín, "¿En qué sentido es relativista David Hume?" *Contrastes. Revista Internacional de Filosofía*, XV, 2010, 7-24.

centraremos en *The English Malady* de Cheyne; en el segundo en las *Reflexiones críticas sobre la poesía y sobre la pintura* de Du Bos; en los tercero y cuarto en su obra estética, junto con sus reflexiones sobre la anatomía, la unidad y sustancialidad del sujeto y la materialidad del pensamiento, respectivamente; y en el quinto concluiremos calificando la estética de Hume de heterónoma, identificando la ayuda que puede prestarnos en la reubicación del arte y la estética actualmente en marcha.

1. LA FISIOLÓGÍA DE CHEYNE

Hume conoció durante su juventud algunas obras de Cheyne, comenzando por sus *Philosophical Principles of Natural Religion* (1705), un tratado sobre la hipótesis de un Dios diseñador inteligente ejemplificada con el cuerpo humano. Este libro fue recomendado en la Universidad de Edimburgo por Robert Stewart cuando Hume fue su alumno, formando parte de su Biblioteca Fisiológica⁵. Cheyne describió de nuevo al ser humano como una máquina bien diseñada en *The English Malady* (1733), que seguramente Hume también leyó.

Cheyne definió “la enfermedad inglesa” como una combinación de depresión y obesidad que debía afrontarse con dieta y ejercicio⁶. Hume la sufrió y la afrontó así, como muestra su carta de 1734 a un médico de identidad desconocida (Hume 1932, 12-18). Hoy parece aceptarse que fue Cheyne⁷, asignación acorde con el modelo de relación del cuerpo con la mente que Hume aceptó y la terapéutica que siguió. Redactada un año después de publicarse *The English Malady* y justo antes de su viaje a Francia, su importancia reside en la evidencia que proporciona sobre el convencimiento del joven Hume de que la relación del cuerpo humano con su entorno condicionaba nuestro comportamiento mental y emocional. Conviene, pues, describir el modelo fisiológico de Cheyne antes de aplicarlo a la interpretación de la estética de Hume.

5 GUERRINI, Anita, *Obesity and Depression in the Enlightenment: The Life and Times of George Cheyne*. Norman: University of Oklahoma Press, 2000, 89. Sobre la Biblioteca Fisiológica y Hume, véase BARFOOT, Michael, “Hume and the Culture of Science in the Early Eighteenth Century”. En STEWART, M. A. (ed.), *Studies in the Philosophy of the Scottish Enlightenment*. Oxford: Clarendon, 1991, 152-160.

6 GUERRINI, *Obesity and Depression in the Enlightenment*, 144-149.

7 BURTON, John Hill, *Life and Correspondence of David Hume I*. Edinburgh: William Tait, 1846, 42-43; MOSSNER, E. C., “Hume’s Epistle to Dr. Arbuthnot, 1734: The Biographical Significance”. *Huntington Library Quarterly*, VII, 1944, 135-152; WRIGHT, John P., “Dr. George Cheyne, Chevalier Ramsay, and Hume’s Letter to a Physician”. *Hume Studies*, 29, 2003, 125-141; WRIGHT, John P., *Hume’s A Treatise of Human Nature: An Introduction*. Cambridge: Cambridge University Press, 2009, 16-19.

Cheyne pertenecía a una generación de médicos escoceses que estaban explorando la relación entre el cuerpo, los nervios, el cerebro y la mente⁸. Como otros que aceptaron la teología del diseño, se aproximó al cuerpo humano desde la metáfora mecánica, analizándolo como un mecanismo compuesto por piezas infinitamente pequeñas, dada la divisibilidad infinita de la materia (Cheyne 1733, 4-5 y 92-93)⁹. Entre ellas estaban el cerebro y los nervios, descritos con la metáfora del tejido y el hilado que pusieron de moda los microscopistas, quienes vieron el cuerpo animal como haces de fibras tejidas que formaban los órganos (Cheyne 1733, 60-61)¹⁰. Estos haces precisaban cierta tensión, la cual necesitaba un grado de humedad o relación entre sólidos y líquidos (Cheyne 1733, 10). Esta humedad dependía del tipo de partículas que nuestro cuerpo absorbía de su entorno mediante el aire y la alimentación, las cuales adoptaban tamaños y formas concretas y se relacionaban según las leyes gravitatorias newtonianas (Cheyne 1733, 10, 41 y 47)¹¹.

La constitución personal, la alimentación, los hábitos vitales y el medio ambiente definían el tamaño y la forma de las partículas corporales y estas el grado de humedad, determinando la tensión fibrilar que influía sobre nuestra situación emocional y mental. Cheyne utilizó esto para identificar enfermedades y describir terapias que Hume conocía, algo cuya vinculación con *La norma del gusto*, donde remitió las diferencias del gusto a perturbaciones orgánicas, es sugerente. Cheyne afirmó que el sentimiento surgía de

[...] el impulso, movimiento o acción de los cuerpos que presionan, suave o violentamente, sobre los extremos de los nervios, de la piel o de las otras partes del cuerpo, las cuales, por su estructura o *mecanismo*, transmiten este movimiento al principio *sentiente* del cerebro (Cheyne 1733, 71)¹².

Sobre esta base desarrolló una tipología de formas emocionales y mentales en función de las situaciones corporales de las que eran síntomas. Sirva la siguiente cita como un ejemplo extensible a la experiencia estética:

Aquellos que tartamudean, balbucean, hablan bajo y con dificultad, pierden su voz sin estar resfriados y se vuelven mudos, sordos o ciegos sin ningún accidente o

8 WRIGHT, John P., "Metaphysics and physiology: Mind, body and the animal economy in eighteenth-century Scotland", en STEWART, M. A. (ed.), *Studies in the Philosophy of the Scottish Enlightenment*. Oxford, Clarendon, 1991, 251-301.

9 RISKIN, Jessica, *The Restless Clock: A History of the Centuries-Long Argument over What Makes Living Things Tick*. Chicago, The University of Chicago Press, 2016, 81-82.

10 ISHIZUKA, Hisao, *Fiber, Medicine, and Culture in the British Enlightenment*. New York, Palgrave Macmillan, 2016, 29-52.

11 KING, Lester S., "George Cheyne, Mirror of Eighteenth-Century Medicine". *Bulletin of the History of Medicine*, 48, 4, 1974, 528-533.

12 Las traducciones de *The English Malady* son nuestras.

destemplanza aguda, son ágiles, rápidos y apasionados, poseedores de nervios débiles y de un alto grado de sensibilidad; son pensadores ágiles, sienten más fácilmente placer y pena y tienen más viva imaginación (Cheyne 1733, 105).

Que uno fuera un pensador sensible e imaginativo dependía de la posición y contextura de los nervios entre los órganos de nuestro cuerpo. Este comportamiento mental permitía elaborar diagnósticos e identificar causas orgánicas que explicasen nuestras respuestas emocionales e imaginativas. Basándose en esto, Cheyne catalogó tipos mentales en función de sus cualidades orgánicas. Por ejemplo:

[...] la división usual del ser humano en *pensadores ágiles, lentos y no pensadores* no carece de base en la naturaleza y la verdadera filosofía. Las personas de nervios finos y débiles pertenecen por lo general a la primera clase: la actividad, movilidad y delicadeza de sus órganos intelectuales los hace así y, por lo tanto, debilita y relaja los *órganos materiales* de las facultades intelectuales (Cheyne 1733, 182).

Cheyne estaba describiendo la relación del pensamiento con el cuerpo, problema al que los médicos escoceses contemporáneos dieron diferentes soluciones¹³. Presentó la suya con una metáfora, entonces común¹⁴, que comparaba al cuerpo humano con un instrumento musical manejado por un músico alojado en el cerebro:

Estando las fibras menos interrumpidas o rotas, especialmente los *nervios*, el movimiento que se les comunica sirve para trasladar la impresión, que reciben de los objetos exteriores o de las fibras musculares, al *sensorio* del cerebro y desde aquí al principio sentiente o *músico*, así como desde este a los órganos (Cheyne 1733, 64).

Se preguntó si “no tendría el *principio sentiente* su asiento en algún lugar del cerebro en el que terminan los nervios, como el *organista* dentro de su órgano” (Cheyne 1733, 88)¹⁵. Esta curiosa imagen hacía pensar en un sujeto autónomo separado de su cuerpo que recogía las fibras corporales, procesaba su sonido y componía como música nuestro comportamiento cotidiano. En esto

13 WRIGHT, “Metaphysics and physiology: Mind, body and the animal economy in eighteenth-century Scotland”, 255-275.

14 ISHIZUKA, *Fiber, Medicine, and Culture in the British Enlightenment*, 218-228; GOUK, Penelope, “Musik and the Nervous System in Eighteenth-Century British Medical Thought”, en KENNAWAY, James (Ed.), *Music and the Nerves 1700-1900*. New York, Palgrave Macmillan, 2014, 44-71; KENNAWAY, James, *Bad Vibrations: The History of the Idea of Music as Cause of Disease*. New York, Routledge, 2016, 23-38.

15 SARAFIANOS, Aris, “«Subsiding Passions and the Polite Arts of Healing»: Music and Images of the Medical Profession in «Moderate Enlightenment», en KENNAWAY, James (Ed.), *Music and the Nerves 1700-1900*. New York, Palgrave Macmillan, 2014, 118-151.

Hume no lo siguió, si bien esta metáfora se puede rastrear en el *Tratado* sometida a una significativa transformación, como veremos.

2 EL AMBIENTALISMO DE DU BOS

Habiendo leído a Cheyne, Hume inició su viaje por Francia con un instrumento apropiado para apreciar las *Reflexiones críticas sobre la poesía y sobre la pintura* (1719) de Du Bos¹⁶, quien analizó la determinación ambiental del gusto con una radicalidad sin precedentes¹⁷. Leyó esta obra antes de 1733 y seguramente en 1734 tuvo relación personal con su autor en París¹⁸, de manera que su contacto con las *Reflexiones* fue intenso durante la génesis del *Tratado*¹⁹, afirmando Sertoli que “los momentos y la secuencia del discurso de Hume se corresponden con las páginas de las *Reflexiones*”²⁰.

Du Bos afirmó que el público no juzgaba las obras de arte con el razonamiento sino con el sentimiento, al que debían someterse las reglas de la crítica, que solo daban cuenta de lo que previamente gustó o disgustó (Du Bos 2007, 313; Du Bos 2015, 425). Puso un ejemplo culinario con el que afirmó, anticipándose a Francis Hutcheson, la existencia de un sentido interior:

Hay en nosotros un sentido destinado a juzgar el mérito de estas obras [...] Es ese sexto sentido que hay en nosotros sin que percibamos sus órganos; es la parte de nosotros mismos que juzga sobre la impresión que padece y que, sirviéndome de la terminología de Platón, se pronuncia sin consultar la regla ni el compás. Eso es lo que comúnmente se denomina el sentimiento (Du Bos 2007, 313-314; Du Bos 2015, 426).

Que el valor artístico dependiera de un sentimiento interior infalible introducía el problema de la indecidibilidad del gusto, que Du Bos afrontó recurriendo a

16 COURTOIS, Jean-Patrice, “La teoría de los climas en Montesquieu, Hume y Voltaire. (Un problema de gramática histórica del Siglo de las Luces)”, *Araucaria. Revista Iberoamericana de Filosofía, Política y Humanidades*, 36, 2016, 141-148.

17 MAZZOCUT-MIS, Maddalena, “Du Bos e la teoría climatica”, en RUSSO, Luigi (Ed.), *Jean-Baptiste Du Bos e l'estetica dello spettatore*. Palermo, Centro Internazionale Studi di estetica, 2005, 103-118; LOMBARD, Alfred, *L'abbe Du Bos, un initiateur de la pensée moderne*. Paris, Hachette, 1913, 242-254.

18 MOSSNER, *The Life of David Hume*, 71.

19 Sobre la conexión con Du Bos véase JONES, Peter, *Hume's Sentiments: The Ciceronian and French Context*. Edinburgh: Edinburgh University Press, 1982, 93-135; y TOWNSEND, Dabney, *Hume's Aesthetic Theory. Taste and Sentiment*. New York, Routledge, 2014, 76-85.

20 SERTOLI, Giuseppe, “Du Bos e l'estetica inglese del Settecento”, en RUSSO, *Jean-Baptiste Du Bos e l'estetica dello spettatore*, 149.

la relación de nuestro cuerpo con su entorno y a la convergencia estadística de la evaluación histórica de las obras de arte, que marcaba una unidad de tendencia que él explicó mediante nuestra identidad corporal. Un argumento similar condujo a Hume a dar relevancia al crítico, pero Du Bos se la concedió a la opinión pública:

Siempre es el sentimiento del público el que prevalece [...] El público futuro, permítaseme esta expresión, que juzgará por sentimiento como lo ha hecho el público contemporáneo, siempre será de la opinión de este (Du Bos 2007, 329; Du Bos 2015, 449).

Con el tiempo crecerá la veneración hacia las obras del pasado (Du Bos 2007, 344; Du Bos 2015, 472), traduciéndose su duración en aumento de valor. Para explicar eso Du Bos recurrió a la dependencia de nuestros órganos corporales con respecto a su entorno físico, reconduciendo la diversidad del gusto a la homogeneidad de los procesos naturales. Consideró menos importante el “medio ambiente moral” (gobierno, población, ética, política o religión) porque, dependiendo de la cambiante voluntad de los sujetos, era una coyuntura “afortunada” que operaba “sin hacer en la naturaleza ningún cambio físico” (Du Bos 2007, 250; Du Bos 2015, 332-333), es decir, estable. Concedió más relevancia al “medio ambiente físico”, cuyos efectos eran más eficaces, afirmando que “las personas atribuyen, a menudo, a las causas morales efectos que pertenecen a las causas físicas” (Du Bos 2007, 255 y 268; Du Bos 2015, 340 y 359). Tales consideraciones se basaban en el modelo del cuerpo humano como una máquina inserta en un medio ambiente climático y terreno que determinaba su respuesta física, emocional y mental. Du Bos creía que nuestros pensamientos variaban como la vegetación:

[...] por el encadenamiento y la dependencia recíproca que existe entre todas las partes del mundo material, las diferencias de clima que se dejan sentir en las plantas deben extenderse y producir algún efecto también en los cerebros (Du Bos 2007, 257; Du Bos 2015, 342).

También desarrolló la hipótesis causal de que el clima de un lugar determinaba al suelo, el suelo al aire y este al cerebro de sus habitantes:

[...] las mismas cualidades del aire dependen de la cualidad de las emanaciones de la tierra que envuelve el aire; así, el aire que la envuelve será diferente según la composición de la tierra. Y las emanaciones de la tierra, en cuanto es un cuerpo mixto en que se producen fermentaciones continuas, no pueden ser siempre de la misma naturaleza en una determinada comarca; estas emanaciones, sin embargo, no pueden variar sin cambiar la temperatura del aire y sin alterar de alguna manera sus cualidades. Por tanto, en virtud de estas vicisitudes, en algún momento tienen que producirse cambios en el espíritu y en el humor de las personas de un determinado

país, en el que debe haber siglos más favorables que otros a la formación física de los niños. Así, en Francia unas generaciones serán más espirituales que otras [...] por la acción de la misma causa que hace que los años no sean igualmente templados y que los frutos de una cosecha valgan más que los de otra (Du Bos 2007, 285; Du Bos 2015, 385).

Esta hipótesis de las emanaciones se conectaba con el desarrollo internacional de la microscopía, que estaba descubriendo organismos insospechados y condujo a Du Bos a pensar que había animálculos de tamaño increíblemente pequeño por todas partes:

[...] los físicos prueban también que el aire está lleno de una infinidad de pequeños animales y de sus espermatozoides. Esto ya es suficiente para concebir sin dificultad que el aire debe estar sujeto a una infinidad de alteraciones resultantes de la mezcla de los corpúsculos incluidos en su composición (Du Bos 2007, 286; Du Bos 2015, 386).

Du Bos situaba esos animálculos en el aire y, mediante él, en nuestro cerebro²¹. Las personas nacían con el mismo cerebro, pero el ambiente físico y sus cambios históricos determinaban que se desarrollase generando unas estructuras intelectuales y emocionales distintas, en las que las artes y el gusto tenían una presencia variable. El gusto y su aparente indecidibilidad eran signos de estos procesos físicos, concluyendo Du Bos que “nuestro espíritu señala el estado actual del aire con una exactitud próxima a la de los barómetros y termómetros” (Du Bos 2007, 287; Du Bos 2015, 387). Hume conoció esto y su influencia es perceptible desde el *Tratado*²².

Acabamos de presentar el marco fisiológico (Cheyne) y ambiental (Du Bos) que consideramos crucial en la interpretación de la estética de Hume, a cuyo desarrollo y análisis dedicaremos los siguientes apartados.

3. DEL TRATADO DE LA NATURALEZA HUMANA A LA NORMA DEL GUSTO

En Francia Hume comenzó a preparar el *Tratado de la naturaleza humana*²³, donde definió la belleza como una “disposición de las partes tal que, sea por la *originaria constitución* de nuestra naturaleza, por *costumbre*, o por *capricho*,

21 WILSON, Catherine, *The Invisible World. Early Philosophy and the Invention of the Microscope*. Princeton, New Jersey, Princeton University Press, 1995, 70-102 y 140-175.

22 YOUNG, James O. y CAMERON, Margaret, “Jean-Baptiste Du Bos’ *Critical Reflections on Poetry and Painting* and Hume’s *Treatise*”. *British Journal of Aesthetics*, 58, 2, 124-125.

23 WRIGHT, *Hume’s A Treatise of Human Nature: An Introduction*, 15-36.

es apropiada para producir en el alma placer y satisfacción”, siendo su causa “la relación con el yo” (Hume 1988, 417-418 y 420; Hume 2002, 195 y 196). En el *Tratado* dominó la relación con el sujeto, pero Hume apuntó a la originaria constitución de nuestra naturaleza (no al capricho) para paliar la indecidibilidad del gusto que implicaba.

En el *Tratado* y en la *Investigación sobre los principios de la moral* (1751) ese esquema se complementaba con la comodidad, la simpatía y la imaginación²⁴, derivándose la belleza del agrado que sentíamos cuando contemplando un útil nos poníamos con la imaginación, mimética y simpatéticamente, en el lugar del usuario²⁵. No estamos ante algo desinteresado²⁶, pues la usabilidad era la fuente de la complacencia estética²⁷, dominando el sentido somático de la experiencia del confort²⁸.

En la *Investigación* Hume extendió sus ejemplos a las bellas artes²⁹, lo que le permitió conceder más importancia relativa al principio estético subjetivo, aunque en términos generales el esquema era el mismo (Hume 1991, 165; Hume 1983, 86). Mantuvo el principio utilitarista, explicando que en el juicio de gusto se tienen en cuenta los sentimientos de la persona afectada, que “comunican al espectador toques parecidos de dolor y placer” (Hume 1991, 92; Hume 1983, 46). Los sentimientos del espectador son correlatos subjetivos de los del actor que usa el objeto; por ejemplo, un vestido que dificulte sus movimientos es desagradable para este, pero también para quien lo observa. La afirmación de que “solo de los sentimientos del espectador surge la elegancia y la belleza” se acompañaba de un planteamiento simpatético que subordinaba la indecidibilidad del gusto al bienestar orgánico (Hume 1991, 166; Hume 1983, 87). Desechamos, pues, la

24 KORSMEYER, Carolyn, “Hume and the Foundations of Taste”. *Journal of Aesthetic and Art Criticism*, 35, 2, 1976, 206-209.

25 Para Halberstadt era un principio de objetividad, en HALBESTADT, William H., “A Problem in Hume’s Aesthetics”. *The Journal of Aesthetics and Art Criticism*, 30, 2, 1971, 213; véase COSTELLOE, *Aesthetics and Morals in the Philosophy of David Hume*, 3. Sobre la cualidad mimética de la imaginación en Hume véase COSTELLOE, Timothy M., *The Imagination in Hume’s Philosophy: The Canvas of the Mind*. Edinburgh, Edinburgh University Press, 2018, 7-14.

26 TOWNSEND, *Hume’s Aesthetic Theory. Taste and Sentiment*, 165.

27 KIVY, *The Seventh Sense*, 291.

28 LÓPEZ LLORET, Jorge, “De la utilidad de la belleza: Argumentos sobre el clasicismo en la estética de David Hume”. *Daimon Revista Internacional de Filosofía*, 28, 2003, 25-40; AUGUSTO SOLÓRZANO, D. I., “David Hume: La belleza que suscita la comodidad”. *Iconofacto*, 2, 3, 2006, 61-72.

29 KIVY, *The Seventh Sense*, 292.

interpretación de la utilidad humeana como desinterés³⁰, pues era crucial que la satisfacción corporal del actor fuera real³¹.

Con *La norma del gusto* (1757) Hume llegó a la madurez estética. Comenzó vinculando la indecidibilidad del gusto con su condición subjetiva y recurrió a parámetros fisiológicos para evitar cualquier planteamiento antinómico. Como introducción citaremos el famoso episodio narrado por Sancho Panza:

Dieron a dos de mis familiares a probar del vino añejo de una cuba, que se suponía de excelente cosecha, pidiéndoles su parecer sobre el estado del vino. El uno lo probó con la punta de la lengua y tras una reflexión, dijo que el vino sería bueno si no fuera porque sabía a piel curtida. El segundo se lo acercó a la nariz y tras ciertas precauciones, dio su veredicto favorable, aunque dijo que podía distinguir fácilmente un cierto sabor a hierro [...] Con el tiempo, al limpiar la cuba, hallaron en ella una llave pequeña pendiente de una correa de piel curtida (Hume 2008, 254; Hume 1994, 234-235).

En este ejemplo hay tres elementos: 1) el reconocimiento común de la bondad del vino (fase afectiva); 2) el gusto y el olfato de los catadores (fase perceptiva); y 3) la llave con la correa de piel en la cuba llena de vino, que tiene la capacidad de afectar a los sentidos, generando el reconocimiento de la excelencia³². El episodio defendía la intersubjetividad estética porque cada catador aportaba algo para reconstruir un referente común³³, pero creemos más importante que la experiencia del sujeto dependa de la repuesta corporal al entorno a través de leyes que los conectan³⁴. Con esta parábola Hume presentaba su “teoría causal del gusto”³⁵,

30 SACCAMANO, Neil, “Aesthetically Non-Dwelling: Sympathy, Property, and the House of Beauty in Hume’s *Treatise*”. *Journal of Scottish Philosophy*, 9, 1, 2011, 37-58; STOLNITZ, Jerome, “On the Origin of «Aesthetic Disinterestedness». *The Journal of Aesthetics and Art Criticism*, 20, 2, 1961, 131-143, no trató a Hume en su análisis del desinterés.

31 Véase PERRICONE, Christopher, “The Body and Hume’s Standard of Taste”. *The Journal of Aesthetics and Art Criticism*, 53, 4, 1995, 371-378. Perricone se centró en el humor.

32 SHELLEY, James R., “Hume and the Nature of Taste”. *The Journal of Aesthetics and Art Criticism*, 56, 1, 1998, 33-34.

33 FRANZINI, Elio, *La estética del siglo XVIII*. Madrid: Visor, 2000, 140; SHELLEY, James R., “Hume’s Double Standard of Taste”. *The Journal of Aesthetics and Art Criticism*, 52, 4, 1994, 442-443.

34 KULENKAMPFF, Jens, “The Objectivity of Taste: Hume and Kant”. *Nous*, 24, 1990, 99.

35 SHINER, Roger A., “Hume and the Causal Theory of Taste”, *Journal of Aesthetics and Art Criticism*, 54, 3, 1996, 239-243; COSTELLOE, *Aesthetics and Morals in the Philosophy of David Hume*, 2.

cuyo tema no era la belleza sino la capacidad, primero perceptiva y después imaginativa³⁶, de los jueces para reconocerla y argumentarla³⁷.

Hume rechazó la indecidibilidad del gusto presente en la tesis de que “todo sentimiento es bueno, porque el sentimiento solo hace referencia a sí mismo” (Hume 2008, 249; Hume 1994, 229-230). Casi todos los argumentos de *La norma del gusto* remiten a la constitución orgánica del sujeto, a la que se accedía identificando “lo que suele complacer en todos los países y épocas” (Hume 2008, 251; Hume 1994, 231), siguiendo con ello los métodos de la estadística. El ejemplo obligado era Homero, cuya excelencia resultaba tan obvia en el siglo XVIII como más de dos mil años atrás (Hume 2008, 253; Hume 1994, 233). La estadística histórica indicaba inductivamente la convergencia de los sujetos hacia algo más que subjetivo, algo que pertenecía a los “avatares de la vida” compartida (Hume 2008, 96; Hume 1994, 3), sugiriendo causalidades, no casualidades:

Lo que depende de pocas personas en gran medida puede atribuirse a la casualidad, o a causas secretas y desconocidas; lo que surge de un gran número de personas, puede explicarse por causas fijas y conocidas (Hume 2008, 144; Hume 1994, 112).

Divergiendo de Du Bos, Hume conectó la convergencia estadística con el crítico porque solo él, “acostumbrado a ver, examinar y sopesar las varias creaciones admiradas en diferentes edades y naciones, puede estimar los méritos de una que se le presenta” (Hume 2008, 258; Hume 1994, 238). Introdujo la dimensión fisiológica de la experiencia estética dentro de esta reivindicación social del crítico, entendido como un agente real y no ideal³⁸. Aunque solo él poseía suficiente sensibilidad innata, optimizada mediante el entrenamiento histórico, para paliar “los defectos de los órganos internos” que debilitaban “la influencia de esos principios generales de los que depende nuestro sentimiento de belleza” (Hume 2008, 254; Hume 1994, 234), cualquiera podría acceder al gozo de unos órganos sanos, lo que reducía la indecidibilidad del gusto a un estado transitorio de enfermedad y hacía depender la idoneidad del juicio de la aptitud perceptiva resultante de la salud orgánica³⁹. Era una realidad empírica compartida por todos, idéntica en la salud y diferente en la enfermedad: “todas las criaturas pueden encontrarse en un es-

36 Para GRACYCK, Theodore, “Delicacy in Hume’s Theory of taste”. *Journal of Scottish Philosophy*, 9, 1, 2011, 1-16, era una capacidad de la imaginación.

37 MARSHALL, David, “Arguing by Analogy: Hume’s Standard of Taste”. *Eighteenth Century Studies*, 28, 3, 1995, 328-329.

38 ROSS, Stephanie, “Humean Critics: Real or Ideal?” *The British Journal of Aesthetics*, 48, 1, 2008, 20-28.

39 DURÀ-VILÀ, Víctor, “Courage in Art Appreciation: A Humean Perspective”. *British Journal of Aesthetics*, 54, 1, 2014, 80-82.

tado sano y uno enfermo, y solo el primero nos proporciona la verdadera norma del gusto y del sentimiento” (Hume 2008, 253; Hume 1994, 233). Puesto que el gusto dependía de la “disposición de los órganos” (Hume 2008, 250; Hume 1994, 230), era una manifestación epidérmica cuya indecidibilidad aparente estaba determinada por la situación contextual de nuestro cuerpo, cuyos defectos orgánicos pasajeros afectarían también al crítico si estuviera en la situación del que yerra. Si minimizamos esta presencia del cuerpo en el planteamiento de Hume podría parecer que el crítico naturalizaba el privilegio social burgués⁴⁰, pero introduciendo el cuerpo su propuesta resultará más compleja: el privilegio, construido socialmente⁴¹, permitía optimizar una respuesta cuyos elementos no eran sociales sino fisiológicos, basados en el cuerpo común. Esto hacía viables unos “principios universales de base empírica”⁴², la convergencia de la crítica y el público como algo históricamente factible⁴³.

Lo normal era el error y la indecidibilidad aparente del gusto, pero su interpretación en términos de salud y enfermedad orgánicas permitía elaborar un diagnóstico y proponer como terapia universal una cultura de la aptitud para la que el crítico se preparaba consultando la Historia, con cuyos registros controlaba sus juicios⁴⁴. Como sujetos fisiológicos enraizados en la naturaleza y la historia podíamos afinar progresivamente nuestros órganos, poniéndolos en disposición de discriminar mejor en situaciones en las que las cualidades objetivas “pueden encontrarse en pequeño grado, o se mezclan y confunden” (Hume 2008, 255; Hume 1994, 235). Hume afirmó que “los principios generales del gusto son uniformes y dependen de la naturaleza humana” y que la perfección de nuestros sentidos residía en que fueran “capaces de percibir correctamente los objetos más insignificantes” (Hume 2008, 256; Hume 1994, 236). La norma del gusto no era un principio apriorístico sino la posibilidad de converger en el proceso histórico gracias a la identidad de nuestra naturaleza⁴⁵.

40 SHUSTERMAN, Richard, “The Scandal of Taste: Social Privilege as Nature in the Aesthetic Theories of Hume and Kant”. *The Philosophical Forum*, 20, 3, 1989, 213-221.

41 STRADELLA, Alessandra, “The Fiction of the Standard of Taste: David Hume on the Social Constitution of Beauty”. *The Journal of Aesthetic Education*, 46, 4, 2012, 41-44.

42 DICKIE, George, *El siglo del gusto. La odisea filosófica del gusto en el siglo XVIII*. Madrid, Antonio Machado, 2003, 238 y 252.

43 SHELLEY, James, “Hume and the Joint Verdict of True Judges”. *The Journal of Aesthetics and Art Criticism*, 71, 2, 2013, 148-149.

44 BACESKI, Tina, “Hume on Art Critics, Wise Men, and the Virtues of Taste”, *Hume Studies*, 39, 2, 2013, 233-256.

745 FRIDAY, Jonathan, “Hume’s Sceptical Standard of Taste”. *Journal of the History of Philosophy*, 1998, 36, 4, 554-557.

4. EL CONDICIONAMIENTO FISIOLÓGICO DEL GUSTO

La estética de Hume ha de conectarse con su interés por la anatomía, su análisis crítico de la sustancialidad y la unidad del sujeto y sus consideraciones sobre la materialidad del pensamiento.

En una carta a Hutcheson de 1739, al defenderse de la crítica de que el *Tratado* no estimulaba la causa de la virtud, Hume se definió como anatomista (Hume 1932, 32-33), usando unos argumentos que introdujo después en el *Tratado*:

El anatomista no debe emular nunca al pintor [...] sin embargo, el anatomista se halla admirablemente dotado para aconsejar al pintor; es más, resulta incluso impracticable descollar en este último arte sin la ayuda del primero. Necesitamos tener un conocimiento exacto de las partes, y de su situación y conexión, antes de poder dibujar con elegancia y corrección (Hume 1988, 821; Hume 2002, 395)⁴⁶.

En *La norma del gusto* Hume adoptó esta analogía anatómica⁴⁷, significativa por dos motivos: 1) porque con el término “anatomía”, como sus contemporáneos, se refería también a la fisiología⁴⁸; y 2) por su alta estima hacia los planteamientos fisiológicos, tanta como para afirmar que “el examen de nuestras sensaciones pertenece más a los anatomistas y filósofos de la naturaleza que a la filosofía moral” (Hume 1988, 51; Hume 2002, 11). Aunque no rebasó los límites disciplinares entre la filosofía de la naturaleza y la filosofía moral, aceptó la dependencia de las percepciones con respecto a la fisiología, pues “dependen de causas físicas y naturales” y su examen sería competencia de “las ciencias de la anatomía y la filosofía natural” (Hume 1988, 388; Hume 2002, 181). Adicionalmente, al tratar sobre las relaciones de semejanza, contigüidad y causalidad afirmó que podría profundizarse en sus causas diseccionando el cerebro, “mostrando por qué cuando concebimos una idea los espíritus animales corren por todas las huellas contiguas, despertando a las otras ideas relacionadas con la primera” (Hume 1988, 116-117; Hume 2002, 44). Hume ya había utilizado los espíritus animales para describir su propio estado físico y emocional en la carta a Cheyne (Hume 1932, 17).

46 En el *Abstract* (1740) del *Tratado* se indicaba que “propone anatomizar la naturaleza humana de un modo regular”, en HUME, David, *Abstract*. Barcelona: Humanitas, 1983, 100 (Hume 2002, 407); véase MAZZA, Emilio, “«Cloathing the Parts again»: The Ghost of the *Treatise* in *The Standard of taste*”, en BABICH, Babette (ed.), *Reading David Hume's Of The Standard of Taste*. Berlin, Walter de Gruyter, 2019, 281-299.

47 IMMERWAHR, John, “The Anatomist and the Painter: The Continuity of Hume's *Treatise* and *Essays*”. *Hume Studies*, 17, 1, 1991, 1-14, interpretó que en el *Tratado* Hume adoptó la perspectiva del anatomista, pero en los *Ensayos* la del pintor.

48 CUNNINGHAM, Andrew, *The Anatomist Anatomis'd. An Experimental Discipline in Enlightenment Europe*. Farnham, Ashgate, 2010, 156-164.

Para definir mejor su respuesta a la aparente indecidibilidad del gusto recurriremos a sus reflexiones sobre la identidad personal, la sustancialidad del sujeto y la materialidad del pensamiento, con las que puso en duda la autonomía del sujeto estético. Esta autonomía dependía de la posibilidad de distanciarse del propio cuerpo, proviniendo de la radicalidad con la que Descartes separó las sustancias pensante y extensa: “¿Qué soy, entonces? Una cosa que piensa”⁴⁹, que puede “fingir que no tenía cuerpo y que no había mundo ni lugar alguno en el que me encontrase”⁵⁰, pudiendo separarse de la animalidad de la extensión. Solo el ego humano se distanciaba de su cuerpo, posibilitando la autonomía estética. Hume recurrió a la anatomía comparada para negar la no corporalidad del pensamiento con la misma radicalidad con la que Descartes la afirmó, estableciendo que entre la mente humana y animal la diferencia era de grado:

Los anatomistas suelen añadir a sus observaciones y experimentos sobre cuerpos humanos otras pruebas realizadas con animales, deduciendo de la concordancia de estos experimentos un argumento adicional en favor de alguna hipótesis particular. De hecho, es cierto que, cuando la estructura de las partes de los animales es igual que la del hombre, y la operación de estas partes es también la misma, las causas de esta operación no podrán ser diferentes, de modo que todo lo que descubramos que es verdad en los unos podrá considerarse sin duda alguna como cierto en los otros (Hume 1988, 449-450; Hume 2002, 211).

Hume aceptaba la legitimidad y utilidad de la fisiología experimental a partir del método comparado, incluso lo propuso para comparar la mente y las pasiones humanas y animales⁵¹. No solo las pasiones eran las mismas, sino que sus causas también eran “iguales en los animales y en nosotros, con la natural excepción de nuestro superior conocimiento y entendimiento” (Hume 1988, 450-451; Hume 2002, 211-212). Esta identificación del cuerpo y la mente de los animales con los de los humanos daba lugar a dos hipótesis:

1. Que el animal era sustancia pensante que podía distanciarse de su cuerpo y, con ello, acceder a la autonomía, incluida la estética.
2. Que el humano era sustancia extensa como el animal y su mente no diferiría de su cuerpo. Esto negaba la propuesta cartesiana e introducía serias dudas sobre la posibilidad de la autonomía.

49 DESCARTES, René, *Meditaciones metafísicas* (1641). Oviedo, KRK, 2004, 148.

50 DESCARTES, René, *Discurso del método* (1637). Madrid: Tecnos, 2003, 47.

51 Es la primera de las analogías de MENDUS, Susan, “Personal Identity: The Two Analogies in Hume”. *Philosophical Quarterly*, 30, 118, 1980, 65-63 (la segunda compara la mente con el Estado).

Hume optó por la segunda, lo que tuvo notables consecuencias para la estética, pues lo que se dice de pasiones como el orgullo y la humildad se puede decir del placer y displacer suscitados por los objetos. Estas conclusiones se reforzarán si las unimos con sus reflexiones sobre la identidad personal y la materialidad del pensamiento, que introduciremos retomando la metáfora musical de Cheyne. Hume le dio un giro radical prescindiendo del músico, como en el *Tratado* en relación con la simpatía:

Las mentes de los hombres son similares en sentimientos y operaciones y no hay ninguna que sea movida por una afección de la que, en algún grado, estén libres las demás. Del mismo modo que cuando se pulsán por igual las cuerdas de un instrumento el movimiento de una se comunica a las restantes, así pasan fácilmente de una persona a otra las afecciones, originando los movimientos correspondientes en toda criatura humana (Hume 1988, 765; Hume 2002, 368).

Hume recurrió a la teoría de la armonía que Jean-Philippe Rameau elaboró en su *Génération harmonique* (1737), según la cual un sonido produce automáticamente los sonidos de su acorde⁵², complementándola posiblemente con el mecanicismo de la *Medicina Musica* (1729) de Richard Browne, para quien los nervios respondían simpatéticamente a las vibraciones del aire producidas por los instrumentos musicales⁵³, con lo cual la respuesta simpatética a los sentimientos ajenos tenía una condición reproductiva⁵⁴. No comparó la mente humana con un músico autónomo que recogía y pulsaba las cuerdas de su instrumento, sino con las cuerdas mismas sometidas a las leyes de la acústica⁵⁵, reduciendo la condición del músico al conjunto de sonidos resonantes no unificados por un sujeto que los trascendiera.

Hume también usó la metáfora de la mente como teatro:

La mente es una especie de teatro en el que distintas percepciones se presentan en forma sucesiva; pasan, vuelven a pasar, se desvanecen y mezclan en una variedad infinita de posturas y situaciones. No existe en ella con propiedad ni *simplicidad* en un tiempo, ni *identidad* a lo largo de momentos diferentes (Hume 1988, 357; Hume 2002, 165)⁵⁶.

52 RAMEAU, Jean-Philippe, *Génération harmonique*. Paris, Prault, 1737, Propositiones IV a X, 4-7.

53 BROWNE, Richard, *Medicina Musica: or a Mechanical Essay on the Effects of Singing, Musick, and Dancing on Human Bodies*. London, John Cooke, 1729, 35-36.

54 TOWNSEND, *Hume's Aesthetic Theory. Taste and Sentiment*, 99 y 218, nota 19.

55 ISHIZUKA, *Fiber, Medicine, and Culture in the British Enlightenment*, 222-223.

56 SCHWERIN, Alan, *Hume's Labyrinth: A Search for the Self*. Newcastle upon Tyne, Cambridge Scholars Publishing, 2012, 327-329.

Nuevamente dio un sesgo peculiar a una metáfora entonces usual. La metáfora teatral nos podría definir como actores, espectadores o dramaturgos, lo cual presupondría un sujeto unificador, pero para Hume la mente no era sino los acontecimientos que sucedían en el escenario; no construía coherentemente una trama, como el dramaturgo; tampoco unificaba los momentos que constituían un papel, como el actor; ni, finalmente, interpretaba un sentido a partir de lo observado, como el espectador. Esto es importante para aproximarse a su reducción fisiológica del sujeto de la estética, pues este no transcendía su sentimiento, que era el extremo interior de sus fibras corporales. Era su “*bundle theory of mind*”, según la cual los “seres humanos no son sino un haz o colección de percepciones” (Hume 1988, 356; Hume 2002, 165) o, más radicalmente, un montón o colección (*heap or collection*) de las mismas (Hume 1988, 300; Hume, 2002, 137)⁵⁷. Esta teoría afirmaba que “cuando vuelvo mi reflexión sobre *mí mismo* [...] no puedo percibir nunca otra cosa que las percepciones” (Hume 1988, 830; Hume 2002, 399), rechazando con ello el ego unificado cartesiano. De esta teoría del “anti-cogito” resultaba que el yo no era decisivo en la determinación de la experiencia estética porque se reducía al placer o displacer, ambos condicionados orgánicamente⁵⁸. Lo decisivo era el cuerpo.

La crítica a la inmaterialidad del alma tiene las mismas implicaciones, realizando las connotaciones materialistas de la reducción del sujeto a fisiología. Hume consideró que no había nada tan rebatible como la afirmación de que “es imposible que el pensamiento pueda ser causado en ningún caso por la materia” (Hume 1988, 348; Hume 2002, 161), adoptando una posición definida como “materialismo escéptico” o “vitalista”⁵⁹. Aplicando las conclusiones de su análisis de la causalidad, concluyó que “la materia y el movimiento pueden ser considerados en muchas ocasiones como causa del pensamiento” (Hume 1988, 351-352; Hume 2002, 163).

En su ensayo *Sobre la inmortalidad del alma* afirmó que ningún principio ajeno a la experiencia hacía saber “si la materia, por su estructura u orden, no puede ser la causa del pensamiento”, aproximándose a la hipótesis materialista:

57 Disentimos de STRADELLA, Alessandra, “The Dramatic Nature of Our Selves: David Hume and the Theatre Metaphor”. *Literature and Aesthetics*, 20, 2, 2010, 157, para quien la metáfora teatral conectaba el “haz de percepciones” con la conciencia cotidiana unificada. En palabras de STRAWSON, Galen, *The Evident Connexion: Hume on Personal Identity*. Oxford, Oxford University Press, 2011, 118, “bundles are just bundles”.

58 Sobre el “anti-cogito” véase AINSLIE, Donald C., *Hume’s True Scepticism*. Oxford, Oxford University Press, 2015 192-197.

59 BUCKLE, Stephen, “Hume’s Sceptical Materialism”. *Philosophy*, 82, 322, 2007, 553-578; WILSON, Catherine, “Hume and Vital Materialism”, *British Journal for the History of Philosophy*, 24, 5, 2016, 1002-1021.

“todo es común entre el alma y el cuerpo. Los órganos de uno son los órganos del otro; la existencia, por lo tanto, del uno, debe depender de la del otro” (Hume 2008, 284 y 289; Hume 1994, 591 y 596). Gracias a esto Yolton situó a Hume entre los que podrían aceptar la materialidad del pensamiento⁶⁰. Lo mismo que del pensamiento podría decirse del sentimiento y, con él, de la experiencia estética, de manera que el estado subjetivo de placer o displacer sería algo correspondiente con los órganos de nuestro cuerpo afectados por su entorno.

Tal fue la solución de *La norma del gusto* a la aparente indecidibilidad del gusto: tarde o temprano, por su identidad fisiológica los gustos convergerían, como anticipaba la estadística. Con la estadística del proceso histórico, convergente en la figura del crítico⁶¹, ratificó que el gusto dependía de la relación orgánica con nuestro entorno e, identificando al sujeto con su cuerpo, redujo las diferencias a disfunciones orgánicas superables con el entrenamiento adecuado. Si “todo conocimiento se reduce a la probabilidad” (Hume 1988, 269; Hume 2002, 122)⁶², mientras más se ratifique una opción estética más probable será en el futuro, hasta volverse casi una certeza, como mostraba el ejemplo de Homero⁶³. Se trata de la teoría de la convergencia de las opiniones⁶⁴, que Hume planteó genéricamente en su *Investigación sobre el conocimiento humano* (1748) y que se podría aplicar a la estética:

[...] cuando se ha encontrado que efectos distintos se siguen de causas que, al parecer, son exactamente iguales, todos estos efectos distintos han de presentarse a la mente al trasladar el pasado al futuro, y deben entrar en nuestra consideración cuando determinamos la probabilidad del acontecimiento. Aunque optemos por el efecto que hemos encontrado más frecuentemente y estemos convencidos de que este se dará, no hemos de pasar por alto otros efectos, sino que hemos de asignar a cada uno de ellos un peso y valor determinados según haya sido más o menos frecuente (Hume 2015, 106; Hume 1993, 38).

El argumento fisiológico, histórico y estadístico de *La norma del gusto* conectaba a Hume con la fisiología de Cheyne y el ambientalismo de Du Bos. Justificó la viabilidad de un estándar del gusto explicando las desviaciones en términos funcionales de salud y enfermedad, neutralizando la indecidibilidad subyacente a

60 YOLTON, John W., *Thinking Matter: Materialism in Eighteenth-Century Britain*. Oxford: Basil Blackwell, 1983, 49-63.

61 NOXON, James, “Hume’s Opinion of Critics”. *The Journal of Aesthetics and Art Criticism*, 20, 2, 1961, 161-162.

62 GOWER, Barry, “Hume on Probability”. *The British Journal for the Philosophy of Science*, 42, 1, 1991, 1-19.

63 TOWNSEND *Hume’s Aesthetic Theory. Taste and Sentiment*, 169-173.

64 HOWSON, Colin, *Hume’s Problem: Induction and the Justification of Belief*. Oxford, Oxford University Press, 2003, 208-210.

la remisión de la experiencia estética al sujeto. La de Hume fue una estética de la heteronomía, para la que el gusto era algo que formaba parte de las respuestas de nuestro cuerpo al entorno en el que estaba inmerso.

CONCLUSIÓN: HUME Y LA HETERONOMÍA ESTÉTICA

Arthur D. Danto afirmó que en la segunda mitad del siglo XX se produjo el final del arte porque en él todo se había hecho posible y no se podían proponer discursos legitimadores⁶⁵. El discurso legitimador por excelencia en Occidente fue el del arte como imitación de la realidad⁶⁶, que procedente de la tradición greco-latina se mantuvo hasta finales del siglo XIX, no sin discordancias⁶⁷. El discurso no mimético se inició a finales del siglo XVIII, cuando Adam Smith, afirmando que “la melodía y la armonía de la música instrumental [...] no significan ni sugieren nada”⁶⁸, desplazó el paradigma pictórico-mimético hacia el musical-abstracto, abriendo una vía que culminó cuando Wassily Kandinsky recurrió a la música para legitimar la pintura abstracta en *De lo espiritual en el arte* (1911).

La teoría mimética englobaba muchas disciplinas y no era excluyente⁶⁹. Pese al deseo de establecer jerarquías sociales, como en los parangones renacentistas, las disciplinas más “bajas” también imitaban la realidad mediante procedimientos reglados. Esto fue así hasta el siglo XVIII y la sociedad georgiana lo aceptaba, incluido Hume⁷⁰. Hemos visto el rango tan amplio de ejemplos a los que este recurrió, desde los útiles confortables hasta la épica de Homero⁷¹. Las patas de una silla Chippendale imitaban las garras animales, la coronación de un escritorio Hepplewhite la cornisa de un templo griego o los motivos en las telas las flores, formando con los cuadros de William Hogarth, la arquitectura de Robert Adam o las novelas de Samuel Richardson parte de la misma trama vital⁷².

65 DANTO, Arthur C., *Después del fin del arte* (1997). Barcelona, Paidós, 1999, 75-76.

66 DANTO, Arthur C., “The Artworld”. *The Journal of Philosophy*, 61, 19, 1964, 571-584.

67 TATARKIEWICZ, *Historia de seis ideas*, 301-310.

68 SMITH, Adam, *Ensayos filosóficos*. Madrid: Pirámide, 1998, 203.

69 TATARKIEWICZ, *Historia de seis ideas*, 43-45.

70 TOWNSEND, *Hume’s Aesthetic Theory. Taste and Sentiment*, 87 y 94; TOWNSEND, Dabney, “Hume’s Aesthetic Move: The Legitimation of Sentiment”. *The European Legacy. Towards New Paradigms*, 18, 5, 2013, 554-555.

71 JONES, Peter, “Hume on the Arts and «The Standard of Taste» Texts and Contexts”, en NORTON, David Fate y TAYLOR, Jacqueline Anne (Eds.), *The Cambridge Companion to Hume*. Cambridge: Cambridge University Press, 2007, 414-425.

72 El mejor estudio de este aspecto de la cultura georgiana es el de BREWER, John, *The Pleasures of the Imagination. English Culture in the Eighteenth Century*. London, Harper Collins, 1997.

Según Larry Shiner⁷³, la diferenciación del arte de lo que no es arte se remonta al siglo XVIII. Cuando estaba en marcha surgió el relato de justificación más relevante de los siglos XIX y parte del XX, el de la autonomía de la estética y el arte. Aunque el relato de la autonomía tampoco ha tenido una aceptación homogénea, desde Friedrich Schiller hasta Martin Heidegger, pasando por los grupos de *Arts and Crafts* en la segunda mitad del siglo XIX, el ingreso de varios pintores abstractos en la Bauhaus o las críticas de Walter Benjamin y John Dewey, por indicar solo algunos disidentes⁷⁴, mantuvo su hegemonía hasta la Segunda Guerra Mundial, siendo en la actualidad uno más entre otros.

El arte puede ser lo que quiera un grupo humano, pero nos preguntarnos si habría relatos legitimadores apropiados para la situación actual. Creemos que un candidato sería el relato de la heteronomía, planteado por autores como Gianni Vattimo o Jacques Rancière⁷⁵. A diferencia del de la autonomía, el de la heteronomía es incluyente y asume la pluralidad de los matices éticos, políticos, utilitarios y comerciales de nuestra realidad cotidiana. La heteronomía acaba con la distinción del arte con respecto a otras disciplinas que construyen dicha realidad y exige la inmersión del trabajo artístico en esta, que no es la realidad del museo ni del ego cartesiano, sino la de nuestro cuerpo en sus complejos entornos cotidianos. Nosotros remitimos el relato de la heteronomía a Hume, dejando de lado la alargada sombra de la autonomía kantiana, cuyo ciclo histórico ha concluido.

La autonomía kantiana se basaba en dos momentos del juicio de gusto: el primero afirmaba que no era sensorial ni ético y el tercero que no tenía relación con los fines utilitarios⁷⁶. Con esto estaba distanciándose de Hume⁷⁷, quien hizo depender la experiencia estética de nuestro cuerpo y la vinculó con los objetos utilitarios⁷⁸. Hume no utilizó la palabra “heteronomía”, pero el rechazo kantiano a sus presupuestos hace interesante que se le aplique un término que permite identificar una tradición que desde él ha llegado a nuestros días, cuando la autonomía ha desaparecido ante el aumento de los referentes de la experiencia estética y nuestra necesidad de hacer cotidiano al arte.

73 SHINER, Larry, *La invención del arte. Una historia cultural*. Barcelona, Paidós, 2004, 119.

74 SHINER, *La invención del arte*, 302-362.

75 MURCIA SERRANO, Inmaculada, “La heteronomía de la estética en la filosofía de Gianni Vattimo”. *Aisthesis*, 49, 2011, 205-216; LARREA, Felipe, “Sobre el régimen estético de las artes en Jacques Rancière: heteronomía, sensorium y montaje”. *Pananbí*, 3, 2016, 116-119.

76 KANT, Immanuel, *Crítica del discernimiento* (1790). Madrid, Alianza, 2012, 247-260 y 303-309.

77 MARTÍN PRADA, Juan, “David Hume y el juicio estético”. *Revista de Filosofía*, 73, 2017, 259-279, especialmente 277; INFANTE DEL ROSAL, Fernando, *La autonomía del diseño*. Valencia, Universitat de València, 2018, 22-28.

78 TOWNSEND, *Hume's Aesthetic Theory. Taste and Sentiment*, 108-111.

No minimizamos la importancia de la revolución kantiana, pero en estética ya es algo solo histórico, cuya naturaleza no conceptual, no emocional, no sensorial y no utilitaria resulta tan desencarnada (“hedonismo castrado” dijo Adorno)⁷⁹ que nos sirve de poco. Si bien durante mucho tiempo Kant triunfó sobre Hume⁸⁰, este ha tomado la revancha por haber promovido lo emocional, utilitario, corporal y sensorial, lo que lo convierte en un referente de la alternativa heterónoma. En realidad, el referente sería la cultura georgiana a la que pertenecía, la última que vivió la heteronomía del arte sin problema antes del advenimiento de la autonomía en 1790, cuando el arte empezó a separarse de nuestra vida común. Visto así, a diferencia de Jones, quien creía que el contexto en el que Hume escribió era profundamente diferente del nuestro⁸¹, creemos que Hume y su cultura están más próximos a nosotros que Kant.

La historia de la estética y el arte en clave heterónoma, a partir de nuestro cuerpo inserto en su entorno físico y social, conduce a una pregunta simple: ¿y si el arte, en lugar de acabarse, se hubiera encarnado de nuevo y vuelto heterónimo? En este caso lo relevante no es que todo sea posible, sino que vuelve a escuchar a nuestro cuerpo y a enriquecer nuestra existencia colectiva cotidiana, con la misma naturalidad que en la época georgiana, pero según nuestras peculiaridades sociales e históricas, con unos medios tan poderosos de intervención en nuestros entornos que hubieran hecho las delicias de los georgianos. Esta reencarnación del arte como un ámbito no privilegiado entre otros, politizado o no, moralizado o no, en nuestros objetos de uso, en nuestras casas, en nuestras calles, en cualquier parte, dirigiéndose de nuevo a nuestros cuerpos en relación, ya no construiría espacios de institucionalización excluyente, sino que haría nuestra vida común más humana y consciente:

Sin instituciones, sin Estética, sin espacios exclusivos, sin la ventaja de la excepción, es más, resistiendo su oposición, los diseñadores y creadores de lo cotidiano invertimos un gran esfuerzo en suscitar una percepción generalizada que haga que la gente disfrute de las acotaciones del mundo. No solo hacemos el mundo más operativo, más accesible o más bello, lo hacemos más consciente⁸².

79 ADORNO, Theodor W., *Teoría estética* (1970). Madrid, Akal, 2004, 18.

80 INFANTE, *La autonomía del diseño*, 21.

81 JONES, Peter, “Hume on the Arts and «The Standard of Taste» Texts and Contexts”, 445.

82 INFANTE, *La autonomía del diseño*, 138.

REFERENCIAS BIBLIOGRÁFICAS

- ADORNO, Theodor W., *Teoría estética* (1970). Trad. Jorge Navarro Pérez. Madrid, Akal, 2004.
- AINSLIE, Donald C., *Hume's True Scepticism*. Oxford, Oxford University Press, 2015.
- ARRIETA URTIZBEREA, Agustín, “¿En qué sentido es relativista David Hume?” *Constrastes. Revista Internacional de Filosofía*, XV, 2010, 7-24.
- AUGUSTO SOLÓRZANO, D. I., “David Hume: La belleza que suscita la comodidad”. *Iconofacto*, 2, 3, 2006, 61-72.
- BACESKI, Tina, “Hume on Art Critics, Wise Men, and the Virtues of Taste”, *Hume Studies*, 39, 2, 2013, 233-256.
- BARFOOT, Michael, “Hume and the Culture of Science in the Early Eighteenth Century”. En STEWART, M. A. (Ed.), *Studies in the Philosophy of the Scottish Enlightenment*. Oxford, Clarendon, 1991, 152-160.
- BREWER, John, *The Pleasures of the Imagination. English Culture in the Eighteenth Century*. London, Harper Collins, 1997.
- BROWNE, Richard, *Medicina Musica: or a Mechanical Essay on the Effects of Singing, Musick, and Dancing on Human Bodies*. London, John Cooke, 1729.
- BUCKLE, Stephen, “Hume's Sceptical Materialism”. *Philosophy*, 82, 322, 2007, 553-578.
- BURTON, John Hill, *Life and Correspondence of David Hume I*. Edinburgh, William Tait, 1846.
- CHEYNE, George, *The English Malady: Or, a Treatise of Nervous Diseases of all Kinds*. London, Strahan, 1733.
- COSTELLOE, Timothy M., *Aesthetics and Morals in the Philosophy of David Hume*. New York, Routledge, 2009.
- , *The British Aesthetic Tradition*. Cambridge, Cambridge University Press, 2013.
- , *The Imagination in Hume's Philosophy: The Canvas of the Mind*. Edinburgh, Edinburgh University Press, 2018.
- COURTOIS, Jean-Patrice, “La teoría de los climas en Montesquieu, Hume y Voltaire. (Un problema de gramática histórica del Siglo de las Luces)”, *Araucaria. Revista Iberoamericana de Filosofía, Política y Humanidades*, 36, 2016, 131-163.
- CUNNINGHAM, Andrew, *The Anatomist Anatomis'd. An Experimental Discipline in Enlightenment Europe*. Farnham, Ashgate, 2010.
- DANTO, Arthur C., “The Artworld”. *The Journal of Philosophy*, 61, 19, 1964, 571-584.
- , *Después del fin del arte* (1997). Trad. Elena Neerman Rodríguez. Barcelona, Paidós, 1999.
- DESCARTES, René, *Discurso del método* (1637). Trad. Eduardo Bello Reguera. Madrid, Tecnos, 2003.
- , *Meditaciones metafísicas*. Trad. Vidal Peña. Oviedo, KRK, 2004.
- DICKIE, George, *El siglo del gusto. La odisea filosófica del gusto en el siglo XVIII*. Trad. Francisco Calvo Garzón, Madrid, Antonio Machado, 2003.

- DU BOS, Jean-Baptiste, *Reflexiones críticas sobre la poesía y sobre la pintura*. Trad. Josep Monter. Valencia, Universitat de València, 2007.
- , *Réflexions critiques sur la poésie et sur la peinture*. Paris, Beaux-Arts de Paris, 2015.
- DURÀ-VILÀ, Víctor, “Courage in Art Appreciation: A Humean Perspective”. *British Journal of Aesthetics*, 54, 1, 2014, 77-95.
- FRANZINI, Elio, *La estética del siglo XVIII*. Trad. Francisco Campillo. Madrid, Visor, 2000.
- FRIDAY, Jonathan, “Hume’s Sceptical Standard of Taste”. *Journal of the History of Philosophy*, 1998, 36, 4, 545-566.
- GOUK, Penelope, “Music and the Nervous System in Eighteenth-Century British Medical Thought”, en KENNAWAY, James (Ed.), *Music and the Nerves 1700-1900*. New York, Palgrave Macmillan, 2014, 44-71.
- GOWER, Barry, “Hume on Probability”. *The British Journal for the Philosophy of Science*, 42, 1, 1991, 1-19.
- GRACYK, Theodore, “Delicacy in Hume’s Theory of taste”. *Journal of Scottish Philosophy*, 9, 1, 2011, 1-16.
- GUERRINI, Anita, *Obesity and Depression in the Enlightenment: The Life and Times of George Cheyne*. Norman, University of Oklahoma Press, 2000.
- HALBESTADT, William H., “A Problem in Hume’s Aesthetics”. *The Journal of Aesthetics and Art Criticism*, 30, 2, 1971, 209-214.
- HOWSON, Colin, *Hume’s Problem: Induction and the Justification of Belief*. Oxford, Oxford University Press, 2003.
- HUME, David, *The Letters of David Hume I: 1727-1765*. Ed. J. Y. T. Greig. Oxford, Oxford University Press, 1932.
- , *Abstract*. Trad. Alicia Olabuenaga. Barcelona, Humanitas, 1983.
- , *An Enquiry Concerning the Principles of Morals*. Ed. de J. B. Schneewind. Indianapolis, Hackett, 1983.
- , *Tratado de la naturaleza humana*. Trad. Félix Duque. Madrid, Tecnos, 1988.
- , *Investigación sobre los principios de la moral*. Trad. Gerardo López Sastre. Madrid, Espasa-Calpe, 1991.
- , *An Enquiry Concerning Human Understanding*. Ed. Eric Steinberg. Indianapolis, Hackett, 1993.
- , *Essays: Moral, Political, and Literary*. Ed. Eugene F. Miller. Indianapolis, Liberty Fund, 1994.
- , *A Treatise of Human Nature*. Ed. por David Fate y Mary J. Norton. Oxford, Oxford University Press, 2002.
- , *Ensayos morales y literarios*. Trad. Estrella Trincado Aznar. Madrid, Tecnos, 2008.
- , *Investigación sobre el conocimiento humano*. Trad. Jaime de Salas Ortueta Madrid, Alianza, 2015.
- IMMERWAHR, John, “The Anatomist and the Painter: The Continuity of Hume’s *Treatise and Essays*”. *Hume Studies*, 17, 1, 1991, 1-14.

- INFANTE DEL ROSAL, Fernando, *La autonomía del diseño*. Valencia, Universitat de València, 2018.
- ISHIZUKA, Hisao, *Fiber, Medicine, and Culture in the British Enlightenment*. New York, Palgrave Macmillan, 2016.
- JONES, Peter, "Hume on the Arts and «The Standard of Taste» Texts and Contexts", en NORTON, David Fate y TAYLOR, Jacqueline Anne (Eds.), *The Cambridge Companion to Hume*. Cambridge, Cambridge University Press, 2007, 414-446.
- JONES, Peter, *Hume's Sentiments: The Ciceronian and French Context*. Edinburgh, Edinburgh University Press, 1982.
- KANT, Immanuel, *Crítica del discernimiento*. Trad. Roberto R. Aramayo y Salvador Mas. Madrid, Alianza, 2012.
- KENNAWAY, James, *Bad Vibrations: The History of the Idea of Music as Cause of Disease*. New York, Routledge, 2016.
- KING, Lester S., "George Cheyne, Mirror of Eighteenth-Century Medicine". *Bulletin of the History of Medicine*, 48, 4, 1974, 517-539.
- KIVY, Peter, *The Seventh Sense. Francis Hutcheson and Eighteenth-Century Aesthetics*. Oxford, Oxford University Press, 2003.
- KORSMEYER, Carolyn, "Hume and the Foundations of Taste". *Journal of Aesthetic and Art Criticism*, 35, 2, 1976, 201-215.
- KULENKAMPPF, Jens, "The Objectivity of Taste: Hume and Kant". *Nous*, 24, 1990, 93-110.
- LARREA, Felipe, "Sobre el régimen estético de las artes en Jacques Rancière: heteronomía, sensorium y montaje". *Panambí*, 3, 2016, 109-128.
- LOMBARD, Alfred, *L'abbe Du Bos, un initiateur de la pensée moderne*. Paris, Hachette, 1913.
- LÓPEZ LLORET, Jorge, "De la utilidad de la belleza: Argumentos sobre el clasicismo en la estética de David Hume". *Daimon Revista Internacional de Filosofía*, 28, 2003, 25-40.
- MARSHALL, David, "Arguing by Analogy: Hume's Standard of Taste". *Eighteenth Century Studies*, 28, 3, 1995, 323-343.
- MARTÍN PRADA, Juan, "David Hume y el juicio estético". *Revista de Filosofía*, 73, 2017, 259-279.
- MAZZA, Emilio, "«Cloathing the Parts again»: The Ghost of the *Treatise* in *The Standard of taste*", en BABICH, Babette (ed.), *Reading David Hume's Of The Standard of Taste*. Berlin, Walter de Gruyter, 2019, 281-299.
- MAZZOCUT-MIS, Maddalena, "Du Bos e la teoría climatica", en RUSSO, Luigi (Ed.), *Jean-Baptiste Du Bos e l'estetica dello spettatore*. Palermo, Centro Internazionale Studi di Estética, 2005 103-118.
- MENDUS, Susan, "Personal Identity: The Two Analogies in Hume". *Philosophical Quarterly*, 30, 118, 1980, 61-68.
- MOSSNER, E. C., "Hume's Epistle to Dr. Arbuthnot, 1734: The Biographical Significance". *Huntington Library Quarterly*, VII, 1944, 135-152.

- MURCIA SERRANO, Inmaculada, "La heteronomía de la estética en la filosofía de Gianni Vattimo". *Aisthesis*, 49, 2011, 205-216.
- NOXON, James, "Hume's Opinion of Critics". *The Journal of Aesthetics and Art Criticism*, 20, 2, 1961, 157-162.
- PERRICONE, Christopher, "The Body and Hume's Standard of Taste". *The Journal of Aesthetics and Art Criticism*, 53, 4, 1995, 371-378.
- RAMEAU, Jean-Philippe, *Génération harmonique*. Paris, Prault, 1737.
- RIBEIRO, Brian, "Hume's Standard of Taste and *De Gustibus Sceptic*", *British Journal of Aesthetics*, 47, 1, 2007, 16-28.
- RISKIN, Jessica, *The Restless Clock: A History of the Centuries-Long Argument over What Makes Living Things Tick*. Chicago, The University of Chicago Press, 2016.
- ROSS, Stephanie, "Humean Critics: Real or Ideal?" *The British Journal of Aesthetics*, 48, 1, 2008, 20-28.
- SACCAMANO, Neil, "Aesthetically Non-Dwelling: Sympathy, Property, and the House of Beauty in Hume's *Treatise*". *Journal of Scottish Philosophy*, 9, 1, 2011, 37-58.
- SARAFIANOS, Aris, "«Subsiding Passions and the Polite Arts of Healing»: Music and Images of the Medical Profession in «Moderate Enlightenment», en KENNAWAY, James (Ed.), *Music and the Nerves 1700-1900*. New York, Palgrave Macmillan, 2014, 118-151.
- SCHWERIN, Alan, *Hume's Labyrinth: A Search for the Self*. Newcastle upon Tyne, Cambridge Scholars Publishing, 2012.
- SERTOLI, Giuseppe, "Du Bos e l'estetica inglese del Settecento", en RUSSO, Luigi (Ed.), *Jean-Baptiste Du Bos e l'estetica dello spettatore*. Palermo, Centro Internazionale Studi di Estetica, 2005, 139-158.
- SHELLEY, James R., "Hume's Double Standard of Taste". *The Journal of Aesthetics and Art Criticism*, 52, 4, 1994, 437-445.
- , "Hume and the Nature of Taste". *The Journal of Aesthetics and Art Criticism*, 56, 1, 1998, 29-38.
- , "Hume and the Value of the Beautiful". *The British Journal of Aesthetics*, 51, 2, 2011, 213-222.
- , "Hume and the Joint Verdict of True Judges". *The Journal of Aesthetics and Art Criticism*, 71, 2, 2013, 145-154.
- SHINER, Larry, *La invención del arte. Una historia cultural*. Trad. de Eduardo Hyde y Elisenda Lulisbert. Barcelona, Paidós, 2004, 119.
- SHINER, Roger A., "Hume and the Causal Theory of Taste", *Journal of Aesthetics and Art Criticism*, 54, 3, 1996, 237-249.
- SHUSTERMAN, Richard, "The Scandal of Taste: Social Privilege as Nature in the Aesthetic Theories of Hume and Kant". *The Philosophical Forum*, 20, 3, 1989, 211-229.
- SMITH, Adam, *Ensayos filosóficos*. Trad. Carlos Rodríguez Braun. Madrid, Pirámide, 1998, 203.
- STOLNITZ, Jerome, "On the Origin of «Aesthetic Disinterestedness»". *The Journal of Aesthetics and Art Criticism*, 20, 2, 1961, 131-143.

- STRADELLA, Alessandra, "The Dramatic Nature of Our Selves: David Hume and the Theatre Metaphor". *Literature and Aesthetics*, 20, 2, 2010, 154-167.
- STRADELLA, Alessandra, "The Fiction of the Standard of Taste: David Hume on the Social Constitution of Beauty". *The Journal of Aesthetic Education*, 46, 4, 2012, 32-47.
- STRAWSON, Galen, *The Evident Connexion: Hume on Personal Identity*. Oxford, Oxford University Press, 2011.
- STROUD, Barry, *Hume*. London, Routledge, 1977.
- TATARKIEWICZ, Wladyslaw, *Historia de seis ideas*. Trad. Francisco Rodríguez Martín. Madrid, Tecnos, 1990.
- TOWNSEND, Dabney, "Hume's Aesthetic Move: The Legitimation of Sentiment". *The European Legacy. Towards New Paradigms*, 18, 5, 2013, 552-562.
- , *Hume's Aesthetic Theory. Taste and Sentiment*. New York, Routledge, 2014.
- WILSON, Catherine, *The Invisible World. Early Philosophy and the Invention of the Microscope*. Princeton, New Jersey, Princeton University Press, 1995.
- , "Hume and Vital Materialism", *British Journal for the History of Philosophy*, 24, 5, 2016, 1002-1021.
- WRIGHT, John P., "Metaphysics and physiology: Mind, body and the animal economy in eighteenth-century Scotland", en STEWART, M. A. (Ed.), *Studies in the Philosophy of the Scottish Enlightenment*. Oxford, Clarendon, 1991, 251-301.
- , "Dr. George Cheyne, Chevalier Ramsay, and Hume's Letter to a Physician". *Hume Studies*, 29, 2003, 125-141.
- , *Hume's A Treatise of Human Nature: An Introduction*. Cambridge, Cambridge University Press, 2009.
- YOLTON, John W., *Thinking Matter: Materialism in Eighteenth-Century Britain*. Oxford, Basil Blackwell, 1983.
- YOUNG, James O. y CAMERON, Margaret, "Jean-Baptiste Du Bos' *Critical Reflections on Poetry and Painting* and Hume's *Treatise*". *British Journal of Aesthetics*, 58, 2, 119-130.

DEL PAGANISMO A LA GRACIA: UNA INTERPRETACIÓN FILOSÓFICA DE LA EXISTENCIA RELIGIOSA EN EL PENSAMIENTO DE S. KIERKEGAARD

*FROM PAGANISM TO GRACE: A PHILOSOPHICAL INTERPRETATION
OF RELIGIOUS EXISTENCE IN THE THOUGHT OF S. KIERKEGAARD*

ÁNGEL ENRIQUE GARRIDO MATURANO

Doctor en filosofía

Investigador principal

CONICET – Instituto de Investigaciones Geohistóricas (IIGHI), Resistencia, Argentina

Chaco/Argentina

hieloypuna@hotmail.com

ORCID: 0000-0002-0509-6692

Recibido: 2/09/2020

Revisado: 23/04/2021

Aceptado: 6/09/2021

Resumen: El artículo analiza la relación entre las formas del paganismo en el pensamiento de Kierkegaard y los tipos de religiosidad (A y B) distinguidas por Climacus en el *Postscriptum*. A través de una hermenéutica pre-confesional aplicada a los discursos edificantes distingue la necesidad de Dios como criterio que separa el paganismo irreligioso del que puede existir en la religiosidad A. Luego explicita la aceptación de la gracia como criterio distintivo entre ambas formas de paganismo y la religiosidad consumada. Finalmente determina la noción de gracia y elucida en qué sentido su aceptación concede a la religiosidad una cierta significación cósmica.

Palabras clave: Absoluto, existencia, gracia, paganismo, religiosidad.

Abstract: The article analyzes the relationship between modes of paganism in Kierkegaard's thought and the types of religiousness (A and B) distinguished by Climacus in the *Postscriptum*. By means of a pre-confessional hermeneutics applied to some edifying discourses, he distinguishes the need of God as a criterion that separates nonreligious paganism from a paganism that can exist in religiousness A. The article goes on to explain the acceptance of grace as the criterion that distinguishes between the two modes of paganism and a consummate religiousness. Finally, it determines the notion of grace and elucidates the sense in which its acceptance provides religiousness with a certain cosmic meaning.

Keywords: Absolute, existence, grace, paganism, religiousness.

INTRODUCCIÓN

Como es sabido, en el *Postscriptum*¹ el autor pseudónimo Johannes Climacus afirma que existen dos tipos fundamentales de religiosidad, la cristiana o religiosidad “B”, y la religiosidad natural o “A”. Esta última, en la que el propio Climacus sitúa su existencia², “sólo tiene como presupuesto la universal naturaleza humana”³ y es experimentada “toda vez que la relación con Dios es descubierta por el existente en la interioridad de la subjetividad”⁴. Tal descubrimiento se produce cuando el sujeto reconoce en su interioridad estar en relación con algo absoluto, que Climacus denomina genéricamente “relación a una felicidad eterna (a la inmortalidad o a la vida eterna)”⁵. De esta religiosidad “A” se afirma explícitamente que ella “puede existir en el paganismo”⁶. Por otra parte, en *Los lirios del campo y las aves del cielo*, más específicamente en la segunda serie de los discursos edificantes que componen dicha obra, titulada “Las preocupaciones de los paganos”, Søren Kierkegaard sostiene que el pagano ha conseguido matar dentro de sí el pensamiento de Dios y, con él, “todos los sentimientos y emociones que como mensajeros del Dios nos lo evocan”⁷. Las reflexiones que aquí comienzan intentarán, por un lado, determinar y distinguir estas dos formas fundamentales que asume el paganismo en Kierkegaard⁸; y, por otro, mostrar en

1 En el caso del *Postscriptum* hemos tomado como referencia para el citado la traducción alemana de B. y S. Diderichsen por considerarla una traducción excelente. (Hay también una buena traducción española: KIERKEGAARD, Søren, *Post Scriptum. No científico y definitivo a Migajas Filosóficas*. (Trad. J. Teira Lafuente y N. Legarreta Bilbao). Salamanca, Sígueme: 2011). Para las demás obras de Kierkegaard nos hemos regido por las traducciones al español. En todos los casos hemos confrontado los textos citados con la siguiente edición danesa de las obras de Kierkegaard: *Søren Kierkegaards Skrifter*. (Edición del Søren Kierkegaard Forskningscenter). Copenhague: Gad, 1994-2008. Sigla: SKS. En todos los casos también se indica la correspondencia entre la traducción citada y los SKS según número de volumen y de página.

2 Cfr. KIERKEGAARD, Søren, *Philosophische Brosamen und Unwissenschaftliche Nachschrift*. 4ª ed. (Trad. alemana de B. y S. Diderichsen). München: Deutscher Taschenbuch Verlag, 2015, 762. (SKS 7, 506). Sigla: UN.

3 UN, 764. (SKS 7, 508).

4 UN, 765. (SKS 7, 509).

5 UN, 764. (SKS 7, 508).

6 UN, 761. (SKS 7, 506).

7 KIERKEGAARD, Søren, *Los lirios del campo y las aves del cielo*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2007, 133. Sigla: LC. (SKS 10, 75).

8 Muy recientemente ha aparecido en nuestro idioma un estudio integral sobre el paganismo en el danés: VIÑAS VERA, Ángel, “Notas sobre el paganismo en Kierkegaard”. *Anales del Seminario de Historia de la Filosofía*, 37 (2), 2020, 259-268. El autor distingue tres formas de paganismo. En primer lugar, el pagano “no sólo es aquel que vive sin Dios en el mundo, sino aquel que puede vivir sin estar en sí mismo” (*op. cit.*, 262), es decir, aquel que tiene por nada tanto su propio espíritu y su propia libertad como a Dios que lo puso como el espíritu libre que él es. A este hombre, que se ha negado a sí y a Dios, sólo le queda el mundo, por lo que su existencia se agota en las preocupaciones

qué medida la contradicción entre el pseudónimo y el autor en lo que a sus juicios sobre el paganismo respecta es sólo aparente, toda vez que cada uno de ellos se refiere a fenómenos distintos, que, junto con la religiosidad “B”, constituyen estaciones diferentes a lo largo de la vía que conduce de la negación de la religiosidad a la realización consumada de esta dimensión constitutiva de la existencia humana. Dentro de este contexto problemático en el que habrá de moverse la investigación, cabe señalar que ella persigue dos objetivos específicos. En primer lugar, ofrecer los criterios de distinción entre ambas formas de paganismo y entre el paganismo en general y la religiosidad “B”. En este sentido nos regimos por la hipótesis de que estos criterios pueden ser caracterizados a partir de una hermenéutica personal de los textos kierkegaardianos que apele no a la obra filosófica pseudónima, sino a los discursos edificantes que llevan la firma del propio autor. A nuestro modo de ver, el hecho de que dichos criterios surjan de obras edificantes no impide que, desde tales obras, pueda reconstruirse una interpretación filosófica personal no confesional y, por tanto, atingente a todo hombre en tanto tal, de las diferencias entre las formas del paganismo y entre el paganismo y la religiosidad consumada. Finalmente, el artículo aspira a poder elucidar el sentido que asume el fenómeno de la *gracia* en el marco de la realización de la religiosidad humana, mostrando en qué medida en virtud de este fenómeno la religiosidad adquiere –para un autor cuya filosofía no sin razones ha sido calificada como un pensamiento de la subjetividad y de la interioridad– una cierta dimensión cósmica.

Una serie de aclaraciones metodológicas tal vez no sean superfluas para evitar posibles malentendidos. En lo que al método respecta, aquí se trata de reconstruir la articulación entre las tesis fundamentales de ciertos discursos edificantes, consagrados a las preocupaciones de los paganos, a la necesidad de Dios o a la dádiva perfecta y el agradecimiento, y los tipos de religiosidad distinguidos en el *Postscriptum*. Con el término *reconstruir* se mienta el hecho de que la articulación

mundanas, hasta tal punto que termina convirtiéndose en esclavo de su propio deseo de ejercer poder sobre el mundo. En segundo lugar, “el paganismo cristiano”. Respecto de éste hace bien el propio autor en citar *El instante* para recordarnos que, en él, «el cambio en relación con el paganismo es que todo permaneció inalterado, pero tomó el predicado de cristiano» (*op. cit.*, 264). Finalmente el paganismo ejemplificado por Sócrates, el cual entiende el ser humano como interioridad “que puede vivirse en relación a algo que no es él, a la verdad, al bien”. (*op. cit.*, 266). A diferencia de Viñas Vera, no nos proponemos aquí hacer un análisis integral y exhaustivo del uso del término pagano en la entera obra de Kierkegaard, sino que abordamos el fenómeno para encontrar, desde un punto de vista filosófico, los criterios que permitan distinguir las formas fundamentales de paganismo entre sí y éstas respecto de una religiosidad consumada. Sin embargo, me parece que las tres figuras que señala el estudioso pueden reducirse a dos. En realidad, como la propia cita de *El instante* lo ratifica, no hay ninguna diferencia *esencial* en el modo de vivir la religiosidad entre las dos primeras formas de paganismo, salvo la hipocresía de la segunda. Y respecto del paganismo socrático, en su búsqueda de una verdad o bien incondicionales, no difiere en nada del pagano que puede existir en la religiosidad “A”. De allí la afirmación de que, en realidad, sólo hay *esencialmente dos* formas de paganismo.

propuesta, si bien no está expresa en los discursos kierkegaardianos, está implícita en ellos como una significación potencial susceptible de ser desplegada por obra de una interpretación que tome como eje de referencia para la lectura la búsqueda de la respuesta a la siguiente pregunta: “¿qué significa ser pagano y, consecutivamente, qué significa existir en el mundo de modo religioso?” En la medida en que se propone explicitar y configurar lo que permanece implícito en la obra kierkegaardiana como mera posibilidad de sentido en función de este eje de referencia, el método de lectura se comprende a sí mismo como *hermenéutico*. Por ser hermenéutico y desplegar una *posibilidad* de sentido implícita, él no pretende ni ser fiel a Kierkegaard ni reivindicar para la lectura propuesta el calificativo de “única correcta” ni el de “completa”. De allí que mi interpretación de Kierkegaard no aspire ser la única posible, sino una plausible y coherente. De lo que acabamos de afirmar se desprende que no se trata aquí de una hermenéutica de tipo filológico ni de intención teológica, sino de lo que podríamos llamar una hermenéutica *filosófica de la facticidad*. En efecto, no me interesa tanto comprender de modo intrínseco los discursos de Kierkegaard ni las implicaciones teológicas o bíblico-exegéticas de su fe a través del análisis erudito de los textos, cuanto describir la esencia de los distintos modos fácticos de existencia a través de los cuales se trasunta la religiosidad constitutiva del hombre *en tanto tal*. Por ello hablamos de una hermenéutica pre- (no pro ni anti) confesional de la facticidad y de un enfoque filosófico y no filológico o teológico. Finalmente agregaré que esta hermenéutica responde a una intención *fenomenológico-trascendental*. Jér me de Gramont, refiriéndose a los discursos edificantes de Kierkegaard, sugiere “reconocer como decididamente fenomenológico el hecho de que el *cómo* de la existencia permita acceder a la cosa misma”⁹. Si esto es así, éste ejercicio hermenéutico, que piensa el significado de la cosa misma, en este caso la relación religiosa, en correlación con un modo de experimentar la propia existencia se mantiene, también, en un sentido amplio, en el orden fenomenológico. Puede agregarse que se trata de una fenomenología trascendental, toda vez que ella intenta mostrar las condiciones últimas de posibilidad sobre la base de las cuales es posible comprender el *factum* de una existencia religiosa.

A la intención de interpretar *filosóficamente* los distintos modos de paganismo y de existencia religiosa como categorías constitutivas de la existencia humana en general y, más aún, de hacer tal cosa recurriendo a discursos edificantes podría oponer un lector crítico el argumento de que resulta incoherente reconstruir desde una perspectiva filosófica discursos que son explícitamente confesionales y de intención casi homilética. Permítaseme responder, con G. Thonhauser, que, aunque

9 DE GRAMONT Jérôme, “Kierkegaard. Pequeña fenomenología de la donación”. *Open Insight*, Vol 5 (2014), 79-192, aquí: 85.

ciertamente es menester reconocer el rol central que la espiritualidad cristiana y el comentario evangélico juega en los discursos, “sin embargo no por ello debemos privarnos de percibir también otros contenidos y tendencias en ellos y leerlos desde otras perspectivas”¹⁰. Pues, aunque no se destaquen en ellos menciones a la historia de la filosofía ni advirtamos la intención de desarrollar argumentos de índole filosófica, “la problemática de los discursos toma su punto de partida en cuestionamientos filosóficos”¹¹. Permítaseme, además, recordarle a este imaginario lector crítico, esta vez de la mano de F. Harrits, que los discursos, antes que predicaciones que proclamen enseñanzas teológicas, son invitaciones al lector a volverse hacia sí mismo y comprenderse en función de su lectura de lo allí escrito. Observa Harrits que “de los prefacios y también de muchos pasajes de los discursos surge el hecho de que ellos fueron escritos, hablados y pensados como partituras, compuestas precisamente para aquel que quisiera leerlas en voz alta y les diera, así, su propia voz, de modo que el lector se pueda escuchar hablar a sí mismo y, tal vez, también entenderse a sí mismo en lo dicho”¹². En efecto, para quien lee en voz alta, se disuelve la oposición entre su propia palabra y la versión definitiva de un discurso escrito ya impreso y publicado. “Su carácter como escrito colapsa y se convierte en aquello que en un doble sentido ya era y es: una palabra que nos vuelve hacia nosotros mismos”¹³. Si esto es así, no veo por qué la voz que se les confiera a los discursos a través de una lectura que se apropie de ellos “en voz alta” no pueda ser una filosófica, ni tampoco razón por la cual la vuelta hacia sí mismo, hacia la verdad de la propia existencia, no pueda hacerse, igualmente, en clave filosófica.

1. EL PAGANISMO IRRELIGIOSO

La naturaleza de este tipo de paganismo es expuesta por Kierkegaard, con notoria agudeza, en los discursos sobre las preocupaciones de los paganos. En el primero de esos discursos, consagrado a la preocupación de la pobreza, se observa que, a diferencia del pájaro, que vive del pan que le toque en suerte cada día sin afanarse por el de mañana, y más aún del cristiano, que pide y agradece, porque

10 THONHAUSER, Gerhard, *Über das Konzept der Zeitlichkeit bei Søren Kierkegaard mit ständigem Hinblick auf Martin Heidegger*. Freiburg/München: Alber, 2011, p. 167.

11 *Loc. cit.*

12 HARRITS, Fleming, “Wortwortlichkeit des Geistes. Über Søren Kierkegaards Sprachverständnis in den Reden von 1843-1845”. En: CAPPELØRN, Niels Jørgen y DEUSER, Hermann (Eds.), *Kierkegaard Studies. Yearbook 2000*. Berlin/ New York: Walter de Gruyter, 2000, 121-134; aquí: 133.

13 *Loc. cit.*

confía en que del mañana se hará cargo, de un modo u otro, la providencia divina, el pagano vive preguntándose: “¿Qué comeré; ?, y ¿qué beberé?, ¿hoy?, ¿mañana?, ¿pasado mañana?, ¿este invierno?, ¿la primavera próxima?, ¿cuándo llegue a viejo?, ¿yo y los míos? y toda la nación?; ¿qué comeremos y beberemos?”¹⁴. Y no es que se lo pregunte en un momento de tribulación, arrepintiéndose luego de vivir obsesionado por llenar graneros. No. Como él existe sin Dios y no tiene nadie a quien pedir ni cree que haya nada que agradecer, se lo pregunta a cada momento; no cesa jamás de preguntárselo; y vive con el ilusorio convencimiento de que por sí mismo *puede asegurar* cómo será el día de mañana. Cree posible tener la certeza de que los graneros estarán ahítos, y piensa, por tanto, que, en última instancia, de llenarlos depende el sentido de su existencia y su relación con el cosmos. Su función en el universo queda, pues, reducida a alcanzar un cierto dominio del mundo que le permita por un momento la ilusión de que también mañana podrá satisfacer sus deseos y necesidades – lo que Kierkegaard llama su “cuestión vital”. En términos del propio filósofo: “Estima que ocupa un puesto significativo en el cosmos porque se ocupa exclusivamente de esta cuestión vital; (...)”¹⁵. El pagano, en su forma más irreligiosa, no es, entonces, el politeísta, por ejemplo, en éste nuestro subcontinente sudamericano, el aborigen que confía en que la tierra lo proveerá y que, por ello, aún sumido en el trabajo agotador, vive siempre en el hoy, gozando ahora los bienes que el hoy puede ofrecerle, o incluso sufriendo las penurias que ahora le toquen en suerte, sin perder por ello la confianza en que mañana acabarán y sin dejar tampoco de agradecer a la tierra divina que hasta hoy le ha dado el pan. El pagano, para Kierkegaard, no es, primordialmente, el hombre de la religiosidad natural. Es el que existe sin Dios alguno en el mundo¹⁶ y que, por ello, cree que él puede asegurarse del curso del mundo, puede apropiarse, por un esfuerzo de la voluntad y el conocimiento, del control del mundo y ponerlo a servicio de satisfacer sus deseos, por arbitrarios que estos fuesen. Es víctima, entonces, de un doble absurdo que, tarde o temprano, habrá de pagar. Primero del absurdo objetivo de creer que su conocimiento y su esfuerzo finitos puedan dominar la infinitud y complejidad de relaciones que dictan el curso de los sucesos. En otros términos, es víctima del engaño de creer que él puede ser el señor del mundo y decidir cuál es el significado que el cosmos habrá de tener para su existencia, en vez de reconocer que es el cosmos indomeñable el que nos empuja a tener en él un cierto significado. Segundo, es víctima del absurdo subjetivo

14 LC, 90. (SKS 10, 30).

15 LC, 90. (SKS 10, 30).

16 A mi modo de ver, la expresión “sin Dios en el mundo” debe ser entendida constantemente en un doble sentido: sin un Dios para mí en el mundo, cuya providencia deba pedir y agradecer y en cual se salvaguarde el sentido último de mi existencia; y en un mundo sin Dios como principio que articula y estructura el decurso del universo más allá del poder humano.

de creer que es dueño y señor de sus deseos y necesidades, de que podrá alguna vez acallarlos y satisfacerlos por completo. En otros términos, ignora que nada que pueda alcanzar el día de mañana es absoluto y que, por tanto, ninguna cosa podrá satisfacerlo absolutamente, por lo que siempre habrá un nuevo deseo que turbará de inquietud la aurora. ¿Y cuál es el castigo con el que inexorablemente habrá de pagar el hombre o la sociedad que vivan sumidos en este doble engaño? El de sucumbir a la tentación. ¿Cuál tentación? Kierkegaard es claro al respecto: “La de vivir para afanarse como un esclavo; la tentación consiste en perderse a sí mismo, perder su alma, cesar de ser hombre y de vivir como un hombre para, en vez de ser más libre que el ave, quedar abandonado de Dios y ser un esclavo más miserable que el bruto”¹⁷. Se trata, a fin de cuentas, de no existir libremente, de no decidir cada mañana cómo habrá de salir al mundo y cuál, de entre las posibilidades que se me ofrecen, habrá hoy de realizar en él, así como el pájaro decide cada día hacia dónde habrá de emprender el vuelo. Nada de ello. El castigo consiste en vivir “en esa maldita esclavitud (...) con el corazón sobrecargado por el cuidado de la subsistencia y contaminado de avaricia a pesar de la pobreza en que se está”¹⁸. En síntesis, la primera característica determinante de este paganismo, que vive sin Dios en el mundo, es negar que un Poder otro que el mío dictamina el curso del cosmos y que él tiene, además, en sus manos mi destino, para, en lugar de ello, afirmar que puedo poseer el mundo y asegurarme del día de mañana.

Pero el hombre que vive en el mundo sin Dios no sólo cree que puede apoderarse del mundo, sino que hay una segunda característica esencial suya, que Kierkegaard describe en otro de los discursos consagrados a las preocupaciones de los paganos, a saber, el dedicado a “La preocupación de la temeridad”. Se trata del espejismo de ser el dueño de mi propio ser y capaz de realizarlo por mí mismo. Es una doble ilusión. Primero, porque yo soy puesto como espíritu libre capaz de relacionarse a sí en su relación con lo otro, pero no soy el origen de ésta, mi condición espiritual. Y segundo, porque no está en mi poder ser plenamente el sí mismo que quiero ser. Para ello dependo de lo otro con lo que me relaciono, lo cual, en última instancia, no está bajo mi control. Lo único que puedo realizar con certeza por mí mismo es morir. En este sentido el pagano es lo contrario al pájaro y al lirio que “*quieren siempre lo que Dios quiere y hacen siempre la voluntad de Dios*”¹⁹. Así, por ejemplo, el pájaro, surcando los cielos, realiza aquello para lo cual ha sido constituido y, de este modo, siempre quiere y hace la voluntad de Dios. Y cuando, fatigado de tanto volar, comprende que hoy no puede recorrer distancias tan largas y viene a posarse en una rama o un peñasco, entonces,

17 LC, 93. (SKS 10, 33).

18 LC, 93. (SKS 10, 33).

19 LC, 128. (SKS 10, 70).

siendo el sí mismo finito que él es, cumple con la voluntad de Dios, “puesto que esto era cabalmente lo que Dios quería que hiciese”²⁰. Pero si este pagano se opone al pájaro que, inocente e ignorante, cumple la voluntad divina y realiza del mejor modo a su alcance las posibilidades y *el modo de existencia* que le ha sido concedido, más aún se opone al cristiano que “todo lo toma de la gracia divina”²¹, en relación con la cual “no tiene voluntad propia”²². A diferencia del pájaro y también del cristiano, el pagano vive “como si fuese su propio señor, el propio artífice de su dicha, que se tiene que cuidar de todo y también de quien todo depende (...)”²³. Viviendo de tal suerte, niega a Dios y pretende sustituirlo en lo que a Él le es propio. Esta pretensión es precisamente la *temeridad* de aquel que no sólo ignora a Dios, sino que *cree* que puede prescindir de Él para realizarse a sí mismo. Su creencia en sí –su autoafirmación y empoderamiento- es incredulidad en Dios. Su osadía, la absurda temeridad del que se lanza a una misión que supera en mucho sus fuerzas ¿Y cuál es el castigo con el que inexorablemente habrá de pagar el hombre o la sociedad que vivan sumidos en el engaño de creer que son artífices de su vida y que pueden por sí mismos realizarse de un modo consumado? La angustia. En ellos no hay ni puede haber paz ninguna, pues, por mucho que lo nieguen, saben que Dios puede en cualquier momento arrebatarles todo nuevamente. “Por eso el pagano está con todo derecho a merced de la angustia; porque en realidad nunca sabe exactamente a merced de quien está. ¿No es esto angustioso?”²⁴.

Podemos reseñar las dos características que definen al pagano que existe en el mundo sin Dios. Él cree que puede poner el futuro del mundo bajo su control y, correlativamente, afirma que es dueño de sí y que puede realizarse a sí mismo según le plazca. No ignora, sino que niega, que, en última instancia, no puede nada y para todo depende de la gracia. Pero, por más que lo niegue, su angustia delata que él no puede nada. No puede nada al principio, puesto que no es el origen de su propia condición ontológica de espíritu libre, la cual no es el fruto de su poder, sino que *le ha sido dada* . No puede nada al final, porque, haga lo que haga y quiere lo que quiera, toda realización de sí de la que sea capaz desemboca en la muerte. Y tampoco puede, en un sentido absoluto, nada entre el principio y el fin, pues no está en su poder manejar las infinitas variables que disponen el curso de la realidad efectiva y, por tanto, tampoco lo está saber con certeza a dónde los conducen, a él y a la propia realidad, el ejercicio de sus poderes. Por no aceptar

20 LC, 128. (SKS 10, 70).

21 LC, 131. (SKS 10, 73).

22 LC, 131. (SKS 10, 73).

23 LC, 133. (SKS 10, 75).

24 LC, 134. (SKS 10, 76).

su nada y confiarse a la gracia, en virtud de la cual puede ser más que sí mismo, el pagano queda disminuido en su ser a su propia representación de sí²⁵. Este pagano irreligioso, que no sólo vive en el mundo sin Dios, sino que niega que tanto él como el mundo necesiten de Dios, no es el politeísta sencillo que sabe que su destino está en manos de la tierra y el cielo y que él nada puede respecto del poder que maneja cielo y tierra. Este pagano irreligioso responde a una concepción de la subjetividad, propia de nuestros días, según la cual el ser humano es capaz de determinar por sí mismo decisivamente la evolución y el sentido de la realidad en su conjunto. Se trata de una concepción que ha llevado hasta la locura la voluntad de poder humana y su pretensión de control sobre el curso de todas las cosas, de tal modo que el hombre ha terminado cayendo en el pecado satánico: creerse Dios. En efecto, como bien observa C. Möller, “los hombres quieren ser como Dios, cuando planean el futuro como un tiempo que estuviera a su disposición”²⁶.

2. PAGANISMO Y RELIGIOSIDAD A

Del paganismo irreligioso distingue Climacus en el *Postscriptum* otro tipo de pagano, a saber, aquel que interiormente experimenta un pathos religioso. Tal cosa ocurre “cuando el individuo se relaciona a una felicidad eterna en la expresión más decisiva del pathos existencial”²⁷. Ahora bien, para comprender cabalmente tal determinación de esta otra forma de paganismo es menester tener en cuenta dos cosas. Primero que, para Climacus, el pathos existencial resulta de

25 En realidad, el pagano que quiere realizarse por sí mismo plenamente termina restringiendo su propia realización a su finitud y limitaciones, en lugar de permitir que aquello que le sale al encuentro –la gracia– le abra a horizontes por él insospechados y para él inalcanzables. Por eso el pagano es un hombre encerrado en sí mismo, que vive en la soledad del héroe trágico. En este sentido no se puede sino coincidir con R Purkardthofer cuando afirma que “sólo a través del anonadamiento (*Vernichtung*) de las representaciones autorrestrictivas y autoimpuestas (...) accede el hombre «a ser sí mismo»: desde la mano de Dios”. PURKARTHOFER, Richard, *Kierkegaard*. Stuttgart: Reclam, 2005, 99.

26 HEYMEL, Michael/ MÖLLER, Christian, *Das Wagnis, ein Einzelner zu sein. Glauben und Denken Søren Kierkegaards am Beispiel seiner Reden*. Zürich: Theologischer Verlag, 2013, 122. Coincido en un todo con Möller cuando contrapone la relación con el porvenir que se rige por una actitud de advento (*adventischer Art*) a la que lo hace por una actitud de índole futurística (*futurischer Art*). “La actitud del preocupado es de índole futurística y pretende ejercer el poder sobre el porvenir considerado como el *futurum* que aparentemente está a su disposición. La actitud del hombre sereno y confiado (*Gelassenen*) es de la índole del advento, porque comprende el provenir como el advenir de la eternidad de Dios en el tiempo”. *Op. cit.*, 121.

27 UN, 761. (SKS 7, 506).

una relación del individuo con un *télos* absoluto²⁸. Se trata de una relación que transforma por entero su existencia, toda vez que la encauza y supedita hacia ese absoluto con el cual se relaciona. En segundo lugar, que dicha relación es un *pathos*, esto es, que el sujeto la padece y que el absoluto con el cual se relaciona no es, como en el caso del paganismo irreligioso, una representación de su voluntad o intelecto, sino un padecimiento que ha sido puesto en él, pero del cual él no es el origen²⁹. Si tenemos en cuenta estos dos aspectos, entonces bien podemos concluir que lo propio de este tipo de paganismo radica en que el individuo, en vez de negar, asume su propio padecimiento originario y, consecuentemente, también la relación con un fin absoluto como la expresión decisiva de su existencia, esto es, como aquello hacia lo cual debieran orientarse y en función de lo cual cobrarían sentido todas las acciones a través de las cuales expresa su interioridad. Este paganismo, a diferencia del anterior, no se opone a la religiosidad B, que se consume en el verdadero cristiano, puesto que la religiosidad que en él se experimenta como *pathos* existencial y que Climacus denomina religiosidad A “debe hallarse primero en el individuo antes de que pueda hablarse de una conciencia de éste de la dialéctica propia de B”³⁰. En efecto, quien niega que sea un constitutivo esencial de su existencia la relación con un poder absoluto que lo trasciende, mal podrá tener fe en Cristo como aquel absoluto en el cual depositar el sentido entero de su existencia. Por ello insiste Climacus en advertir “cuan insensato es que un hombre sin *pathos* pretenda relacionarse con lo esencialmente cristiano, puesto que antes de que pueda hablarse de estar en situación de hacerse consciente de ello, él debe primero existir en la religiosidad A”³¹. En este contexto surgen inevitablemente dos preguntas. Primero: ¿cómo habrá de entenderse esta relación que sostiene el pagano partícipe de la religiosidad A con un *télos* absoluto? Segundo: ¿por qué esta relación, a pesar de separarlo del pagano irreligioso, lo mantiene dentro del paganismo? La expresión más *concreta* de la respuesta a estas preguntas puede encontrarse, según la interpretación aquí propuesta, en los discursos de Kierkegaard y, en particular, en aquel que se titula “Necesitar de Dios es la suprema perfección del hombre”. En este sentido los discursos edificantes, firmados por el propio Kierkegaard, podrían ser considerados como un estadio más

28 “Si el *télos* absoluto, a través de su relación con la existencia de un individuo, no la transforma absolutamente, entonces el individuo no se relaciona con *pathos* existencial, sino con *pathos* estético”. UN, 559. (SKS 7, 353).

29 “Una felicidad eterna es justamente un eterno presupuesto para todo individuo que lo precede en la inmanencia” UN, 225. (SKS 7, 93).

30 UN, 761. (SKS 7, 506). Quizás convenga aquí precisar que la religiosidad cristiana o B es dialéctica por ser paradójica y absurda, es decir, por afirmar como verdad absoluta lo que, desde el punto de vista de la razón es absolutamente contradictorio, a saber, que el Dios infinito y eterno se halla encarnado en, un hombre finito e histórico: Cristo.

31 UN, 761. (SKS 7, 506).

avanzado en la comprensión concreta de lo que significa existir religiosamente que el alcanzado por Climacus en el *Postscriptum*³².

Pues bien, el criterio que determina la existencia del pagano de la religiosidad A y que, además, lo distingue decisivamente de aquel otro que niega a Dios y niega, por tanto también, cualquier absoluto en el que pueda sostenerse y al que pueda referir su existencia finita, está, a mi modo de ver, condensado en el título de uno de los discursos edificantes de 1844: “Necesitar de Dios es la suprema perfección del hombre”. El pagano que existe en la religiosidad A ha alcanzado la suprema perfección que puede alcanzar un hombre por sí mismo, a saber, reconocer que, por sí mismo, tal cual vimos antes, no puede *absolutamente* nada, y que, por tanto, necesita de un Poder absoluto que llamamos Dios. Su relación con el *télos* absoluto, de la que carece el pagano irreligioso, no consiste, por tanto, en la fe convencida y confiada en *ese télos*, sino en el reconocimiento en la interioridad de su *necesidad* de él. Por eso este tipo de paganismo está presupuesto por el cristianismo, pues el que se engaña a sí mismo y piensa que puede vivir sin un absoluto que fundamente sus fines relativos o aquel otro que toma por absolutos fines que en el fondo sabe que son relativos, mal puede llegar a tener fe en Dios como *télos* absoluto. Debe entenderse el título del discurso de modo textual: necesitar de Dios no es la suprema perfección en absoluto, ni tampoco el *télos* absoluto. Necesitar de Dios es la suprema perfección *del hombre*, es decir, la suprema perfección que el hombre puede alcanzar por sí mismo, reflexionando sinceramente sobre la constitución de su propia existencia. Necesitar de Dios es, pues, la conclusión a la que llega el hombre capaz de “conocerse a sí mismo en relación consigo mismo”³³. Ésta es la conclusión del pagano Sócrates, que, intentando conocerse a sí mismo, descubre que no detenta ni la verdad ni el bien absoluto que busca para fundar en él su existir. De allí que bien pueda calificarse

32 El propio Climacus admite que el *Postscriptum* no ha sido escrito por alguien que pueda llamarse cristiano, sino por quien existe en la religiosidad A y se halla, por tanto, personalmente interesado en un bien supremo o una felicidad eterna. Por eso mismo Climacus se pregunta por el cristianismo, dado que éste promete al cristiano dicha felicidad (Cf. *UN*, p. 831; *SKS* 7, 560). Al contrario, los discursos edificantes sí son “discursos cristianos”. Annemarie Pieper nos recuerda que “el nombre histórico Johannes Climacus se deriva de la palabra griega «klimax», en latín «scala paradisi»” PIEPER, Annemarie (Ed.), *Kierkegaard: Die Hauptwerke. Ein Lesebuch*. Tübingen: Narr Francke Attempto, 2018, 12. En concordancia con lo declarado por Climacus en el *Postscriptum* y con la etimología de su propio nombre podría decirse que las afirmaciones del pseudónimo representan un peldaño en la comprensión de la existencia religiosa que conduce a otro superior, asociable a los discursos edificantes y efectivamente cristianos firmados por el propio Kierkegaard.

33 KIERKEGAARD, Søren, *Escritos de Søren Kierkegaard. Volumen 5: Discursos edificantes. Tres discursos para ocasiones supuestas*. (Trad. D. González). Madrid: Trotta, 2010, 307. (*SKS* 5, 305). Sigla: *DE*.

este tipo de paganismo, como lo hace Viñas Vera, de paganismo socrático³⁴. El ejercicio del socratismo le permite a Kierkegaard afirmar que, “si un hombre no se conoce a sí mismo en el sentido de saber que él mismo no es capaz de nada, entonces no advierte, en su significado profundo, que Dios existe”³⁵. Una cosa es, en efecto, condición de la otra. Quien cree falsamente que él lo puede todo y no se percató de que, en un sentido absoluto, no puede nada, no necesita de Dios ni se relaciona fácticamente, esto es, en su propio existir actuante, con Él. Razón por la cual, para él, Dios no existe. Aunque este hombre mencione una y otra vez el nombre de Dios, aunque lo invoque en ocasiones, Dios no existe para él. Por el contrario, aquel que no “ha llegado a conocer a Dios a través de las épocas pasadas de la historia”³⁶, es decir, que no tiene ningún conocimiento objetivo ni doctrinario acerca de Dios, pero que, en cambio, “sabe para sí mismo que no es capaz de nada, ése tiene cada día y a cada instante la deseada e incontestable *oportunidad* de experimentar que Dios vive”³⁷. De él puede decirse que “está en *condiciones* de captar el primer secreto de la verdad, que él mismo no es nada en absoluto, tanto como que el hecho de necesitar de Dios es su suprema perfección”³⁸. Estamos ahora en condiciones de responder las dos cuestiones en torno de las cuales giró este análisis de la relación entre religiosidad A y paganismo. A la pregunta que inquiría por la índole de la relación que sostiene el pagano, partícipe de la religiosidad A, con Dios como *télos* absoluto, habrá que responder que ella no es otra que la *necesidad*. Es precisamente esta necesidad o, para ser más precisos, la apropiación o asunción consciente de esta necesidad de Dios, lo que lo distingue esencialmente del pagano irreligioso. Y a la pregunta que cuestionaba por qué esta relación, a pesar de separarlo del pagano irreligioso, lo mantiene dentro del paganismo, no cabe sino responde apelando a las palabras señaladas en cursiva de las dos citas precedentes. Necesitar de Dios nos ofrece la *oportunidad* de experimentar que Dios vive, no la experiencia viva de Dios³⁹. Nos coloca en *condiciones* de captar el primer secreto de la verdad, pero no implica necesariamente confiar en la verdad. Por ello mismo puede decirse que

34 “Con Sócrates Kierkegaard descubre que ser pagano es una posibilidad de vivir apasionado y en tensión continua hacia el bien perfecto”. VIÑAS VERA, Ángel, *op. cit.*, 266.

35 *DE*, 314. (SKS 5, 312).

36 *DE*, 314. (SKS 5, 313).

37 *DE*, 315. Cursivas mías. (SKS 5, 313).

38 *DE*, 318. Cursivas mías. (SKS 5, 316).

39 En este sentido Michael Bjergso afirma que en la religiosidad A “Dios no es el Dios viviente uno y trino, sino el principio supremo que fundamenta una idealidad ética: que el hombre a cada instante debe relacionarse de modo relativo a lo relativo y de modo absoluto a lo absoluto” BJERGSØ, Michael, *Kierkegaards deiktiske Theologie. Gottesverhältnis und Religiosität in den erbaulichen Reden*. (Trad. Alemana de Krista-Maria y Hermann Deuser). Berlin/ New York: Walter de Gruyter, 2009, 91.

necesitar de Dios constituye la presuposición para creer en Dios o, lo que es lo mismo, existir en la religiosidad A constituye el presupuesto para llegar a ser un verdadero cristiano, pero que ella, en sí misma, no es aún la religiosidad auténtica y consumada. El pagano de la religiosidad sigue siendo, entonces, pagano, puesto que él no confía en la realidad efectiva de aquel Dios que necesita. ¿Quién es, concretamente, este pagano? Otra vez sería un despropósito identificarlo con las religiones politeístas. Este pagano es todo hombre que se rebela contra el discurso del empoderamiento de sí, de la autorrealización y de la autoafirmación que caracteriza particularmente a nuestra sociedad contemporánea. Es todo hombre que comprende el sinsentido de creer que se puede tener todo bajo control y que, consecuentemente, es posible afirmar que el futuro está en nuestras manos. Es aquel individuo inquisitivo que eleva su grito contra una sociedad que se afana en convencerse de que, al fin y al cabo, lo único que verdaderamente importa es obtener más y más placer y, por ello, corre, angustiada y enloquecida, de un goce efímero al otro, cueste lo que cueste la obtención de tales goces. En síntesis, es todo aquel hombre que se conoce a sí mismo y que, por eso, vive en tensión hacia y en búsqueda de un absoluto que no encuentra en sí, pero que necesita. Precisamente por ello el signo de su existencia es el sufrimiento. Él sólo “experimenta la felicidad eterna mediada a través del sufrimiento de su propia nulidad en comparación de la medida absoluta”⁴⁰. El sufrimiento es, así, la consecuencia inevitable del que existe detenido en la religiosidad A. Ésta no puede convertirse en el estadio final de las etapas de la vida. Es menester llegar a B. El paso es un salto e implica una decisión libre: aceptar la gracia.

3. LA GRACIA Y LA RELIGIOSIDAD B

¿Cómo es posible pasar de necesitar a Dios a confiar en Dios? ¿Cómo un *télos* absoluto, ideal e inmanente al sujeto, puede convertirse en un Dios viviente fuera del sujeto y actuante en la realidad efectiva? ¿Cómo se sube el peldaño que asciende de la religiosidad A a la B? La respuesta usual a esta pregunta es de naturaleza teológica y confesional. Puede sintetizarse, por ejemplo, en la afirmación de A. Pieper: “La condición para ello es la aceptación de Jesucristo como Dios”⁴¹. Por supuesto que existen múltiples pasajes que avalan esta interpretación. Así, por ejemplo, explícitamente afirma Climacus que la religiosidad B consiste en

40 BANSER, Ann-Kathrin/ BODE, Philipp, *Selbstwerden. Über das Selbst als Aufgabe und die Möglichkeiten seiner Realisierung bei Søren Kierkegaard*. Würzburg: Königshausen&Neumann, 2018, 95.

41 PIEPER, Annemarie, *Søren Kierkegaard*. München: Beck, 2014, 95.

la relación del individuo con Dios como algo externo a sí mismo a través de la paradoja⁴². Y la paradoja, imposible de pensar, no es otra que la encarnación del Dios eterno en la persona humana e histórica de Cristo⁴³. Sin embargo, Climacus, en la conclusión del *Postscriptum*, hace también una advertencia: “Lo reconocible inmediato es paganismo: todas las solemnes declamaciones de que se trata en verdad de Cristo y de que Cristo es el verdadero Dios no ayudan mucho en cuanto acaban en lo reconocible inmediato”⁴⁴. En efecto, si todo se redujera a declarar que Cristo es Dios se correría el riesgo de reconvertir a Dios en un dogma objetivo y cognoscible de un modo inmediato. Ahora bien, un dogma tal es una representación humana, con lo que Dios quedaría convertido en una determinación del sujeto. Con ello, en lugar de avanzar a la religiosidad B, caeríamos por debajo de la A, puesto que aquella necesita de Dios, pero no lo encuentra dentro de sí, mientras que ahora parece habérselo encontrado en un conocimiento objetivo propio. Por lo tanto, y aun concediendo que la condición para pasar a la religiosidad B es la aceptación de Cristo, debiera volver a preguntarse, sobre todo desde una perspectiva filosófica, ¿qué significa tal aceptación? ¿Cuál es la existencia concreta que manifiesta en la realidad efectiva que nos hallamos ante un individuo que ha dado el paso hacia la forma más alta de religiosidad: la B?

Indicaciones decisivas las ofrece el propio Kierkegaard en uno de sus discursos edificantes⁴⁵ de 1843, titulado “Toda dádiva buena y perfecta viene de arriba”. Allí se afirma no sólo que toda dádiva buena y perfecta viene del “Padre de las luces”, sino que la dádiva buena y perfecta “es aquello por lo cual doy gracias a

42 Cfr. UN, 766. (SKS 7, 510).

43 Cfr. UN, 766. (SKS 7, 510).

44 UN, 811. (SKS 7, 545).

45 En principio parecería paradójal recurrir a discursos cristianos para intentar explicitar de modo filosófico el significado concreto y vital de la religiosidad B. Sin embargo, por este camino no hago sino avanzar en una huella avistada por H. Deuser. El estudioso, luego de afirmar que Kierkegaard abre en los discursos “una comunicación religiosa propia”, observa que ésta “es establecida en el contexto cristiano, pero su significación religiosa implica, a la vez, el hecho de que la mediación vitalmente útil de lo (religioso) incondicionado es posible y realizable”. Y lo es para todo hombre en tanta tal. De allí que la religiosidad descrita en los Discursos no tenga por qué reducirse al cristianismo confesional sino que ella, continúa Deuser, “caracteriza y trasunta un aspecto creativo constitutivo, aunque a la vez no-empírico e inasible, de la existencia humana y lo presenta en su efectividad real”. DEUSER, Hermann, “Christologische Motive in den *Christlichen Reden* (dritte und vierte Abteilung)”. En: CAPPELØRN, Niels Jørgen, DEUSER, Hermann y SÖDERQUIST, Brian K. (Eds.), *Kierkegaard Studies. Yearbook 2007*. Berlín/ New York: Walter de Gruyter, 2007, 293-317; aquí: 316. En una nota al trabajo que acabamos de citar Deuser señala, además, que establecer “en qué medida las determinaciones de Kierkegaard de los *estadios* y de la *religiosidad A/B* coinciden con este concepto de comunicación religiosa real debe ser constatado por separado”. *Loc. cit.* (nota 82). El presente trabajo representa precisamente un primer y provisional intento en tal dirección, a saber, establecer una relación posible entre las religiosidades A y B y la comunicación religiosa real de los *Discursos edificantes*.

Dios”⁴⁶. ¿En qué medida es relevante esta afirmación en el contexto que nos ocupa? Ante todo recordemos el modo en que se caracteriza al verdadero cristiano, cuando se opone su figura a la del pagano irreligioso. De este verdadero cristiano, que sin dificultad podríamos asimilar al individuo que existe acorde con la religiosidad B, no se dice que anda por ahí dando voces y declarando que ha aceptado a Jesús como su señor y salvador, sino algo quizás mucho más profundo y seguro mucho más difícil. De él se dice que “vigila y vigila sin cesar la voluntad de Dios, [pues] sólo desea *contentarse con la gracia divina*”⁴⁷. Aquel que, como los lirios y los pájaros, no quiere otra cosa que el cumplimiento de la voluntad divina, no le interesa ni que se realicen sus deseos ni pide que Dios “le ampare de un modo particular distinto del querido por Él”⁴⁸. Lo único que pide es la gracia que Dios tenga a bien enviarle. De allí que Kierkegaard insista en que el cristiano “lo primero que tiene que aprender es a contentarse con la gracia divina”⁴⁹. Ser cristiano, vivir acorde con la religiosidad B, significa, entonces, aceptar la gracia. Y hacerlo es lo mejor que puede hacer el hombre, pues, al fin y al cabo, “toda dádiva buena y perfecta viene de arriba”. ¿Cuál es esta dádiva? Kierkegaard responde que, cuando al cristiano se le preguntá por está dádiva perfecta, él responde “aquello por lo que doy gracias”. Y agrega que el verdadero cristiano “debe agradecerlo todo”. ¿No hay aquí una suerte de círculo? ¿No se determina el agradecimiento por la dádiva perfecta y la dádiva perfecta por el agradecimiento? Y, si hay que agradecerlo todo, ¿no termina siendo todo una dádiva perfecta? De esta perplejidad sólo se sale si entendemos correctamente la idea de que “toda dádiva buena y perfecta viene de arriba”. Dios no da algo que uno pueda identificar como bueno y perfecto por tal y cual razón, por lo que por ello habría que dar gracias. Lo que Dios da, esto es, el modo en que Dios deja que suceda lo que sucede, es bueno y perfecto *por haberlo dado Él*, aun cuando la realidad vaya en contra de la representación humana acerca de lo que sería lo mejor. Por ello insiste Kierkegaard en que, quien ha interpretado cabalmente la sentencia evangélica, no se angustió por determinar qué es lo bueno que viene de Dios y por lo cual habría que dar gracias, sino tan sólo llegó a “aprender una sola cosa: dar siempre gracias a Dios, y de esa manera llegar a entender una sola cosa: que todas las cosas redundan en el bien de quienes aman a Dios”⁵⁰. No es el hombre quien sabe ni puede determinar qué es lo mejor, lo en última instancia perfecto para sí mismo y el mundo entero, por lo tanto debe agradecerlo todo, incluso la menesterosidad y los deseos negados. Escribe Kierkegaard: “Y cuando la parte que te fue concedida fue escasa ¿diste

46 DE, 64. (SKS 5, 51).

47 LC, 131. (SKS 10, 73).

48 LC, 131. (SKS 10, 73).

49 LC, 131. (SKS 10, 73-74).

50 DE, 64. (SKS 5, 51).

entonces gracias a Dios? Y cuando la parte que te fue concedida fue sufrimiento, ¿diste entonces gracias a Dios? Y cuando lo que pediste te fue negado, ¿diste entonces gracias a Dios?”⁵¹. El agradecimiento debe, incluso, “agradecer también por el destino”⁵², por amargo que éste fuese.

Ahora bien, ¿qué sentido puede tener desde un punto de vista filosófico una concepción que hace radicar la religiosidad en la aceptación de la gracia, esto es, en agradecer que la realidad fáctica se dé tal cual ella se da por provenir de Dios; en agradecer incluso el destino? Estimo que de los análisis del discurso kierkegaardiano bien puede concluirse que la religiosidad auténtica no significa sojuzgarse a ciertos dogmas, autoridades y ritos. Si ser religioso, incluso ser en un sentido fundamental un verdadero cristiano, significa contentarse con la gracia, entonces la forma consumada de la religiosidad implica aceptar con modestia nuestra universal condición humana; y aceptar, por tanto también, que no sabemos qué es lo mejor, que no podemos determinar decisivamente ni el futuro ni el curso del cosmos y que, tal vez, lo mejor sea doloroso para nosotros. Una aceptación tal sólo puede conducir a cualquier hombre, sea cual fuere su confesión, a reconocer que eso perfecto o absoluto que necesita tiene que *recibirlo* “desde arriba”, esto es, desde un Poder que lo trasciende. Ahora bien, aquello que el hombre recibe “desde arriba” y que resulta articulado por un Poder que sin duda lo trasciende es el conjunto de la realidad que le sale al encuentro: el rostro del cosmos. Frente a ella el hombre religioso toma una decisión. A diferencia del pagano que quiere amoldar la realidad a sus deseos, él se contenta y se conforma con lo que se da, con la gracia, y, por ello, agradece por la realidad tal cual ella se manifiesta y se nos ofrece a cada instante. Semejante agradecimiento sólo es posible sobre la base de la *confianza* en que el Poder infinitamente superior al humano, que hace que la realidad sea, que permite que se conforme y articule tal cual lo hace, y que vuelve posible que haya un hombre como él capaz de experimentarla, *hace siempre lo mejor*. Por ello, para Kierkegaard, tener verdadera fe religiosa, como la que tuvo Moisés cuando se dirigió al Faraón, consciente de que él no podía nada y de que todo había que encomendárselo al Señor, significa “creer que, suceda lo que suceda, sucederá lo mejor”⁵³. Una religiosidad sustentada en una fe como ésta, es decir, una fe inescindible de la gracia y del agradecer la gracia, no es un llamado a la resignación, sino precisamente un llamado a la confianza en que tiene sentido “creerlo amorosamente todo”⁵⁴. Se trata de una fe difícil, pero posible

51 DE, 64. (SKS 5, 51).

52 DE, 68. (SKS 5, 55).

53 DE, 305. (SKS 5, 303).

54 KIERKEGAARD, Søren, *Las obras del amor*. (Trad. D. Gutiérrez Rivero). Salamanca: Sígueme, 2006, 283. (SKS 9, 236).

para todo hombre con la condición de que cada uno tome la decisión singular de aceptar que, finito como es, no puede saberlo ni controlarlo todo, y de que, por tanto, habrá de agradecer, es decir, asumir y comprometerse con la facticidad que le ha tocado en suerte. Entonces habrá también de *ajustar* su propia existencia a esa facticidad. ¿De qué modo? De modo tal que, existiendo, contribuya, desde las posibilidades fácticas y finitas que le fueron dadas, a que cada ser con el que se relaciona y él mismo puedan, en virtud de esa relación, desplegar cada uno del mejor modo posible su ser. Eso y no otra cosa es “creerlo amorosamente todo”. Amorosamente cree en todo el lirio, cuando se mece al viento y, haciéndolo, torna más hermoso el campo, sin esperar nada a cambio. Amorosamente cree en todo el pájaro, cuando despliega sus alas y vuelve más bello el cielo, sin pedirle al cielo que le devuelva algo. ¿Existir de tal modo, existir agradeciéndolo y amándolo todo, es recibir la dádiva perfecta? ¿No lo amenaza al lirio la siega y al pájaro el cazador? ¿Ocurre en verdad siempre lo mejor a lo que debería ajustarme? ¿Incluso esta plaga, incluso esta escasez? ¿Tiene sentido creerlo amorosamente todo? No hay modo de saberlo. El hombre, en términos absolutos, no sabe ni puede saber qué es lo mejor. Sólo puede confiar en que esto que, en términos relativos (en la temporalidad) no le parece lo mejor, en términos absolutos (en la eternidad) de los que él no dispone, sí lo sea; y, entonces, seguir agradeciéndolo todo y creyendo amorosamente en todo. Sin embargo, la duda no puede ser acallada. Por ello la existencia religiosa no es un estadio definitivo, sino que debe reconquistarse a cada instante. Por ello en la misma persona pueden convivir el pagano y el cristiano. Por ello es tan difícil tener fe y agradecerlo todo. Ante esta dificultad, ¿habrá de convenirse que la fe sólo puede ser una fe ciega que no tiene ningún otro motivo que el absurdo para la esperanza? No creo que éste sea el único camino para entender a Kierkegaard, más aún luego de la lectura de los discursos. La fe en que, suceda lo que suceda, sucederá lo mejor; la fe que lleva a la aceptación de la gracia como criterio que distingue a la religiosidad consumada, no tiene pruebas definitivas, pero sí tiene aquellos mismos motivos para correr el riesgo de optar por ella que los que Kierkegaard señala para alegrarse cada día de ser el hombre que nos *ha sido dado ser*:

“¿Acaso no será tampoco ningún motivo de alegría el que hayas nacido, que existas, que consigas «hoy» lo necesario para subsistir; que hayas nacido hombre, que puedas ver, oír, oler, gustar, tocar? ¿Que el sol brille para ti, y que por ti, cuando el sol se cansa, aparezca la luna y se enciendan las estrellas?”⁵⁵.

La fe, entonces, no surge del absurdo arbitrario. Surge del agradecimiento conmovido ante el milagro de la vida y del universo. Comprenderlo nos ayudaría

55 LC, 192. (SKS 11, 43).

a vivir de modo fructífero en el día de hoy todo lo que hoy podamos vivir (*ajustarnos*), en vez de convertir el día de hoy –el único del que a fin de cuentas disponemos– en la espera ansiosa y angustiada del pagano por el ilusorio día de mañana.

4. A MODO DE CONCLUSIÓN: LA GRACIA Y EL COSMOS⁵⁶

Si la primera forma de paganismo podía definirse como existir en el mundo sin Dios, la relación religiosa de quien se contenta con la gracia no es existir con Dios sin el mundo, sino *existir en el mundo con Dios*. No otra cosa es ajustarse a la situación fáctica que nos es dada, con la convicción interior –la fe, comprendida como confianza y no como conocimiento de lo sobrenatural– de que lo que se da y nuestro rol en lo que se da forma parte de un sentido absoluto, por más que no podamos determinar ni conocer ese sentido. A diferencia del pagano, el hombre verdaderamente religioso no se quiere salvar a sí mismo en el mundo; y, a diferencia del fanático, no quiere salvar al mundo por sí mismo, sino que, contento con la *gracia*, se dirige humildemente al mundo que le es dado para existir en él del modo más provechoso que le fuese posible. “Entonces se entrega a su trabajo, contento con él, ama a su esposa, contento con ella, educa a sus hijos que son su alegría, ama a sus semejantes, se goza con la vida. Y si se exige más de él, Dios [a través de la propia realidad que le sale al encuentro A G- M] se lo dará a entender a las claras, y en tal caso Él lo ayudará”⁵⁷. Vivir con Dios en el mundo significa, pues, contentarse con el milagro de ser hombre y de poder participar, con humildad pero activamente, en la tarea continua de la creación y recreación del universo.

Para ilustrar el sentido cosmológico que tiene, para una perspectiva no confesional, contentarse con la gracia resulta paradójicamente significativo el que quizás sea el paradigma del fenómeno en el cristianismo, a saber, el concederle a una sencilla joven de Galilea ni más ni menos que la gracia de convertirse en la madre de Dios. No dejan de llamar la atención las palabras con las que el ángel Gabriel

⁵⁶ Recalco aquí lo que ya había sido anticipado en la introducción: las presentes conclusiones, al igual que la determinación de los criterios distintivos entre las formas fundamentales del paganismo, no resultan deducciones inevitables de la obra de Kierkegaard, ni pretenden ser la exposición de lo que Kierkegaard realmente dijo en los *Discursos Edificantes* ni, menos aún, la exposición literal de lo afirmado por Climacus en el *Postscriptum*. Ellas son el fruto de una interpretación personal surgida de la puesta en diálogo libre de ese grupo de obras. Lo cual me parece legítimo en un artículo que, como éste, no quiere ser exegético, sino filosófico-hermenéutico.

⁵⁷ KIERKEGAARD, Søren, *Ejercitación del cristianismo*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2009, 88. (SKS 12, 79).

anuncia este acontecimiento: “Alégrate María, llena eres de gracia”. (Lc 1, 28). La gracia, entonces, más que una carga o un sufrimiento es, en el sentido profundo, una apelación a la alegría. María Benedetta Curi observa que el ángel formula esta apelación “en la dimensión verbal del imperativo, como si fuese el mandamiento de alegrarse por un don absolutamente gratuito”⁵⁸. Con este paradigma cristiano en mente, volvamos por un momento a la concepción de la gracia que operó como trasfondo del presente ejercicio hermenéutico. Hemos dicho aquí que ella se le ofrece a todos los seres que viven sobre la tierra y bajo el cielo, en cuanto consiste, para una visión no confesional, en la posibilidad que *recibimos* y en modo alguno originamos, de que nuestro “ser-en-el-mundo” tenga un sentido para el devenir de todo lo que es (observemos que no otra cosa ha sido la gracia ofrecida a María). Consecuentemente, aquello a través de lo cual la gracia se ofrece es el mundo, es decir, el cómo del acontecer del conjunto de la realidad que nos sale al encuentro. Ahora bien, puede decirse que este ofrecimiento por cierto se produce en imperativo, porque la gracia implica una obligación. Se trata de una obligación que naturalmente puede ser desobedecida, pero, como en el caso de cualquier otro imperativo, esto trae como consecuencia un castigo. El castigo es la insignificancia y la esterilidad de la vida. El sinsentido y la consecuente búsqueda desesperada del pagano por una justificación para su existencia. Su obediencia, por el contrario, implica la alegría de una vida con sentido, esto es, de una vida que sí puede realizarse a sí misma, toda vez que, como dice Anti-Climacus en las palabras finales de *La Enfermedad mortal*, ella “se apoya lúcidamente en el Poder que la fundamenta”⁵⁹. No en su voluntad, sino en el Poder que le ha dado la realidad efectiva que le toca en suerte. Según la fe cristiana, María comprendió perfectamente lo que los tiempos exigían de ella y asumió esta exigencia. Esto transformó su vida por completo y la transformación permitió que se pusiera al servicio del acontecimiento del nacimiento de Cristo. El ejemplo de María muestra que la aceptación de la gracia implica siempre una transformación tal de la propia existencia, que la lleva a ser fecunda, a convertirse en pro-genitora de una nueva emergencia del ser y del sentido. Sin embargo, este sentido que nace o emerge ya no es sólo una cuestión mía, como el nacimiento de Cristo no fue tan sólo una cuestión que afectó a María, sino que concierne al mundo todo. El sentido del universo, aunque tan sólo se trate de su más modesto destello, puede acontecer, por un instante, también a través mío, si asumo la gracia que hoy me es ofrecida

58 CURI, Maria Benedetta, “Nella grazia di un nuovo pensare. La vocazione mariale dell’ontologia trinitaria”. *Quaderni di Inshibboleth*, N. 1 (2021), 129-144; aquí: 130.

59 KIERKEGAARD, Søren, *La enfermedad mortal*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2008, 168. (SKS 11, 242).

y la exigencia que ahora ella implica. ¿Qué otra razón más profunda puede haber para alegrarse?

REFERENCIAS BIBLIOGRÁFICAS

- BANSER, Ann-Kathrin/ BODE, Philipp, *Selbstwerden. Über das Selbst als Aufgabe und die Möglichkeiten seiner Realisierung bei Søren Kierkegaard*. Würzburg: Königshausen&Neumann, 2018.
- BJERGSØ, Michael, *Kierkegaards deiktische Theologie. Gottesverhältnis und Religiosität in den erbaulichen Reden*. (Trad. Alemana de Krista-Maria y Hermann Deuser). Berlin/ New York: Walter de Gruyter, 2009.
- CURI, Maria Benedetta, "Nella grazia di un nuovo pensare. La vocazione mariale dell'ontologia trinitaria". *Quaderni di Inshibboleth*, N. 1 (2021), 129-144.
- DE GRAMONT Jérôme, "Kierkegaard. Pequeña fenomenología de la donación". *Open Insight*, Vol 5 (2014), 79-192.
- DEUSER, Hermann, "Christologische Motive in den *Christlichen Reden* (dritte und vierte Abteilung)". En: CAPPELØRN, Niels Jørgen, DEUSER, Hermann y SÖDERQUIST, Brian K. (Eds.), *Kierkegaard Studies. Yearbook 2007*. Berlín/ New York: Walter de Gruyter, 2007, 293-317.
- HARRITS, Fleming, "Wortwortlichkeit des Geistes. Über Søren Kierkegaards Sprachverständnis in den Reden von 1843-1845". En: CAPPELØRN, Niels Jørgen y DEUSER, Hermann (Eds.), *Kierkegaard Studies. Yearbook 2000*. Berlin/ New York: Walter de Gruyter, 2000, 121-134.
- HEYMEL, Michael/ MÖLLER, Christian, *Das Wagnis, ein Einzelner zu sein. Glauben und Denken Søren Kierkegaards am Beispiel seiner Reden*. Zürich: Theologischer Verlag, 2013.
- KIERKEGAARD, Søren, *Søren Kierkegaards Skrifter*. (Edición del Søren Kierkegaard Forskningscenter). Copenhague: Gad, 1994-2008.
- , *Las obras del amor*. (Trad. D. Gutiérrez Rivero). Salamanca: Sígueme, 2006.
- , *Los lirios del campo y las aves del cielo*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2007.
- , *La enfermedad mortal*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2008.
- , *Ejercitación del cristianismo*. (Trad. D. Gutiérrez Rivero). Madrid: Trotta, 2009.
- , *Escritos de Søren Kierkegaard. Volumen 5: Discursos edificantes. Tres discursos para ocasiones supuestas*. (Trad. D. González). Madrid: Trotta, 2010.
- , *Philosophische Brosamen und Unwissenschaftliche Nachschrift*. 4ª ed. (Trad. alemana de B. y S. Diderichsen). München: Deutscher Taschenbuch Verlag, 2015.
- PIEPER, Annemarie, *Søren Kierkegaard*. München: Beck, 2014.
- PIEPER, Annemarie (Ed.), *Kierkegaard: Die Hauptwerke. Ein Lesebuch*. Tübingen: Narr Francke Attempto, 2018.
- PURKARTHOFER, Richard, *Kierkegaard*. Stuttgart: Reclam, 2005.

THONHAUSER, Gerhard, *Über das Konzept der Zeitlichkeit bei Søren Kierkegaard mit ständigem Hinblick auf Martin Heidegger*. Freiburg/ München: Alber, 2011.

VIÑAS VERA, Ángel, "Notas sobre el paganismo en Kierkegaard". *Anales del Seminario de Historia de la Filosofía*, 37 (2), 2020, 259-268.

ESCRIBIR FILOSOFÍA: KIERKEGAARD Y SU *FORORD*

WRITING PHILOSOPHY: KIERKEGAARD AND HIS FORORD

ÁNGEL VIÑAS VERA

Doctor en Filosofía
Profesor Asistente
Universidad Loyola Andalucía,
Sevilla/España
avin@uloyola.es
ORCID: 0000-0002-6793-2748

Recibido: 19/04/2021

Revisado: 16/08/2021

Aceptado: 6/09/2021

Resumen: Kierkegaard lanzó al mundo, bajo el pseudónimo de Nicolaus Notabene, el libro *Forord*. Ahí podemos encontrar una reflexión rica sobre la tarea de escribir y hacer filosofía. Lo que aquí pretendemos es, estudiando este libro, mostrar la idea de filosofía que el autor danés nos expresa así como el temple, o estado de ánimo con el que se debe escribir filosofía. Además, pondremos en relación este peculiar libro con otras reflexiones sobre los prefacios en Nietzsche y Foucault. Al hacerlo, reivindicaremos una manera de escribir filosofía que no acabe con ella misma, hoy amenazada no sólo por los índices de calidad objetiva o por los gobiernos de turno.

Palabras clave: Kierkegaard, prefacio, filosofía, Stemning, escribir, leer.

Abstract: Kierkegaard launched the book *Forord* into the world under the pseudonym Nicolaus Notabene. There we can find a rich reflection on the task of writing and doing philosophy. What we intend to do here is, by studying this book, to show the idea of philosophy that the Danish author expresses as well as the temper, or state of mind with which one should write philosophy. Furthermore, we will relate this peculiar book to other reflections on prefaces in Nietzsche and Foucault. In doing so, we will claim a way of writing philosophy that does not do away with philosophy itself, today threatened not only by objective quality indices or by the governments of the day.

Keywords: Kierkegaard, preface, philosophy, Stemning, writing, reading.

INTRODUCCIÓN

En el año 1844, en la librería Reitzel salía un libro con un título peculiar: *Forord*¹, es decir, *Prólogos*. El autor que lo firmaba también parecía singular, Nicolaus Notabene. El subtítulo de este indicaba también la naturaleza del texto: *Lectura ligera para ciertos estamentos según tiempo y ocasión*. Todo resultaba realmente llamativo, desde el título y el contenido donde aparecen ocho prólogos precedidos de un prólogo, pasando por el subtítulo que nos habla de una lectura ligera, de entretenimiento, popular, nada científica, y, por supuesto, el nombre del que la escribe. Todo apunta a un texto que no va a pasar a los anales de la filosofía. En todo caso, quedará como una nota breve y algo irónica sobre la tarea de escribir y la filosofía misma.

El objetivo de este texto es mostrar el aporte de este libro tan peculiar, y no de los más estudiados de Kierkegaard², en el contexto actual de la filosofía. Hoy en día contemplamos atónitos al espectáculo, no siempre edificante, de algunos autores buscando editoriales y empresarios buscando financiación para publicar revistas. A veces, encontramos escritores que seleccionan editoriales y que se autocensuran porque esas publicaciones no tienen impacto social, así como editoriales que tienen que aplicar filtros porque se acumulan más y más artículos pendientes. No siempre nos encontramos este espectáculo y hay editoriales, autores, revisores y empresarios que invierten tiempo, dinero y esfuerzo en seguir traduciendo, publicando ensayos e investigaciones. La ANECA³ contempla y anima este espectáculo y los filósofos debemos seguir preguntándonos qué sentido tiene escribir filosofía. El texto de Kierkegaard aporta indicaciones sobre cómo afrontar estas tareas. El irónico pensador danés nos ayudará a cómo situarnos

1 KIERKEGAARD, S., *Søren Kierkegaards Skrifter*, Gads Forlag, 1997-2013, tomo 4. A partir de ahora se citará SKS seguido del tomo y la página correspondiente. También se tienen los escritos de Kierkegaard en edición digital, <http://sks.dk/forside/indhold.asp>. Este libro, en concreto, está en SKS 4, 365ss. También cotejamos la edición francesa, Kierkegaard, Søren *Œuvres complètes*, Paris: Éditions de l'Orante, 1966-1986, en el tomo VII, 259-326 (A partir de ahora OC seguido del número romano del tomo y la página). La edición española es *Migajas filosóficas. El concepto de angustia. Prólogos*, Madrid: Trotta, 2016, 279-350. La citaremos después de la cita danesa indicando el título y la página.

2 Uno de los grandes estudios consagrados a este libro es PERKINS, R.L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions (International Kierkegaard Commentary)*, Georgia: Mercer University Press, 2007. Aquí Podemos leer: "Prefaces is both the shortest text and the one least attended to in the reception of the pseudonymous works of Søren Kierkegaard.", Lasse Horne Kjældgaard, "The Age of Miscellaneous Announcements: Paratextualism in Kierkegaard's *Prefaces* and Contemporary Literary Culture", 7.

3 Cf. El editorial de *Razón y Fe*, 2014, 269, 1383, 5-13. Este texto, elaborado por la profesora Olga Belmonte ayuda a contextualizar el momento actual que estamos viviendo con un análisis riguroso y crítico.

ante el sistema, de ayer y de hoy. Al traer a nuestro debate actual este libro tan especial, que conecta como veremos con otros autores que han abordado la singularidad de los prefacios, se quiere apuntalar la necesaria revisión en profundidad de la tarea urgente, necesaria y crítica de hacer filosofía.

Comenzaremos situando el texto en la producción kierkegaardiana e indicando otros textos relevantes de la filosofía consagrados a los prefacios. Seguiremos abordando el contenido del libro del pensador danés. Al hacer esto mostraremos cómo entiende el prefacio N.N. y cómo ironiza sobre el sentido de escribir y leer filosofía. Heiberg, Hegel y demás pensadores irán apareciendo en sus análisis de manera directa o velada⁴. Todo para intentar, tanto por la forma de escribir como por el contenido de lo escrito, que otra manera de hacer filosofía sea posible, porque son posibles otras maneras de ser filósofo. Lo que está en juego no es la evaluación de la filosofía, sino el que dejemos de hacerla los que nos dedicamos a ella. Aquí también vale la máxima socrática. Lo malo en su sentido eminente no es lo que nos hacen, sino el mal que hacemos. Los filósofos, o al menos los que deseamos serlo, debemos seguir pensando lo que se debe y no podemos dejarnos enredar en mercados, en lógicas del marketing o de cualquier sistema político o ideológico.

1. SOBRE PREFACIOS Y PREFACIOS

Es necesario recordar, para situar el texto que nos ocupa, que en el mismo año 1844, Kierkegaard había lanzado al mundo bajo sus pseudónimos otros libros: *Migajas filosóficas*⁵ y el mismo día que salió *Prólogos, El concepto de angustia*⁶. Además, con su propia firma, había entregado al mundo nueve Discursos edificantes. Sabemos, además, que el citado autor no paró de escribir miles de páginas a pluma en esta época.⁷ Es bueno consignar este elemento porque una de las cuestiones que Notabene se planteará en sus prólogos es por qué escribir⁸.

4 BRUN, J. "Introduction", OC 7, XXVII-XXIX donde muestra los autores con los que dialoga Kierkegaard en este libro.

5 SKS 4, 213ss. Johannes Climacus, *Philosophiske Smuler, Migajas Filosóficas*, 26ss.

6 SKS 4, 309ss. Vigilius Haufniensis, *Begrebet Angest, El concepto de angustia*, 125ss

7 Cf. CARLISLE, C. *El filósofo del corazón. La inquieta vida de Søren Kierkegaard*, Barcelona: Taurus, 2021, en especial para ver el contexto vital de estos años tan prolíficos, el capítulo 11 titulado "Cómo angustiarse", 219-235.

8 "To be an author, therefore, is to dig deep into one's own inner struggles, deliberated upon without regard for publishing markets or public opinion.", CRITES, S. "The Unfathomable Stupidity of Nicolaus Notabene". En: PERKINS, R. L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions (International Kierkegaard Commentary)*, Mercer University Press, 2007, 30.

Otra acotación se hace necesaria consignar: la referencia que Kierkegaard hará a este libro dos años después en su *Post-scriptum no científico y definitivo a 'Migajas filosóficas'*. Aquí se da una ojeada a la literatura danesa que se ha venido publicando en los últimos años. Johannes Climacus nos dice que *Prólogos* es “en lystig lille Bog” (SKS 7, 245)⁹, es decir, un gracioso librito. Este librito es alegre, con buen humor, divertido que además es popular y que puede venir bien en alguna ocasión o tiempo.

Antes de abordar el contenido de este libro, es necesario subrayar algunos elementos más del contexto de la autoría kierkegaardiana mostrando el aporte de este libro. La pregunta qué es la filosofía y cómo escribirla está presente desde su tesis consagrada a la ironía y a Sócrates. La primera gran obra, *O lo uno o lo otro*¹⁰ es un libro que recoge múltiples y minuciosos análisis de lo estético y lo ético. Esta obra es un ejemplo de lo decisivo que es el cómo a la hora de escribir filosofía: no se puede escribir igual si se narra la vivencia del enamorado, de las damas de la pena o de la comunidad de los “muertos”, o la vida de un magistrado y casado. El lector es puesto en una situación peculiar porque debe evitar juzgar lo que se entrega en la vida estética con las vivencias éticas. Y si no fuera suficiente con los matices y diversas experiencias vitales que se nos narran en las dos primeras partes de esta obra, se añade al final un escrito consagrado a un tema religioso donde se nos indica que sólo la verdad que edifica es verdad. Esta obra está escrita teniendo en cuenta a sus múltiples y posibles lectores. No tanto para decirles lo que quieren escuchar o ajustando el discurso a la tiranía del público, sino por insinuar caminos de vida en función al momento existencial que viva cada uno. ¿Se le debe escribir filosofía de la misma manera a un joven enamorado, a una mujer apenada porque su amor ha fracasado, al que está disfrutando de los placeres musicales, al que está casado y está sacando su matrimonio adelante, etc.? Kierkegaard, desde el principio de su autoría, sabe que es decisivo lo que se escribe, pero es capital el cómo. Nuestro libro, *Forord*, es una continuación de este camino. Su principal aporte es el diálogo con un interlocutor nuevo: lo escritores, editoriales, censores, y el mundo de la escritura. Su segundo aporte es la relación entre escribir y estado de ánimo, como después veremos. Las intuiciones que aparecen *Forord* tendrán continuidad en *Post-Scriptum* al mostrar que la verdad es subjetividad, entre otras cosas, porque “la verdad es la incertidumbre objetiva sostenida en la apropiación de la interioridad más apasionada; tal es la más excelsa verdad

9 SKS 7, 245. Johannes Climacus, *Afsluttende uvidenskabelig Efterskrift til de filosofiske Smuler. Post Scriptum No científico y definitivo a «Migajas filosóficas»*, Trotta, 2010, 269.

10 SKS 2, Victor Eremita *Enten Eller. O lo uno o lo otro*. Trotta, Madrid, 2006.

que hay para un existente.” (SKS, 7, 186)¹¹. ¿Cuál es el modo de existir que nos puede ayudar a alcanzar más y mejor la verdad, las verdades más profundas sobre el ser humano, Dios y el mundo? Kierkegaard nos ha mostrado con su autoría que es imprescindible una filosofía segunda para poder alcanzar las grandes cuestiones esenciales de la filosofía primera. El análisis de la subjetividad, y en concreto los análisis sobre cuál es el modo de escribir y ejercer filosofía, es un tema central de estos libros. Veamos lo que nos aporta el libro editado por Notabene.

Lo primero que quisiera subrayar es que este libro de prefacios viene precedido de otro prefacio, de un prefacio a los prefacios. ¿Consideramos a este prefacio del libro un traidor a los demás prefacios liberados de libros? ¿Qué hace N.N. en este prefacio tan peculiar? Veamos lo que nos dice el prefacio de este libro.

Él hace bastantes referencias en su texto al contexto de su época. En relación a los prefacios nos dice lo siguiente: “En la ciencia moderna el prólogo ha sido herido de muerte.”¹² Notabene nos deja caer que la ciencia actual no parece precisar del prefacio o prólogo. Posiblemente, como así ha sido recogido en los comentarios daneses a su obra¹³, se refiere al archiconocido prólogo que aquí citamos:

Una explicación, tal como es costumbre darla en el prefacio de un libro sobre el fin que el autor se propone en él, así como sobre los motivos y sobre la relación que el autor cree que ese libro guarda con otras formas anteriores o contemporáneas de tratar ese mismo objeto, no sólo parece superflua cuando se trata de un libro de filosofía, sino que a causa de la naturaleza de la cosa parece incluso inadecuada y contraproducente. Pues lo que en un prefacio fuese pertinente decir acerca de la filosofía, como por ejemplo el *contar* en qué corriente se sitúan el libro y el punto de vista que en él se adopta, o el *explicar* cuál es el contenido en general y cuáles sus resultados, lo que no sería sino un trabar afirmaciones o aseveraciones acerca de lo verdadero, que hablarían desde este o aquel lugar, eso, digo, no podría considerarse el modo y manera como hay que exponer la verdad filosófica¹⁴.

11 SKS 7, 245. Johannes Climacus, *Afsluttende uvidenskabelig Efterskrift til de filosofiske Smuler. Post Scriptum No científico y definitivo a «Migajas filosóficas»*, Trotta, 2010, 205.

12 SKS 4, 468. Nicolaus Notabene, *Forord. Prólogos*, p. 282.

13 Cf. Søren Kierkegaards *Skrifter Kommentarer til Gjentagelsen, Frygt og Bæven, Filosofiske Smuler, Begrebet Angest, Forord*, Gads Forlag, 2012, pp. 531-627, en especial el comentario desde la página 565. Comentario realizado por Johnny Kondrup. Se recogen ideas de este comentario en la “Introducción” del libro *Migajas filosóficas. El concepto de angustia. Prólogos*, Editorial Trotta, 2016, en especial las páginas 20-24. Así mismo, y con algo más de detalle, lo retoma PARCERO, O., “La emancipación del Prólogo en la *Fenomenología del espíritu*”, *Anales del Seminario de Historia de la Filosofía*, 561, 33. 2, 2016, 561-582.

14 HEGEL, G. W., *Fenomenología del Espíritu*, Valencia: Pre-Textos, 2006, 111. En el año 1807, como es sabido, salió al mundo este libro que marcó la historia de la filosofía.

Este prólogo de Hegel, así como los escritos de Heiberg, están en el trasfondo de la escritura de N.N.¹⁵. En este texto citado, Hegel parece mostrar que el prólogo no es la mejor manera de exponer la verdad filosófica, que el prólogo y filosofía no casan bien. Además, se considera por parte de los que hacen la ciencia filosófica, que los pensadores anteriores con sus prefacios mueven a risa o a lástima. Quizás este libro de Notabene podría incluirse en esta categoría. Esto nos va remitiendo a que habría que distinguir entre prólogos y prólogos, como nos sugiere el texto kierkegaardiano. Hay algunos que son superfluos, insignificantes para el desarrollo del libro. Éste se escribe y posteriormente se decide si se le hace un prólogo. Éste no cumple una función más que de súbdito del libro y sin relevancia. Además, si el libro quiere comenzar por el principio sin principio, o por la nada, no sería muy adecuado que en su marcha hubiera un texto introductorio cuya naturaleza sería más bien extraña o inquietante con lo que se pretendiera realizar. De ahí que Notabene nos proponga la posibilidad de hablar de otro tipo de prólogos: los emancipados.

Ahora bien, si también sin tema se puede tener ganas de escribir un prólogo, entonces se ve fácilmente que no es necesario que este trate sobre algún tema, pues en tal caso el propio prólogo se convierte en libro, y la cuestión sobre el prólogo y el libro queda desplazada. Un prólogo así, el prólogo emancipado, no debe pues tener tema alguno sobre el que tratar, sino tratar sobre nada, y en la medida en que parezca ocuparse de algo y tratar sobre algo, debe ser una ilusión y un movimiento fingido¹⁶.

Este tipo de prólogos nos dejan algo perplejos. ¿Qué nos quiere decir al afirmar que no trata sobre un tema concreto o que puede ser un movimiento fingido? ¿Qué es, pues, un prólogo si su determinación no viniera del tema que desarrolla? Él nos muestra una intuición, una sospecha: “Et Forord er Stemning”¹⁷, es decir, un prólogo es un estado de ánimo, un temple, una disposición, una manera, un cómo más que un qué. Escribir un prólogo así es como “llamar a alguien a su puerta para gastar una broma, como pasar por delante de la ventana de una muchacha y mirar para los adoquines, es como dar bastonazos al aire o al viento...”¹⁸. El prólogo es una insignificancia que puede dejarnos entrever la manera de ser de quien escribe, o la manera de concebir la escritura de quien lo firma. ¿Cómo sería el que escribe prólogos emancipados? “Él va y viene entre la gente como un bufón en invierno y un payaso en verano, es un hola y adiós en una

15 Para un estudio más detallado de la diatriba con Heiberg pueden verse, además de los textos citados, STEWART, J., “The Polemic with Heiberg in Prefaces”, en *Kierkegaard's Relations to Hegel Reconsidered*, Cambridge: Cambridge University Press, 2003, 419-447.

16 SKS 4, 469. Nicolaus Notabene, *Forord. Prólogos*, 282.

17 SKS 4, 469. Nicolaus Notabene, *Forord. Prólogos*, 283.

18 *Ibid.*

persona, siempre feliz y despreocupado, contento consigo mismo, todo un frívolo inútil, sí, una persona inmoral.”¹⁹ El que escribe puede ser considerado inmoral, pero igual nos ayuda a redescubrir la tarea tan singular de escribir. Porque una de las cosas que igual no hemos pensado suficientemente es en qué consiste ésta, por qué escribimos, cómo escribimos y lo que en ella se muestra de nuestra manera de entender la existencia. No hablamos de cualquier escribir, sino de este escribir que quiere abordar los misterios de la vida y el ser humano. Este es el tema de este artículo.

Termino esta acotación sobre el sentido del prólogo o prefacio en este libro con un dato curioso. La mujer de Notabene le prohibió que escribiera libros, como se nos dice en el prólogo de este libro²⁰. ¿Por qué? “Ser escritor cuando uno está casado, dice ella, es una infidelidad manifiesta, todo lo contrario de lo que dijo el pastor.”²¹ Ella llegó hasta a quemarle manuscritos y notas, porque veía que su marido le podía ser infiel con la escritura. Llegaron a un pacto: él podía escribir prólogos. Ella consideraba que si le prohibía escribir libros, tampoco podría llegar a escribir prólogos porque van unidos. Él comprendió, sin embargo, que podrían escribirse prólogos sin necesidad de escribir el libro. Es lo que hizo. Escribir ocho prólogos.

19 SKS 4, 470. Nicolaus Notabene, *Forord. Prólogos*, 283. ¿Por qué es inmoral, nos podemos preguntar? “No va a la Bolsa a hacer rentable su dinero sino que se limita a pasar por delante, no habla en asambleas generales porque el aire está demasiado encarcelado, no propone un brindis en ninguna sociedad porque ha de anunciarse con varios días de antelación; no rinde honores al Sistema; no cotiza a la deuda pública, y ni siquiera la toma en serio; va por la vida como un aprendiz de zapatero silbando por la calle, aunque el que necesite las botas esté esperando por ella.” Notabene puede que nos esté hablando indirectamente de un hombre que paseaba asiduamente por el puente de Knippelsbro. En uno de los extremos de ese puente está la Bolsa, en la calle de su propio nombre *Børsgade*. Kierkegaard se dibujó a sí mismo en el citado puente en una carta a Regina Olsen, su amor. Por este puente, y por la Bolsa, paseaba muchas veces camino a la casa de los Olsen y también de camino a donde hizo su examen en homilética el 12 de enero de 1841, en la Holmens Kirke. La imagen a la que nos referimos está en SKS 28, 215. *Breve, Dedikationer*.

20 Robert L. Perkins nos habla de la posible relación entre, por un lado Jantipa-Sócrates, y la mujer de N.N., por otra. Esta parodia, esta batalla de sexos es otro ingrediente más de la ironía y lo socrático en *Prefacios*. Cfr. Perkins, R.L. “Reading Kierkegaard’s *Prefaces* with “Continual Reference to Socrates”. En: PERKINS R. L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions (International Kierkegaard Commentary)*, Georgia: Mercer University Press, 2007, 115.

21 SKS 4, 474. Nicolaus Notabene, *Forord. Prólogos*, 287.

2. SOBRE OTROS PREFACIOS LUMINOSOS: NIETZSCHE Y FOUCAULT

Dejemos a un lado la diferenciación técnica, si fueran realmente relevante, entre prefacios y prólogos. Lo singular de un prefacio no es fácil de describir. Como bien nos indica Cossutta: “El prefacio es un espacio intermedio entre la exterioridad y la interioridad”²². El prefacio se relaciona con el resto del libro, con el mundo exterior, con las lógicas de otros prefacios del mismo autor, con la relevancia que se le de en ese momento.

Hay prefacios aparentemente irrelevantes que lo que hacen es sintetizar lo que viene detrás o simplemente enunciar algún tema central del libro. Los prefacios de Kierkegaard no forman parte de este grupo. Ni siquiera los tan enjundiosos prefacios de sus discursos edificantes o de su obra pseudónima. En el libro que nos ocupa no ha lugar un prefacio sometido al texto posterior. En ese sentido se parece más a dos tipos de prefacios que me gustaría traer a colación por la cercanía con el tono kierkegaardiano. Uno será el prólogo de Nietzsche en la *La ciencia jovial* y por otro lado el de Foucault en su segunda edición de la *Historia de la locura*²³. ¿Por qué traer a estos autores y obras? En primer lugar, porque ellos coinciden con Kierkegaard en una crítica a la posibilidad de escribir un sistema. Los tres autores son críticos con la posibilidad de pensar el todo desde una supuesta racionalidad moderna que todo lo abarca. En segundo lugar, porque en los tres autores hay un compromiso con el ser humano singular que no se debe dejar atrapar por el todo, por la idea que mata a la vida, o lo normalizado y sus discursos de poder. En tercer lugar, porque los tres han reflexionado sobre el prefacio liberado, es decir, sobre la importancia de este. Al hacerlo, no sólo reflexionan sobre el prefacio y su singularidad, sino sobre el escribir y, en concreto, sobre la escritura filosófica. Los tres, dentro de sus semejanzas y diferencias, nos ayudan a seguir pensando el ideal de la filosofía y el oficio de escribir.

El texto de Nietzsche pone en valor la necesidad de encontrar un verdadero lector. Su manera de entender la escritura está íntimamente conectada con su manera de comprender la lectura filosófica. Este prólogo a la segunda edición de *La ciencia jovial* nos indica en su primera frase que “tal vez este libro necesita más de un prólogo”²⁴. La peculiaridad festiva de este libro así lo exige. En el citado prefacio, Nietzsche vincula salud y filosofía incluso llegando a ironizar sobre el

22 COSSUTTA, F., “La préface en philosophie : une approche discursive », *Argumentation et Analyse du Discours*, 22 (2019), 2.

23 Para una relación de los prefacios kierkegaardianos con Descartes, Derrida y Freud, puede verse FERGUSON, H. “Before the Beginning: Kierkegaard’s Literary History”. En: PERKINS, R. L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions (International Kierkegaard Commentary)*, Georgia: Mercer University Press, 2007, 41-65.

24 NIETZSCHE, F. *La ciencia jovial*, Madrid: Gredos, 2010, 309.

papel supuestamente salúfero de su obra: “Pero dejemos al Señor Nietzsche: ¿qué nos importa que el Señor Nietzsche recobre de nuevo la salud?”²⁵. ¿Quién escribe este prefacio? ¿Este desdoble es totalmente ajeno al juego de la pseudonimia de Kierkegaard? Pero lo más interesante de este prefacio, en lo que se refiere al tema que nos ocupa, es la relación que establece entre sentido de la filosofía, escritura y lectura.

Nietzsche se hace cuestión del por qué hacemos filosofía, por qué filosofamos, qué esperamos con esta actividad. Algunos, dirá, filosofan por necesidad, otros por lujo. Pero lo que no podemos hacer es dejar de hacer filosofía. Y esto supone: “Nosotros, los filósofos, no somos libres de separar el cuerpo del alma, como lo hace el pueblo; aún menos libres de separar el alma del espíritu. No somos ranas pensantes ni aparatos de objetivación o de registro, con las entrañas heladas –nosotros continuamente tenemos que parir nuestros pensamientos desde nuestro dolor y proveerles maternalmente de todo cuanto hay en nosotros de sangre, corazón, fuego, placer, pasión, tormento, conciencia, destino, fatalidad”²⁶. El filósofo no puede liberarse de su deber, no puede convertirse en un instrumento objetivador ni de la verdad ni del bien. El filósofo no puede estar helado por dentro, muerto en vida. Su dolor y su pasión internas hacen de parteras de sus pensamientos. El filósofo no es un ser abstracto que piensa la verdad con su razón objetiva al margen de su carne y huesos. Por eso a Nietzsche le repugna que se convierta este fuego y pasión en un mercadeo: “¡Con qué malicia prestamos atención desde ahora al gran ajeteo de la feria anual, con el que el ‘hombre culto’, el habitante actual de la ciudad se deja violentar por medio del arte, de los libros y de la música en su búsqueda de ‘gustos espirituales’ y con la ayuda de bebidas alcohólicas!”²⁷. Estas dos anotaciones hacen que Nietzsche pida un lector diferente. Es cierto que tenemos que encontrar este texto en un fragmento no retenido en la segunda edición de este libro. Nietzsche, como nos indica Cossutta, necesita otro tipo de lector: “Pero el objetivo de este prefacio debe ser otro que aquel de recodar a mis lectores lo que son las virtudes de un lector: buena voluntad, reflexión, atención, comprensión, perspicacia; eso sería pura simulación, si yo me limitara a eso”²⁸. La apelación directa a sus posibles lectores supone no sólo una mirada hacia ellos, sino hacia el escritor. Depende de cómo se entienda la tarea de escritor, se comprende la manera de leer y viceversa. Porque un libro no es un hecho objetivo que tiene una lectura una y única. Aquí es donde aparece

25 *Op.cit.* 310.

26 *Op.cit.* 313.

27 *Op. cit.* 314.

28 COSSUTTA, F., “La préface en philosophie : une approche discursive », *Argumentation et Analyse du Discours*, 22 (2019), 3.

la relevancia de la apropiación, es decir, de una actividad filosófica que nace libre y se despliega libremente. Lo que está en juego también es la libertad, es decir, que el escritor y el lector sean convocados no el uno al otro, ni siquiera al libro, sino a un juego de relaciones abiertas y pendientes de libertades encontradas.

¿Qué nos dice Foucault en su famoso libro sobre la historia de la locura? Todos sabemos que a la segunda edición introdujo un nuevo prólogo. Este texto breve es, como indica Cossutta, “un meta-prefacio anti-prefacios”²⁹. Foucault busca con este prefacio liberar al texto de su tiranía. Algunos piensan que escribir un prólogo o un libro es lo siguiente: “Así se escribe el prólogo, primer acto por el cual empieza a establecerse la monarquía del autor, declaración de tiranía: mi intención debe ser vuestro precepto, plegareis vuestra lectura, vuestros análisis, vuestras críticas, a lo que yo he querido hacer.”³⁰ Es posible escribir un libro queriendo que este diga sólo y exclusivamente lo que el autor le imponga. El prólogo de Foucault quiere liberar al texto del autor, y de la tiranía de los lectores sabios que imponen o usan el texto a su antojo. Frente a una manera de entender la escritura que pasa por considerar al libro como su hijo que debe ser siempre respetado y buscada la intención del escritor como norma absoluta de lectura, él propone en la misma página otra posibilidad:

Quiero que este objeto- acontecimiento, casi imperceptible entre tantos otros, se re-copie, se fragmente, se repita, se imite, se desdoble y finalmente desaparezca sin que aquel a quien le tocó producirlo pueda jamás reivindicar el derecho de ser su amo, de imponer lo que debe decir, ni de decir lo que debe ser. En suma, quiero que un libro no se dé a sí mismo ese estatuto de texto al cual bien sabrán reducirlo la pedagogía y la crítica; pero que no tenga el desparpajo de presentarse como discurso: a la vez batalla y arma, estrategia y choque, lucha y trofeo o herida, coyuntura y vestigios, cita irregular y escena respetable³¹.

Un libro que tenga el derecho sobre sí a ser batalla y arma a la vez. Que no quede reducido a un mero texto inamovible. El prólogo es, en este caso, además de un anti-prólogo, una manera de animar a leer el libro, a estimular al lector, a comprender mejor al autor, y a batallar con el texto como acontecimiento que puede ser para el lector. Como podremos ver en seguida, a la hora de exponer las ideas que Kierkegaard desarrolla en los diversos prefacios de *Forord*, no están muy lejos los planteamientos de estos autores.

La breve cata, debido a los límites y objeto de este estudio, en Nietzsche y Foucault nos indican caminos transitados donde el prefacio es enfatizado como

29 *Op. cit.* 9.

30 FOUCAULT, M., *Historia de la locura I*, México: Fondo de Cultura económica, 1998, 3.

31 *Op. cit.* 3.

un texto peculiar. Más bien habría que hablar que un prefacio puede ser, siguiendo la línea de Foucault, un objeto-acontecimiento. El prefacio no sólo llama a la liberación del texto, como nos ayuda a ver Foucault, sino que en el texto de Kierkegaard será llevado a la máxima expresión al verse desvinculado por completo del libro. En su texto nos encontramos ocho prefacios que no tienen libros. Tampoco podríamos hablar, en propiedad, de prólogos en búsqueda de libros. Aquellos han quedado liberados del libro y hablan por sí mismos. Lo que muchas veces se escribe al final de todo, lo que muchas veces es imperceptible o minusvalorado, cobra fuerza por sí mismo. Y son ellos los que plantearán temas muy filosóficos que tienen mucho que ver con su propia esencia: el origen del filosofar, la posibilidad de que un sistema necesite prólogos, cómo leer y escribir, y por ello, qué significa ser filósofo. Un prefacio es un estado de ánimo no sólo porque él lo refleja o lo anima, lo muestra indirectamente o lo convoca, sino porque es escrito con un estado de ánimo del escritor que invita libremente al del lector. ¿Cuál es el estado de ánimo con el que N.N. escribe estos prefacios? Acometamos el análisis más detallado de los mismos.

3. DIVERSAS MANERAS DE ESCRIBIR FILOSOFÍA

Son muchas las indicaciones e insinuaciones que nos da el libro de Kierkegaard sobre el acto de escribir. Notabene nos habla de ellos en todos los prólogos. En el fondo, nos está indicando que es necesario mirar de nuevo la tarea de escribir porque es un acto cargado de responsabilidad. Un prólogo manifiesta una manera de posicionarse ante el hecho de escribir, el ser escritor y la manera de entender al lector.

N.N. nos habla de diversas maneras de entender la escritura. Una de ellas puede venir marcada por querer tener en cuenta lo que necesita el tiempo presente. El buen escritor sería aquel que sabe dar a su tiempo lo que éste necesita. Así, el escritor juega un papel social muy relevante. Aquel debe ser capaz de conocer la realidad de su tiempo y, desde ahí, entregarle un libro que le ayude. Esto quedaría, evidentemente, reflejado en su prefacio. Para hacer esto con total coherencia es bueno estudiar también cuál es el mejor momento para entregar el libro al público. La ironía y el humor aparecen en cada línea de este primer prefacio. Notabene indica que es una opinión generalizada en Copenhague que el mejor momento para regalar libros es en Año Nuevo. Porque ahí se suelen comprar libros. La escritura, la tarea del escritor, y por ende la del lector, están sometidos a su tiempo. Esto tiene su aporte y su significatividad, su relevancia y utilidad. Ahora bien, según Notabene, queda una incógnita por resolver: ¿cómo saber lo que necesita nuestro tiempo?

Para esa pregunta juega un papel esencial el editor y el crítico o recensor. Gracias a estos, continúa Notabene, el autor puede saber qué escribir, qué se necesita y cuándo es el mejor momento. Ellos le iluminan, gracias a su experiencia y sapiencia. Esto supone que el crítico tiene un papel muy relevante. No tanto porque lea los libros que salen, sino porque conoce el sentir de su tiempo, lo que necesita su tiempo, continúa la ironía kierkegaardiana. Los críticos, según Notabene, no siempre leen los libros. ¿Qué es un recensor? “Un salteador que se lanza sobre un libro que acaba de salir, un charlatán que se agarra fuerte a un escrito para hacer sitio a sus reflexiones y conseguir que sean escuchadas, un altivo don nadie que “aprovecha la ocasión” de un libro nuevo para decir algo.”³² Es curioso como describe al recensor. En el fondo es alguien que aprovecha que ha salido un libro para dar rienda suelta a sus posiciones personales. Utiliza el libro para mostrar sus ideas o su querencia. El ánimo con el que afronta la lectura no es para leer, sino para volver a decir lo que él ya piensa. No se lee para aprender o entrar en diálogo o para batallar contra uno mismo o el mundo, sino para aprovechar el libro para mostrar lo que ya se piensa. El libro no es considerado en su importancia, sino usado como medio. Hay maneras de escribir donde lo relevante no son los libros. No tendría por qué ser así. Notabene no dice cómo deben ser todos los recensores, sino como son de hecho algunos. Estos lectores-charlatanes son los que ayudan, o no, a que un libro ocupe un lugar en el mundo de los lectores (*Læseverdenen*). No importa tanto si el recensor lo ha leído o no. Hay críticas que se hacen sin leer el libro. Y hay críticos que leen la crítica del recensor. Y la gente habla en la plaza o en la academia de la crítica al recensor que no ha leído el libro. El libro, pues, ha quedado olvidado en un rincón. Lo relevante es lo que de él se ha ido diciendo³³. La conclusión demoledora última es la siguiente: “Cuando de un modo tan bello la literatura y el periodismo colaboran codo con codo, ¡entonces sí que resplandece ante nosotros el futuro de Dinamarca!”³⁴. La filosofía acaba siendo periodismo, así como otras ciencias. Esto es una posibilidad no muy lejana de lo que puede pasarnos hoy. La manera de entender el escribir al servicio de lo que el tiempo presente necesite puede derivar en estos caminos.

Pero escribir también puede ser analizado desde otro punto de vista que no tiene por qué ser contradictorio con el enunciado. Escribir puede ser entendido como algo para eruditos. Es lo que se desarrolla en el prólogo número VI. Es curioso porque en este prefacio liberado se nos dice que todavía no hay en la

32 SKS 4, 484. Nicolaus Notabene, *Forord. Prólogos*, 297.

33 Puede uno ver paralelismos entre lo que Kierkegaard critica con ironía y la realidad que a veces se descubre en nuestro presente sobre publicaciones, acreditaciones, índices de calidad, etc.

34 SKS 4, 485. Nicolaus Notabene, *Forord. Prólogos*, 299.

literatura danesa un sermón edificante para eruditos³⁵. Y eso se echa en falta, por lo visto. Hay gente que critica que nadie haya escrito sermones para esas personas tan peculiares que son los eruditos, los formados, los cultos. ¿Qué tipo de escritos serían los que serviría a estas personas formadas? ¿Cómo sería el libro del que este prólogo formaría parte? En el libro se recogerían sermones que:

Están atravesados por un esfuerzo, un pensamiento continuo, una tendencia sistemática. El sermón individual no busca valer o significar algo en sí mismo, sino, apuntando constantemente por encima de sí mismo, llevar al lector hacia la totalidad, que es lo único en lo que puede edificarse el erudito³⁶.

Aquí se nos apunta que el escritor culto busca la totalidad o el sistema. Sólo en él puede ser edificado. Ni siquiera el cristianismo parece que le edificaría. Para el erudito es bien poco lo singular o el individuo particular ya que lo decisivo es lo general, lo común³⁷. Porque lo que realmente edifica a este tipo de personas es lo siguiente:

Para el erudito, por el contrario, es realmente bien poco ocuparse de una persona particular, incluso aunque esta fuera él mismo. El erudito no quiere ser importunado cuando se dispone a edificarse, no quiere recordar todo lo insignificante, las personas particulares, a sí mismo; pues olvidar todo esto es justamente lo edificante³⁸.

Así es el erudito, *dannede*, el formado, forjado, cultivado, etc. Es el que no quiere recordar lo insignificante que es una persona particular. Esta manera de escribir puede llevar a plantearnos cómo entender ese libro ya que está al servicio del Sistema. ¿Sería el último libro? ¿Tendría sentido seguir escribiendo después? ¿No sería más útil aprender a leer puesto que ya se ha llegado al final? Es más, ¿tendría sentido seguir aprendiendo a escribir? N.N. lo dice de manera clara: “En cuanto este haya salido, las posteriores generaciones no tendrán siquiera necesidad de aprender a escribir, pues no habrá ya nada más que escribir, sino solo que leer: el Sistema”³⁹. Esto nos lleva a tener que analizar con más detenimiento la acción de escribir y cómo comprenderla. Porque es bien cierto que, si se entiende el escribir de esta manera, es porque también se le da una importancia decisiva

35 Nótese que Notabene no nos habla de discursos edificantes, *opbyggelige Taler*, sino de sermones-escritos edificantes, *Opbyggelsesskrift*.

36 SKS 4, 493. Nicolaus Notabene, *Forord. Prólogos*, 309.

37 “La vida de la comunidad, la grandiosa disposición del Sistema, lo puramente humano, cosas estas que no tientan al individuo a pensar en sí mismo o a querer consumir algo, sino que solo lo edifican en la medida en que las reflexiona, todo esto es en el presente escrito objeto de consideración.” SKS 4, 494. Nicolaus Notabene, *Forord. Prólogos*, 310.

38 *Íbid.*

39 SKS 4, 478. Nicolaus Notabene, *Forord. Prólogos*, p. 292.

a la mediación como clave en el proceso de creación literaria. Pero puede que Notabene nos muestre otras posibilidades de entender la escritura, otro temple desde el que escribir. Igual el libro necesita ser emancipado también, y no sólo el prólogo. Escribir no es hablar, y de ahí puede que emerjan muchas de sus posibilidades y debilidades. El diálogo con otro impide que todo se pueda decir. El libro es demasiado frágil. El papel lo soporta todo. Aunque no debería ser así⁴⁰.

4. SOBRE LOS FILÓSOFOS

La categoría de la mediación es picoteada en los dos últimos prólogos. Hegel, o al menos lo hegelianos daneses están, también aquí, explícitamente presentes⁴¹. Escribir en nuestro tiempo, ¿es fácil o difícil? “Escribir en nuestro tiempo es lo más fácil que hay si, siguiendo la costumbre, uno elige diez más viejos que traten sobre el mismo tema y a partir de ellos escribe sobre ese mismo tema un undécimo.”⁴² Escribir sería un movimiento del espíritu que media, que realiza la superación y recuperación de lo mejor del pasado. Mediar sería así entendido como un libro que supera, conserva y crítica al anterior realizando una nueva síntesis superadora. El libro quedaría, entonces, enredado en este movimiento del espíritu que se va desplegando. Y esto, con una carga irónica importante, afecta también a la filosofía: “Si soy tan estúpido que la filosofía no me puede resultar comprensible, lo opuesto a esto es que la filosofía es tan lista que no puede comprender mi estupidez; estos opuestos son mediados en una unidad superior, esto es, en una estupidez común”⁴³.

Pero esta tarea es ardua y pesada. Es complicado, por no decir irónico, el entender que un libro que se concibe a sí mismo como mediación de los anteriores deba ser mediado a su vez. ¿Cuándo entonces tendríamos el libro, el del sistema? Como esto es casi imposible, se toma el atajo de la promesa. El escritor escribe un libro donde promete que en diez años dará un sistema de Dogmática, o de

40 La ironía, tema central de la tesis de Kierkegaard consagrada a Sócrates, aparece en esta obra de manera constante. Aunque esta obra no es del todo irónica, hay momentos, como el que comentamos, donde ella pone al desnudo la supuesta verdad y bondad del escribir y publicar. Véase KIERKEGAARD, S. *De los papeles de alguien que todavía vive. Sobre el concepto de ironía*, Madrid: Trotta, 2006 (primera reimpresión), 81ss.

41 “Of course, the principal mark of anti-Hegelian motif is the very form of the work itself: a collection of prefaces, not a system of philosophy.”, PERKINS, R. L., “Reading Kierkegaard’s Prefaces with “Continual Reference to Socrates””. En: PERKINS, R.L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions (International Kierkegaard Commentary)*, Georgia: Mercer University Press, 2007, 111.

42 SKS 4, 497. Nicolaus Notabene, *Forord. Prólogos*, 313.

43 SKS 4, 518. Nicolaus Notabene, *Forord. Prólogos*, 332.

estética. La promesa es suficiente para el recensor y crítico, para la gente. Aquí el escritor es alguien que va a ayudar a otros a salir de su ignorancia, su tristeza o su falta de pericia. Así dicho, el libro sería un manual que el que sabe entrega al ignorante para que salga de la misma. El problema es que es el libro undécimo y no sabe uno si será el último.

Pero igual todo esto no es suficiente para el individuo singular que quiere dicha y salvación ahora, en su presente. ¿Cómo se entendería aquí la tarea del escritor? Notabene nos dice: “Escribir un libro para ayudar a otros, cuando él sabe que esencialmente todo hombre es por igual inteligente, y sobre todo, que en lo esencial todo hombre debe ayudarse a sí mismo.”⁴⁴ En tono socrático, como todo el libro, se nos deja caer que para lo esencial, para lo que tiene que ver con la felicidad y el bien, el ser humano aprende de una manera misteriosa y diferente. Es posible entender la escritura de otra manera porque es posible otra filosofía, otra manera de vivir. Uno puede escribir por una necesidad personal, como ya nos indicaba Nietzsche, porque está hablando consigo mismo, porque necesita hablar con otros para juntos llegar a más verdad y bondad. En este sentido “uno escribe un libro como del árbol brota su copa, y si alguien se complace, tanto mejor.”⁴⁵ El lector aquí no es considerado ni menor de edad en lo que se refiere al conocimiento ni abocado a ser un nuevo autor o escritor. Trata al otro como un ser humano particular. El autor escribe porque sale de sí y lo necesita. En otro prólogo de otro pseudónimo del mismo año se nos dice lo siguiente:

Quien se dispone a escribir un libro hace bien, según mi opinión, en pensar un poco en el asunto sobre el que quiere escribir. Tampoco le hará daño, si le resulta posible, ponerse al tanto de lo que se ha escrito antes acerca del mismo asunto. Si por ese camino se topa con alguien que de modo exhaustivo y satisfactorio consideró alguno de sus aspectos, seguramente ha de alegrarse como lo hace el amigo con el novio cuando oye la voz de este. Si lo hace con toda calma y con el fervor del enamoramiento, el cual busca siempre la soledad, entonces no necesita nada más; entonces acaba de escribir su libro con la frescura con la que el pájaro canta su melodía y, si hay alguien que saca provecho o halla alegría en él, tanto mejor; entonces lo publica sin apenas desvelos, sin darse la importancia de quien cree darlo todo por terminado o de quien supone que en su libro han de ser bendecidas todas las generaciones de la tierra⁴⁶.

44 SKS 4, 501. Nicolaus Notabene, *Forord. Prólogos*, 317.

45 SKS 4, 501. Nicolaus Notabene, *Forord. Prólogos*, p. 317. Las resonancias de tantos prefacios a los discursos edificantes donde se habla del discurso como algo que está ahí libremente para el querido lector que quiera apropiárselo son evidentes. Queda para otros trabajos el vincular todos los prefacios de la obra kierkegaardiana.

46 SKS 4, 313. Vigilius Haufniensis, *Begrebet Angest*. Tomado de *Migajas filosóficas. El concepto de angustia. Prólogos*. Madrid: Trotta, 2016, 129.

Es una cita larga pero enjundiosa. Se encuentra en un libro escrito el mismo año aunque con pseudónimo diferente. En ella se nos dice, como también Notabene nos indica, que para escribir primero hay que pensar bien el tema y leer lo que se ha escrito. Notabene no hace una apología de la escritura espontánea o efervescente que diga lo que se quiera o lo primero que se nos ocurra. Vigilius Haufniensis nos indica que hacer el trabajo de escudriñar y leer supone una actitud humilde de pensar que con uno no ha comenzado la historia de la verdad⁴⁷. Seguidamente, se hace imprescindible el temple o estado de ánimo o de humor con el que se escribe. Es decir, es bueno que el escritor se ría un poco de sí mismo. Porque una cosa es que él escriba y de alimento verdadero, aunque vete tú a saber quién podría decirlo, y otra muy distinta que el lector encuentre edificación o alimento en lo que se lee. En el primer caso, el escritor se concibe a sí mismo como alguien que enseña porque conoce. En el segundo, el lector es alguien que recibe un texto y que, por su manera de acoger el mismo, éste se convierte en vida, o no, para él. Lo hermoso de este proceso es que algo que uno puede escribir en soledad acompañada por otros se pueda convertir, por la apropiación del lector, en vida para él. Nada de darse importancia a uno mismo, ni considerar la obra como la definitiva después de la cual poco habría que añadir. Prólogos, como casi toda la obra kierkegaardiana, es comunicación indirecta donde el lector juega un papel decisivo⁴⁸. Esta manera de entender el proceso de escritura y lectura no queda aquí sujeto a la mediación. No es un libro que forma parte del esquema tesis-antítesis-síntesis. Porque ni el autor se vive así, ni la obra es presentada desde ahí, ni el lector se ve abocado a leerlo así. Lo escrito, sea prólogo o libro, el escritor y el lector quedan emancipados y liberados del Sistema que los aniquila y que los usa-deshecha. Lo que empezó siendo un proceso insignificante de liberar al prólogo del libro va dando paso a una manera de comprender al individuo sin-

47 “En esta perspectiva, leer es cosa seria, a pesar de que lo serio no esté allá donde se le busca habitualmente: lo serio no está unido a las dificultades que presenta la comprensión de un contenido, sino a la dificultad de apropiarse de ese contenido, de hacerlo suyo.” H. Politis, *Lire, écrire, publier: l'économie kierkegaardienne du livre*, Paris : Société Française de Philosophie - Bulletin de la Société Française de Philosophie, 2011/1, 27.

48 La comunicación indirecta es la manera en que Kierkegaard escribió casi toda su obra, como ha quedado mostrado en LEGARRETA, N. *Método y objeto en la autoría de Kierkegaard o el arte de pescar con señuelos de pluma*, Universidad Pontificia Comillas, Madrid 2013. La obra pseudónima no es indirecta porque no la firme Kierkegaard, sino porque la verdad es subjetividad, la apropiación personal es decisiva y la obra no puede ser acogida como una verdad objetiva que debe ser pensada, creída o defendida. En los *Discursos edificantes* lo indirecto se hace hiperbólico ya que el Cristo en su manera de hacerse presente al individuo impide que se pueda dar la comunicación directa. Además, la obra kierkegaardiana no se presenta como dogma a ser creído o comentario autorizado de un pastor. Permítaseme remitir a mi tesis doctoral *S. Kierkegaard: Una teoría del cielo*, Universidad Pontificia Comillas, 2017, 17ss. <https://repositorio.comillas.edu/xmlui/bitstream/handle/11531/18502/TD00247.pdf?sequence=1&isAllowed=y>.

gular en su tarea de escribir. Fruto de esta liberación tanto el autor como el lector se convierten en lectores, aprendices que necesitan pensar y hablar y actuar para seguir viviendo. Kierkegaard escribe su obra de tal manera que él se convierte en el primer lector de la suya. No sólo por las conexiones que él establece entre sus diversos pseudónimos y discursos edificantes, sino porque toda la obra está escrita con un temple. Ese estado de ánimo con el que está escrito es de quien no sabe la verdad, de quien la busca porque la necesita. Por eso dirá que Sócrates es su modelo⁴⁹. Él no escribe otro sistema, sino que ante la abundancia de estos decide volver a comenzar la aventura filosófica por el principio, por aquel que no escribió nada, por aquel que hablaba con sus conciudadanos buscando más verdad. Sócrates no era un sofista. Kierkegaard no es un filósofo del sistema disfrazado. La filosofía no es una actividad que surge en el ocio, sino una necesidad vital de primera magnitud.

El ser filósofo va apareciendo cada vez más en los prólogos, si es que no ha estado presente desde el principio. El tema del sufrimiento surge en los últimos prólogos conectados con esta manera de entender la escritura⁵⁰. Un sufrimiento que tiene dos grandes dimensiones: el sufrimiento del que se compadece de todo dolor humano: “Poder compadecerse de la flaqueza del ser humano es el verdadero principio del conocimiento”⁵¹. El que actúa, sea en la escritura o en otras acciones que impliquen a los demás, descubre que la acción es una maestra misteriosa. En ella podemos vivirnos muy atentos al daño que podamos producir a los otros. Este texto, que remite al versículo de la Biblia Hb 4, 15, es el origen del conocimiento. No es el asombro ni la admiración la que está en la raíz de este camino filosófico, sino el compadecerse de todo hombre y mujer. Su escritura y su oficio de escritor quedan totalmente redimensionados.

El sufrimiento es el secreto divino de la existencia mundana, que solo en el sufrimiento se está en trato con Dios, y que por ello si alguien es excluido del conocimiento, lo es solo aquel afortunado al que la fortuna ha hecho tan frívolo que no sufre nada en absoluto, cosa de la que el serio siempre se guardará, incluso aunque lo

49 “La única analogía que tengo por delante es Sócrates; mi tarea es una tarea socrática, revisar el sentido de ser cristiano. Yo no me considero a mí mismo un cristiano (manteniendo libre el ideal), pero puedo poner en evidencia que los otros lo son menos aún.”, *Øieblikket* Nr. 10., SKS 13, 405. *El instante*, Trotta, 2006, 189.

50 Remitimos para un análisis más completo del sufrimiento a Johannes anti-Climacus *Sydommen til Døden (La Enfermedad mortal)* o el ya citado *Begrebet Angest (El concepto de angustia)*. También creemos conveniente completar esos análisis con *Opbyggelige Taler i forskjellig Aand (Discursos edificantes en diversos espíritus)* en especial la tercera parte de este libro que es *Lidelseernes Evangelium (El Evangelio de los sufrimientos)*, así como al libro *Christelige Taler (Discursos cristianos)*. De estos dos últimos libros no contamos con traducciones directas desde el danés al castellano.

51 SKS 4, 502. Nicolaus Notabene, *Forord. Prólogos*, 318.

externo y el destino no lo hiciesen sufrir, porque en tal caso él mismo sabrá obtener el sufrimiento por simpatía⁵².

Incluso el que no sufre porque la vida le haya podido tratar muy bien, se compadece y sufre por simpatía con los demás porque se hace cargo de sus dolores. En el sufrimiento el hombre puede aprender. No se hace aquí una apología de cualquier sufrimiento, sino de aquel que está en relación con el mismo que actúa. Notabene no está hablando del sufrimiento de dolor de cabeza al intentar escribir algo muy sesudo. Tampoco del que pasa noches enteras pariendo su texto. Es, más bien, del que escucha y acoge lo que se le dice en el día a día y se lo toma muy en serio. No es la contraposición escribir para sí o para la gente donde se juega la manera de comprender aquella. El que quiere vivir su escritura con humildad, sin querer salvar o aleccionar a los otros, no vive al margen de los demás. Se toma tan en serio sus sufrimientos que no los utiliza ni para escribir ni para hacerse un nombre en el mundo de los lectores o escritores. El que escribe así sabe que el primer deber es intentar entenderse a sí mismo, saber por qué actúa y dice lo que dice. Así Notabene recupera el hablar cotidiano y popular, nótese el subtítulo de este libro, no para caer en las redes del parloteo de la plaza que endiosa o condena a autores y libros. Filósofo es el que habla con los otros sabiendo la gravedad de lo que esto supone. ¿Qué es lo que debe pensar y entender el que así se vive? Una de las lecciones más duras, y sobre la que hay que pensar detenidamente, nos la da el siguiente apunte:

Aquel que en su concepción de la vida y de las vicisitudes de la vida no ha alcanzado a saber qué puede aguantar una seria prueba y qué puede, como una fina capa de hielo, resistir solo el rápido paso de un patinador, lógicamente no podrá nunca lograr seriedad en sí mismo o en su labor. A cada persona le corresponde solo ocuparse de sí misma, y aquel que se cuida de mantenerse en esto tendrá una sólida base sobre la que pisar que no le decepcionará⁵³.

¿Qué concepción de la vida es capaz de soportar las pruebas más atroces y permanecer en pie? De esto es de lo que se trata de pensar y escribir. Hay maneras de pensar y escribir sobre la vida que no aguantan ni una duda ni un sufrimiento ni una prueba. He ahí un criterio de filosofía como responsabilidad radical ante la verdad y ante las personas que leen. Una verdad que pueda sostenernos cuando la prueba rompa los cimientos de nuestra vida. En este camino uno descubre, aunque no sea el descubrimiento último y definitivo, algo muy sencillo, pero enormemente decisivo:

52 *Ibid.*

53 SKS 4, 503. Nicolaus Notabene, *Forord. Prólogos*, 319.

Lo que cualquiera puede saber es lo que quiere, lo que puede, lo que ha hecho, todo lo cual es adecuado para una seria reflexión que por su parte regenerará su alma en una nueva seriedad... Aquel que por tanto vaya a ser mi lector deberá estar de acuerdo conmigo en que, aunque en nuestros tiempos la ciencia lo ha concluido todo, desgraciadamente ha olvidado la cuestión principal... que esté convencido de que a cada uno solo le corresponde ocuparse de sí mismo⁵⁴ y que esto es lo primordial, solo este es mi lector; pero por esto mismo no puedo realmente saber si no estará ya por delante de mí... Nunca se me podría ocurrir arrogarme la importancia de que fuese necesario para otros que yo escribiese mi libro⁵⁵.

CONCLUSIONES

El objeto de nuestro estudio era poner en valor el libro *Prólogos* de Kierkegaard para repensar el escribir filosofía hoy. Al hacerlo, hemos querido mostrar sus vínculos con otras maneras de entender los prólogos o prefacios. Kierkegaard entiende que en la manera de escribir, y no solo por lo que escribimos, manifestamos nuestra manera de entender la filosofía, de escribirla y hacerla. Hay maneras de escribir, como hemos visto, que parecen incitar a que no se escriba más porque en ese escrito está la verdad para ser leída y enseñada. Kierkegaard entiende que la tarea de escribir y leer van de la mano, ya que considera que el escritor debe ser lector de lo mismo que escribe con ansias de aprender para poder vivir, existir. En esta filosofía se escribe, si se me permite la expresión, porque es una obligación. No es un ocio o un camino para llegar a ser alguien. Se escribe para dialogar con uno mismo, con los otros en aras a buscar más verdad por la que vivir y morir. Escribir no es la tarea más importante, ya que la seriedad comienza siempre con la acción. En esta, está en juego nuestra vida. El misterio de la acción es enorme. Las aporías que en ella se nos presentan son variadas a lo largo de la vida.

Querer pretender que un libro aclare y nos saque de las mismas, puede que sea una actitud inadecuada. Esta diferenciación no le llevó a Kierkegaard a dejar de escribir. De múltiples y variadas maneras, con pseudónimos y con su propia firma, no dejó de dirigirse al mundo. La escritura filosófica siempre estará encajada en ese dilema radical en relación con la acción ética: no poder dejar de escribir aún a sabiendas que lo esencial siempre viene antes y después de la escritura. La tentación a la que no deberíamos sucumbir es utilizar la vida y menos aún el

54 El término danés, *anvise sig selv*, remite a ocuparse y a encargarse, a estar remitido, señalado a uno mismo. Es algo tan sencillo como que uno tiene consigo mismo una relación de carga, peso, ocupación, que sería bueno vivirla con responsabilidad. No es que no se encargue de los demás, sino que uno tiene un trabajo consigo mismo esencial porque no debería olvidar que existe.

55 SKS 5, 505. Nicolaus Notabene, *Forord. Prólogos*, 320.

sufrimiento de los demás para mayor gloria de nuestra escritura. Tampoco parece razonable, al menos desde el pensamiento kierkegaardiano, someterse a la tiranía de la necesidad de los tiempos marcados por las supuestas autoridades académicas o científicas. El filósofo escribe para el mundo, sus necesidades y dolores, para sus compañeros de camino, pero no escribe conforme el mundo le dicta. Escribir es una tarea ética y pareciera que el temple adecuado es el de la responsabilidad para con la verdad y para con el bien. La escritura tiene sus límites y sus posibilidades para iluminar el obrar humano. La lectura personal es imprescindible en el proceso libre de recepción del libro. Pretender aniquilar el mismo es haber acabado con la escritura. Hay maneras de escribir sobre lo esencial que acaban con el diálogo, lo que Kierkegaard no dejó de hacer con sus compatriotas. De este estamos muy faltos. De aquellas, estamos abrumados.

REFERENCIAS BIBLIOGRÁFICAS

- KIERKEGAARD, Søren, *Søren Kierkegaards Skrifter*. Copenhague: Gads Forlag, 1997-2013.
- , *Migajas filosóficas. El concepto de angustia. Prólogos*. Madrid: Trotta, 2016.
- , *Post Scriptum No científico y definitivo a «Migajas filosóficas»*, Madrid: Trotta, 2010.
- CARLISLE, Clare, *El filósofo del corazón. La inquieta vida de Søren Kierkegaard*. Barcelona: Taurus, 2021.
- PARCERO, Oscar, “La emancipación del Prólogo en la Fenomenología del espíritu”. *Anales del Seminario de Historia de la Filosofía*, 561, Vol. 33, Núm. 2, 2016, 561-582.
- COSSUTTA, Frédéric, “La préface en philosophie: une approche discursive”. *Argumentation et Analyse du Discours*, 22, 2019, 1-16.
- NIETZSCHE, Friedrich, *La ciencia jovial*. Madrid: Gredos, 2010.
- FOUCAULT, Michel, *Historia de la locura*, Volumen 1, México: Fondo de Cultura Económica 1998.
- POLITIS, Hélène, *Lire, écrire, publier: l'économie kierkegaardienne du livre*, Société Française de Philosophie - Bulletin de la Société Française de Philosophie, 2011/1.
- KJÆLDGAARD, Lasse H., “The Age of Miscellaneous Announcements: Paratextualism in Kierkegaard’s Prefaces and Contemporary Literary Culture”. En: PERKINS, L. Robert, (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions* (International Kierkegaard Commentary), Georgia: Mercer University Press, 2007, 7-28.
- CRITES, Stephen, “The Unfathomable Stupidity of Nicolaus Notabene”. En: PERKINS, Robert L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions* (International Kierkegaard Commentary), Mercer University Press, 2007, 29-40.

- HEGEL, Georg Wilhelm Friedrich, *Fenomenología del Espíritu*. Valencia: Pre-Textos, 2006.
- STEWART, Jon, "The Polemic with Heiberg in Prefaces". En: *Kierkegaard's Relations to Hegel Reconsidered*. Cambridge: Cambridge University Press, 2003, 419-447.
- PERKINS, Robert L., "Reading Kierkegaard's Prefaces with "Continual Reference to Socrates". En: PERKINS, Robert L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions* (International Kierkegaard Commentary), Mercer University Press, 2007, 111-138.
- FERGUSON, Harvie, "Before the Beginning: Kierkegaard's Literary History". En: PERKINS Robert L. (ed.), *Prefaces and Writing Sampler/Three Discourses on Imagined Occasions* (International Kierkegaard Commentary), Mercer University Press, 2007, 41-65.

CULTURA Y ESCLAVITUD EN EL PENSAMIENTO DE NIETZSCHE. ¿UN ARISTOCRATISMO MERAMENTE ESPIRITUAL?

*CULTURE AND SLAVERY IN THE THOUGHT OF
NIETZSCHE. A MERELY SPIRITUAL ARISTOCRATISM?*

JORGE POLO BLANCO

Doctor en Filosofía
Facultad de Arte, Diseño y Comunicación Audiovisual
Escuela Superior Politécnica del Litoral (ESPOL)
Guayaquil/Ecuador
polo@espol.edu.ec
hiperbolik1983@hotmail.com
ORCID: 0000-0001-9415-5406

Recibido: 8/03/2021
Revisado: 18/07/2021
Aceptado: 6/09/2021

Resumen: En este trabajo queremos ahondar en cierto ángulo de la filosofía nietzscheana, a saber, los vínculos que el filósofo alemán establece entre arte y poder. Su aristocratismo, lejos de ser meramente espiritual, es también social y político. Nietzsche se nos muestra crudamente diáfano en muchos de sus planteamientos, puesto que llega a sostener que las condiciones óptimas para la emergencia de un gran florecimiento artístico, llevado a término por una selecta y genial minoría, sólo podrán darse si una inmensa mayoría de la población vive sometida a un trabajo esclavo. Para que el “tipo superior de hombre” pueda medrar se precisa de un máximo de explotación de los hombres bajos e inferiores. Semejante tesis nietzscheana, íntimamente relacionada con otros de sus filosofemas, es a nuestro modo de ver descriptiva y normativa.

Palabras clave: Nietzsche, arte, aristocratismo, esclavitud, elitismo.

Abstract: The present study aims at delving into a certain perspective of the Nietzschean philosophy, namely, the links that the German philosopher established between art and power. His aristocratism, far from being merely spiritual, is also social and political. Nietzsche is starkly diaphanous in many of his approaches, considering that he argued that the optimal conditions for the emergence of a great artistic development, carried out by a select and brilliant minority, may only occur if a vast majority of the population lives under slave labor. For the “higher man” to thrive, a maximum level of exploitation of the lower

type of man is required. According to this study, such a Nietzschean thesis, closely related to others of his philosophemes, is descriptive and normative.

Keywords: Nietzsche, art, aristocratism, slavery, elitism.

1. ¿HUBO UN NIETZSCHE ILUSTRADO?

Acertaba Eugen Fink, cuando consignaba en su clásico estudio sobre Nietzsche que este atravesó una fase, en cierto modo, científico-positivista; una fase “ilustrada”, advertía¹. En esta fase intermedia de su biografía intelectual, que comprendería *Humano, demasiado humano, Aurora* y *La gaya ciencia*, se habría puesto en movimiento una psicología destructiva y desenmascaradora. Y así fue, en cierto modo y hasta cierto punto. Sería posible hallar en este periodo una saludable “actitud ilustrada”, podríamos decir, puesto que la trituradora nietzscheana se desplegó a través de una refinada desmitificación de seculares espejismos religiosos; “deconstruyendo” –diríamos hoy– venerables tradiciones sacrosantas. Detrás de las cosas más sublimes y adoradas, anunciaba Nietzsche, se descubren monumentales embustes. Autoengaños con siglos de antigüedad quedaban, por fin, desnudados. Practicó en estas obras, por lo tanto, una labor de cruento desencantamiento.

En el aforismo 9 de *Aurora*, por ejemplo, podemos observar una cierta apología de aquellos “espíritus libres” que acaso consiguen zafarse del peso siniestro y amenazante de la tradición, siendo inmediatamente catalogados por ello como malvados y peligrosos; una defensa de esos espíritus escogidos y extraños que dejan de prestar obediencia a los prejuicios morales aquilatados en atávica e indiscutida costumbre. Es este, tal vez, uno de los pasajes más “ilustrados” de Nietzsche². Pero, no obstante, debemos realizar una observación importante: en estos pasajes presuntamente ilustrados nunca hallaremos una propuesta fundamentada en la “liberación del pueblo”, puesto que a Nietzsche no le preocupan las servidumbres colectivas. Por el contrario, él solo habla de una suerte de “espíritus raros y selectos” que ya no quieren pertenecer al rebaño y, por ello, desechan todos los valores legados por la tradición. La rebeldía así postulada es, en suma, eminentemente individualista. Lo que se propone es un gesto solipsista y egotista de desdén, una irreverencia agresiva pero solitaria, sin que tal cosa se encamine en la dirección de una emancipación colectiva. ¿Es esto último propio de la Ilustración? Evidentemente, no.

1 FINK, Eugen. *La filosofía de Nietzsche*. Madrid: Alianza, 1976, 50-70.

2 NIETZSCHE, Friedrich. *Aurora. Pensamientos sobre los prejuicios morales*. Madrid: Biblioteca Nueva, 2015, 67-69.

Como señaló Ernst Nolte, si en algún momento Nietzsche manejó ciertas nociones que casi lo convertían en un ilustrado vulgar –sumándose a esa promesa que señalaba aquello de que cuando los hombres se hagan científicos se acabarán los sacerdotes y los dioses– lo cierto es que su trayectoria intelectual prosiguió y se afianzó en un sentido radicalmente anti-ilustrado. Esos modernos “librepensadores” no eran más que mórbidos “niveladores”, despreciables “predicadores de la igualdad”, tumefactos apologetas de la “moral del esclavo” y, por ende, despreciables de la vida. Así, Ilustración y Revolución francesa constituían, a sus ojos, un fenómeno (político, cultural y espiritual) ubicado o encuadrado en el secular proceso de la degeneración vital de Occidente. Las revoluciones de signo democratizador (socialismo, anarquismo y sufragismo aparecían como horripilantes jalones del moderno “movimiento democrático”), y todas las ideas que las sustentaban, eran consideradas por Nietzsche culminaciones nauseabundas de esa decadencia cultural en la que estaba embarcado el hombre europeo desde los tiempos del infausto Sócrates³. ¿Compartían Nietzsche y la Ilustración, al menos, el sentimiento anticlerical? Tal vez sí, pero por diferentes motivos. Recordemos que para Nietzsche todo discurso o movimiento de corte igualitarista procedía, en último término, del *ethos* cristiano. Es decir, desde su óptica un revolucionario jacobino, un anarquista o un socialista no constituían sino la perpetuación de la figura del Sacerdote; tal vez presentarían otros ropajes y otros lenguajes, pero todos esos revolucionarios modernos compartían –en su fondo último– el deletéreo “instinto moral” del cristiano.

La propia ciencia, además, terminó siendo objeto de esa “psicología desenmascaradora”, y nuestro filósofo se fue distanciando de las coordenadas intelectuales mostradas en este lapso que, solo muy vagamente, podríamos denominar “ilustrado-positivista”. La “fría” racionalidad también habría de ser, a la postre, un ídolo a destronar, por mucho que ella hubiese prestado un buen servicio a la hora de derribar algunas inveteradas ilusiones de la teología y de la metafísica. Porque la ciencia moderna, que a su juicio conservaba e incluso amplificaba un impulso metafísico de adoración por la Verdad, cercenaba igualmente la “jovialidad” de la vida. Y no solo eso; había un motivo más. Y es que Nietzsche llega a sostener que la ciencia progresa en “alianza” con el movimiento igualitario y democrático...algo terrible, a sus ojos⁴. En base a lo dicho, ¿puede sostenerse verosímilmente que Nietzsche perteneció, siquiera sea en alguna fase de su vida, al movimiento político-cultural e ideológico-filosófico de la Ilustración? Resultaría

3 NOLTE, Ernst. *Nietzsche y el nietzscheanismo*. Madrid: Alianza, 1995, 156-158 y 165-169.

4 NIETZSCHE, Friedrich. *Fragmentos póstumos (1885-1889)*. Volumen IV. Madrid: Tecnos, 2008, 131.

muy complicado responder afirmativamente a semejante cuestión, por mucha admiración que dijera profesarle a Voltaire (al cual dedicó, es cierto, *Humano, demasiado humano*). Además, del propio Voltaire –de todo lo que este representaba o simbolizaba– se burlará en *Más allá del bien y de mal*, cuando exclame sardónicamente: “¡Oh Voltaire! ¡Oh humanidad! ¡Oh estupidez!”⁵.

2. CONTRA LOS “PREDICADORES DE LA IGUALDAD”

No podremos coincidir con Fink en su caracterización de Nietzsche como un “filósofo de la libertad”, al menos que maticemos con mucho cuidado la noción nietzscheana de libertad. En *Así habló Zaratustra*, obra que inaugura el nuevo ciclo, explotarán todas las terribles fuerzas que, como señala el propio Fink, se hallaban contenidas de forma subterránea en las obras de la mencionada fase “científico-positivista”. Zaratustra, que desciende al país de los hombres tras sus años de soledad en la montaña, anuncia la muerte de Dios. Con ello, se estaría anunciando una profundísima liberación, puesto que los hombres ya podrían dejar de ser “camellos”, esto es, dejarían de soportar la pesada y asfixiante carga de los trasmundos eternos. Muerto Dios, y desvanecido el oprobioso sueño del “más allá”, desaparecía esa autoalienación, que ya duraba milenios. Los hombres ya no tendrían que inclinarse con gesto agotado ante la omnipotencia de Dios; ya no tendrían que arrodillarse, exhaustos y famélicos, ante la sublimidad draconiana de las leyes morales. Dejando de ser “camello” para convertirse en “león”, el ser humano se zafaría de lo trascendente, que oprimía y constreñía con fiereza la vida humana en la tierra. Ese contundente *no* a la tiranía de un Dios omnisciente –y a la tiranía de una ley moral presuntamente objetiva–, sin embargo, habría de dar paso a otra figura, la del “niño”, que en este caso ya no se limitaría a negar lo trascendente, sino que alcanzaría un nuevo estado desde el cual se proferiría un contundente y lúdico *sí* al “sentido de la tierra”, poniendo en marcha un nuevo y creativo “querer”, una nueva capacidad de generar valores inéditos⁶.

Aniquilado el “mundo verdadero”, la humanidad podría respirar aliviada y risueña. ¡Ya no existirían sentidos “eternos y últimos”, impuestos coactivamente por una exterioridad absoluta! Y ello sería motivo de una indecible alegría, porque esos nuevos seres humanos que bailotearían y jugarían sobre el cadáver de Dios podrían ahora generar nuevos e inéditos sentidos. Al llegar a ese punto, la osadía

5 NIETZSCHE, Friedrich. *Más allá del bien y del mal. Preludio para una filosofía del futuro*. Madrid: Edaf, 2006, 95.

6 NIETZSCHE, Friedrich. *Así habló Zaratustra*. Madrid: Alianza, 1998, 53-55.

“infantil” del superhombre cristalizaría, según Fink, en un grandioso ensanchamiento de la libertad humana. “El terreno de juego de la libertad es inabarcable si Dios no limita ya al hombre, si esa pared insalvable no cierra ya su camino ascendente, si ya no cae sobre el país humano la sombra inmensa del Señor”⁷. Sin embargo, habríamos de comprobar qué tipo de libertad sería la que adviene con este nuevo amanecer. Pero, para ello, hemos de seguir escuchando otras prédicas de Zaratustra.

Los viejos seres humanos, legados por la tradición socrático-platónica y judeocristiana, se arrastran como seres fracturados y esquizofrénicos, puesto que viven instalados en una escisión improbable y enfermiza: desprecian el cuerpo, al que consideran una prisión absurda y una rémora para el espíritu; reniegan del “mundo sensible”, de lo sensual, puesto que esa dimensión los ata a los más bajo, esto es, a lo terrenal. Hombres fracturados que pretenden “mortificar su carne”, para elevar su alma al reino de lo inteligible, único lugar donde presuntamente reside la Verdad eterna. Ante semejante panorama, Zaratustra se revuelve con frenesí y con asco. Dirigiéndose a los hombres, los exhorta con vehemencia a “permanecer fieles a la tierra”, y asevera que todos los que han venido predicando la existencia de trasmundos ideales y eternos no han resultado ser más que pérfidos “envenenadores” y “despreciadores de la vida”⁸. Pero, llegados a este conocido pasaje, hemos de inquirir algo. Esos “envenenadores” de la humanidad son, a juicio de Nietzsche-Zaratustra, los ascetas, los sacerdotes, los santos, los teólogos, todos los filósofos de estirpe platónica... ¿Y quién más? Oh, claro, los modernos envenenadores de la humanidad se visten con los ropajes de la emancipación social; jacobinos, demócratas radicales y socialistas de todo pelaje que andan por ahí propalando doctrinas deletéreas. En el capítulo titulado “De las tarántulas” el tono polémico se radicaliza, y Zaratustra clama contra todos esos “predicadores de la igualdad” que, mostrando una repulsiva sed de venganza, pretenden derribar todas las jerarquías. Como tarántulas venenosas emergen todos esos revolucionarios que, espoleados por su mezquina impotencia, pretenden sublevarse contra la natural desigualdad que ha de existir entre los poderosos y los débiles⁹. Recubren su propia debilidad con un disfraz llamado “afán de justicia”, tratando con ello de aparentar un virtuosismo social cuando, en realidad, solamente los mueve el resentimiento y la fragilidad. Aquí encontramos prefiguradas la “moral de señores” y la “moral de esclavos”, que Nietzsche desarrollaría con mayor prolijidad en las obras subsiguientes.

7 FINK, Eugen. *op. cit.*, 88.

8 NIETZSCHE, Friedrich. *Así habló Zaratustra*, *op. cit.*, 36.

9 *Ibid.*, 155-158.

Empezó a escribir *Also sprach Zarathustra* en 1883. Sin embargo, el asunto que nos concierne empezó mucho antes. En 1870 su odio antidemocrático –demofóbico, antipopular y antiplebeyo– estaba bien perfilado y consolidado. Ya nunca abandonaría esa posición política. Como señaló Thomas Mann, en su pequeño pero magnífico ensayo sobre Nietzsche, este siempre fue el mismo; no hubo saltos cualitativos en su biografía intelectual. Todo lo que Nietzsche ha representado en la historia del pensamiento estaba ya incoado en sus primeros fogonazos; lo único que fue modificándose, si acaso, es el acento: cada vez más chillón, grotesco y frenético¹⁰. Pero su obra se moverá siempre en unas coordenadas bien definidas: naturalización de la justicia y biologización de la moral. Como veremos en lo que sigue, la animalización nietzscheana del ser humano tiene un objetivo político bien definido. González Varela ha sintetizado ese gesto teórico de un modo formidable, señalando que Nietzsche se vio empujado, ante el ejemplo tremendo de la *Commune*, a promover una crítica reaccionaria del antropocentrismo, mediante la siguiente maniobra: si la formulación revolucionaria de los Derechos del Hombre instituyó un discurso muy duro de igualdad *dentro* del género humano y, al mismo tiempo, un signo absoluto de diferencia con respecto al mundo animal y natural, la ecuación por él propuesta era una completa inversión de dichos términos. Es decir, como un verdadero “anti-Rousseau”, Nietzsche radicalizará la diferencia entre hombre y hombre (genio-esclavo) y reafirmará, a su vez, la continuidad (o la no diferenciación esencial) entre el mundo humano y el mundo animal¹¹. Podríamos representárnoslo así: el pensamiento revolucionario de estirpe rousseauiana tiraba líneas horizontales entre los hombres, y líneas verticales (jerárquicas) entre aquellos y el mundo animal. Nietzsche habría invertido esos términos, trazando una línea horizontal (esto es, una línea de continuidad) entre el género humano y el mundo animal, y múltiples líneas verticales (jerárquicas) en el interior del mundo humano.

Cuando subrayaba de forma descarnada la desigualdad entre los hombres, impugnando con supremo desdén todos los igualitarismos rousseauianos y revolucionarios, Nietzsche albergaba en su cabeza el esbozo de un modelo de sociedad. No resulta creíble sostener que tales cuestiones le traían sin cuidado. Por ejemplo, en los esbozos de un plan para un libro inconcluso, proyectaba dedicar todo un capítulo a teoría del Estado; era el año 1869¹². La forma estatal preferente debiera ser aquella capaz de vehicular una auténtica cristalización de los instintos poderosos, esto es, aquella capaz de permitir una jerarquización aristocrática del orden

10 MANN, Thomas. *Schopenhauer, Nietzsche, Freud*. Madrid: Alianza, 2000, 101-102.

11 GONZÁLEZ VARELA, Nicolás. *Nietzsche contra la democracia*. Barcelona: Montesinos, 2010, 123.

12 *Ibid.*, 112.

social; un organismo político, en definitiva, diseñado para la plasmación “natural” de las diferencias irreductibles entre señores y siervos. En múltiples pasajes, con descarnada y explícita vehemencia, arrojaba vituperios contra los “Derechos Fundamentales del Hombre” y contra la “Dignidad del Trabajo”, consignas altamente deletéreas. Sostenía la imperiosa necesidad de que una masa embrutecida siguiera trabajando para una minoría privilegiada, como comprobaremos a continuación. Funesto fue el momento, señalaba, en el que los esclavos fueron estimulados para elevarse más allá de su condición. El inasequible optimismo de toda “cultura socrática”, y tal es el caso de la moderna Ilustración, conlleva de forma inherente una contradicción irresoluble: prometiendo la felicidad terrenal para todos, y aludiendo a ensueños de fraternidad igualitaria, semejante proyecto socava las bases mismas de su propia existencia; porque los esclavos, en efecto, quieren cobrarse esa promesa y exigen su materialización efectiva. La “proclamación” de derechos debe ser institucionalizada; llegó la hora. Una exigencia “amenazadora y espantosa”, sostiene Nietzsche, teniendo en cuenta la presencia de los movimientos socialistas¹³.

No se trataba, por ende, de una simple “aristocracia intelectual”. Nietzsche señalará abiertamente que los verdaderos “genios” solo pueden emerger dentro de una selecta minoría que domine, autoritariamente, el orden social. De eso se está hablando, ni más ni menos: del orden social. La “gran cultura” debe ser accesible solo a unos pocos, por lo cual es completamente absurdo postular programas educativos para el pueblo trabajador. Pretender ofrecer una suerte de “formación cultural” generalizada constituiría, además, un proyecto verdaderamente anti-germano¹⁴. Véanse, en ese sentido, la serie de conferencias que Nietzsche impartió en Suiza en los primeros meses de 1872¹⁵. En un trabajo que llevaba por título *Sobre el porvenir de nuestras escuelas* –en realidad, son las conferencias suizas mencionadas– aseveró que una verdadera formación (*Bildung*) debe estar propiciada por una determinada institucionalidad pedagógica que permita una adecuada selección y promoción de los “genios”; esto es, un sistema pensado no para alfabetizar a las masas, sino para que florezcan los “individuos superiores”. He ahí el modelo educativo de Nietzsche, su *paideia* eminentemente elitista.

En un prefacio a dichas conferencias, llegó a escribir que “una restricción de la cultura a pocas personas es una ley necesaria de la naturaleza”¹⁶. Sin em-

13 NIETZSCHE, Friedrich. *El nacimiento de la tragedia*. Madrid: Alianza, 1973, 146-147 y 154.

14 GONZÁLEZ VARELA, Nicolás. *op. cit.*, 115, 123, 138, 147 y 148.

15 JANZ, Curt Paul. *Friedrich Nietzsche. 2. Los diez años de Basilea (1869-1879)*. Madrid: Alianza, 1987, 142-145.

16 NIETZSCHE, Friedrich. *Sobre el porvenir de nuestras escuelas*. Barcelona: Tusquets, 2009, 25.

bargo, los modelos pedagógicos imperantes estaban diseñados para arruinar la *excepción*, para adormecerla, embotarla y reducirla en un pozo de nivelación mediocre. Es cierto que en la primera de dichas conferencias encontramos valiosísimas críticas a la especialización extrema del mundo académico, y también a la mercantilización del conocimiento (la adquisición de este con vistas, únicamente, al desarrollo de una buena carrera laboral y pecuniaria). Semejante diagnóstico tiene hoy, lamentablemente, muchísima vigencia¹⁷. Pero lo que no podemos admitir es la “solución” de Nietzsche a semejante deriva. Porque para él suponía un fantástico error pensar que todos los individuos tienen derecho a una misma educación, cuando la realidad es que solo unos pocos están preparados para alcanzar los umbrales de una genuina formación.

Esta pedagogía aristocrática, por lo tanto, presupone que el talento no está distribuido de una forma equitativa, y solo unos poquitos elegidos poseen una naturaleza adecuada para recibir una educación superior. La gran cultura jamás florecerá de la muchedumbre amorfa; esta debe recibir apenas una formación elemental que garantice su inserción dócil en la estructura social. Todos los sueños ilustrados que cristalizan en programas de “educación universal” descansan en un sofisma, esto es, en la absurda creencia de que un “pueblo educado” se emancipará del dolor y de la injusticia. Para Nietzsche, tal cosa no sucederá jamás; es más, ni siquiera sería deseable que así sucediera. Muy al contrario, ciertas aristocracias culturales, y los individuos excepcionales que brotan de ellas, deben acaparar el privilegio de la formación superior. Convertir esta en un derecho universal del pueblo...desembocaría en una ruina espiritual. Un sistema educativo orientado a toda la población sería un aberrante proyecto del todo inadecuado a la naturaleza desigual de los seres humanos. Las mieles de la verdadera cultura están reservadas a un segmento privilegiado; las masas trabajadoras, por el contrario, apenas deben saber lo justo y necesario para seguir realizando sus labores productivas y reproductivas en la base de la pirámide social¹⁸.

El sistema educativo idóneo, asevera Nietzsche con resonancias platónicas (aparece aquí la sombra de la *República*), es aquel que se organiza en torno a la “selección de genios” (“*die Zucht des Genius*”). ¿Puede traducirse *Zucht* como “cría” o “crianza”? Resonancias zoológicas, connotaciones eugenésicas; elitismo indisimulado. Resulta curioso observar cómo los hagiógrafos de Nietzsche ignoran estas conferencias, guardando sobre ellas un atronador silencio (Vattimo o Kaufmann, por ejemplo) ¿Por qué ese empeño en ocultar el verdadero *pathos* político de Nietzsche? En la cuarta de estas conferencias, una vez más, nos habla

17 *Ibid.*, 31-58.

18 ESTEBAN ENGUITA, José Emilio. *El joven Nietzsche. Política y tragedia*. Madrid: Biblioteca Nueva, 2004, 203-205.

de la perentoria necesidad de no hacer colapsar una distancia, a saber, la que media entre esa masa obtusa que se multiplica por instinto (observamos aquí cierta animalización denigrante de las capas populares) y los grandes individuos contemplativos –“ejemplares” robustos– capaces de crear cosas eternas¹⁹. La distancia entre unos y otros *debe* permanecer intacta; jamás debiera ser atenuada o reducida. La cúspide de la pirámide (esta imagen es del propio Nietzsche) no puede ser maculada por manos plebeyas. Las élites culturales, aupadas en esas rutilantes alturas, no pueden ser estorbadas o menoscabadas por un sistema de Instrucción Pública orientado a la “formación del pueblo”. Pero no vayamos a equivocarnos: las “aristocracias espirituales” de Nietzsche tienen que estar sostenidas por un determinado modelo sociopolítico y socioeconómico, por una determinada institucionalidad educativa y por una determinada organización estatal.

También en el párrafo 467 de *Humano, demasiado humano* arremetería contra la instrucción pública ofrecida por el Estado, por considerarla mediocre²⁰. Y se intuye claramente que aquello que la convierte necesariamente en algo deficiente y desprestigiado es, precisamente, su vocación de ser una instrucción tendencialmente universal. Que las masas adquieran cultura y educación es algo nefasto, piensa Nietzsche. Y esta postura se entiende todavía mejor cuando leemos el párrafo 462 de esa misma obra, pues en él dejó consignado que una “utopía” más adecuada consistiría en una cierta organización social en la que los trabajos más duros y penosos fueran acarreados por aquellos que sufren menos, esto es, por los más estúpidos²¹. Aquellos con una capacidad mayor de sufrimiento, precisamente los que portan un espíritu más refinado, han de permanecer exonerados de toda laboriosidad fatigante. Ahora comprendemos mejor su fobia a la “instrucción pública”, puesto que una masa de hombres ignorantes y embrutecidos siempre será necesaria para sustentar la pirámide social. Esto lo dirá una y otra vez. Remarcó, con meridiana claridad, que una civilización superior solo podrá germinar allí donde perviva, de algún modo, una nítida diferenciación entre la casta de los trabajadores (forzados, si es menester) y la casta de los ociosos²².

Como prueba palmaria de que en sus tesis aristocráticas no hay meros componentes “culturales”, “moralizantes” o “estetizantes”, podemos acudir al aforismo 206 de *Aurora*, texto de 1881. Allí, un Nietzsche más tardío y maduro, sigue señalando que el conflicto social que conmueve y conturba a Europa podría solucionarse importando fuerza de trabajo de origen chino, tal vez porque

19 NIETZSCHE, Friedrich. *Sobre el porvenir de nuestras escuelas*, op. cit., 127.

20 NIETZSCHE, Friedrich. *Humano, demasiado humano*. Madrid: Edaf, 2010, 259.

21 *Ibid.*, 258.

22 *Ibid.*, 248.

aquella industriosa mano de obra, caracterizada como “hormigas laboriosas”²³, podría ser ilimitadamente explotada sin que mediasen disputas laborales o cuestionamientos morales. Una propuesta muy dionisiaca, desde luego. En ese mismo parágrafo, que no tiene desperdicio, encontramos alusiones particularmente irritantes. Señala Nietzsche que los obreros europeos, triturados espiritualmente, creen ingenuamente que una mejora salarial podría redimir su pobre existencia; utilizados como herramientas, sometidos a la ignominia de la esclavitud fabril, ya no son capaces de experimentar la profunda vergüenza de su situación. “¡Qué horror creer que con un salario más elevado desaparecería *lo esencial* de su miseria, de su servidumbre impersonal! ¡Que nadie se deje convencer con el argumento de que, aumentando esa impersonalidad, por medio del engranaje de la máquina de una nueva sociedad, podría convertirse en virtud la vergüenza de la esclavitud! ¡Tener un precio por el que se deja de ser persona para convertirse en tornillo! ¿Sois vosotros los cómplices de la actual locura de las naciones que pretenden producir mucho y enriquecerse lo más posible? Vuestra tarea sería presentarles una contrapartida. ¡Qué gran suma de valor interior se desperdicia para lograr ese fin tan superficial! Pero, ¿dónde está vuestro valor interior si no sabéis lo que es respirar con libertad, si apenas tenéis la fuerza suficiente? ¿Si con frecuencia estáis cansados de vosotros mismos, como de una bebida que ha perdido la fragancia?”²⁴. Leyendo descuidadamente este pasaje, pudiera parecer que Nietzsche se lamentaba, como el joven Marx de los *Manuscritos económico-filosóficos*, de la tremenda situación de alienación en la que malvivían los obreros de la industria moderna. Pues es cierto, como el propio Nietzsche advierte, que el problema no se reducía a una cuestión meramente salarial, toda vez que el sometimiento de aquellas gentes alcanzaba límites más profundos. Era la suya, también, una degradación espiritual. Pero, en realidad, nuestro filósofo las culpaba a ellas, a las víctimas.

Desde un olímpico desprecio, observa a los obreros comportarse como estóliidos insectos, corroídos por la envidia y espoleados por el resentimiento; incapaces de comprender su situación, aborregados y sumisos, se hunden cada vez más en el fango de una tenebrosa animalización. ¿Acaso está exhortándolos a la acción emancipadora? ¡Desde luego que no! Con todo su cinismo, les espeta que “presenten una contrapartida”. Pero Nietzsche, durante toda su vida, destiló puro veneno contra todos los movimientos políticos populares y revolucionarios que, en efecto, presentaban una contrapartida al sistema socioeconómico dominante. Recordemos cuánto asco y miedo experimentó ante el episodio de la Comuna parisina. En ese mismo texto, de hecho, arremete una vez más contra los “flautistas

23 NIETZSCHE, Friedrich. *Aurora*, op. cit., 189.

24 *Ibid.*, 188.

socialistas” que enardecen a los obreros con “esperanzas absurdas”²⁵. El partido revolucionario le produce náuseas. ¿Qué salida ofrece a la “cuestión obrera”, más allá del desprecio por las clases trabajadoras? Siente repugnancia ante un espectáculo tan desolador, pero Nietzsche jamás ofrece una sola línea de fuga. Bueno, en realidad sí; en este mismo párrafo recomienda a los obreros que, para escapar de su infame servilismo, se enrolen en alguna aventura colonial para, en tierras salvajes, poder comenzar una nueva existencia, esta vez como “señores”. Emigrar a lugares lejanos, logrando así traspasar su condición de explotados en la de explotadores.

Es indudable que Nietzsche profesaba un odio muy profundo a todo aquello que oliese lo más mínimo a igualitarismo político, y entendía que la moderna democracia no es más que la irrefrenable decadencia del Estado. El llamado “conflicto social”, además, existe porque los patronos modernos no saben mandar. El aforismo 40 de *La gaya ciencia* es verdaderamente delicioso, en ese sentido. Comienza diciendo que cualquier “cultura militar” está muy por encima de la “cultura industrial”. ¿En qué sentido? Los capataces de la moderna industria, incapaces de legitimar moralmente su jerarquía, no pueden infundir el debido respeto en el obrero. Lo dueños de las fábricas no han sabido insuflar en sus trabajadores una elevada moral de la disciplina, y por ello se vuelven discolos y revoltosos. El virus infeccioso del socialismo palpita en el corazón de las masas, concluye Nietzsche con amargura, porque la casta dirigente actual ya no alberga el *pathos* de los auténticos hombres superiores; solo estos podrían impedir el desgarrador conflicto, enseñando a los obreros a sentir y vivir como soldados, con lo cual ya no cuestionarían el orden jerárquico ni experimentarían como penoso su sometimiento²⁶.

3. RADICALISMO ARISTOCRÁTICO

El pensamiento nietzscheano debe ponerse en diálogo con las tradiciones ideológicas y políticas de su tiempo²⁷. Se trata de una filosofía que tiene evidentes derivaciones políticas²⁸. De hecho, consideramos adecuadas las tesis sostenidas por Domenico Losurdo en su imprescindible *Nietzsche, il ribelle aristocratico*. Sostenía de forma muy consistente que la obra nietzscheana solo podía

25 *Ibid.*, 189.

26 NIETZSCHE, Friedrich. *La gaya ciencia*. Madrid: Edaf, 2002, 115-116.

27 OTTMANN, Henning. *Philosophie und Politik bei Nietzsche*. Berlín: Walter De Gruyter, 1987.

28 POLO BLANCO, Jorge. *Anti-Nietzsche. La crueldad de lo político*. Madrid: Taugenit, 2020.

comprenderse como la propia de un intelectual orgánico de la contrarrevolución²⁹. Al igual que otros muchos pensadores y científicos sociales decimonónicos, Nietzsche biologizó la existencia humana y naturalizó el orden social. A su juicio, era la propia pugna vital, en su cruento devenir ininterrumpido, la que debía establecer las posiciones ocupadas por cada cual. Y este escenario así conformado no debería ser trastocado, o “corregido”, bajo el criterio de algún ideal justiciero. Las coincidencias con Herbert Spencer son palmarias³⁰. Inhibir la rapiña violenta, o amortiguar de algún modo la despiadada competición existencial, conllevaría un ruinoso adormecimiento de los nervios últimos del mundo humano. Extender tales principios pacifistas, humanitaristas o igualitaristas a todo el cuerpo social equivaldría a una pura y simple *negación de la vida*. Este corolario es decisivo, a la hora de interpretar su visión general de las cosas.

Reconocer que cualquier otra voluntad es igual –en dignidad– a la mía, piensa Nietzsche, fue siempre una pésima y abominable norma que conducía, en último término, a la disolución de los instintos poderosos. “Hay que ir hasta el fondo de las cosas y prohibirse toda debilidad; vivir es *esencialmente* despojar, herir, violentar lo que es extraño y débil, oprimirlo, imponerle duramente sus formas propias, asimilarlo o, al menos (esta es la solución más suave), explotarlo [...] Será necesariamente voluntad de poder encarnada, deseará crecer y extenderse, acaparar, conquistar la preponderancia, no por yo no sé qué razones morales o inmorales, sino porque *vive*, y porque la vida, precisamente, es voluntad de poder”³¹. Este texto, tantas veces citado y al mismo tiempo tantas veces dulcificado por hermenéuticas de la exoneración, es extremadamente elocuente. Y aquí nos parece del todo razonable, más allá de la hipérbole, una observación que el historiador alemán Franz Mehring, políticamente vinculado a la socialdemocracia, vertió contra Nietzsche ya en 1891. “En frases tan lapidarias escribe Nietzsche la filosofía del capitalismo”³². Pese a que muchos hayan pretendido que su filosofía se deslizaba por etéreas cumbres espirituales, en realidad su quehacer teórico también se enfrascaba en la sucia materialidad del devenir histórico de su tiempo. “Y justo por eso *Más allá del bien y del mal*, que ni filosóficamente ni científicamente vale la tinta con la que ha sido escrito, tiene en lo sociopolítico

29 LOSURDO, Domenico, *Nietzsche, il ribelle aristocratico. Biografia intellettuale e bilancio critico*. Turín: Bollati Boringhieri, 2004.

30 SPENCER, Herbert. *La justicia*. Madrid: La España Moderna, 1904.

31 NIETZSCHE, Friedrich. *Más allá del bien y del mal, op. cit.*, 280.

32 MEHRING, Franz. “Sobre Nietzsche”. En MORALEJA, Alfonso (Coord.), *Nietzsche y la “gran política”. Antídotos y venenos del pensamiento nietzscheano*. Madrid: Cuaderno Gris, Departamento de Filosofía de la Facultad de Filosofía y Letras de la Universidad Autónoma de Madrid, 2001, 47-64, 51.

una gran importancia sistemática”³³. Sin compartir su áspera denostación de la que nosotros sí consideramos una de las obras filosóficas más potentes de la historia occidental, concordamos con Mehring en su observación de que la filosofía nietzscheana tiene una importancia sociopolítica “sistemática”. Algo, esto último, que posteriormente fue negado con asiduidad.

El crítico y filósofo danés Georg Brandes, que mantuvo correspondencia personal con Nietzsche desde finales de 1887 y a lo largo de 1888 (esto es, casi hasta el mismo momento de su brote de salvaje megalomanía y del subsecuente colapso nervioso), sostenía que su filosofía debería ser definida como una suerte de “radicalismo aristocrático”. Y así se lo hizo saber en una carta. Nietzsche, en otra misiva escrita desde Niza en diciembre de 1887, le respondería amistosamente: “La expresión «radicalismo aristocrático», que usted me dirige, me agrada. Permítame decirle que es lo más fuerte que de mí se ha dicho”³⁴. ¿Y qué significaba para Brandes semejante concepto? En primer lugar, una “dureza” moral que tenía que oponerse a toda esa sensiblería “inglesa” y utilitaria que pretendía extender la felicidad para “el mayor número” posible de seres humanos; o, de forma correlativa, reducir el dolor a “cuantos más hombres” mejor. Esa predominancia meramente cuantitativa de la masa plebeya le resultaba, a Nietzsche, completamente insoportable³⁵. ¿No debiera reservarse la gran felicidad solo a los mejores, a los genios sobresalientes capaces de alzarse e imponerse en la gloria de su propia fuerza? Brandes, que simpatizaba con las ideas nietzscheanas (divulgó sus ideas en Dinamarca y en toda Escandinavia, estando Nietzsche todavía vivo y cuerdo), compartía plenamente aquella tesis según la cual la moral propia de los esclavos germinaba a partir del puro resentimiento: la impotencia se había convertido en bondad; la debilidad se había transformado en virtud. A “no poder vengarse” lo denominaron “no querer vengarse”. Nunca hubo jugada más capciosa. Inventaron la “Justicia”, sacra palabra escrita con mayúscula. Idearon la “Justicia en sí”. ¡Esta había sido la monumental y astuta operación de los débiles! De tal forma, con semejante maniobra, se había consumado una secular derrota del espíritu aristocrático.

El asunto de la justicia, contrariamente a lo sostenido por muchos, era esencial en el pensamiento de Nietzsche. “No se puede, de ningún modo, hablar de injusticias cometidas por la casta superior. Porque no existe justicia ni injusticia en sí. Para él, el hecho de matar a alguien, de aplastarlo, de explotarlo, de anularlo, no constituye una injusticia, no puede constituirlo, por no ser la vida misma en

33 *Ibid.*, 52.

34 BRANDES, Georg. *Un ensayo sobre el radicalismo aristocrático*. México: Sexto Piso, 2008, 78.

35 *Ibid.*, 36.

su esencia, en sus funciones primordiales, más que destrucción, explotación, reducción a la nada. El derecho no podrá constituir jamás otra cosa que un estado de excepción; es decir, una restricción impuesta al instinto esencialmente vital, cuyo fin es el poder”³⁶. Creemos que Brandes comprendió bien lo que se estaba jugando con semejantes propuestas; lo comprendió mucho mejor, de hecho, que ulteriores hermeneutas del siglo XX. Y nos parece digna de mención esa alusión al derecho, pues éste quedaría comprendido como otra de las cristalizaciones de esa “astucia” manejada por los débiles. El derecho de un orden político democrático, en la visión nietzscheana, no sería más que una vil artimaña de supervivencia tejida por las castas inferiores. El derecho, en suma, sería un “estado de excepción” opuesto a la esencia de la vida, que no es sino eterna lucha. Una vida que en Nietzsche ya no equivalía meramente a “voluntad de vivir” (*Wille sum Leben*), a “lucha por la existencia” (*Struggle for life*) o a un conatus que desea “permanecer en el ser” casi por mera inercia. No; la vida es mucho más que todo eso. Porque en el pensamiento del Dioniso alemán no se trata únicamente de sobrevivir a cualquier precio, de forma adaptativa. “Según él, la lucha se establece, no por la vida, la vida desnuda, sino por el poder”³⁷. La vida es, en su consistencia última, una pujanza irreprimible hacia el acrecentamiento del propio poder. Lo absolutamente determinante, aquello en torno a lo cual bascula todo lo demás, es precisamente el *Macht*.

Las ilusiones progresistas quedan completamente aniquiladas en la filosofía nietzscheana, calcinadas de raíz. Su amigo epistolar lo expresó de nuevo con plástica nitidez, sin ornamentos: “El gozo que inspira a Nietzsche la lucha en tanto que lucha, completamente opuesto a la manera de ver inspirada por el humanitarismo moderno, es muy característico. Para él, la magnitud de un progreso se mide por la importancia de los sacrificios que exige. Una higiene que mantiene vivos a millones de seres débiles e inútiles, que hubieran debido morir, no es un progreso verdadero. Un tranquilo bienestar medio asegurado a un número lo mayor posible de criaturas miserables, que en nuestros días se llaman seres humanos, no representa tampoco un progreso grande y verdadero”³⁸. Las utopías del porvenir con las que Nietzsche soñaba tenían que ver con la “cría” de un género de hombres superiores, dominantes y grandiosos, pero jamás con el progreso de las “mayorías sociales”. Deseaba hombres capaces de movilizar instintos salvajes, sin la mediación abúlica del “remordimiento” y sin la introyección putrefacta del “sentimiento de culpa” (ese mal que se infligen a sí mismos los que odian la vida, los que no soportan su belleza doliente, los que no son capaces de asumir su ten-

36 *Ibid.*, 48.

37 *Idem.*

38 *Ibid.*, 49.

sión violenta e inmisericorde). Hombres, así lo anhelaba, que supieran restaurar la bondad del “afán de dominio”, ese mismo que fue valorado o estigmatizado como “malvado” por parte de una moral demasiado aberrante que se tornó hegemónica en la cultura occidental.

4. NO SE PUEDE MODIFICAR LO INMODIFICABLE

Si, para socialistas y anarquistas, la explotación se encuentra asentada en determinadas condiciones históricas de tipo socioeconómico, que son contingentes y por ende susceptibles de ser transformadas o, al menos, corregidas en sus efectos más perniciosos, para Nietzsche la explotación de unos a manos de otros *no* es fruto de una sociedad desequilibrada, injusta o corrompida. No, esa explotación es para nuestro filósofo un dato *inamovible*, un factor natural en el orden eterno del mundo, algo que pertenece a la esencia misma de la *vida*. Pero debemos comprender que dicha concepción no deja de ser paradójica en alguien que había venido, esgrimiendo su martillo genealógico, a “historizarlo” todo. Porque resulta que no; resulta que nos las habemos con una realidad cuya naturaleza no es contingente o fruto de una cierta disposición de estructuras históricas: la explotación y la dominación. Ellas constituyen un hecho bruto y primordial, insuprimible.

Nietzsche sumerge todos los dispositivos culturales en un proceso de fluidificación, porque nada es eterno y todo ha ido emergiendo y cristalizando con el paso del tiempo...a excepción de las realidades mencionadas. Resulta que la explotación y el dominio constituyen una solidez eterna, un fondo primordial no-histórico, una ley de hierro transhistórica que siempre ha estado presente y que siempre lo estará. “Forma parte del concepto de lo viviente el que tenga que crecer, el que aumente su fuerza y en consecuencia el que tenga que asimilar fuerzas extrañas. Se habla, bajo la ofuscación producida por narcóticos morales, de un derecho del individuo a defenderse. En el mismo sentido, también es lícito hablar de su derecho al ataque, pues *ambos* –y el segundo no menos que el primero– son necesidades para todo lo viviente. El egoísmo agresivo y el defensivo no son cosas que dependan de la elección o, en absoluto, de la «libre voluntad», sino que pertenecen a la *fatalidad* de la vida misma”³⁹. Tal descripción de la “lógica profunda de lo viviente” adquiere connotaciones profundas y fatales, porque no se trata de un rasgo meramente “psicológico” o “antropológico”: es la *vida*, que funciona así, nos guste o no nos guste. De hecho, el asunto adquiere tintes ontológicos. La única especificidad de las sociedades modernas, señalará, es que en ellas el “afán

39 NIETZSCHE, Friedrich. *Fragmentos póstumos sobre política*. Madrid: Trotta, 2004, 192.

de poder” (que es en lo que la vida consiste de forma ineluctable) ha cristalizado *ahora* en el afán de acumular dinero y riqueza material⁴⁰.

Prestemos minuciosa atención, pues Nietzsche nos está diciendo que cuando el anarquista, “como vocero de capas *en declive* de la sociedad”, reclama con indignación más “justicia” o “igualdad de derechos”, en realidad no están sabiendo comprender *por qué* tienen ellos que sufrir⁴¹. Esas amargas quejas, que no son más que una pulsión vengativa, en realidad proceden de la debilidad y la impotencia. El revolucionario, asevera Nietzsche, atribuye el malestar y el sufrimiento a otros; es decir, busca culpables para explicar el propio padecimiento. En ese sentido, cristianos, anarquistas y socialistas comparten una misma estructura moral e intelectual. “El cristiano y el anarquista –ambos son *décadents*. –Pero incluso cuando el cristiano condena, denigra y mancilla al mundo, lo hace a partir del mismo instinto con el que el obrero socialista condena, denigra y mancilla a la *sociedad*: el «juicio final» mismo es también el dulce consuelo de la venganza –la revolución, como la aguarda el trabajador socialista, sólo que pensada como algo más lejano...El mismo «más allá»”⁴². En otro lugar, y con extremo sarcasmo, Nietzsche abunda en la misma tesis, aseverando que todos aquellos que buscan “profundas injusticias” en el plano social no son sino impotentes resentidos cuyo único recurso, para no admitir su ignominiosa debilidad y su *autoculpable* postulación, es buscar una causalidad extrínseca que dé cuenta de sus males. Desde la perspectiva nietzscheana, por lo tanto, es enteramente improcedente buscar una responsabilidad en las estructuras políticas y económicas del orden vigente, a la hora de explicar el sufrimiento y la condición miserable de amplias capas de la población. Los socialistas recurrirían siempre al “instinto cristiano”; ese sería su más fino ardid. También el neoliberalismo, como bien sabemos a estas alturas de la historia, ha querido convencernos de que el individuo es siempre el *único* responsable de su propia situación. Nietzsche nos está espetando en la cara que el “sistema” no es el causante del padecimiento atroz de las capas populares; el sistema económico-político jamás es responsable del sojuzgamiento de las masas, ni de la mísera pobreza sufrida por millones de seres humanos. Culpar al orden social de todos esos “males”, añade, es propio de criaturas enfermas, impotentes y resentidas.

Por otro lado, en *Zur Genealogie der Moral* se observa una estrecha conexión entre la moral aristocrática y la potencia física predispuesta a la lucha. “Los juicios de valor caballeresco-aristocráticos tienen como condición previa una

40 NIETZSCHE, Friedrich. *Aurora*, op. cit., 186.

41 NIETZSCHE, Friedrich. *Crepúsculo de los ídolos o cómo se filosofa con el martillo*. Madrid: Biblioteca Nueva, 2002, 113.

42 *Ibid.*, 114.

poderosa corporalidad, una salud floreciente, rica, incluso desbordante, junto con lo que es la causa de su conservación: la guerra, la aventura, la caza, la danza, los juegos de lucha y en general cuanto implique un obrar fuerte, libre y con ánimo alegre⁴³. Esa alegría de la potencia vital, esa nobleza que combatía y guerreaba sin cortapisas, fue calumniada por la espiritualidad de los “sacerdotes”, consumándose así una transvaloración infame: los débiles consiguieron que sobre esos valores guerreros pesara la sombra de lo “malo”. Semejante transmutación procedería del “odio de la impotencia”, una jugada magistral ejecutada por esos enfermos que no soportaban aquella esplendorosa y cuántuple ecuación por medio de la cual lo bueno, lo noble, lo poderoso, lo bello y lo feliz aparecían como valores enteramente equivalentes, consustanciales los unos a los otros. Y a partir de entonces, una moral transmutada empezó a decir: “[...] «sólo los desgraciados son buenos; sólo los pobres, impotentes y bajos son buenos; los que sufren, los indigentes, los enfermos, los feos son los únicos píos, los únicos bienaventurados a los ojos de Dios, sólo para ellos hay bienaventuranza, mientras que vosotros, los nobles y potentes, vosotros sois por toda la eternidad los malvados, los crueles, los lascivos, los insaciables, los impíos, y vosotros seréis también eternamente los desdichados, los malditos y condenados!»”⁴⁴. Nietzsche, a renglón seguido, asevera de forma enigmática: “Ya se sabe *quién* ha heredado esta transvaloración judía...”⁴⁵. Se refiere al movimiento democrático en general y al movimiento socialista en particular, como mórbidos herederos seculares de aquella primera decadencia. Suscribimos, en ese sentido, la valoración general de Georg Lukács. “Lo coherente y lo sistemático radica precisamente en el contenido social de su pensamiento: en la lucha contra el socialismo”⁴⁶. Su interpretación de la obra nietzscheana es, en algunos aspectos, reduccionista. Sin embargo, consideramos que el filósofo húngaro acertó en lo esencial.

La metáfora zoológica es verdaderamente sintomática, y debemos escucharla: “Que los corderos vean con malos ojos a las grandes aves rapaces nada tiene de extraño: sólo que eso no es razón alguna para tomarles a mal a las grandes aves rapaces que cojan pequeños corderos [...] Exigir de la fuerza que *no* se manifieste como fuerza, que *no* sea un querer debelar, un querer someter, un querer enseñorearse, una sed de enemigos y resistencias y triunfos, es tan absurdo como exigir de la debilidad que se manifieste como fuerza”⁴⁷. Corregir ese desequilibrio

43 NIETZSCHE, Friedrich. *La genealogía de la moral. Un escrito polémico*. Madrid: Edaf, 2000, 63-64.

44 *Ibid.*, 65.

45 *Ibid.*

46 LUKÁCS, Georg. *El asalto a la razón. La trayectoria del irracionalismo desde Schelling hasta Hitler*. Barcelona: Grijalbo, 1968, 322.

47 NIETZSCHE, Friedrich. *La genealogía de la moral. Un escrito polémico*, op. cit., 80.

de fuerzas, a través de una torticera maniobra por medio de la cual se consigue imputar “maldad” al inexorable hecho de que un ave rapaz *debe* atrapar con sus garras un cordero, es un ardid de los débiles corderos. Los oprimidos y los pisoteados, dice Nietzsche, llevados de “la astucia de la impotencia”, consiguieron establecer que “bueno” era el que no ataca, el que no hiere a nadie, o el que no juega para ganar y destruir al contrincante⁴⁸. De su extremada debilidad –de su impotencia vital– emergió esa “moralidad de corderillos”, esa moral de esclavos: convirtieron la debilidad en un mérito, en un signo de bondad.

Seres vengativos y reactivos, socialistas de toda laya, urdieron con inquina “la conjura de los dolientes contra los bien plantados y victoriosos”⁴⁹. ¿Qué es lo que pretenden? “Andan por entre nosotros como reproches en persona, como advertencias que se nos hacen, como si la salud, la gallardía, la fuerza, el orgullo, la sensación de poder fuesen de suyo cosas viciosas por las que algún día se tuviese que hacer penitencia, una amarga penitencia”⁵⁰. Estos envenenadores de la vida siempre llevan en la boca la palabra “justicia”, dice Nietzsche, pero no representan sino la lucha de lo enfermo contra lo sano: “Todos estos son hombres del resentimiento, todos estos desafortunados y agusanados fisiológicamente” inyectan dosis de “mala conciencia” en los poderosos, que sucumben cuando empiezan a avergonzarse de su propia fuerza y de su propia felicidad, que claudican cuando empiezan a cargar en su propia conciencia el sufrimiento de los débiles y derrotados⁵¹. ¡Pero la fuerza victoriosa, así lo grita una y otra vez, jamás debiera rebajarse a un estado de conmiseración auxiliadora! El privilegiado puede y debe ganar, sin ejercer de “médico” de lo más bajo. El poderoso debe mantenerse a una más que prudente *distancia* de lo enfermo y malogrado; no puede ensuciarse ni reblandecerse con elementos de “compasión”⁵². Solo el socialismo, como un pestífero cristianismo modernizado, y también cierto humanitarismo burgués reformista, pueden aspirar a *reducir* el carácter de lucha inmisericorde que impregna la esencia misma de la vida. La cosmovisión nietzscheana se acopla ergonómicamente a la cosmovisión del liberalismo económico más extremo, toda vez que únicamente en el marco agonístico de un radical *laissez-faire* –en el que prevalezca la más cruda competencia y, por ende, la victoria inapelable del más fuerte– pudiera encauzarse, quizás, esa despiadada voluntad de poder que late, según Nietzsche, en el fondo terrible de la vida.

48 *Ibid.*, 81.

49 *Ibid.*, 193.

50 *Ibid.*, 194.

51 *Ibid.*, 195.

52 *Ibid.*, 196.

Ese mismo esquema, no lo perdamos de vista, debe extenderse a las colectividades. “En esto, lo mismo vale si se toma en consideración un individuo o un cuerpo viviente o una «sociedad» que se esfuerza por ascender. El derecho a castigar (o la autodefensa social) sólo es válido en el fondo por un mal uso de la palabra «derecho»: un derecho se adquiere a través de contratos, pero defenderse y protegerse no descansa sobre la base de un contrato. Cuando menos, con muy buen sentido un pueblo está autorizado a calificar de derecho –acaso el derecho al crecimiento– su necesidad de conquista, su avidez de poder, ya sean satisfechas mediante las armas, el litigio, el tráfico comercial o la colonización. Una sociedad que repudia definitivamente su *instinto* para la guerra y la conquista es una sociedad decadente, que está madura para la democracia y el régimen de tenderos... Sin duda, en la mayor parte de los casos los convenios que aseguran la paz son meros anestésicos”⁵³. El texto es harto elocuente, descarnado. Nietzsche extiende el principio último de todo lo viviente, que es pura voluntad de poder, a esos organismos vivos o “cuerpos vivientes” llamados sociedades. Una colectividad, por lo tanto, tiene “derecho” a incrementar sus fuerzas; tiene “derecho” a crecer vitalmente, a expandirse. Pero no hablamos de relaciones jurídicas o económicas; son, más bien, relaciones de pura fuerza. La colectividad más fuerte tiene derecho a dominar, y por ello puede y debe absorber a las débiles (anexionarlas, quizá). Conquistar mediante la guerra, colonizar violentamente, apoderarse de otras colectividades; ello no depende de una decisión racional, de un cálculo frío, sino que forma parte de la naturaleza misma de todo lo vivo.

Su desasosiego, al tomar conciencia de la “decadencia europea”, tiene que ver con esa concepción vitalista. Entiende, desde dicho esquema, que Europa es una suerte de organismo que ha perdido sus instintos; cansada, atrofiada y exánime, ya no es capaz de mostrar una verdadera voluntad de dominio. Por ello, en un momento dado, Nietzsche llega a decir que el engrandecimiento de Rusia sería deseable, pues ante semejante amenaza exógena quizás Europa por fin reaccionaría, despertaría y volvería a *revitalizar* su voluntad de poder. Es decir, sueña con “una agravación tal del peligro ruso que Europa tuviese que convertirse ella misma en amenazadora, es decir, forjarse, mediante el órgano de una nueva casta dominante europea, *una voluntad única*, una prolongada y formidable voluntad capaz de perseguir un fin a la escala de varios milenios; así pondría fin a la comedia que ya ha durado demasiado, la división de Europa en pequeños estados y su pelaje abigarrado de dinastías y democracias. El tiempo de la pequeña política ha acabado ya; el próximo siglo ya acarreará la lucha por la dominación universal, la *obligación* de hacer una gran política”⁵⁴. Está proponiendo una suerte de Imperio

53 NIETZSCHE, Friedrich. *Fragmentos póstumos sobre política*, op. cit., 192.

54 NIETZSCHE, Friedrich. *Más allá del bien y de mal*, op. cit., 199.

Europeo todopoderoso, expansivo y beligerante. Un Imperio que nada tendría que ver con la democracia, desde luego. Que la “pequeña política” dé paso a la “gran política”. *Re-vitalizar* Europa, esto es, insuflarla de un nuevo poderío político, agresivo y atacante.

Un asunto, el de la “gran política”, que también aparece en un fragmento de diciembre de 1888. En él, Nietzsche apuntaba que la fisiología habría de convertirse en ama y señora de todas las otras cuestiones; y con ese mismo criterio, habría de levantarse un poder lo suficientemente fuerte, que fuera capaz de *crear* una nueva humanidad. Un poder que mostrara implacable dureza con todo lo degenerado y parasitario, con todo lo que conduce a la ruina vital... Crear un “partido de la vida” para llevar a término esa “gran política”⁵⁵. Revitalizar Europa, por ende. El pensamiento de Nietzsche terminó siendo (tras superar su inicial fase teutómana o germanófila) de corte “paneuropeísta”. Es más, Félix Duque enfatiza que la voz nietzscheana, finalmente, se levantó *contra* Alemania y *a favor* de Europa⁵⁶. Soñaba con una Europa unida, en la que por fin quedasen sepultadas las rencillas intestinas; una Europa que dejase atrás los chovinismos patrioterros anclados en la glorificación del terruño local⁵⁷. Aunque, apunta Duque, debe reconocerse –en honor a la verdad– que Nietzsche jamás se desprendió por completo de ciertas referencias a la higiene racial⁵⁸. Y, además, nunca abandonó del todo su querencia por la “caracterología” de los diferentes espíritus nacionales⁵⁹. En ese sentido, el “elemento inglés” le resultará muy aborrecible durante toda su vida mientras que, con el “elemento francés”, mostrará mayor ambivalencia. La cultura meridional de Italia o España le provoca apuntes más o menos pintorescos. Y el “elemento alemán”, que en muchos momentos de juvenil romanticismo pudo exaltar de forma chovinista, le inspirará en sus últimos años pocas simpatías. En *Ecce homo* señalará que el “espíritu alemán” es una indigestión que no da lugar a nada; y añadirá pocas páginas después que, a donde llega Alemania, se corrompe la cultura⁶⁰. Dirá, en ese mismo texto autobiográfico, que sus instintos más profundos son extraños a todo lo alemán⁶¹. Pero estaba mintiendo, al decir tal cosa, pues hubo un tiempo en el que sí creyó ciegamente en la preeminencia de la cultura alemana.

55 NIETZSCHE, Friedrich. *Fragmentos póstumos (1885-1889)*. Volumen IV, op. cit., 773-774.

56 DUQUE, Félix. *Los buenos europeos. Hacia una filosofía de la Europa contemporánea*. Oviedo: Nobel, 2003, 103-105.

57 *Ibid.*, 86-91.

58 *Ibid.*, 90.

59 *Ibid.*, 91-95.

60 NIETZSCHE, Friedrich. *Ecce homo*. Madrid: Alianza, 2002, 43 y 49.

61 *Ibid.*, 52 y 132.

5. ¿QUÉ RÉGIMEN SOCIOPOLÍTICO SE REQUIERE PARA LA EMERGENCIA DEL AUTÉNTICO GENIO ARTÍSTICO?

Es crucial, a la hora de interpretar y valorar el conjunto de la intervención nietzscheana, comprender adecuadamente el asunto del arte y el poder⁶². En muchos de sus fragmentos póstumos, Nietzsche aparece extremada y crudamente diáfano en sus planteamientos, y puede sostener sin ambages que los auténticos despliegues artísticos *únicamente* tendrán ocasión de darse allí donde la gran mayoría de la comunidad humana se encontrase atornillada a la ignominia del trabajo. Para que el “tipo superior de hombre” pueda medrar, se precisa de un máximum de explotación de los hombres bajos e inferiores. Pero mucho cuidado con este asunto, porque estas proposiciones no son meramente descriptivas, esto es, un modo de comprender lo que fue el pasado de la humanidad; muy al contrario, son al mismo tiempo aseveraciones normativas acerca de las condiciones que toda sociedad *debe* cumplir para que dentro de ella pueda germinar una gran cultura. En 1871 lo decía con absoluta rotundidad: “Para que exista el fundamento de un gran desarrollo artístico, la inmensa mayoría tiene que ser sometida, mediante la esclavitud, al servicio de una minoría [...] A sus expensas, por medio de un trabajo excedente, aquella clase privilegiada es apartada de la lucha por la existencia, para crear entonces un nuevo mundo de necesidades. Según esto, hemos de comprender adecuadamente, cuando caracterizamos las horrosas condiciones fundamentales de toda formación, que la esclavitud pertenece a la esencia de una cultura [...] La pobreza de la masa que vive penosamente se ha de aumentar todavía para hacer posible a un número de hombres olímpicos la producción del mundo del arte. Aquí se encuentra la fuente de aquella rabia mal disimulada que los comunistas y los socialistas, y también sus descendientes más pálidos, la raza blanca de los liberales de todas las épocas, han alimentado contra el arte, pero también contra la antigüedad clásica”⁶³. Este pasaje es un excelente botón de muestra.

Una masa humana sobreexplotada es *imprescindible* para el surgimiento esplendoroso de la alta cultura, aseveraba Nietzsche. Y apostillaba, sin escrúpulo aparente, que la masa de los que “viven penosamente” debe aumentar. Porque su aristocratismo cultural, en realidad, nos hablaba de la nostalgia por una época en la que los valores culturales constituían un privilegio de clase, un excelso reino del que estaban excluidas (y deberían seguir estándolo por siempre) las masas plebeyas. Los antiguos griegos eran lo suficientemente honrados como para reconocer

62 DE SANTIAGO GUERVÓS, Luis Enrique. *Arte y poder. Aproximación a la estética de Nietzsche*. Madrid: Trotta, 2003.

63 NIETZSCHE, Friedrich. *Fragmentos póstumos sobre política, op. cit.*, 85.

abiertamente la necesidad de la esclavitud, el único humus a partir del cual podían crecer las bellas flores de la cultura. Sin el concurso de los esclavos, no podría generarse esa energía excedente que una élite creativa precisa para alumbrar una cultura inmortal y profunda. Y proclamar, como hacían los revolucionarios modernos, una democratización social destinada a obtener el “bienestar” del mayor número posible, semejante proyecto, nos habría de conducir a un infecto lodazal de mediocridad del que jamás volverían a surgir genios o grandes personalidades. Nietzsche estaba diciendo algo tan grave como lo siguiente: la justicia distributiva y la cultura genuina son *incompatibles*. ¿Cómo soslayar, en nuestra lectura de Nietzsche, un dictamen tan demoledor? No podemos ignorar todo este asunto, si es que deseamos enjuiciar cabalmente la intervención estético-política de Nietzsche.

En la primera de sus *Consideraciones intempestivas* había arremetido duramente contra David Strauss (un pensador igualmente odiado por Wagner; de hecho, resulta altamente probable que fuera este quien encargara a Nietzsche la elaboración del destructivo texto) porque encarnaba a la perfección ese talante ilustrado, conformista, acomodaticio y suavemente pequeñoburgués que, glorificando el trabajo y el cálculo, había convertido el mundo en una fría y gigantesca máquina. Es verdad que el jovencísimo Nietzsche, en su camino hacia el ateísmo, fue muy influido por la obra *La vida de Jesús*, del propio Strauss. Pero este pensador se había convertido, ahora, en la perfecta encarnación del “erudito filisteo”; en el prototipo espiritualmente decadente del Imperio recién fundado, pues este avanzaba a pasos agigantados por la senda de la modernización capitalista y de la industrialización⁶⁴. Las esperanzas acerca de un resurgir trágico del verdadero espíritu alemán, se convirtieron en cenizas; y ello, paradójicamente, tras la victoria militar sobre Francia. Una victoria en el campo de batalla que en absoluto se correspondió con un resurgir cultural de “lo alemán”. La fuerza portentosa del soldado prusiano no tuvo su equivalente en el plano espiritual. Emponzoñados de utilitarismo progresista, los artífices de tal mundo habían hecho del arte un ornamento hermoso, un accesorio aliviador...pero secundario. Nietzsche, que en esto mostraba mucha consonancia con las sensibilidades románticas, no podía soportar que el arte ocupara un lugar tan relegado y tan mezquino. Porque el arte era un abismo, no un pasatiempo o un entretenimiento diseñado para distraer las horas muertas del descanso dominical burgués; el arte, en su concepción, era una suerte de compuerta que, al abrirse, podía conectarnos con las profundidades trágicas y dolientes del ser. La cultura, para su florecimiento y rejuvenecimiento, necesitaba sumergirse de nuevo en ese elemento dionisiaco-heraclitiano y trágico-poético.

64 NIETZSCHE, Friedrich. *Consideraciones intempestivas*, 1. David Strauss, el confesor y el escritor. Madrid: Alianza, 2006.

Los coqueteos de Nietzsche con “lo otro de la razón” aparecen en demasiadas ocasiones; más en sus obras juveniles, es cierto, pero también en sus textos de madurez. No se produce, en ese sentido, un “corte epistemológico” tan abrupto entre el Nietzsche romántico y el posromántico.

La gran cultura requiere primordialmente de un elemento de impostergable crueldad: la esclavitud. El modelo social de la Antigüedad clásica, en ese sentido, es ensalzado sin pudor. A Nietzsche le horroriza la idea de que los esclavos tomen conciencia de su situación, y lo escribe sin tapujos: “Nada hay tan horrible como un estado bárbaro de esclavos que ha aprendido a considerar su existencia como una injusticia y se dispone a tomar la venganza no sólo para sí, sino también para todas las generaciones”⁶⁵. Espartaco sería una de las máximas pesadillas de Nietzsche. Pero debemos entenderlo también en clave contemporánea, porque el verdadero miedo se lo producen todos los “Espartacos” modernos, esto es, las masas obreras organizadas y portadoras de una férrea conciencia política. Esa irrupción es una crisis amenazante, y los culpables son todas aquellas sabandijas “demócratas” que han inoculado en las multitudes explotadas un atroz veneno: la conciencia de que estaban sufriendo como consecuencia de una injusticia objetiva contra la cual podían rebelarse. Pretender aliviar el dolor de las masas es oponerse al devenir doloroso, cruel y trágico de la vida; pero es, ante todo, socavar la base imprescindible de un orden social en el que aún puedan florecer grandes genios. El gran arte solo emerge sobre las espaldas explotadas de la gran mayoría. También la guerra era, a su juicio, un medio fabuloso para retomar el contacto con aquel magma trágico-poético. Podemos detectar en su destemplado pensamiento una extraña y subterránea conexión entre la guerra y el gran arte. Ambas realidades nos conectarían con las profundidades “elementales” de la vida, pues ambas podrían lanzarnos a una disolución extática y dionisiaca.

Ese pasaje nietzscheano evoca la inolvidable séptima tesis sobre la historia de Walter Benjamin, aquella en la que señalaba que detrás de todo “documento de cultura” se esconde un “documento de barbarie”. Pero, mientras Benjamin subrayaba la necesidad de “cepillar la historia a contrapelo”, precisamente para restituir la voz a todos aquellos que fueron derrotados, masacrados y silenciados tras las mascaradas de la “alta cultura”, Nietzsche *celebraba* que eso hubiera sucedido así; es más, prescribía que así *debiera* ser siempre. Porque los humildes deben permanecer como una mayoría silenciosa y sumisa; su mugre plebeya no debe tratar de escalar posiciones en la jerarquía social y cultural. Sin la existencia de una masa esclavizada por el trabajo extenuante, no habrá ocasión para el surgimiento de esa

65 SAFRANSKI, Rüdiger. *Nietzsche. Biografía de su pensamiento*. Barcelona: Círculo de Lectores, 2001, 74.

gran cultura que solo una élite –levantada sobre el lomo fatigado de una mayoría inculta y trabajadora– puede producir.

El escritor polaco Tadeusz Borowski lo expresaría de forma desgarrada en *Nuestro hogar es Auschwitz*: “Ahora sé qué elevado precio pagaron otros en la Antigüedad. ¡Qué crimen espantoso fueron las pirámides de Egipto, los templos y estatuas griegas! ¡Cuánta sangre tuvo que derramarse sobre las calzadas romanas, las fortificaciones fronterizas y los edificios de las ciudades!”⁶⁶. Pero ese “elevado precio”, pensaba Nietzsche, siempre debe pagarse. ¡Siempre! Porque no hay verdadera cultura sin el dolor fatigado de las mayorías trabajadoras.

En el otoño de 1877 lo dejaba consignado de forma explícita, cuando escribía que “puesto que tiene que hacerse mucho trabajo duro y rudo, también hay que conservar a los hombres que se someten al mismo, esto es, en la medida en que las máquinas no *pueden* ahorrar este trabajo. Si la necesidad y el refinamiento de una formación superior penetran en la clase obrera, entonces ésta ya no podrá hacer ese trabajo sin sufrir desmesuradamente. Un obrero desarrollado hasta tal punto aspira al ocio y no exige ningún alivio del trabajo, sino la liberación del mismo, es decir: quiere cargárselo a otro. Quizá podría pensarse en una satisfacción de sus deseos y en una importación masiva de pueblos bárbaros de Asia y África, de modo que el mundo civilizado tuviese perpetuamente a su servicio al mundo incivilizado y de esta manera se considerase la falta de cultura precisamente como obligación de servidumbre”⁶⁷. El texto es, se mire como se mire, de una sinceridad escalofriante. No sabemos de qué manera podría una “izquierda nietzscheana” justificar alegóricamente estas despiadadas palabras. Señala la impertinencia de ofrecer formación educativa a la clase obrera europea, toda vez que alguien tendrá que hacer los trabajos más “duros y rudos” (esos que ni siquiera las máquinas pueden hacer); y, obviamente, un trabajador con un nivel educativo más alto albergará necesariamente otras aspiraciones; surgirán en él nuevos horizontes. Semejante posibilidad horroriza a Nietzsche, toda vez que se requiere de una masa de trabajadores incultos que hagan el trabajo duro sin levantar la frente de suelo. Decía en otro fragmento póstumo (noviembre de 1887-marzo de 1888) que no terminaba de comprender qué pretendía hacerse con el trabajador europeo; si se desea una cierta meta, se han de querer también los medios para lograrla. Esto es, si los esclavos son imprescindibles, ¿cómo diablos han tenido la infeliz ocurrencia de otorgar una mayor formación a esos obreros que, por mor de

66 BOROWSKI, Tadeusz. *Nuestro hogar es Auschwitz*. Barcelona: Alba, 2004, 59.

67 NIETZSCHE, Friedrich. *Fragmentos póstumos sobre política*, op. cit., 140.

ella, experimentan su situación como menesterosa y, por ende, exigen cada vez más derechos? Así se lamentaba Nietzsche⁶⁸.

Una cantidad muy importante de seres humanos tendrán que cumplir el papel de bestias de carga, para que una sociedad pueda funcionar de manera saludable. Hablamos, por lo tanto, de un clasismo elitista indisimuladamente atroz. En el otoño de 1880 perseveraba en ese *miedo*: “La objeción fundamental contra el socialismo es que quiere instituir la ociosidad para la gente común”⁶⁹. Tal vez por “instituir la ociosidad” se refería Nietzsche a los reclamos del movimiento obrero para reducir esas interminables jornadas laborales de doce y catorce horas... No eduquen a los obreros, advertía, pues entonces ellos aspirarán a liberarse de un trabajo esclavo y embrutecedor; querrán escapar del “reino de la necesidad”, por emplear la expresión de Marx. Tal liberación no debiera producirse jamás, pensaba Nietzsche. Pero la cosa no queda ahí, y nuestro filósofo saca a relucir una miserable mentalidad colonialista (tan extendida, es verdad, en buena parte de la intelectualidad europea decimonónica), señalando que si la clase obrera autóctona se niega ya a seguir realizando esos trabajos sin rechistar (reclamando mejoras salariales y educación para sus hijos, por ejemplo), una buena “solución” sería importar mano de obra (barata y dócil) de los pueblos “bárbaros” de Asia. Si la clase obrera europea se niega a *servir* a la élite europea, entonces las relaciones de dominación habrán de tomar otro rumbo: ahora serán los pueblos incivilizados los que se someterán a los pueblos civilizados.

Además, señalaba Nietzsche que “como espectador, uno se engaña sobre los sufrimientos y las privaciones de las capas inferiores del pueblo, porque los mide involuntariamente de acuerdo con la medida del propio sentimiento, como si uno mismo se pusiese, con su cerebro muy excitable y capaz de sufrir, en la situación de aquéllas. En verdad, los sufrimientos y las privaciones aumentan con el crecimiento de la cultura del individuo. Las capas inferiores son las más obtusas; mejorar su situación significa aumentar su capacidad de sufrir”⁷⁰. Otro texto cruel, en el que se burla incluso de aquella actitud filantrópica que pretende empatizar con el dolor de las “capas inferiores del pueblo”; esa empatía no es posible, porque las élites y el populacho trabajador se hallan en planos enteramente incommunicados e incommensurables. Y, añade, es inútil otorgar cultura a esas clases populares (definitivamente lerdas), porque en tal caso lo único que se lograría es aumentar su capacidad de sufrir. Condenadas a trabajar sin descanso, siendo esa su única función, viven mejor sin tener una formación que les haga comprender

68 NIETZSCHE, Friedrich. *Sabiduría para pasado mañana. Antología de fragmentos póstumos (1869-1889)*. Madrid: Tecnos, 2009, 310-311.

69 NIETZSCHE, Friedrich. *Fragmentos póstumos sobre política, op. cit.*, 146.

70 *Ibid.*, 140.

la fatídica objetividad de su situación. Solo se conseguiría perturbar y desequilibrar la estructura jerárquica del orden socioeconómico, si los sueños igualitaristas y filantrópicos se llevasen a término.

El asunto de la propiedad también ocupaba su atención, desde luego, y se preguntaba de forma retórica si acaso podía llegar a repartirse con justicia. Los socialistas venían preconizando desde hacía tiempo la eliminación de la propiedad o, más exactamente, la transición a una posesión comunitaria. Para ellos, dice Nietzsche, el viejo mandamiento no debiera dictar “No hurtad”, sino “No poseed”. Es muy sencillo proclamar “lotes de tierra iguales para todos”, advierte; sin embargo, “¡cuánta amargura producen las divisiones y los desgarros imprescindibles en ese reparto, y la pérdida de la antigua y venerable propiedad!, ¡cuánta piedad ofendida y sacrificada! Al borrar los límites que se paran las tierras, borramos también la moralidad”⁷¹. Utilizando argumentos económicos de sabor liberal, sostenía Nietzsche que si la tierra fuese *común* se perderían los cultivos y las cosechas, puesto que el hombre solo cuida lo que es *suyo*, lo que posee de manera exclusiva. Por el contrario, la posesión colectiva conduciría de manera inevitable a la pura e inútil inoperancia. La melodía utópica de Platón, que hablaba de la eliminación de la propiedad privada como camino expedito para la eliminación del egoísmo, seguía siendo entonada por los socialistas modernos, sin que estos pudieran comprender lo contraproducente de tales delirios⁷². Aquí observamos, por lo tanto, a un Nietzsche muy en consonancia con el *homo oeconomicus* de la economía política burguesa. En el párrafo 452 de *Humano, demasiado humano* reconoce que la distribución actual de la propiedad es el resultado o la cristalización de innumerables violencias perpetradas durante largo tiempo. El pasado de la civilización es un cúmulo de esclavitudes y engaños. Sin embargo, y al contrario de lo que piensan los socialistas, no podemos destruir o revertir semejante herencia por decreto; no es posible –ni deseable– imponer violentamente un nuevo reparto de la propiedad. Nietzsche, que en muchísimas ocasiones no ha tenido ningún empacho a la hora de glorificar lo violento, recula cuando dicha agresividad puede ponerse al servicio de una transformación radical de las relaciones de propiedad. Porque, además, los no poseedores no albergan ningún privilegio moral con respecto a los poseedores⁷³.

71 NIETZSCHE, Friedrich. *El caminante y su sombra*. Madrid: Marte, 1988, 148.

72 *Ibid.*, 149.

73 NIETZSCHE, Friedrich. *Humano, demasiado humano*, op. cit., 253.

REFERENCIAS BIBLIOGRÁFICAS

- BOROWSKI, Tadeusz, *Nuestro hogar es Auschwitz*. Barcelona: Alba, 2004.
- BRANDES, Georg, *Un ensayo sobre el radicalismo aristocrático*. México: Sexto Piso, 2008.
- DE SANTIAGO GUERVÓS, Luis Enrique, *Arte y poder. Aproximación a la estética de Nietzsche*. Madrid: Trotta, 2003.
- DUQUE, Félix, *Los buenos europeos. Hacia una filosofía de la Europa contemporánea*. Oviedo: Nobel, 2003.
- ESTEBAN ENGUITA, José Emilio, *El joven Nietzsche. Política y tragedia*. Madrid: Biblioteca Nueva, 2004.
- FINK, Eugen, *La filosofía de Nietzsche*. Madrid: Alianza, 1976.
- GONZÁLEZ VARELA, Nicolás, *Nietzsche contra la democracia*. Barcelona: Montesinos, 2010.
- JANZ, Curt Paul, *Friedrich Nietzsche. 2. Los diez años de Basilea (1869-1879)*. Madrid: Alianza, 1987.
- LOSURDO, Domenico, *Nietzsche, il ribelle aristocratico. Biografia intellettuale e bilancio critico*. Turín: Bollati Boringhieri, 2004.
- LUKÁCS, Georg, *El asalto a la razón. La trayectoria del irracionalismo desde Schelling hasta Hitler*. Barcelona: Grijalbo, 1968.
- MANN, Thomas, *Schopenhauer, Nietzsche, Freud*. Madrid: Alianza, 2000.
- MEHRING, Franz, "Sobre Nietzsche". En MORALEJA, Alfonso (Coord.), *Nietzsche y la "gran política". Antídotos y venenos del pensamiento nietzscheano*. Madrid: Cuaderno Gris, Departamento de Filosofía de la Facultad de Filosofía y Letras de la Universidad Autónoma de Madrid, 2001, 47-64.
- NIETZSCHE, Friedrich, *El nacimiento de la tragedia*. Madrid: Alianza, 1973.
- , *El caminante y su sombra*. Madrid: Marte, 1988.
- , *Así habló Zaratustra*. Madrid: Alianza, 1998.
- , *La genealogía de la moral. Un escrito polémico*. Madrid: Edaf, 2000.
- , *Ecce homo*. Madrid: Alianza, 2002.
- , *Crepúsculo de los ídolos o cómo se filosofa con el martillo*. Madrid: Biblioteca Nueva, 2002.
- , *La gaya ciencia*. Madrid: Edaf, 2002.
- , *Fragmentos póstumos sobre política*. Madrid: Trotta, 2004.
- , *Consideraciones intempestivas, I. David Strauss, el confesor y el escritor*. Madrid: Alianza, 2006.
- , *Más allá del bien y de mal. Preludio para una filosofía del futuro*. Madrid: Edaf, 2006.
- , *Fragmentos póstumos (1885-1889). Volumen IV*. Madrid: Tecnos, 2008.
- , *Sabiduría para pasado mañana. Antología de fragmentos póstumos (1869-1889)*. Madrid: Tecnos, 2009.

- , *Sobre el porvenir de nuestras escuelas*. Barcelona: Tusquets, 2009.
 - , *Humano, demasiado humano*. Madrid: Edaf, 2010.
 - , *Aurora. Pensamientos sobre los prejuicios morales*. Madrid: Biblioteca Nueva, 2015.
- NOLTE, Ernst, *Nietzsche y el nietzscheanismo*. Madrid: Alianza, 1995.
- OTTMANN, Henning, *Philosophie und Politik bei Nietzsche*. Berlín: Walter De Gruyter, 1987.
- POLO BLANCO, Jorge, *Anti-Nietzsche. La crueldad de lo político*. Madrid: Taugenit, 2020.
- SAFRANSKI, Rüdiger, *Nietzsche. Biografía de su pensamiento*. Barcelona: Círculo de Lectores, 2001.
- SPENCER, Herbert, *La justicia*. Madrid: La España Moderna, 1904.

SIMONE WEIL: LA ATENCIÓN Y LA ACCIÓN COMO RECONOCIMIENTO DE LA EXISTENCIA

*SIMONE WEIL: ATTENTION AND ACTION
AS RECOGNITION OF EXISTENCE*

MARÍA DEL SOL ROMANO
Doctora Internacional en Filosofía
Doctora Canónica en Filosofía
Departamento de Humanidades
Universidad Iberoamericana
Puebla/México
msromano.09@gmail.com
ORCID: 0000-0001-9800-7240

Recibido: 30/04/2021
Revisado: 27/08/2021
Aceptado: 6/09/2021

Resumen: La extrema desdicha provoca degradación social y desarraigo. Quienes la experimentan, no solamente son inexistentes, anónimos y despreciables para la sociedad, sino que también son cómplices de su propia desdicha. De acuerdo con la filósofa francesa Simone Weil, los desdichados se liberan de su estado de inexistencia cuando alguien les dirige una “atención creadora”. Mediante esta atención, que es inspirada por el amor divino, se reconoce y afirma la existencia del otro, particularmente la de los desdichados. Para ello, es fundamental consentir a renunciar a un “yo” –que siempre está empeñado en ser el centro del mundo–, con el fin de mirar a los desdichados y determinar sus verdaderas necesidades. Es entonces como puede actuarse en favor de ellos por medio de una “acción no actuante”, vaciada de la marca del “yo” y que permita la existencia del otro.

Palabras clave: acción no actuante, amor, atención creadora, consentimiento, desdicha, renuncia.

Abstract: Extreme affliction provokes social degradation and uprootedness. Those who experience it are not only non-existent, anonymous and despicable for society, but they also become accomplices of their own affliction. According to the French philosopher Simone Weil, the afflicted are liberated from their status of non-existence when someone turns a “creative attention” to them. Through this attention, which is inspired by divine love, the existence of the other, particularly that of the afflicted, is recognized and affirmed. To achieve

this, it is essential to consent to renounce the “self” –which is always determined to be the center of the world–, in order to turn the gaze to those who are afflicted and identify their real needs. This is how one can act in favor of them through a “non-acting action”, emptied of the mark of the “self” and which permits the existence of the other.

Keywords: non-acting action, love, creative attention, consent, affliction, renunciation.

“Ce n’est pas seulement l’amour de Dieu qui a pour substance l’attention. L’amour du prochain, dont nous savons que c’est le même amour, est fait de la même substance”¹

INTRODUCCIÓN

La atención es un tema que ocupa un lugar central en el pensamiento y en la vida de la filósofa francesa Simone Weil² (1909-1943). Para ella, la atención no se limita a mantener una postura física, los músculos tensionados o a llevar a cabo un prolongado y agotador trabajo de concentración. Se trata de un concepto que en la autora es original³ y alcanza diferentes niveles desde el discursivo, el intuitivo, hasta un nivel superior que es el de la oración⁴. De este modo, la atención puede ser estudiada desde diferentes perspectivas, además de que –como observa Joël Janiaud– “es la clave de una relación auténtica con lo real y, por consiguiente, se relaciona con importantes cuestiones cognitivas, espirituales, éticas y políticas”⁵.

1 WEIL, S., “Réflexions sur le bon usage des études scolaires en vue de l’amour de Dieu”, [1942]. En: *Œuvres complètes*, t. IV: *Écrits de Marseille*, vol. 1: *Philosophie, science, religion, questions politiques et sociales*. Paris: Gallimard, 2008, 261-262. En adelante se usará la abreviatura OC, tomo, volumen y página. Para este artículo se han utilizado algunas ediciones antiguas de la obra weiliana en francés, así como los tomos y volúmenes de la edición crítica francesa publicados bajo el nombre de *Œuvres complètes*. Hay traducción castellana del texto weiliano aquí citado a cargo de A. López y M. Tabuyo y puede encontrarse en WEIL, S., *A la espera de Dios*. Madrid: Trotta, 2009.

2 En adelante se escribirá a lo largo del texto “Weil”, salvo en los títulos de libros y artículos, así como en las citas textuales que hagan referencia al nombre completo de la autora.

3 A este respecto, debe señalarse que Weil recibió principalmente de su maestro Alain (Émile Chartier) una gran influencia en su pensamiento filosófico. Sin embargo, con el paso del tiempo, varias de las nociones transmitidas por su maestro, como es el caso de la atención, llegarán a tener en Weil un enfoque propio, más extenso y a la vez original.

4 Véase WEIL, S., “Condition première d’un travail non servile”, [1942], OC, IV 1, 427.

5 “L’attention [...] est la clé d’une relation authentique au réel et se voit par là même associée à des enjeux cognitifs, spirituels, éthiques et politiques considérables”. JANIAUD, J., “Simone Weil et l’attention”. En: DELSOL, C. (Ed.), *Simone Weil*. Paris: Cerf, 2009, 169.

Weil dedica su vida a prestar atención a la realidad⁶ y para ella “la atención debería preocupar tanto a los dirigentes políticos como a los educadores”⁷. Y concretamente respecto al ámbito educativo, en sus “Réflexions sur le bon usage des études scolaires en vue de l’amour de Dieu”⁸, la autora destaca la importancia de formar la facultad de atención en los estudios, en donde los ejercicios escolares tienen como principal finalidad el entrenamiento de la atención⁹. Durante los estudios se prepara al espíritu para una atención al vacío, la cual lo conduce a su vez a un nivel de atención superior que se orienta hacia Dios: el de la oración¹⁰. Quien accede a estos niveles de atención, con el tiempo no solamente alcanza un progreso espiritual, sino también moral, lo que comporta un compromiso social y político, especialmente hacia quienes sufren los destructores efectos de la desdicha.

Para Weil la atención y la acción están íntimamente unidas, ya que, como señala en uno de sus escritos de juventud, “si la acción no estuviera precedida por la atención, no sería acción, sino movimiento del cuerpo en el sueño; para el alma sería pasión”¹¹. Más aún, desde la perspectiva weiliana “estar atento a la desdicha, es actuar para impedir la o mitigarla”¹². Esto manifiesta también su incesante preocupación por los problemas relativos a la condición humana que “exige que se haga un esfuerzo para mantener el más elevado nivel de conciencia posible”¹³.

En tal sentido, es preciso subrayar que Weil no solamente reflexionó sobre la atención, sino que también la vivió. Desde su juventud, la autora tuvo siempre una vocación de compromiso hacia el mundo, particularmente hacia los más desdichados, hacia aquellos que a causa de la injusticia social no les es reconocida su

6 “Jusqu’au bout de sa vie, Simone Weil fait effort pour exercer son attention sur la réalité du monde”. JANIAUD, J., “Simone Weil et l’attention”, 179.

7 “L’attention devrait préoccuper les dirigeants politiques ainsi que les éducateurs”. JANIAUD, J., “Simone Weil et l’attention”, 177.

8 Weil escribe este texto para los estudiantes de Montpellier pertenecientes al movimiento Jeunesse étudiante chrétienne (JEC) con quien su amigo, el dominico Joseph-Marie Perrin, estaba en contacto.

9 “Bien qu’aujourd’hui on semble l’ignorer, la formation de la faculté d’attention est le but véritable et presque l’unique intérêt des études. La plupart des exercices scolaires ont aussi un certain intérêt intrinsèque; mais cet intérêt est secondaire”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 256. Del mismo modo, en otro texto recuerda que: “Les exercices scolaires n’ont pas d’autre destination sérieuse que la formation de l’attention”. WEIL, S., “Condition première d’un travail non servile”, [1942], OC, IV 1, 426.

10 Véase ROMANO, M.S., “Simone Weil: atención y oración”. En *Cauriensia*, XI, 2016, 697-711.

11 “Si l’action n’était pas précédée d’attention, elle ne serait pas action, mais mouvement du corps dans le sommeil; pour l’âme, elle serait passion”. WEIL, S., “Que la seule action est la pensée”, [1926], OC, I, 316.

12 JANIAUD, J., *Simone Weil: L’attention et l’action*. Paris: PUF, 2002, 120.

13 “La condition humaine exige que l’on fasse effort pour maintenir un niveau de conscience le plus élevé possible”. JANIAUD, J., “Simone Weil et l’attention”, 179.

dignidad. Su atención hacia estos no consistía en contemplar a distancia su dolor, sino en estar realmente a su lado y actuar con el fin de encontrar los medios que pudieran liberarlos de la degradación social. Y, mientras que en su juventud su compromiso era característicamente social y político, a medida que el itinerario espiritual de Weil se intensifica, fundamentalmente después de su experiencia de Cristo en 1938¹⁴, su reflexión sobre la atención y la acción alcanzará una mayor fuerza y profundidad.

En el presente artículo se considerarán esencialmente algunas de las obras weilianas de “madurez”¹⁵, que justamente corresponden con el período más espiritual de Weil, como es el caso del tiempo que pasó entre Marsella y Londres¹⁶. Asimismo, se tiene como objetivo abordar el importante vínculo entre la atención desde su dimensión creadora y la acción desde su perspectiva “inactiva” que, como se verá, ofrecen la posibilidad de mirar y reconocer la existencia de la realidad exterior al “yo”, en especial la de aquella que es invisible y anónima para la sociedad: la de los desdichados.

Por ello, en primer lugar, se abordará el tema de la desdicha, particularmente la que tiene como causa el factor social, ya que provoca la degradación social y el desarraigo de quien está sometido a ella. Luego, se verá que es a través de una atención creadora por la que se reconoce la existencia del otro. Y, finalmente, se mostrará que para posibilitar la existencia del otro se requiere no actuar persiguiendo los propios fines, sino por una necesidad de acción que es inspirada por el amor divino.

1. LA ATENCIÓN COMO MIRADA DE LA DESDICHA

La desdicha (*le malheur*) tiene como origen la dimensión social y provoca que el mal penetre en lo más profundo de la persona hasta hacerle perder el sentido de su propia dignidad¹⁷. A este respecto, es importante destacar la experiencia que tuvo Weil al trabajar como obrera en las fábricas de Alsthom, J. J. Carnaud et Forges y Renault (1934-1935), en donde además de vivir de cerca la opresión social, pudo constatar que la desdicha más profunda no proviene de la naturaleza, sino de la injusticia social. Esto es, un sufrimiento físico que proviene

14 Véase, por ejemplo, la “autobiografía espiritual” que Weil dirige al padre Perrin. WEIL, S., “Autobiographie spirituelle”, [1942]. En: *Attente de Dieu*. Paris: Fayard, 1966, 35-62. En adelante se usará la abreviatura *AD*.

15 Se les denomina de “madurez” debido a que Weil murió cuando tenía tan solo 34 años.

16 Se utilizarán particularmente los textos escritos entre 1941 y 1943.

17 En este artículo se hará referencia a la desdicha que tiene como origen el factor social.

de la naturaleza, por muy penoso que sea, solo causa una “media desdicha” (*un demi-malheur*)¹⁸ en tanto que no destruye el reconocimiento de la dignidad de quien la padece. Mientras que la desdicha que se deriva de lo social acaba con este reconocimiento y somete a quien la sufre a la condición de esclavo.

Y aunque el sufrimiento físico puede estar presente en la desdicha que tiene como raíz el factor social, esta consiste primordialmente en un “desarraigo de la vida, [en] un equivalente más o menos atenuado de la muerte”¹⁹. Por esta razón, independientemente de que la desdicha alcance o no otras dimensiones, si es tocada por el factor social, arrastra a quien la sufre a la degradación social. De ahí que, en términos de la autora:

Solo hay verdadera desdicha si el acontecimiento que se ha adueñado de una vida y la ha desarraigado, la alcanza directa o indirectamente en todas sus partes, social, psicológica, física. El factor social es esencial. No hay verdaderamente desdicha donde no hay degradación social en alguna de sus formas o conciencia de tal degradación²⁰.

De acuerdo con esto, la desdicha más radical es la que proviene del ámbito social puesto que degrada socialmente y provoca a quien la padece una sensación de privación de su dignidad. Esto es, además del rechazo social, la desdicha produce la pérdida de la conciencia del propio valor, así como del sentido de pertenencia a un grupo social. Se trata de un sentimiento de desprecio hacia uno mismo, que aparta del mundo, impide crecer en la vida social, oprime, ahoga y desarraiga. Por eso, Weil sin duda está en lo cierto cuando asegura que la desdicha es algo que “se apropia del alma y la marca”, es “la marca de la esclavitud”²¹. Es un sufrimiento que –en palabras del filósofo Maurice Blondel– “destroza una vida

18 Sobre esto, hay que señalar que la propia Weil padeció la mayor parte de su vida de terribles migrañas. Se trataba de un sufrimiento físico que podría ser un evidente signo de desdicha. No obstante, después de experimentar la opresión social durante su trabajo como obrera en la fábrica, reflexiona sobre lo que es verdaderamente la desdicha y determina que el sufrimiento físico es una “media desdicha” comparado con la destrucción a la que puede conducir el mal causado por el factor social. Véase WEIL, S., “Autobiographie spirituelle”, [1942], *AD*, 41-42.

19 “Le malheur est un déracinement de la vie, un équivalent plus ou moins atténué de la mort”. WEIL, S., “L’amour de Dieu et le malheur”, [1942], *OC*, IV 1, 347.

20 “Il n’y a vraiment malheur que si l’événement qui a saisi une vie et l’a déracinée l’atteint directement ou indirectement dans toutes ses parties, sociale, psychologique, physique. Le facteur social est essentiel. Il n’y a pas vraiment malheur là où il n’y a pas sous une forme quelconque déchéance sociale ou appréhension d’une telle déchéance”. WEIL, S., “L’amour de Dieu et le malheur”, [1942], *OC*, IV 1, 348.

21 WEIL, S., “L’amour de Dieu et le malheur”, [1942], *OC*, IV 1, 347. Weil experimentó esta “marca de la esclavitud”, particularmente en su experiencia obrera, como señala André Devaux: “La ‘marque’ de l’esclavage, physique et social, elle l’a portée, imprimée en son cœur et en son corps”. DEVAUX, A., “Malheur et compassion chez Simone Weil”. *Cahiers Simone Weil*, VIII, 4, 1985, 390.

sin matarla”²², puesto que el desdichado (*le malheureux*) adquiere la condición de un esclavo que “no existe, ni a los ojos del amo ni a sus propios ojos”²³, es decir, no es reconocido socialmente como una persona y no puede “escapar al desprecio y al odio de la sociedad”²⁴.

El desdichado se convierte, de este modo, en víctima del desprecio de la sociedad y –siguiendo a Weil– “este desprecio, esta repulsión, este odio, se vuelven en el desdichado contra sí mismo, penetran en el centro de su alma, y desde ahí tiñen con su tonalidad venenosa el universo entero”²⁵. La degradación social, por tanto, corrompe y provoca simultáneamente un desprecio hacia uno mismo y hacia los otros. El desdichado se siente desarraigado del mundo y encadenado de por vida a su propia desdicha, por eso se vuelve su cómplice y no busca los medios para librarse de ella²⁶. Como advierte la autora, la “complicidad obstaculiza todos los esfuerzos que pudiera hacer para mejorar su suerte; hasta le impide buscar los medios de liberarse, a veces incluso le impide desear el liberarse”²⁷.

La complicidad que tiene el desdichado en relación con su propia desdicha revela que la mirada que este dirige hacia sí mismo está completamente distorsionada, hasta el punto de considerarse despreciable e indigno de todo reconocimiento social. Y, a la vez, quienes no están sumergidos en la desdicha manifiestan una actitud de indiferencia y de desprecio hacia los desdichados y tal desprecio –como declara Weil– “es lo contrario de la atención”²⁸. Paralelamente a esto, “la misma incapacidad de prestar atención a la desdicha, que impide la compasión de quien ve a un desdichado, impide la gratitud del desdichado que es socorrido”²⁹. Esto

22 BLONDEL, M., *Œuvres complètes*, I: *Les deux thèses*. Paris: PUF, 1995, 363.

23 “Un esclave n’existe pas, ni aux yeux du maître, ni à ses propres yeux”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 293.

24 Como indica Miklós Vetö, aludiendo a Weil: “cette déchéance est ressentie plus particulièrement par les mendiants, les prisonniers, les prostituées, toutes sortes d’êtres humains qui sont continuellement en butte à la méfiance et au mépris de la société. Que ces gens aient vraiment commis des crimes ou non, ils ne peuvent échapper au mépris et à la haine de la société”. VETÖ, M., *La métaphysique religieuse de Simone Weil*. Paris: L’Harmattan, 1997, 77.

25 “Ce mépris, cette répulsion, cette haine, chez le malheureux, se tournent contre lui-même, pénètrent au centre de l’âme, et de là colorent de leur coloration empoisonnée l’univers tout entier”. WEIL, S., “L’amour de Dieu et le malheur”, [1942], OC, IV 1, 350.

26 “On aspire à sortir du malheur [...]. Mais, à un certain degré de malheur, aspire-t-on à en sortir?” WEIL, S., “Cahier II”, [1941], OC, VI 1, 222. Hay traducción castellana de C. Ortega en WEIL, S., *Cuadernos*. Madrid: Trotta, 2001.

27 “Cette complicité entrave tous les efforts qu’il pourrait faire pour améliorer son sort; elle va jusqu’à l’empêcher de rechercher les moyens d’être délivré, parfois même jusqu’à l’empêcher de souhaiter la délivrance”. WEIL, S., “L’amour de Dieu et le malheur”, [1942], OC, IV 1, 351.

28 “Le mépris est le contraire de l’attention”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 296.

29 “La même incapacité de faire attention au malheur qui empêche la compassion chez celui qui voit un malheureux empêche la gratitude chez le malheureux secouru. La gratitude suppose la

se debe precisamente a que la desdicha no es un objeto de atención, sino de aversión, porque como asevera la autora, “la tendencia de la naturaleza humana es la de no prestar atención a los desdichados”³⁰, pues simplemente “no se tiene compasión por las cosas totalmente destruidas”³¹. En virtud de ello Weil afirma que “el gran enigma de la vida humana no es el sufrimiento, es la desdicha”³², dado que, al igual que la verdad, la desdicha es muda³³.

En la medida en que la desdicha es muda, el desdichado es invisible para la sociedad, no es escuchado, está sometido al silencio y al anonimato, como si se tratara de un ser inexistente. Quien está totalmente inmerso en la desdicha es abandonado por los otros. Como puede leerse, por ejemplo, en el libro de Job del Antiguo Testamento, cuando a este le advienen una serie de males es abandonado incluso por sus parientes. Es también el caso de Cristo en la Cruz que fue abandonado por sus amigos por el temor que estos tenían de arriesgar sus vidas³⁴. Hasta en los libros de historia se encuentran las grandes hazañas de los vencedores, mientras que los vencidos han sido borrados de sus páginas³⁵. Por esta razón, Weil acierta al considerar que prestar atención a quienes están completamente sumergidos en la desdicha es prácticamente un milagro ya que –como indica– “la capacidad de prestar atención a un desdichado es cosa muy rara, muy difícil; es casi un milagro; es un milagro. Casi todos los que creen tener esta capacidad no la tienen. La calidez, el impulso del corazón, la piedad, no son suficientes”³⁶.

A pesar de la dificultad que implica prestar atención a la desdicha, esta no puede negarse, es una realidad que existe y la atención puede orientarse hacia ella³⁷. Es más, quienes no están sometidos a la desdicha tienen la obligación moral de hacer todo lo que esté a su alcance para ayudar a los desdichados, puesto que los desdichados necesitan que su dignidad sea reconocida. En palabras de Weil,

capacité de sortir de soi et de contempler son propre malheur du dehors dans toute sa laideur. C'est trop affreux”. WEIL, S., “Cahier XV”, [1942], OC, VI 4, 231.

30 “Parce que le penchant de la nature humaine est de ne pas faire attention aux malheureux”. WEIL, S., “L'Enracinement”, [1943], OC, V 2, 162.

31 “On n'a pas de compassion pour les choses totalement détruites”. WEIL, S., “L'Enracinement”, [1943], OC, V 2, 288.

32 “La grande énigme de la vie humaine, ce n'est pas la souffrance, c'est le malheur”. WEIL, S., “L'amour de Dieu et le malheur”, [1942], OC, IV 1, 348.

33 Véase WEIL, S., “Carnet de Londres”, [1943], OC, VI 4, 364.

34 Véase WEIL, S., “L'Enracinement”, [1943], OC, V 2, 288.

35 Véase WEIL, S., “L'Enracinement”, [1943], OC, V 2, 289.

36 “La capacité de faire attention à un malheureux est chose très rare, très difficile ; c'est presque un miracle ; c'est un miracle. Presque tous ceux qui croient avoir cette capacité ne l'ont pas. La chaleur, l'élan du cœur, la pitié n'y suffisent pas”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 262.

37 WEIL, S., “Carnet de Londres”, [1943], OC, VI 4, 364.

“no necesitan otra cosa en este mundo más que hombres capaces de prestarles atención”³⁸ y así queden libres de la exclusión o del “anatema social” al que han sido reducidos. Se requiere, por tanto, de una compasión que “consiste en prestar atención al desdichado y en ir hacia él a través del pensamiento”³⁹ y de una atención capaz de dirigir una mirada compasiva a todo ser humano que sufre.

La atención no es hacer el bien al desdichado como una satisfacción personal o para ser visto a los ojos de los otros como alguien piadoso. Tampoco es darle una limosna porque, siguiendo a Weil, “la limosna, cuando no es sobrenatural, es semejante a una operación de compra. Compra al desdichado”⁴⁰. Una auténtica atención implica mirar al desdichado sin prejuicios y reconocerlo como un ser humano que, si bien ha sido alcanzado por la desdicha, es un semejante. Como destaca Weil, “es saber que el desdichado existe, no como una unidad en una colección, no como un ejemplar de la categoría social denominada ‘desdichado’, sino como un hombre, exactamente semejante a nosotros, que fue un día golpeado y marcado, con una marca inimitable, por la desdicha”⁴¹.

Solamente cuando se mira al desdichado como a un semejante se pueden realmente determinar sus verdaderas necesidades y de esta manera rescatarlo del estado en el que se encuentra. Como indica la autora “esta mirada es, ante todo, una mirada atenta, en la que el alma se vacía de todo contenido propio para recibir en sí al ser que mira, tal cual es, en toda su verdad. Solo es capaz de ello quien es capaz de atención”⁴². Se trata entonces de alcanzar “la plenitud del amor al prójimo [que] consiste simplemente en ser capaz de preguntarle: ‘¿Cuál es tu tormento?’”⁴³,

38 “Les malheureux n’ont pas besoin d’autre chose en ce monde que d’hommes capables de faire attention à eux”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 262.

39 “La compassion consiste à faire attention au malheureux et à se transporter en lui par la pensée”. WEIL, S., “Cahier XV”, [1942], OC, VI 4, 230.

40 “L’aumône, quand elle n’est pas surnaturelle, est semblable à une opération d’achat. Elle achète le malheureux”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 292.

41 “C’est savoir que le malheureux existe, non pas comme unité dans une collection, non pas comme un exemplaire de la catégorie sociale étiquetée ‘malheureux’, mais en tant qu’homme, exactement semblable à nous, qui a été un jour frappé et marqué d’une marque inimitable par le malheur”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 262.

42 “Ce regard est d’abord un regard attentif, où l’âme se vide de tout contenu propre pour recevoir en elle-même l’être qu’elle regarde, tel qu’il est, dans toute sa vérité. Seul en est capable celui qui est capable d’attention”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 262.

43 “La plénitude de l’amour du prochain, c’est simplement d’être capable de lui demander : ‘Quel est ton tourment ?’” WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 262. Hay que señalar a este respecto la alusión que hace Weil a la obra del siglo XIII *La Queste del Saint Graal*, véase la nota 18 de la obra weiliana referida. También es importante destacar la obra *Parzival* de Wolfram von Eschenbach en donde la pregunta “¿cuál es tu tormento?”, que Parzival dirigiera a su tío Anfortas, pondría fin al gran sufrimiento que padecía este último. Véase WEIL, S., “Cahier XI”, [1942], OC, VI 3, 359-361.

en “amar al ser amado tal y como es, y querer recrearlo”⁴⁴, así como de actuar como el Buen Samaritano que reconoce la existencia del otro, le presta verdaderamente atención y lo auxilia.

2. UNA ATENCIÓN QUE ES CREADORA

A partir de la lectura de la parábola del Buen Samaritano⁴⁵, a la que Weil recurre en diversas ocasiones, puede comprenderse la profundidad de lo que verdaderamente significa dirigir una mirada atenta y compasiva hacia lo que se considera invisible, lo que nadie es capaz de mirar. En este relato del Evangelio, un sacerdote, un levita y un samaritano pasan por donde yace un hombre que ha sido golpeado y despojado de sus bienes por unos ladrones. Mientras que los dos primeros personajes, al ver al hombre herido, se apartan de él y siguen su camino, el samaritano se muestra compasivo y lo socorre. A diferencia del sacerdote y del levita, el samaritano no se limita a mirar al hombre herido y luego seguir su camino, tampoco es movido por un simple sentimiento de piedad. El samaritano presta verdaderamente atención a aquel hombre desvalido, se compadece de él y lo reconoce como prójimo, aun cuando se trata de un desconocido. A esta atención Weil la denomina “atención creadora” (*attention créatrice*):

La atención creadora consiste en prestar realmente atención a lo que no existe. La humanidad no existe en la carne anónima e inerte al borde del camino. El samaritano que se detiene y mira, presta sin embargo atención a esa humanidad ausente, y los actos que siguen dan testimonio de que se trata de una atención real⁴⁶.

La atención creadora no es una tentativa de hacer salir algo a partir de la nada, es decir, de manera *ex nihilo*. La atención para Weil “no crea nada, sino que reconoce la creación”⁴⁷. El Buen Samaritano es el perfecto ejemplo para comprender mejor el concepto weiliano de “atención creadora”. Puesto que es quien no se mantiene a distancia y actúa como verdadero prójimo al compadecerse, prestar atención y reconocer una humanidad muda y ausente, esto es, una realidad inexistente para los demás y que está herida por la desdicha. Su atención

44 “Aimer l'être aimé tel qu'il est, et vouloir le recréer”. WEIL, S., “Cahier XI”, [1942], OC, VI 3, 344.

45 Véase Lc 10, 25-37.

46 “L'attention créatrice consiste à faire réellement attention à ce qui n'existe pas. L'humanité n'existe pas dans la chair anonyme inerte au bord de la route. Le Samaritain qui s'arrête et regarde fait pourtant attention à cette humanité absente, et les actes qui suivent témoignent qu'il s'agit d'une attention réelle”. WEIL, S., “Formes de l'amour implicite de Dieu”, [1942], OC, IV 1, 293.

47 SAINT-SERNIN, B., *L'action politique selon Simone Weil*. Paris: Cerf, 1988, 99.

es inspirada por un auténtico amor al prójimo, que es “el amor que descende de Dios al hombre”⁴⁸ o –en términos de Weil– por un “amor sobrenatural al prójimo” que consiste en un “intercambio de compasión y gratitud que se produce como un relámpago entre dos seres, de los cuales uno está dotado y el otro privado de la condición de persona humana”⁴⁹. El amor orienta al samaritano a ver lo invisible⁵⁰ a mirar lo que la autora denomina “un trozo de carne desnuda, inerte y sangrante, al borde de un foso, sin nombre, del que nadie sabe nada”⁵¹.

Desde esta perspectiva, Cristo mismo está implícitamente presente en este acto de profunda atención, en el que no se ama al prójimo de un modo abstracto o impersonal, sino que se le dirige “un amor personal”⁵², por el que se le mira y ama tal cual es. En tal sentido, Weil precisa que “dondequiera que los desdichados son amados por sí mismos, Dios está presente”⁵³. Y, a pesar de que ni quien es compasivo ni el desdichado tengan conciencia de la presencia de Cristo, el amor divino está igualmente presente en el intercambio de compasión y gratitud, ya que estas últimas –siguiendo a Weil– “descienden de Dios, y cuando se encuentran en una mirada, Dios está presente en el punto en que las miradas se encuentran”⁵⁴.

De acuerdo con la autora, “solo Dios puede prestar atención a un desdichado”⁵⁵, “solo Dios tiene el poder de pensar realmente lo que no es”⁵⁶. Y, a este respecto, puede afirmarse junto con ella que, por participación, quienes están llenos de su amor son también capaces de atención. Contrariamente, para que-

48 “L’amour du prochain est l’amour qui descend de Dieu vers l’homme. Il est antérieur à celui qui monte de l’homme vers Dieu”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294.

49 “Le Christ nous a enseigné que l’amour surnaturel du prochain, c’est l’échange de compassion et de gratitude qui se produit comme un éclair entre deux êtres dont l’un est pourvu et l’autre privé de la personne humaine”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 291.

50 “L’amour voit l’invisible”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 293.

51 “Un peu de chair nue, inerte et sanglante au bord d’un fossé, sans nom, dont personne ne sait rien”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 291.

52 “Dieu n’est pas présent, même s’il est invoqué, là où les malheureux sont simplement une occasion de faire le bien, même s’ils sont aimés à ce titre. Car alors ils sont dans leur rôle naturel, dans leur rôle de matière, de choses. Ils sont aimés impersonnellement. Et il faut leur porter dans leur état inerte, anonyme, un amour personnel”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294.

53 “Partout où les malheureux sont aimés pour eux-mêmes, Dieu est présent”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294.

54 “La compassion et la gratitude descendent de Dieu, et quand elles s’échangent en un regard, Dieu est présent au point où les regards se rencontrent”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294-295.

55 “Dieu seul peut faire attention à un malheureux”. WEIL, S., “Cahier XV”, [1942], OC, VI 4, 230.

56 “Dieu seul a ce pouvoir, de penser réellement ce qui n’est pas”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 293-294.

nes no tienen presente en su alma el amor divino, la desdicha es algo aberrante y repulsivo, es algo que les causa horror. Su “instinto de conservación”⁵⁷ les impide prestar atención a los desdichados y les mueve a huir tanto de la desdicha como de los desdichados para evitar ser tocados y destruidos por ella. En consecuencia, Weil advierte que “todos desprecian en mayor o menor grado a los desdichados, aunque casi nadie tenga conciencia de ello”⁵⁸. Incluso, “los propios desdichados experimentan el mismo estremecimiento de horror ante sí mismos”⁵⁹.

Por todo esto, únicamente el amor divino es capaz de mirar aquello que es inexistente e invisible a los ojos humanos. Únicamente un amor como el de Cristo redime y puede auténticamente preguntar “¿cuál es tu tormento?”⁶⁰. Así pues, puede decirse que el amor divino actúa en el desdichado a través de quien obra como el Buen Samaritano, el cual –consciente o no de la presencia del amor de Dios en su alma– al prestar una atención creadora, por la que reconoce la humanidad de quien ha sido golpeado y herido por la desdicha, se convierte en un medio por el que Dios puede orientarse hacia los desdichados. De esta manera, en consonancia con Weil, será posible “mirarlos verdaderamente con una mirada distinta de la que se dirige a los objetos, escuchar verdaderamente su voz como se escucha una palabra”⁶¹.

Para que el amor de Dios sea transmitido a los desdichados, se necesita que el alma de quien actúa como el samaritano consienta a ser el medio por el que se transmita este amor. Y, a pesar de que esta alma humana sea, como especifica la autora, “la última, la más miserable, la más deforme, Dios se precipita hacia ella para poder, a través de ella, mirar y escuchar a los desdichados”⁶². Esto pone

57 Véase WEIL, S., “Cahier XV”, [1942], OC, VI 4, 231.

58 “Excepté ceux dont le Christ occupe toute l’âme, tout le monde méprise plus ou moins les malheureux, quoique presque personne n’en ait conscience”. WEIL, S., “L’amour de Dieu et le malheur”, [1942], OC, IV 1, 350.

59 “Les malheureux eux-mêmes éprouvent le même frémissement d’horreur devant eux-mêmes”. WEIL, S., “La personne et le sacré”, [1943]. En: *Écrits de Londres et dernières lettres*. Paris: Gallimard, 1957, 35. En adelante se usará la abreviatura *EL*. Hay traducción castellana de M. Larrauri en WEIL, S., *Escritos de Londres y últimas cartas*. Madrid: Trotta, 2000.

60 “Le Christ seul peut demander à un homme : quel est ton tourment?” WEIL, S., “Cahier XI”, [1942], OC, VI 3, 361. Como se ha señalado anteriormente, cuando Weil escribe la pregunta “¿cuál es tu tormento?” hace referencia al *Parzival* de Wolfram von Eschenbach. Véase nota 43.

61 “Seul Dieu présent en nous peut réellement penser la qualité humaine chez les malheureux, les regarder vraiment d’un regard autre que celui qu’on accorde aux objets, écouter vraiment leur voix comme on écoute une parole”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294.

62 “Dieu a hâte de descendre vers les malheureux. Dès qu’une âme est disposée au consentement, fût-elle la dernière, la plus misérable, la plus difforme, Dieu se précipite en elle pour pouvoir à travers elle regarder, écouter les malheureux”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 294.

de manifiesto que, quien se muestra atento y compasivo hacia su prójimo, quien reconoce la existencia de los desdichados consiente, incluso sin saberlo, en ser un medio por el que Dios mire y escuche a los desdichados. Y, además, la atención, compasión y amor que comunica a estos seres que nadie ve ni escucha, implica también un acto de renuncia, una negación de sí mismo, por la que por un momento la mirada se aparta del “yo”, de las propias necesidades y deseos personales, para dirigirla a los otros y afirmar su existencia.

En palabras de Weil, “negándose a sí mismo, uno se hace capaz, con Dios, de afirmar al otro por una afirmación creadora. Uno se da a sí mismo en rescate por el otro. Es un acto redentor”⁶³. La negación de sí mismo que supone la renuncia al “yo” egoísta para afirmar la existencia del otro es lo que la autora denomina “descreación” (*décréation*), por la que –en imitación de la kénosis divina– el alma humana renuncia por amor al “yo” egoísta⁶⁴. A través de este acto de vaciamiento, de renuncia al “yo”, principalmente, con la ayuda de la gracia⁶⁵, el alma puede prepararse y adquirir la capacidad para prestar atención a los desdichados, una atención que es creadora, por medio de la cual se reconoce en ellos una existencia que les ha sido arrebatada, así como rescatarlos del desarraigo, de la exclusión y de la degradación social en la que están profundamente sumergidos.

La atención creadora por la que se fija la mirada en alguien exterior a uno mismo, especialmente en alguien que no tiene voz ni es visible en la vida social – porque la extrema desdicha destruyó su “yo”⁶⁶ – entraña una renuncia al egoísmo, a la propia personalidad, una negación del “yo”. Mediante esta renuncia, Weil declara que “el hombre acepta una disminución al concentrarse en un gasto de energía que no aumentará su poder, que solamente hará existir otro ser distinto a él, independiente de él”⁶⁷, libre y autónomo. La renuncia al egoísmo se convierte, entonces, en un acto de auténtica generosidad, puesto que, “sin esperar ninguna

63 “En se niant soi-même, on devient capable après Dieu d’affirmer un autre par une affirmation créatrice. On se donne en rançon pour l’autre. C’est un acte rédempteur”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 292.

64 Aunque en el presente artículo no se desarrollarán desde una perspectiva weiliana los términos de “descreación” y “kénosis”, es importante subrayar que la descreación consiste para la autora en la imitación humana de la kénosis divina por la que Dios, por amor, renuncia a la creación para dejarla existir: “La Création est de la part de Dieu un acte non pas d’expansion de soi, mais de retrait, de renoncement”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 291.

65 “Seule l’opération surnaturelle de la grâce fait passer une âme à travers son propre anéantissement jusqu’au lieu où se cueille l’espèce d’attention qui seule permet d’être attentif à la vérité et au malheur”. WEIL, S., “La personne et le sacré”, [1943], EL, 36.

66 “Rien n’est pire que l’extrême malheur qui du dehors détruit le je, puisque dès lors on ne peut plus le détruire soi-même”. WEIL, S., “Cahier VII”, [1942], OC, VI 2, 461.

67 “L’homme accepte une diminution en se concentrant pour une dépense d’énergie qui n’étendra pas son pouvoir, qui fera seulement exister un être autre que lui, indépendant de lui”. WEIL, S., “Formes de l’amour implicite de Dieu”, [1942], OC, IV 1, 292.

compensación, sin ningún motivo personal, se acepta ser disminuido en favor de la existencia independiente de otro ser que no sea uno mismo”⁶⁸.

Cuando se renuncia a que el “yo” ocupe el centro de todo y se reconoce a los otros como existentes, con una vida independiente de la propia, se posibilita que el desdichado pase de la no existencia a la existencia, recupere el reconocimiento de su dignidad y se le arraigue en este mundo. De modo que, a la luz del análisis de Weil, “ya no se está en el centro del mundo, se da a partir de ahora un consentimiento amoroso y activo a la coexistencia de otro ser con nosotros”⁶⁹. Si se considera seriamente el planteamiento de la autora y si el “yo” fuera realmente capaz de abandonar su afán por proyectarse en todo lo que le rodea, se posibilitaría la afirmación de la existencia del otro y se le reconocería como una realidad semejante a la propia, una realidad que tiene la misma capacidad de relación y la misma necesidad de ser amado.

3. HACIA UN RECONOCIMIENTO DE LA EXISTENCIA: ATENCIÓN CREADORA Y ACCIÓN NO ACTUANTE

De todo lo anterior se desprende que la atención favorece una transformación interior y un compromiso con lo exterior al “yo”, ya que conduce a salir de uno mismo, a renunciar al propio egoísmo para orientar la mirada hacia la realidad. La atención es entonces, como claramente señala Bertrand Saint-Sernin, un “acto perfecto del pensamiento y del corazón, por el que un ser se encuentra a la vez desposeído de sí mismo y arraigado en una realidad distinta a él mismo, la del universo o Dios”⁷⁰. De este modo, se posibilita el reconocimiento y la afirmación de lo que no es el “yo”, principalmente, de quienes sufren la extrema desdicha y viven bajo el yugo tanto de la degradación social como de la privación del reconocimiento de su dignidad.

Adicionalmente a esto y en consonancia con el planteamiento de Weil, es importante destacar que la verdadera atención no consiste simplemente en mirar cómodamente y a distancia la desdicha. Esta facultad, si bien provoca un desgaste

68 “Dépensant et se dépensant pour un malheureux sans s’attendre à aucune compensation, sans aucun motif personnel, on accepte d’être diminué en faveur de l’existence indépendante d’un autre être que soi-même”. VETÖ, M., *La métaphysique religieuse*, 28.

69 “On n’est plus au centre du monde, on donne désormais son consentement amoureux et actif à la coexistence d’un autre être avec nous”. VETÖ, M., *La métaphysique religieuse*, 28.

70 Como indica Bertrand Saint-Sernin “Simone Weil nomme ‘attention’ cet acte parfait de la pensée et du cœur, par où un être se trouve à la fois dépossédé de lui-même et enraciné dans une autre réalité que lui-même, univers ou Dieu”. SAINT-SERNIN, B., *L’action politique*, 96.

de energía por el esfuerzo que implica salir de uno mismo para dirigirse a una realidad exterior, contiene una capacidad generadora impulsada por el amor. Así pues, la atención tiene como fuente y objeto “el amor como potencia activa”⁷¹, lo que posibilita que se irradie una nueva fuente de energía, que ya no tiene como origen al “yo”, sino que ha sido transformada, mejorada, renovada y es apta para ser transmitida a una existencia anónima con el fin de rescatarla y socorrerla.

Una atención inspirada e impulsada por el amor precisa de un compromiso personal, de un cambio y mejoramiento de sí mismo. Es más, de acuerdo con Saint-Sernin, “la atención solo se hace posible mediante un cambio de estado, una ascesis, casi una conversión”⁷². Cuando tal modificación y cambio de sí mismo se realiza, se posibilita, por un lado, renunciar a un “yo” que siempre intenta proyectarse en el mundo y en los otros, lo cual, siguiendo a Miklós Vetö, “no es más que una crispación feroz, que paraliza y aplasta a los seres y a las cosas que encuentra”⁷³. Y, por otro lado, se adquiere la capacidad de dirigir hacia la realidad “una atención al vacío que espera que algo aparezca, se revele, se manifieste”⁷⁴. Una atención que para Weil consiste en una mirada y no en un apego, porque “el apego es un fabricante de ilusiones”, mientras que a través de la mirada “la ilusión se disipa, lo real aparece”⁷⁵.

Tal atención –caracterizada por una mirada por la que lo real aparece– es un “esfuerzo negativo” en el que el pensamiento “debe estar vacío, en espera, sin buscar nada, pero estar listo para recibir en su verdad desnuda el objeto que va a penetrar en él”⁷⁶. Una atención por la que el alma se dispone a esperar pacientemente a que la realidad se manifieste sin ningún tipo de proyección por parte del

71 “L’attention a donc tout à la fois pour source et pour objet l’amour comme puissance agissante”. SAINT-SERNIN, B., *L’action politique*, 98.

72 “L’attention n’est rendue possible que par un changement d’état, une ascèse, presque une conversion. Quand cette transformation intime s’opère, le relief et la constitution de la réalité deviennent visibles: les forces et les structures apparaissent, les rouages de la mécanique sociale sortent de l’ombre. L’empire de la force se révèle dans son unité, sa nécessité et son indifférence au bien; en même temps, les besoins de l’âme, universels, nécessaires eux aussi, et tournés vers le bien prennent consistance”. SAINT-SERNIN, B., *L’action politique*, 98.

73 “Le moi n’est qu’une crispation féroce, paralysant et broyant les êtres et les choses qu’il rencontre”. VETÖ, M., *La métaphysique religieuse*, 46.

74 “L’attention qui se propose d’apprendre quelque chose est toujours une attention à vide qui attend que quelque chose apparaisse, se révèle, se manifeste”. VETÖ, M., *La métaphysique religieuse*, 46.

75 “Et si à travers les modifications on garde le regard orienté sur telle chose, en fin de compte l’illusion se dissipe, le réel apparaît. La condition est que l’attention soit un regard et non un attachement. L’attachement est fabricant d’illusions, et quiconque veut le réel doit être détaché”. WEIL, S., “Cahier VII”, [1942], OC, VI 2, 458.

76 “La pensée doit être vide, en attente, ne rien chercher, mais être prête à recevoir dans sa vérité nue l’objet qui va y pénétrer”. WEIL, S., “Réflexions sur le bon usage des études”, [1942], OC, IV 1, 260.

“yo” es inseparable de una acción “que no llevaría la marca de un yo, y que no contradiría el orden del universo”⁷⁷. Esta acción que debe ser “justa y razonable” es “una expresión, no del individuo, sino de la situación misma y, más precisamente, de la necesidad inscrita en la situación”⁷⁸.

A esta acción, Weil la denomina “acción no actuante” (*action non-agissante*)⁷⁹, en la que se debe “actuar no *para* una finalidad, sino *por* una necesidad”⁸⁰. Es una acción en la que se actúa sin dejar la marca del “yo”, que no espera nada a cambio, tampoco aspira a una recompensa por la energía gastada al prestar atención. Y, concretamente frente a la desdicha, su única finalidad es la de determinar cuáles son las verdaderas necesidades de quienes están expuestos a ella para actuar en consecuencia y sin dejar rastro del “yo”. La noción de “acción no actuante” entraña entonces “una cierta ‘formalización’ del concepto de la acción vaciada de sus elementos ‘personales’, y denota solo una cierta sucesión de movimientos físicos que deben ser realizados en obediencia a algún tipo de necesidad”⁸¹.

Así pues, no se actúa en vista de un propósito concreto o por un interés personal, por lo que es esencial “no atarse ni a los actos ni a sus frutos”⁸², porque no son una posesión ni una proyección del “yo”. Y, por otra parte, quien actúa no es la causa original de la acción, sino que es un medio por el que se transmite una inspiración recibida. Para Weil la fuente de esta inspiración que mueve a “contemplar lo no contemplable (la desdicha de otro), sin huir”⁸³ y que anima a actuar por una necesidad en favor del desdichado es de origen divino. Esto es, la inspiración

77 “Une action qui ne porterait pas la marque d’un moi, et qui ne contredirait pas l’ordre de l’univers”. JANIAUD, J., *Simone Weil: L’attention et l’action*, 85.

78 “Il faudrait que l’action juste et raisonnable s’impose d’elle-même à partir d’une situation donnée: elle serait une expression, non de l’individu, mais de la situation même et, plus précisément encore, de la nécessité inscrite dans la situation”. JANIAUD, J., *Simone Weil: L’attention et l’action*, 85-86.

79 Es importante destacar la relación del concepto “descreación” con el de “acción no actuante”. Miklós Vető aludiendo a Weil indica: “Le propre de l’état décrété est une ‘activité passive’ que Simone Weil appelle ‘l’action non-agissante’”. VETŐ, M., *La métaphysique religieuse*, 121.

80 “Agir non pour un objet, mais par une nécessité. Je ne peux pas faire autrement. Ce n’est pas action mais une sorte de passivité. Action non agissante”. WEIL, S., “Cahier IV”, [1941], OC, VI 2, 123. La autora subraya.

81 “L’action non-agissante [...] implique une certaine ‘formalisation’ du concept de l’action vidée de ses éléments ‘personnels’, et dénote seulement une certaine succession de mouvements physiques qui doivent être accomplis par obéissance à quelque sorte de nécessité”. VETŐ, M., *La métaphysique religieuse*, 122.

82 Como Weil apunta siguiendo a la Bhagavad-gītā: “ne s’attacher ni aux actes ni à leurs fruits”. WEIL, S., “Cahier IV”, [1941], OC, VI 2, 83.

83 “Contempler le non-contemplable (le malheur d’autrui), sans fuir”. WEIL, S., “Cahier XII”, [1942], OC, VI 3, 385.

es una de las formas como Dios está presente en el mundo⁸⁴, al mismo tiempo que es un don de amor, una gracia que es “lo más exterior y lo más interior” a uno⁸⁵. La inspiración la recibe quien está dispuesto a renunciar a su egoísmo, a adoptar una actitud de espera para que la realidad se manifieste y a dirigir una mirada compasiva hacia los desdichados.

En consonancia con lo dicho hasta ahora, debe señalarse que lo que impulsa la acción de quien posee esta inspiración que viene de fuera del “yo” no es propiamente la voluntad humana, tampoco es una imposición divina por la que se obliga a actuar en contra de la voluntad y de la libertad humanas. Se trata de un consentimiento del alma que –en respuesta al deseo de bien inscrito naturalmente en ella y a la gracia– acoge libremente y por amor el bien que es exterior a ella⁸⁶. En este sentido, siguiendo a la autora, “no podemos transformarnos a nosotros mismos, solo podemos ser transformados, pero no podemos serlo más que si lo queremos”⁸⁷. Luego, mediante ese mismo acto de consentimiento, que “es esencial para la obediencia, así como para el amor”⁸⁸, el alma acepta libremente ser movida por la inspiración que recibe como una gracia. Esto le impulsa a actuar renunciando a los frutos de su acción, así como a transmitir por obediencia –tal y como lo ha recibido– el bien que le ha inspirado a actuar⁸⁹.

Por esta razón, para Weil “el consentimiento humano es algo sagrado” puesto que “es lo que el hombre concede a Dios”⁹⁰ por amor y de una manera completamente libre. Desde esta perspectiva, puede afirmarse que quien consiente a actuar sin dejar una huella o marca del “yo” y acepta “ser un instrumento de contacto

84 “Trois présences de Dieu: de création, d’ordonnance, d’inspiration”. WEIL, S., “Cahier XII”, [1942], *OC*, VI 3, 385. Véase WEIL, S., “Cahier VII”, [1942], *OC*, VI 2, 485.

85 “La grâce est à la fois ce qui nous est le plus extérieur et le plus intérieur”. WEIL, S., “Théorie des sacrements”, [1943]. En: *Pensées sans ordre concernant l’amour de Dieu*. Paris: Gallimard, 1962, 138. En adelante se usará la abreviatura PSO. Hay traducción castellana de A. López y M. Tabuyo en WEIL, S., *Pensamientos desordenados*. Madrid: Trotta, 1995.

86 “Le bien ne nous vient que du dehors, mais il ne pénètre en nous que le bien auquel nous consentons”. WEIL, S., “Théorie des sacrements”, [1943], *PSO*, 138.

87 “Nous ne pouvons nous transformer nous-mêmes, nous ne pouvons qu’être transformés, mais nous ne pouvons l’être que si nous voulons bien”. WEIL, S., “Théorie des sacrements”, [1943], *PSO*, 138.

88 “Le consentement est essentiel à l’obéissance comme à l’amour”. WEIL, S., “Luttons-nous pour la justice?”, [1943], *EL*, 47.

89 “Action non agissante. [En dehors des devoirs naturels,] ne jamais faire un pas au-delà de ce à quoi on est irrésistiblement poussé, le bien même n’étant plus tel s’il n’est pas accompli par obéissance”. WEIL, S., “Cahier VI”, [1942], *OC*, VI 2, 362.

90 “Le consentement humain est chose sacrée. Il est ce que l’homme accorde à Dieu. Il est ce que Dieu vient chercher comme un mendiant auprès des hommes”. WEIL, S., “Luttons-nous pour la justice?”, [1943], *EL*, 47.

entre el prójimo y Dios”⁹¹, se vuelve un puente entre el Creador y su creación, un medio por el que se transmite el amor divino a los otros. Un amor que no es posesión, sino renuncia, que crea, redime e inspira. Este amor alienta al alma humana a renunciar a su propio egoísmo y a reconocer la creación. Concretamente, la exhorta a reconocer y afirmar la existencia de quienes sufren la extrema desdicha. Por tal motivo, el amor divino invita al alma humana a dirigir una atención creadora a quienes están sometidos al silencio y al anonimato. De igual forma, la impulsa a actuar en el mundo con una acción no actuante, una acción por la que se consiente libremente y por amor a “dejar de ser”, a dejar de proyectar el “yo” en el mundo para permitir la existencia del otro. Desde la óptica weiliana, esta es precisamente la manera como Dios ama a sus criaturas:

El Amor está del lado de la no acción, de la impotencia. El Amor, que consiste en querer que algo sea, en no querer tocarlo. Dios nos ama así; de otro modo, dejaríamos inmediatamente de existir. Quedaríamos aniquilados. Consentir por amor a dejar de ser, como debemos hacer, no es aniquilación, sino traslado vertical a la realidad que es superior al ser⁹².

CONCLUSIÓN

En virtud de lo expuesto a lo largo de este artículo, hay que destacar que para Weil la desdicha –particularmente la que tiene su origen en la dimensión social– conduce a quien la padece al desarraigo, a la degradación social e incluso al desprecio de sí mismo. El desdichado es completamente invisible para la sociedad, está sujeto al silencio y al anonimato, como si se tratara de un ser inexistente. Esto imposibilita prestarle atención e incluso el propio desdichado es completamente incapaz de escapar de su desdicha, ya que es cómplice de ella. Para salir de esta condición, el desdichado necesita que se le reconozca como un semejante y se le dirija una atención por la que se afirmen su existencia y su dignidad, de modo que pueda liberarse de la exclusión y del desarraigo. Esta atención debe ser creadora

91 “Être un instrument de contact entre le prochain et Dieu, comme le porte-plume entre moi et le papier”. WEIL, S., “Cahier VII”, [1942], OC, VI 2, 485. “Être poussé par Dieu vers le prochain comme le crayon est appuyé par moi sur le papier”. WEIL, S., “Cahier XII”, [1942], OC, VI 3, 385.

92 “Action non agissante. [...] L’Amour est du côté de la non-action, de l’impuissance. L’Amour, qui consiste à aimer que quelque chose soit, à ne pas vouloir y porter la main. Dieu nous aime ainsi; autrement nous cesserions immédiatement d’exister. Nous serions anéantis. Consentir par amour à ne plus être, comme nous devons faire, n’est pas anéantissement, mais transport vertical dans la réalité supérieure à l’être”. WEIL, S., “Cahier X”, [1942], OC, VI 3, 251-252.

y no consiste en una extensión del “yo”, sino en consentir a renunciar al propio egoísmo para permitir la existencia del otro.

Asimismo, se sea o no consciente de la presencia de Dios en el alma, “cuando se ayuda al prójimo, no se actúa para Dios, sino por Dios, como si Dios se encarnara en el acto de amor”⁹³. Este acto de amor consiste en consentir a renunciar al “yo”, en mirar y reconocer al otro, en transmitir una energía que se proyecte hacia el exterior. Y, en última instancia, “cuando en la manera de actuar hacia las cosas y los hombres, o simplemente en la manera de mirarlos, surgen virtudes sobrenaturales”⁹⁴ es porque Dios está verdaderamente presente en el alma humana. Como indica Weil, “Dios está presente, Cristo está presente allí donde se produce de un hombre a otro un acto de virtud sobrenatural”⁹⁵.

Lo anterior muestra, en definitiva, la preocupación que siempre tuvo Weil por la condición humana y su afán por buscar los medios para construir una sociedad más justa en la que sea reconocida la dignidad de todas las personas, sin excepción. Ante esto debe subrayarse que la reflexión y acción de Simone Weil tiene mucho que ofrecer no solamente a la filosofía francesa, sino también al pensamiento universal, puesto que las condiciones de opresión social y de esclavitud que ella misma experimentó, siguen estando presentes en la actualidad. Su postura no es solamente la de una invitación a pensar sobre los problemas relativos a la condición humana. Weil anima a actuar frente a las diversas situaciones de injusticia y a vivir realmente en una plenitud del amor al prójimo que se alcanza una vez que se mira a este en toda su verdad para determinar cuáles son sus verdaderas necesidades y, mediante una “acción no actuante”, permitir y reconocer su existencia.

REFERENCIAS BIBLIOGRÁFICAS

- BLONDEL, M., *Œuvres complètes, I: Les deux thèses*. Paris: PUF, 1995.
- DEVAUX, A., “Malheur et compassion chez Simone Weil”. *Cahiers Simone Weil*, VIII, 4, 1985, 386-402.
- JANIAUD, J., *Simone Weil: L’attention et l’action*. Paris: PUF, 2002.

93 “Lorsqu’on aide le prochain, on n’agit pas pour Dieu, mais par Dieu, comme si Dieu s’incarnait dans le geste d’amour”. JANIAUD, J., *Simone Weil: L’attention et l’action*, 88.

94 “Quand dans la manière d’agir à l’égard des choses et des hommes, ou simplement dans la manière de les regarder, il apparaît des vertus surnaturelles, on sait que l’âme n’est plus vierge, qu’elle a couché avec Dieu”. WEIL, S., “Cahier XIV”, [1942], *OC*, VI 4, 189.

95 “Dieu est présent, le Christ est présent partout où s’opère d’un homme à un homme un acte de vertu surnaturelle”. WEIL, S., “Cahier XIV”, [1942], *OC*, VI 4, 188.

- , “Simone Weil et l’attention”. En: DELSOL, C. (ed.), *Simone Weil*. Paris: Cerf, 2009, 169-179.
- ROMANO, M.S, “Simone Weil: atención y oración”. *Cauriensia*, XI, 2016, 697-711.
- SAINT-SERNIN, B., *L’action politique selon Simone Weil*. Paris: Cerf, 1988.
- VETÖ, M., *La métaphysique religieuse de Simone Weil*. Paris: L’Harmattan, 1997.
- WEIL, S., *Écrits de Londres et dernières lettres*. Paris: Gallimard, 1957.
- , *Pensées sans ordre concernant l’amour de Dieu*. Paris: Gallimard, 1962.
- , *Attente de Dieu*. Paris: Fayard, 1966.
- , *Œuvres complètes, t. I: Premiers écrits philosophiques (1925-1930)*. Paris: Gallimard, 1988.
- , *Œuvres complètes, t. VI: Cahiers, vol. 1: (1933 – septembre 1941)*. Paris: Gallimard, 1994.
- , *Œuvres complètes, t. VI: Cahiers, vol. 2: (septembre 1941 – février 1942)*. Paris: Gallimard, 1997.
- , *Œuvres complètes, t. VI: Cahiers, vol. 3: (février – juin 1942)*. Paris: Gallimard, 2002.
- , *Œuvres complètes, t. VI: Cahiers, vol. 4: (juillet 1942 – juillet 1943)*. Paris: Gallimard, 2006.
- , *Œuvres complètes, t. IV: Écrits de Marseille, vol. 1: Philosophie, science, religion, questions politiques et sociales*. Paris: Gallimard, 2008.
- , *Œuvres complètes, t. V: Écrits de New York et de Londres, vol. 2: L’Enracinement*. Paris: Gallimard, 2013.

EL CARÁCTER UTÓPICO DE LA FELICIDAD. UNA LECTURA DESDE ORTEGA Y GASSET

*THE UTOPIAN CHARACTER OF HAPPINESS. A
READING FROM ORTEGA Y GASSET*

MORA PERPERE VIÑUALES

Doctora en Filosofía (Universidad de Salamanca)
Máster en Estudios Americanos (Universidad de Sevilla)
Pontificia Universidad Católica Argentina
Buenos Aires/Argentina
moraperpere@uca.edu.ar
ORCID: 0000-0003-4254-1115

Recibido: 1/05/2021
Revisado: 21/07/2021
Aceptado: 6/09/2021

Resumen: El presente trabajo busca analizar el concepto de felicidad en el marco del raciovitalismo orteguiano y precisar en qué sentido debe entenderse el carácter utópico de la misma. Para ello, en primer lugar, se abordará el concepto de felicidad y se explicará qué relación existe entre ésta y la realización del proyecto vital individual. En segundo lugar, se abordará el concepto de utopía en Ortega y se explicará por qué, si bien el filósofo le atribuye a la felicidad un carácter utópico, esto debe ser comprendido, sin embargo, en sentido positivo.

Palabras clave: Autenticidad; Felicidad; Proyecto vital; Utopía; Vida humana.

Abstract: The present work seeks to analyze the concept of happiness within Ortega's ratiovitalism and to specify in what sense its utopian character should be understood. To do this, in first place, the concept of happiness will be addressed and the relationship between it and the realization of the individual life project will be explained. Secondly, the concept of utopia in Ortega will be addressed and it will be determined why, although the philosopher attributes a utopian character to happiness, this must be understood, however, in a positive sense.

Keywords: Authenticity; Happiness; Human Life; Utopia; Vital Project.

INTRODUCCIÓN

En 1906, en el acto de entrega de premios de los Juegos Florales de Valladolid, se leían frente al público las siguientes palabras: «La vida no es un placer, ni una ilusión, sino una larga hilera de esfuerzos para ver si se logra un placer o una ilusión. Las penalidades, las dificultades, los enojos, los problemas, las desesperanzas son tan frecuentes y tan poderosas que la mayor parte de los hombres se rinden, y caídos en un grasiento pesimismo renuncian a la lucha y a sus ilusiones y se contentan con vivir dolorosamente un día tras otro; a esto no le llaméis vivir: llamadlo a lo sumo no morir»¹. Estas palabras, pronunciadas en aquella ocasión por el periodista José Ortega Munilla, habían sido escritas unos días antes por su hijo, José Ortega y Gasset². En aquella oportunidad, el joven filósofo no sólo encantaba con sus palabras a un público que por entonces ignoraba su autoría, sino que además adelantaba en ellas algo en lo cual insistiría a lo largo de toda su vida: que la vida humana requiere de una lucha y un esfuerzo permanente si se la quiere llevar a su propia plenitud. El no realizar aquel esfuerzo, consideraba Ortega, conduce al hombre indefectiblemente a una vida infeliz o, como afirmaba entonces, «a vivir dolorosamente un día tras otro».

Esta llamada a la realización de la propia vida, a la autenticidad individual, será –como Ortega mismo afirmará en 1940–³ el tema más antiguo y constante en su pensamiento. El filósofo insistirá en la necesidad de mantenerse fiel al proyecto vital propio y de realizar el esfuerzo necesario para llegar a ser en la vida efectiva aquel que se es por íntima necesidad. En ese sentido, remarcará una y otra vez la importancia de escuchar aquella voz que llama a cada uno al cumplimiento de su vocación y que lo encamina, así, a su propia felicidad.

En el presente trabajo se analizará qué concepción tiene Ortega de la felicidad y por qué cree que la misma debe ser considerada siempre como el fin de la vida individual. Para ello, en primer lugar, se desarrollará el concepto orteguiano de felicidad, presentándolo como un concepto activo que incita al hombre a realizar el esfuerzo de llevar adelante una vida auténtica. Esto conducirá a cuestionarse por el aparente carácter utópico de la felicidad en la vida humana. Por esta razón, en segundo lugar, será necesario detenerse en el concepto de utopía para Ortega

1 ORTEGA Y GASSET, J., «[Discurso para los juegos florales de Valladolid]» (1906), en ORTEGA Y GASSET, J., *Obras Completas*, VII, Madrid: Taurus-Fundación Ortega y Gasset, 2004-2010, p. 83. En adelante, se citará siempre por esta edición de las *Obras Completas* indicando el título de la obra (y año de publicación de la misma), el tomo en números romanos y las páginas en arábigos.

2 Cfr. ORTEGA Y GASSET, M., *Niñez y mocedad de Ortega*, Madrid: Clave, 1964, p. 25.

3 *La razón histórica [curso de 1940]*, IX, p. 527.

y determinar si la felicidad es realmente asequible al hombre o si, finalmente, consistirá siempre en una meta inalcanzable.

1. LA FELICIDAD COMO FIN DE LA VIDA INDIVIDUAL

Al hacer referencia a la vida humana, Ortega explica que ésta es ante todo un proyecto que cada uno debe realizar si desea vivir de manera auténtica. Este proyecto, que es siempre individual –explica el filósofo en 1932–⁴ no es elegido libremente, sino que constituye el destino de cada uno y le exige su ejecución. La fidelidad de cada hombre a aquel destino determinará el nivel de autenticidad de su vida. De este modo, lo más interesante de la vida humana –explica allí mismo Ortega– no es la lucha de cada individuo con el mundo, sino aquella batalla que libra con su propio proyecto. Y es que el hombre es la única realidad que puede, no sólo decidirse a no ser aquello que por íntimo destino debe ser, sino también decidirse a serlo y, sin embargo, no conseguirlo.

Cuando el cumplimiento del proyecto vital, en efecto, se logra, el hombre siente un bienestar íntimo al que Ortega concibe como *felicidad*. Por el contrario, cuando el mismo no se cumple, el hombre no llega a ser en su vida efectiva aquel que verdaderamente es y queda, así, sumido en la infelicidad⁵.

Ahora bien, a pesar de que la llamada al cumplimiento del proyecto es permanente en Ortega, el filósofo afirma en numerosas oportunidades que esto nunca es posible de manera completa. El hombre nunca alcanza su propia verdad: lograr la efectiva y absoluta coincidencia consigo mismo es, en rigor, una tarea siempre inconclusa. Por esta razón, el ser humano, cuya vida consiste en tener que ser feliz, es al mismo tiempo y siempre parcialmente infeliz⁶. Todo aquello que hace y piensa, lo hace y piensa con el afán de buscar la felicidad. Su vida es pena, pesar,

4 *Goethe desde dentro* (1932), V, p. 124.

5 Esta distinción puede encontrarse principalmente *Goethe desde dentro* (*op. cit.*, pp. 129-133). Sin embargo, Ortega insistió en esta idea en varias de sus obras. Un ejemplo de ello puede encontrarse en 1943, cuando afirma que una vida dedicada a trabajos que alejan al individuo de la realización de su proyecto vital es una «vida que se aniquila y malogra a sí misma», mientras que aquella que se dedica a realizar aquel programa es una «vida que se logra a sí misma –la vida como delicia y felicidad». Por esta razón, agrega allí mismo, «felicidad es la vida dedicada a ocupaciones para las cuales cada hombre tiene singular vocación». [«Prólogo a *Veinte años de caza mayor, del Conde de Yebes*», (1943), VI, p. 273].

6 En este sentido, afirma Aranguren: «Hay instantes en los que, fugazmente, se diría que va a producirse ese “prodigioso fenómeno” de la coincidencia entre nuestra vida proyectada y nuestra vida real, que es en lo que consiste la felicidad. Pero no se produce.» A continuación, considera que esta aspiración del hombre a la felicidad, la exigencia intrínseca de ésta, muestra un punto de apertura de la ética orteguiana a la religión. El pasaje de una a la otra, explica, *puede realizarse*, más allá de

justamente porque antes de ser esto es proyecto y entusiasmo por lograr cumplir con su individual vocación⁷. «El hombre es el único ser que echa de menos lo que nunca ha tenido. Y el conjunto de lo que echamos de menos sin haberlo tenido nunca es lo que llamamos felicidad. De aquí podría partir una meditación de la felicidad, un análisis de esa extraña condición que hace de él el único ser infeliz precisamente porque necesita ser feliz»⁸, explica el filósofo.

A diferencia del hombre, Ortega considera que un ser vivo sin aspiraciones no podría ser infeliz. Una planta, un animal, por ejemplo. Estos seres, desde el momento mismo en que existen ya son aquello para lo cual existen. En este sentido, en 1933 Ortega explica que los animales tienen su modo peculiar de ser felices, mucho más logrado que el del hombre. De hecho, afirma, el animal es normalmente feliz: le alcanza con cubrir sus necesidades biológicas básicas⁹. Esta afirmación, por demás cuestionable, será reformulada por el mismo Ortega en 1940, en su ya citado curso sobre *La razón histórica*. Allí, en cambio, explicará

si Ortega lo realiza o no. (LÓPEZ ARANGUEREN, J. L. *La ética de Ortega*, Madrid: Taurus (tercera ed.), 1966, pp. 74-76).

7 Si bien en el presente trabajo no estudiaremos de manera especial el concepto de vocación, se debe recordar que éste resulta central en la filosofía de Ortega. La vocación será presentada por el filósofo como una llamada a un determinado tipo de vida –aquél que conducirá a cada individuo a cumplir con su proyecto vital–. «Esta llamada hacia un cierto tipo de vida o, lo que es igual, de un cierto tipo de vida hacia nosotros, esta voz o grito imperativo que asciende de nuestro más íntimo fondo es la vocación», explica el filósofo en 1934 («Sobre las carreras», V, 298). Así, por provenir «de nuestro más íntimo fondo», Ortega identifica la vocación con una voz interior, o bien, como afirmará en 1943, como una «vocecita íntima», que «reclama desde secretos y profundos pliegues yacentes en nuestro recóndito ser» («Prólogo a *Veinte años de caza mayor*, del Conde de Yebes» (1943), VI, p.273). En ese sentido –agrega allí mismo– se trata un extraño fenómeno en que «nos llamamos a nosotros mismos». Sin embargo, no debe pensarse que esta voz se conoce por simple introspección del yo. En todo caso, la voz de la vocación sólo resuena en el choque con la circunstancia donde intenta realizarse. En palabras de Ortega, «en el choque enérgico con el fuera brota clara la voz del dentro como un programa de conducta» (*Goethe desde dentro* (1932), V, p. 148, énfasis agregado). Esta concepción de la vocación como imperativo hacia la realización de un modo de vida determinado se puede ver de modo claro cuando, en 1946, Ortega se refiere a la vida de Goya («Sobre la leyenda de Goya», IX, pp. 795-822). Allí explica que «cabría decir que Goya no tenía esa vocación, sino que la vocación le tenía a él, le poseía por completo» (p. 811), y en ese sentido, para el pintor aragonés «le eran sinónimos vivir y pintar» (p. 810). Resulta pertinente señalar también que, si bien la utilización del término «vocación» aparece ya desde los primeros escritos de Ortega, se trata de un concepto que aparecerá formulado como tal a finales de los años '20, momento en que el filósofo presentará su concepción de la vida humana en el marco de su raciovitalismo. Un estudio más detallado sobre este concepto, central en la filosofía de Ortega, puede encontrarse en: Puig, J. V., «El concepto de vocación en el pensamiento de Ortega y Gasset», *Anales valentinos: revista de filosofía y teología*, N° 56, 2002, pp. 349-403. O bien, con relación al fondo insobornable, en ÁLVAREZ GONZÁLEZ, E., «El fondo insobornable: el problema de la autenticidad en Ortega», *Revista de Estudios Orteguianos*, N° 25, 2012, p. 163-183.

8 *Sobre una nueva interpretación de la historia universal. Exposición y examen de la obra de Arnold Toynbee: a study of History* (1948), IX, p. 1368.

9 *Principios de Metafísica según la razón vital. [Lecciones del curso 1933-1934]*, IX, p. 94.

que el hombre envidia muchas veces al animal, no porque éste sea feliz —«el animal no es feliz», afirma esta vez— sino, en todo caso, porque no es infeliz. Y concluye: «El animal, como no se ha propuesto nada imposible, ni es feliz ni es infeliz: coincide con su elemento»¹⁰. En definitiva, lo que Ortega busca explicar es que, sean capaces o no de ser felices, lo propio de estos seres es que, al coincidir naturalmente con aquello que deben ser, se ahorran el esfuerzo —propiamente humano— que implica el cumplimiento del proyecto vital. Al hombre, por su parte, lo constituye la necesidad de ser feliz. Su ser consiste en tener que lograr lo que no tiene ni es. Esta situación de menesterosidad deja de manifiesto una de las ideas centrales de la filosofía de Ortega: el hecho de que, a diferencia de los animales, el hombre es un ser histórico («el hombre no tiene naturaleza, sino que tiene... historia»¹¹, explicará el filósofo). En ese sentido, el ser del hombre es siempre «un movimiento, un ir a ser, un echar de menos algo y movilizarse en su busca, pura y frenético anhelo»¹², explica el filósofo.

Así, la felicidad es, para Ortega, una necesidad humana¹³. Es cierto que no parece ser algo imprescindible en términos biológicos: mientras que el hombre no puede renunciar, por ejemplo, a alimentarse, es evidente que se puede prescindir de la felicidad y continuar con vida. Pero lo cierto es que la sensación de necesitarla estará siempre presente en su vida. Ortega afirma que podría argumentarse

10 *La razón histórica [curso de 1940]*, IX, p. 540.

11 *Historia como sistema* (1941), VI, p. 73. La fecha hace referencia a la publicación de la versión española del libro. Sin embargo, el texto tiene su origen en una serie de artículos que Ortega publicó en *La Nación* entre 1934 y 1935, además de que, también en 1935, el libro se publicó en inglés y alemán. Ortega repetirá esta idea en más de una oportunidad. En *El hombre y la gente [curso de 1939-1940]*, por ejemplo, afirmará que «esta vida humana es historia. (...) El hombre no tiene naturaleza, lo que tiene es historia. Esta razón histórica sería la auténtica metafísica» (IX, p. 361). También en 1935, afirma que «pronto hemos caído en la cuenta de que la razón física tenía que fracasar ante los problemas humanos. Porque el hombre no tiene "naturaleza", no tiene un ser fijo, estático, previo y dado». («Aurora de la razón histórica», V, p. 374, énfasis original). Por otra parte, en consonancia con lo que se viene explicando hasta aquí, podemos recordar que en 1932 explica Ortega: «Para la planta, el animal o la estrella, vivir es no tener duda alguna respecto a su propio ser. (...) Por eso su vida no es drama sino... evolución. Pero la vida del hombre es todo lo contrario: es tener que decidir en cada instante lo que ha de hacer en el próximo» (*Goethe desde dentro*, V, p. 131). Sobre la concepción orteguiana del ser humano como ser histórico en el contexto de su filosofía, véase especialmente ALONSO FERNÁNDEZ, M., «El hombre no tiene naturaleza». Un examen de la metafísica orteguiana», *Revista de Filosofía*, V. 45, N°1, 2020, pp. 69-85.

12 *Principios de Metafísica según la razón vital. [Lecciones del curso 1933-1934]*, IX, p. 94.

13 En efecto, el filósofo afirma allí mismo que al hombre «habría que definirlo como el ente que necesita ser feliz» (*idem*). O bien, al sostener —un momento después— que al ser humano «lo constituye esta necesidad de ser feliz y por lo mismo, su ser consiste inexorablemente en tener que esforzarse en lograr lo que no tiene ni es, lo que ha menester ser». Y es que, como afirma una década después, «hay una vocación general y común a todos los hombres. Todo hombre, en efecto, se siente llamado a ser feliz» («Prólogo a *Veinte años de caza mayor*, del Conde de Yeves» (1943), VI, p. 273).

en su contra que, de ser así, la felicidad no es entonces una necesidad sino un mero deseo como tantos otros. Y, en efecto, la felicidad es un deseo. Pero mientras podemos renunciar a cumplir otros deseos sin que esto suponga un vacío en nuestra vida, la felicidad forma parte de aquellos deseos que, aun si se tuviera que renunciar a satisfacerlos, no se podría jamás renunciar a la realidad a la que aspiran. En este último caso, se podría continuar con vida infelizmente, pero no sería posible renunciar al deseo de la felicidad. La felicidad es, en tanto aspiración, un deseo imprescindible y por lo mismo una necesidad.

Concebir la felicidad de esta manera implica que la misma no es un sentimiento de igual importancia que el resto, sino que sirve, además, como criterio para poder determinar si un individuo está eligiendo el camino más propio o si, por el contrario, se encuentra encaminado hacia una vida inauténtica.

En este sentido, Ortega ofrece dos vías que permiten saber si alguien busca –o ha buscado– la autenticidad de su vida o si, en cambio, ha falsificado su existencia. Uno de ellos consiste en realizar una biografía de la persona y determinar, a partir de la misma, si aquel individuo ha, efectivamente, cumplido con aquel personaje que tenía que ser. Para este fin, se debe primero determinar cuál era la vocación del biografiado y, en segundo lugar, revisar la fidelidad de aquel hombre a su destino singular. Esta opción resulta problemática, en la medida en que supone siempre situarse en la circunstancia del biografiado y realizar el intento de adoptar su punto de vista. Y lo cierto es que la reconstrucción de una perspectiva es siempre hipotética y falible¹⁴. Si bien Ortega lo propone en numerosas ocasiones, también conoce la dificultad e incluso la imposibilidad de realizar aquella biografía ideal. Ya en 1908 afirmaba que «la reconstrucción de un carácter personal no sufrirá jamás garantías de exactitud: por eso una biografía es siempre, al cabo, una labor estética en que el acierto permanece enteramente dudoso»¹⁵.

Sin embargo, existe un segundo criterio que permitiría, a cada instante, determinar si una persona está efectivamente cumpliendo con su proyecto vital o no. El mismo se basa en el gusto o disgusto que una persona siente con las actividades que conforman su vida a cada instante. Estos sentimientos permitirían al individuo reconocer su yo y determinar si se encuentra cumpliendo o no con su programa de vida individual. Cuanta más distancia exista entre la vida efectiva de un hombre y aquello que debería ser, más dolor, angustia, malhumor y vacío sentirá. Uno podría ver en aquel individuo su tiesura, su gesto amargo. El hombre que vive de esa manera ha desoído la voz de su vocación; ha abandonado su proyecto vital, lo ha reemplazado por otro y se ha arrebatado así su propia existencia.

14 Cfr. DE SALAS, J., “Perspectiva y método de salvación en Ortega”, en ZAMORA BONILLA, J. (ed.), *Guía Comares de Ortega y Gasset*, Granada: Comares, 2013, p. 234.

15 «Asamblea para el progreso de las ciencias» (1908), I, p. 185.

La felicidad —o, en todo caso, la infelicidad— se torna así en un criterio para determinar cuándo un individuo está efectivamente cumpliendo con su vocación y cuándo no lo está haciendo. Evidentemente, lo central de este criterio no es que permita determinar cuándo otro hombre se encuentra viviendo de manera inauténtica, sino más bien cuándo lo está haciendo uno mismo con su propia vida. En ese sentido, cada uno, individualmente, puede mirar su propia biografía y ver qué deseos no cumplidos han dejado una sensación de pérdida irreparable, de frustración, de sensación de haber falsificado su existencia. Aquellos sentimientos darán la pauta de que se trataba de deseos que correspondían a la vocación propia. Si, en cambio, se siente que la renuncia pudo haber sido algo dolorosa pero que uno permanece íntegro y sin una renuncia irreparable, entonces se trataba de deseos contingentes o que no respondían a su íntimo ser¹⁶.

Por otro lado, un individuo puede utilizar este mismo criterio con la mirada puesta en el futuro a fin de elegir de manera responsable a qué deseos dar lugar y a cuáles renunciar en función de su propio proyecto vital. A la hora de realizar esta proyección a futuro, el hombre siempre puede servirse de su «experiencia de vida», tal como Ortega la denomina: aquel saber espontáneo que el hombre se va formando sin premeditación —aun contra su voluntad— en la medida en que vive. Esta experiencia forma una parte constitutiva de la vida de cada uno. A medida que el ser humano va adquiriendo este saber —explica Ortega— va modificando su propio vivir. «La vida, viviéndose a sí misma, se va como esclareciendo a sí propia, como descubriendo su propia realidad, y esta averiguación, a su vez, entra a formar parte de la vida, se re-convierte en vida, y así sucesivamente. Es el único saber que es, a la vez y de suyo, vivir»¹⁷, explica. Así, a partir de esta experiencia de vida, y al momento de decidir, el hombre debe tomarse el trabajo de detenerse un instante y colocarse de manera imaginaria en diversas circunstancias que se le presentan como posibles y ver cuáles de ellas parecieran aniquilar o anular su propio yo. Si en una determinada circunstancia proyectada el individuo percibe que allí no tendría posibilidad de cumplir con su proyecto vital sentirá «como si en esa vida quedase amputada en una de sus dimensiones esenciales lo que nos aparece como nuestra verdadera vida. (...) Esa vida que viviésemos nos aparecería como el constante freno, negación, amputación de una parte inalienable de nuestro yo»¹⁸, explica Ortega.

16 BASTIDA FREIJEDO, Á., “Salvación y elegancia de la vida. La Metafísica ética de José Ortega y Gasset”, en ABELLÁN, J. L., LLANO ALONSO, F., CASTRO SÁENZ, A. (coords.), *Meditaciones sobre Ortega y Gasset*, Sevilla: Universidad de Sevilla, 2005, p. 99.

17 *Sobre una nueva interpretación de la historia universal. Exposición y examen de la obra de Arnold Toynbee: a study of History* (1948), IX, p. 1198.

18 «¿Qué es la vida? Lecciones del curso 1930-1931», VIII, p. 436.

Esto último le otorga una importancia aún mayor a la felicidad tomada como criterio, en la medida en que permite evaluar la propia vida a cada instante y corregir el rumbo si es necesario. Resulta clara la insistencia de Ortega en decidir de manera reflexiva y encaminarse a la realización del proyecto vital a fin de llevar adelante una vida auténtica. Y si bien es cierto que cada una de las decisiones tomadas forma necesariamente parte de la trayectoria vital del individuo y, en ese sentido, ninguna decisión es indiferente—¹⁹, eso no quita que, una vez realizado un acto arbitrario o por fuera del destino propio, exista igualmente y de manera frecuente la posibilidad encauzar la vida nuevamente hacia la búsqueda de la propia autenticidad. Si no fuera así, la felicidad e infelicidad, en lugar de consistir en criterios que impulsan al hombre a decidir de manera reflexiva y a evaluar permanentemente sus decisiones para orientarlo, de este modo, hacia su propio perfeccionamiento, se tornarían, en cambio, en un premio o un castigo respectivamente. En «Azorín o primores de lo vulgar» explica Ortega: «Yo creo que todo hombre superior ha tenido esta facultad de asistir a su propia existencia, de vivir un poco inclinado sobre su propia vida, en actitud a la vez de espectador exigente y de investigador alerta, pronto a corregir una desviación o desperfecto, presto al aplauso o al silbido»²⁰. Aplauso y silbido, felicidad e infelicidad respectivamente, serán el criterio que dará aviso al hombre si se encuentra falsificando su íntimo ser.

En este punto debe señalarse que, así como la realización de una biografía ideal resultaba un criterio problemático para determinar la autenticidad de una vida humana, también la felicidad lo es cuando se la toma de esa manera. En primer lugar, porque la posibilidad de que un individuo mida a cada instante y de modo preciso el agrado o desagrado con respecto a las actividades que conforman su vida resulta casi tan ideal como la realización de aquel tipo de biografía al que se hacía referencia. Pero, en segundo lugar, porque tal como han advertido algunos estudiosos, puede resultar extraño que la autenticidad —núcleo de la vida moral para Ortega— se fíe simplemente a un sentimiento²¹. Sin embargo, con respecto

19 Y es que, en la medida en que «el hombre no tiene naturaleza, sino que tiene...historia», explica Ortega, «las experiencias de vida hechas estrechan el futuro del hombre» (*Historia como sistema* (1941), VI, pp. 72-73). Así, cada decisión condicionará la conformación de su propia vida. Aunque el individuo será libre de elegir incluso aquella posibilidad que no vale la pena o que parece un contrasentido en su vida, una vez elegida aquella opción, por insensata que sea, llenará una porción incanjeable de su tiempo vital y, por lo tanto, de su propio ser.

20 «Azorín o primores de lo vulgar» (1917), II, p. 295.

21 Sobre este punto llama la atención José Lasaga en su conocido libro *Figuras de la vida buena* (Madrid: Enigma, 2005, pp. 233-234n). Allí señala que, por esta razón —esto es, por tratarse de un sentimiento— la felicidad parece constituir un criterio frágil en la búsqueda de autenticidad individual. Por su parte, De Salas afirma que el malhumor que Ortega utiliza en ciertos casos como indicador es demasiado general como para señalarlos necesariamente la falta de autenticidad de la propia vida. Este criterio, a su juicio, mezcla aspectos temperamentales con decisiones éticas. Se puede caer en

a este punto –y por lo expuesto hasta aquí–, parece claro que, al referirse a los sentimientos de felicidad, infelicidad, malhumor, vacío o gesto amargo, Ortega no los presenta como una simple emoción aislada ante hechos ocasionales, sino como un modo de estar el individuo en su vida. Es, en definitiva, un modo de relación entre la vida efectiva del hombre –la vida en su conjunto– y el mundo en que se desarrolla²². Por esta razón, aunque pueda resultar problemático en algún aspecto, no se le puede quitar a este criterio la enorme función que tiene.

Ahora bien, si el hecho de encaminarse a la propia felicidad está sujeto a la búsqueda de cumplir con la individual vocación, surge naturalmente la pregunta acerca de la posibilidad de que existan vocaciones perversas o negativas, un punto que el mismo Ortega considera en *Goethe desde dentro*²³ y que ha generado discusión entre los estudiosos de su filosofía²⁴. Si bien en el marco de este trabajo no podemos detenernos de manera detallada y con toda la atención que este tema requiere, resulta necesario como mínimo dejar señalado el problema que esto supone en la concepción orteguiana de la felicidad. Y es que, si existiera este tipo de vocaciones negativas habría que pensar, o bien que el individuo debe cumplir de

el malhumor por muchas otras causas ajenas a la inautenticidad, considera (DE SALAS, J., *op. cit.*, p. 241-242). En una línea similar, Marías afirma también que el malhumor puede darse por otras circunstancias. Sin embargo, en su caso, no considera que este tipo de señales constituyan un criterio demasiado amplio en general, en tanto estas «otras ocasiones» en que el malhumor podría surgir consistirían en condiciones exteriores o acontecimientos muy precisos en la vida del individuo. Por esta razón, afirma, el malhumor habitual es un síntoma casi seguro de inautenticidad (MARÍAS, J., «Ortega ante Goethe», *Cuadernos de la Fundación Pastor*, N° 4, 1961, Madrid, p. 25).

22 En esta misma línea, afirma Cerezo que «el sentimiento es la forma de nuestra apertura estimativa a la circunstancia, al mundo en derredor, el modo primario de estar-en realidad y de experimentar cómo nos va en ella. Sentir-se es conjuntamente sentir lo otro, esto es, padecer la realidad circundante en un determinado temple o disposición anímica (*Stimmung*), de distinta valencia afectiva/valorativa según los casos» (CEREZO, P., «Páthos, éthos, lógos (homenaje a Antonio Rodríguez Huéscar)», *Revista de Estudios Orteguianos*, N° 24, 2012, p. 88).

23 *Goethe desde dentro* (1932), V, p. 130.

24 En su clásico libro sobre *La ética de Ortega* antes citado, Aranguren afirma que la existencia de las vocaciones negativas –en concreto, haciendo referencia a la posibilidad de la vocación de ladrón– sería absurda (ARANGUREN, J.L., *op. cit.*, pp. 68). En una línea similar, aunque de modo menos preciso, parece encontrarse Jesús Conill, quien considera que el contenido más importante de la ética orteguiana es sin duda el imperativo de Píndaro –“llega a ser el que eres”–, pero que, sin embargo, éste no está por encima de las ideas convencionales de bien y de mal. Y afirma que, en la vida del individuo, todo ideal debe siempre impulsar el dinamismo perfeccionador de la misma (CONILL SANCHO, J., “Razón experiencial y ética metafísica en Ortega y Gasset”, *Revista de Estudios Orteguianos*, N°7, 2003, p. 111). Por el contrario, posiciones como las de José Lasaga y Jesús M. Díaz Álvarez sostienen que la existencia de vocaciones perversas o negativas no sería de ninguna manera contradictoria con la filosofía de Ortega. Estos posicionamientos pueden encontrarse en DÍAZ ÁLVAREZ, J. M., “Cuestión de libertad. Ética y filosofía política”, en ZAMORA BONILLA, J. (ed.), *op. cit.*, pp. 265-266. Allí el autor explica que concuerda con Lasaga en que “el imperativo pindárico no puede ser colonizado por el imperativo categórico” y que, por esta razón, el planteamiento orteguiano “se aleja, con todos los problemas que ello supone, de un fundacionalismo en el dominio de la ética”.

todos modos con aquella llamada –en virtud de que no puede renunciar al deseo de felicidad–, o bien, que el individuo que posee una vocación de este tipo debe renunciar a la búsqueda de la felicidad –en tanto continuar con esta búsqueda supondría el intento de realizar efectivamente una vocación perversa–. Tanto una opción como la otra resultan problemáticas. Sin embargo, si bien es cierto que en *Goethe desde dentro* Ortega únicamente se refiere a lo problemático de este asunto, en otros momentos de su obra sí sugiere que sería contradictorio aceptar vocaciones de este tipo. Esto puede verse, por ejemplo, en su artículo “No ser hombre de partido”, donde el filósofo afirma de modo explícito que el mal proviene de la inautenticidad, en tanto no existe la maldad creadora. En palabras de Ortega: «Toda maldad viene de una radical: no encajarse en el propio sino. De ahí que no haya maldad creadora. Todo acto perverso es un fenómeno de compensación que busca el ser incapaz de crear un acto espontáneo, auténtico, que brota de su Destino»²⁵. En ese sentido, ejemplifica allí mismo: «La mentira es un ejemplo particular de acción en que el hombre abandona su verdadero ser». Más allá de éste y otros ejemplos concretos, es cierto que en su obra falta una mayor argumentación al respecto. Sin embargo, consideramos que este tipo de afirmaciones dejan en evidencia que Ortega no consideraba que existieran, en efecto, vocaciones para el mal²⁶.

Se ha dicho anteriormente que la felicidad resulta una necesidad imprescindible en la vida humana, en tanto constituye una meta que la orienta de manera constante al cumplimiento de la vocación. Concebida de esta manera, entonces, la felicidad adquiere en Ortega un carácter utópico. Ahora bien, esta utopía, en la medida en que orienta a los hombres a la autenticidad, no parece tener un sentido negativo. Será necesario, entonces, detenerse en el concepto orteguiano de utopía y determinar en qué sentido, a juicio del filósofo, alcanzar la felicidad resulta utópico para el ser humano.

2. UTOPIA Y FELICIDAD

Uno de los combates que Ortega emprende en sus obras es, precisamente, contra el utopismo racionalista, aquel que había alejado a los hombres del conocimiento concreto del mundo. El error racionalista consistía en afirmar que la realidad tenía una fisonomía propia y que ésta era independiente del punto de

25 «No ser hombre de partido» (1930), IV, p. 309.

26 Sobre este tema me he detenido especialmente en PERPERE VIÑUALES, M., “La posibilidad de una ética de la situación en la filosofía de Ortega y Gasset”, *Revista de Estudios Ortegaianos*, N°40, 2020, pp. 157-165.

vista de cada individuo, pueblo o época. De este modo, resultaba imposible que cualquier punto de vista individual fuera capaz de coincidir con el punto de vista del absoluto. El problema era que por no lograr esta coincidencia se lo consideraba falso. Ortega critica esta postura: para el filósofo la realidad se conoce desde diversos puntos de vista y todos ellos igualmente válidos. La única perspectiva que no es admisible es justamente aquella que pretende ser la única. «Dicho de otra manera: lo falso es la utopía, la verdad no localizada, vista desde «lugar ninguno». El utopista —y esto ha sido en esencia el racionalismo— es el que más yerra, porque es el hombre que no se conserva fiel a su punto de vista, que deserta de su puesto»²⁷, explica Ortega. Este tipo de utopía puede considerarse mortal, afirma unos años más tarde²⁸, en la medida en que implica no hallarse en circunstancia alguna, situación que muy bien define la muerte.

Esto es lo que Ortega considera un falso utopismo: aquel que pretende ser válido para todos los hombres, lugares y tiempos por igual. Y, por sobre todo esto, se presenta además como asequible a los hombres. Ortega critica ampliamente este utopismo, principalmente en «Miseria y esplendor de la traducción». Allí explica que el error de éste consiste principalmente en «creer que lo que el hombre desea, proyecta y se propone es, sin más, posible. Por nada siento mayor repugnancia y veo en él la causa máxima de cuantas desdichas acontecen ahora en el planeta»²⁹. Estas desdichas a las que se refiere Ortega no consisten únicamente en el abandono del punto de vista individual y, por lo tanto, en la falsificación de la propia vida. El problema es mayor en la medida en que, al desertar del punto de vista propio y al buscar, en cambio, valores absolutos y abstractos, se corre el riesgo —de hecho, se está a un paso, afirma Ortega— de exigir la realización de estos valores y de imponer el camino correcto para este fin.

Pero no todo utopismo es falso o negativo. A lo largo de las obras de Ortega pueden encontrarse diversas referencias a un utopismo de carácter positivo que se debe tener igualmente presente. Éste consiste, fundamentalmente, en metas u objetivos que, aun siendo completamente irrealizables, incitan al hombre a la acción. Lo propio de estos objetivos no es simplemente que se presenten como irrealizables, sino que, en tanto imposibles, funcionan como resortes que lanzan al individuo a la acción y a la superación permanente. Esto no significa que el individuo deba negar la realidad y conformarse con que ésta lo mantenga en movimiento. Por el contrario, este tipo de utopía es positiva en tanto orienta la vida del hombre hacia su propia plenitud.

27 *El tema de nuestro tiempo* (1923), III, p. 614.

28 «El estatuto catalán» (1932), V, p. 70.

29 «Miseria y esplendor de la traducción» (1937), V, p. 712.

Ortega utiliza el caso de la traducción para ejemplificar al buen y al mal utopista. Ambos coinciden en un punto básico: consideran que sería deseable corregir la realidad que, a causa de la diferencia de idiomas o lenguas, lleva frecuentemente a los hombres a la incomunicación. Pero existe una diferencia en el planteamiento de uno y otro. Mientras que «el mal utopista piensa que, *puesto* que es deseable, es posible, y de esto no hay más que un paso hasta creer que es fácil»³⁰, sucede que —afirma allí mismo— «el buen utopista, en cambio, piensa que, *puesto* que sería deseable libertar a los hombres de la distancia impuesta por las lenguas, no hay probabilidad de que se pueda conseguir; por tanto, que sólo cabe lograrlo en medida aproximada. (...) Ello abre ante nuestro esfuerzo una actuación sin límites en que siempre cabe mejora». Así, el buen utopista que Ortega describe es aquel que concibe la utopía justamente como un ideal que no podrá nunca realizar o alcanzar por completo, pero al que, sin embargo, sí podrá buscar aproximarse todo lo posible. Este hombre es, en primer lugar, un completo realista y «sólo cuando está seguro de que ha visto bien, sin hacerse la menor ilusión y en su más agria desnudez, la realidad, se resuelve contra ella garboso y se esfuerza en reformarla en el sentido de lo imposible, que es lo único que tiene sentido»³¹, explica Ortega un momento después. La utopía, tomada positivamente, fomentaría en el hombre un esfuerzo sin límites que lo conduciría progresivamente a un mayor perfeccionamiento y a una constante superación.

En el apartado anterior afirmamos que, para Ortega, la vida humana nunca es felicidad por completo, sino que consiste en una búsqueda permanente de la misma. Explicamos que, para el filósofo, no es posible realizar de manera absoluta el proyecto vital o, lo que es lo mismo, que no puede existir una vida humana en la que se dé la perfecta adecuación entre la vocación individual y la vida efectiva. Para reforzar aún más esta idea, pueden recordarse las palabras de Ortega en 1932, cuando afirma que «toda vida es, más o menos, una ruina entre cuyos escombros tendremos que descubrir lo que la persona tenía que haber sido. (...) Todos sentimos nuestra vida real como una esencial deformación, mayor o menor, de nuestra vida posible»³². También en 1937, en esta misma línea afirmaba que «los hombres andan siempre melancólicos, maniáticos y frenéticos, maltraídos por todos estos morbos que Hipócrates llamó divinos. Y la razón de ello está en que los quehaceres humanos son irrealizables. El destino —el privilegio y el honor— del hombre es no lograr nunca lo que se propone y ser pura pretensión, viviente utopía. Parte siempre hacia el fracaso, y antes de entrar en la pelea lleva

30 «Misericordia y esplendor de la traducción» (1937), V, p. 712.

31 *Idem*, p. 713.

32 *Goethe desde dentro* (1932), V, p. 126.

ya herida la sien»³³. Un caso más puede encontrarse aún en 1939, cuando afirma que cada hombre es «un personaje que no se realiza nunca del todo»³⁴. Lejos queda para el individuo, entonces, la posibilidad de una vida completamente feliz.

Sin embargo, más allá de considerar que la felicidad consiste en una utopía para el ser humano, ahora se comprende que esto no debe entenderse en un sentido negativo. Así como, en palabras del filósofo, «el Norte y el Sur nos orientan, sin necesidad de ser ciudades asequibles para las cuales quepa tomar un billete de ferrocarril»³⁵, de la misma manera la felicidad –que no es concebida como un sentimiento aislado sino, como se señaló anteriormente, como un modo de estar el individuo en su vida– guía al ser humano por el camino de su propia realización. La búsqueda de la vida auténtica, de cumplir con el proyecto vital propio, conduce al hombre, entonces, a salir de la quietud, a ponerse en movimiento aun sabiendo que será imposible para él alcanzar por completo y de manera permanente aquella meta a la cual se dirige. El hombre que concibe de esta manera la felicidad, en su búsqueda por alcanzar una vida auténtica, se esforzará por lograrla aun teniendo siempre la conciencia del fracaso.

3. CONSIDERACIONES FINALES

Comenzábamos este trabajo evocando las palabras que el joven Ortega había escrito en 1906 para ser pronunciadas por su padre en el acto de entrega de premios de los Juegos Florales de Valladolid. Allí, el filósofo señalaba que una vida que renuncia a la felicidad no puede, en rigor, ser llamada «vida». Y es que si ella es una actividad que se realiza siempre hacia adelante, el individuo debe proyectar en ese porvenir una meta y una ilusión que lo inciten a permanecer siempre en movimiento.

Ahora bien, si la felicidad tiene siempre un carácter utópico, para que sirva como guía para la acción se debe contar con algún tipo de confianza en que aquella felicidad puede ser vivida. O, como mínimo, en que, en el camino hacia la realización del proyecto vital, aquella felicidad guía a cada uno por el camino adecuado. Este sentimiento de confianza en el futuro es la esperanza³⁶. La vida es permanente quehacer porque es, a la vez, esperanza de alcanzar lo deseado.

33 «Miseria y esplendor de la traducción» (1937), V, p. 707.

34 *Ensismamiento y alteración* (1939), V, p. 541.

35 «El origen deportivo del Estado» (1924), II, p. 706.

36 Un análisis detallado sobre el concepto de «esperanza» en Ortega, se encuentra en el conocido trabajo de Pedro LAÍN ENTRALGO, *La espera y la esperanza. Historia y teoría del esperar humano*, Madrid: Revista de Occidente, 1957, puntualmente en las pp. 416-433.

En ese sentido, pueden recordarse las palabras de Ortega tan sólo cuatro años después del citado discurso en Valladolid. En «Ideas sobre Pío Baroja», el filósofo explica que «el hombre no puede vivir plenamente si no hay algo capaz de llenar su espíritu hasta el punto de desear morir por ello. (...) Sólo nos empuja irresistiblemente hacia la vida lo que por entero inunda nuestra cuenca interior. Renunciar a ello sería mayor muerte que con ello fenecer»³⁷. La esperanza en que el proyecto vital es realizable y que, por lo mismo, la felicidad es asequible al hombre, resulta fundamental en la búsqueda de la propia plenitud. Es probable que por esto mismo Ortega agregara allí mimo: «No he podido sentir nunca hacia los mártires admiración, sino envidia. Es más fácil *lleno de fe* morir, que exento de ella arrastrarse por la vida».

Sin embargo, en aquella larga serie de esfuerzos en que consiste la vida, la gran mayoría de los individuos cae en la desesperanza y acaba rindiéndose. Ello los conduce a falsificar su propia vida, a convertirse en un personaje inferior y, finalmente –como se indicaba al comienzo– a «vivir dolorosamente un día tras otro». Por esta razón, tener presente el carácter utópico de la felicidad resulta decisivo. El individuo que comprende que la importancia de ésta no radica en el hecho de que pueda ser alcanzada sino, en cambio, en ser una meta que impulsa a la acción, podrá emprender la búsqueda de su propia autenticidad –aún en medio de la incertidumbre y la inseguridad que ello conlleva– y transitar ese camino con alegría creadora.

REFERENCIAS BIBLIOGRÁFICAS

- ALONSO FERNÁNDEZ, M., «El hombre no tiene naturaleza». Un examen de la metafísica orteguiana», *Revista de Filosofía*, V. 45, N°1, 2020, pp. 69-85.
- ÁLVAREZ GONZÁLEZ, E., «El fondo insobornable: el problema de la autenticidad en Ortega», *Revista de Estudios Orteguianos*, N° 25, 2012, p. 163-183.
- BASTIDA FREIJEDO, Á., «Salvación y elegancia de la vida. La Metafísica ética de José Ortega y Gasset», en ABELLÁN, J. L., LLANO ALONSO, F., CASTRO SÁENZ, A. (coords.), *Meditaciones sobre Ortega y Gasset*, Sevilla: Universidad de Sevilla, 2005, pp. 55-108.
- CONILL SANCHO, J., «Razón experiencial y ética metafísica en Ortega y Gasset», *Revista de Estudios Orteguianos*, N°7, 2003, pp. 95-117.
- DE SALAS, J., «Perspectiva y método de salvación en Ortega», en ZAMORA BONILLA, J. (ed.), *Guía Comares de Ortega y Gasset*, Granada: Comares, 2013, pp. 231-249.

37 «Ideas sobre Pío Baroja» (1910), II, p. 228.

- DÍAZ ÁLVAREZ, J. M., “Cuestión de libertad. Ética y filosofía política”, en ZAMORA BONILLA, J. (ed.), *Guía Comares de Ortega y Gasset*, Granada: Comares, 2013, pp. 241-286.
- LAÍN ENTRALGO, P., *La espera y la esperanza. Historia y teoría del esperar humano*, Madrid: Revista de Occidente, 1957.
- LASAGA, J., *Figuras de la vida buena*, Madrid: Enigma, 2005.
- LÓPEZ ARANGUREN, J. L., *La ética de Ortega*, Madrid: Taurus (tercera ed.), 1966.
- MARÍAS, J., “Ortega ante Goethe”, *Cuadernos de la Fundación Pastor*, N°4, 1961, Madrid
- ORTEGA Y GASSET, J., *Obras completas*, 10 vols., Madrid: Fundación Ortega y Gasset/Taurus, 2004-2010.
- PERPERE VIÑUALES, M., «La posibilidad de una ética de la situación en la filosofía de Ortega y Gasset», *Revista de Estudios Orteguianos*, N°40, 2020, pp. 157-165.
- PUIG, J. V., “El concepto de vocación en el pensamiento de Ortega y Gasset”, *Anales valentinos: revista de filosofía y teología*, N° 56, 2002, pp. 349-403.

SIGNIFICACIÓN Y LENGUAJE EN *EL SER Y LA NADA*

MEANING AND LANGUAGE IN *BEING AND NOTHINGNESS*

STÉPHANE VINOLO

Doctor en Filosofía

Doctor en Teología

Profesor principal de la Pontificia Universidad Católica del Ecuador

Escuela de Filosofía

Pontificia Universidad Católica del Ecuador

Quito/Ecuador

svinolo@puce.edu.ec

ORCID: 0000-0002-3371-0805

Recibido: 5/06/2020

Revisado: 31/03/2021

Aceptado: 6/09/2021

Resumen: La filosofía del lenguaje de Sartre ocupa apenas algunas páginas en *El ser y la nada*. No obstante, su interés es fundamental dado que cristaliza el problema de la libertad en el momento fenomenológico de la primera filosofía de Sartre. Efectivamente, el sujeto no es arrojado dentro de un mundo virgen de toda significación, sino siempre ya habitado por otros que lo significan mediante significaciones colectivas, el ser-para-otro y el fenómeno de la alteridad. Con el fin de mantener la libertad absoluta del sujeto, Sartre, en diálogo con los pensamientos de Brice Parain y de Jean Paulhan defiende una teoría original del lenguaje que no somete el sujeto al sentido de las palabras, ni a las reglas sintácticas de sus articulaciones. Al hablar, el sujeto crea la gramática y resignifica cada una de las palabras. Así, la filosofía del lenguaje de Sartre evidencia el núcleo fenomenológico de la primera filosofía de Sartre, a través de una libertad absoluta del sujeto que yace en su posición de creador de sentido.

Palabras claves Lenguaje, Brice Parain, Jean Paulhan, Jean-Paul Sartre, significación

Abstract: Sartre's philosophy of language barely occupies a few pages in *Being and Nothingness*. However, its interest is fundamental since it concentrates the problem of freedom in the phenomenological moment of Sartre's philosophy. Indeed, the subject is not thrown into a world free of all significance, but always already inhabited by others who signify it through collective meanings, being-for-others and the phenomenon of otherness. In order to maintain the absolute freedom of the subject, Sartre, in dialogue with the thoughts of Brice Parain and Jean Paulhan, defends an original theory of language that does not subjugate the subject to the meaning of words, nor to the syntactic rules of their articulations. By speaking,

the subject creates the grammar and re-signifies each of the words. Thus, Sartre's philosophy of language reveals the phenomenological core of Sartre's first philosophy, through the subject's absolute freedom that lies in his position as creator of meaning.

Keywords: Language, Brice Parain, Jean Paulhan, Jean-Paul Sartre, signification.

“[...] expresar y ser se identifican”¹.

“Brice Parain es bastante inteligente, sin más. Es un tipo que piensa acerca del lenguaje como Paulhan”².

INTRODUCCIÓN

El problema de la significación y de manera más amplia, el problema del lenguaje, ocupan en filosofía un papel central. Ricoeur identificó este vínculo entre filosofía y lenguaje, y localizó su encuentro dentro del concepto de “esencia”: “El problema de la esencia es idéntico al del lenguaje, al de la denominación”³. Si la filosofía responde, desde su surgimiento griego, a la pregunta “¿qué es?”⁴, el estudio del lenguaje no es solamente una de sus ramas, tal como pueden serlo la estética o la filosofía política, sino lo que posibilita el mismo filosofar⁵. La importancia de la significación es aún más evidente en el campo de la fenomenología ya que ésta se da como tarea explícita, desde Husserl, el dilucidar los sentidos mediante los cuales se dan los fenómenos: “La fenomenología procede aclarando visualmente, determinando y distinguiendo el sentido”⁶. Así, el problema del sentido y de la significación permite alcanzar el centro del gesto fenomenológico ya que tanto la constitución de los objetos intencionales como la del mundo, es decir, la misma fenomenalización, es un proceso de significación.

1 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 511.

2 « Brice Parain est assez intelligent, sans plus. C'est un mec qui pense sur le langage comme Paulhan. », SARTRE, Jean-Paul, « *Lettre à Simone de Beauvoir du 30 avril 1937* ». En : SARTRE, Jean-Paul, *Lettres au castor et à quelques autres, Tome 1, 1926-1939*. Paris : Gallimard, 1983, p. 116.

3 RICCEUR Paul, *Ser, esencia y sustancia en Platón y Aristóteles*, México, Siglo veintiuno editores, 2013, p. 23.

4 VERNANT, Jean-Pierre, *Œuvres. Religions, Rationalités, Politique*. Pais : Seuil, Paris, 2007.

5 BELAVAL, Yvon, *Les philosophes et leur langage*. Paris : Gallimard, 1952.

6 HUSSERL, Edmund, *La idea de fenomenología, cuarta lección*. México: FCE, 1950, 1982, p. 71.

La centralidad de la significación es evidente en la primera filosofía de Sartre, tal como culmina en *El ser y la nada* puesto que, en ésta, el proceso de significación no solo se manifiesta en la relación entre el sujeto y el mundo, en el “circuito de la ipseidad”, sino además en la misma fractura ontológica que yace en el para-sí. A diferencia del ser-en-sí, plenum de ser⁷, el para-sí se distingue como presencia a sí: “El ser de la consciencia en tanto que consciencia consiste en existir a distancia de sí como presencia a sí, y esa distancia nula que el ser lleva en su ser es la Nada”⁸. La presencia implica la dualidad y la posibilidad de ser otra cosa que lo que uno es, una ruptura con su propio ser: “[...] toda presencia a implica dualidad, y por lo tanto, separación, por lo menos virtual”⁹. Pero esta distancia es, además, “significación”, puesto que Sartre la piensa a la luz del concepto de remisión, tal como un signo: “El sí remite, pero remite precisamente al sujeto. Indica una relación del sujeto consigo mismo y esta relación es precisamente una dualidad, pero una dualidad particular, ya que exige símbolos verbales particulares”¹⁰. La posibilidad de la distancia consigo mismo, de reconstruirse mediante nuevos valores y trascendencias es entonces ante todo un problema de ruptura y de remisión, es decir, de significación.

De la misma manera, el circuito de la ipseidad marca la importancia de la significación ya que revela la constitución del mundo a la luz del proyecto originario del sujeto, que no es más que un deseo de auto-fundación. Dado que el para-sí surgió a raíz de un acaecimiento ontológico absoluto, todo sujeto apunta hacia la figura paradójica de un en-si-para-sí, o causa sui, o Dios¹¹: “Cada realidad humana es a la vez proyecto directo de metamorfosear su propio Para-sí en En-sí-Para-sí, y proyecto de apropiación del mundo como totalidad del ser-en-sí, bajo las especies de una cualidad fundamental”¹². Ahora bien, dado que la proyección de sentido sobre el existente bruto permite la aparición de un mundo, la fenomenalización es una valorización vinculada esencialmente con una de las modalidades del sentido: “Es [el valor] como el sentido y el más allá de todo trascender, es como el en-sí ausente que infesta el ser para sí”¹³. El mundo refleja entonces el sujeto puesto que

7 “El en-sí está pleno de sí mismo, y no cabe imaginar plenitud más total, adecuación más perfecta del contenido al continente: no hay el menor vacío en el ser, la menor fisura por la que pudiera deslizarse la nada”, SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 130.

8 *Ibid.*, 135.

9 *Ibid.*, 134.

10 *Ibid.*, 133.

11 “Ser hombre es tender a ser Dios; o, si se prefiere, el hombre es fundamentalmente deseo de ser Dios.”, *Ibid.*, 764.

12 *Ibid.*, 828.

13 *Ibid.*, 154.

aparece de manera valorada, significado a la luz del proyecto de éste¹⁴. El carácter significativo del mundo es de hecho tan importante que el existente bruto nunca puede aparecer en tanto tal, sino exclusivamente significado por un proyecto¹⁵. Por este motivo, en la manifestación de un mundo es imposible distinguir lo que proviene del existente bruto de lo que proviene del sujeto¹⁶. Así, todo el proceso de constitución y de unificación del mundo es un proceso de constitución, es decir, de significación: “El mundo (es) mío porque está infestado por posibles de los cuales son consciencias las consciencias posibles (de) sí que yo soy, y esos posibles, en tanto que tales, le dan su unidad y su sentido de mundo”¹⁷.

El problema de la significación habita entonces la primera filosofía de Sartre, tanto en el problema del sujeto, que en la relación significante que yace entre el para-sí y el existente bruto, o en la totalización que surge a raíz de esta significación: “Llamaremos circuito de la ipseidad a la relación entre el para-sí y el posible que él es, y mundo a la totalidad del ser en tanto que atravesada por el circuito de la ipseidad.”¹⁸ Por esta razón, Sartre desarrolla, en *El ser y la nada*, una teoría de la significación que se despliega bajo la forma de una teoría del lenguaje, pensada en diálogo con dos de sus contemporáneos: Brice Parain y Jean Paulhan¹⁹. Esta teoría presenta un interés fundamental dado que no solo permite entender la fenomenología de Sartre al evidenciar lo que entiende por “sentido” y por “significación”, sino además permite resolver uno de los problemas fundamentales al cual se enfrenta dicha fenomenología. Si toda manifestación, en fenomenología, es cuestión de significación, ¿cómo conciliarla con la libertad absoluta? ¿Acaso el lenguaje no limita la libertad por la precedencia de sentidos y reglas que rigen su uso? Más aún, ¿estas reglas lingüísticas no evidencian la precedencia del otro²⁰?

14 Solo ciertas características del mundo, las cualidades, permiten acceder al fenómeno de ser. Cf. VINOLO, Stéphane, “*El neorrealismo absoluto en El ser y la nada*”. Eidos, 2021, En edición.

15 “[...], el *datum* no aparece jamás como existente bruto y en-sí al para-sí; se descubre siempre como motivo, puesto que no se revela sino a la luz de un fin que lo ilumina.”, SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 662.

16 “[...] la situación, producto común de la contingencia del en-sí y de la libertad, es un fenómeno ambiguo en el cual es imposible al para-sí discernir el aporte de la libertad y del existente bruto.”, *Ibid.*, 662-663.

17 *Ibid.*, 167.

18 *Ibid.*, 164-165.

19 Paradójicamente, las relaciones entre Sartre, Paulhan y Parain, son uno de los temas menos abordados en los estudios sartreanos. Incluso en lengua francesa, la mayoría de estudios se focaliza sobre la relación editorial entre los tres pensadores, o sobre las divergencias entre sus concepciones de la literatura. Aquí, poniéndonos en los pasos de Vincent Descombes y de Raoul Moati, es desde el punto de vista de la pregunta de la significación que queremos analizarlas con el fin de precisar la fenomenología de Sartre.

20 CAMUS, Renaud, *Syntaxe ou l'autre dans la langue*. Paris : POL, 2004.

El problema del lenguaje permite entonces, mediante la significación, precisar los temas más fundamentales de la fenomenología de Sartre.

1. LA PRECEDENCIA DE LAS SIGNIFICACIONES

La afirmación sartreana de la libertad absoluta del sujeto no significa que nada le pueda aparecer en tanto que obstáculo. Si durante un paseo, un derrumbe lleva un peñasco sobre el camino, éste será, en un primer momento, un límite a la libertad puesto que obstaculiza el paso. No obstante, si el ser de la roca permanece fuera del alcance de la libertad ya que no puede hacer que la roca no esté allí²¹, su fenomenalización mediante significaciones determinadas depende del sujeto: “Lo dado en sí como resistencia o como ayuda no se revela sino a la luz de la libertad proyectante. [...] No hay obstáculo absoluto, sino que el obstáculo revela su coeficiente de adversidad [...] en función del valor del fin puesto por la libertad”²². El estar ahí de la roca es contingente, pero su aparecer como obstáculo para acabar el paseo, o en tanto que oportunidad para poder observar el paisaje depende del proyecto del sujeto: “Tal peñasco, que manifiesta una resistencia profunda si quiero desplazarlo, será, al contrario, una ayuda preciosa si quiero escalarle, para contemplar el paisaje”²³. La fenomenalización de la roca, al darse a través de valoraciones y de significaciones es tributaria del sujeto puesto que es a la luz del proyecto que se lanza sobre el existente bruto, que éste se fenomenaliza en tanto que obstáculo o en tanto que oportunidad: “Siendo la intención elección del fin y revelándose el mundo a través de nuestras conductas, la elección intencional del fin revela el mundo, y el mundo se revela tal o cual (en tal o cual orden) según el fin elegido”²⁴. El mundo, entonces, no es más que el reflejo del circuito de ipseidad mediante el cual un sujeto significa un existente bruto que, sin estas significaciones, permanecería invisible²⁵, y que no puede adquirir significaciones sino subjetivas: “Nosotros elegimos el mundo –no en su contextura en-sí, sino en su significación– al elegirnos”²⁶.

21 Salvo que el sujeto se refugie dentro del mundo mágico de las emociones, por ejemplo, desmayándose, lo que anularía la presencia de la roca nihilizando el mundo dentro del cual aparecía: “Podemos concebir ahora en qué consiste una emoción. Es una transformación del mundo.”, SARTRE, Jean-Paul, *Bosquejo de una teoría de las emociones*. Madrid: Alianza, 1971, 2015, p. 69.

22 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, pp. 664-665.

23 *Ibid.*, 655.

24 *Ibid.*, 649.

25 “[...], el *datum* no aparece jamás como existente bruto y en-sí al para-sí; se descubre siempre como *motivo*, puesto que no se revela sino a la luz de un fin que lo ilumina.”, *Ibid.*, 662.

26 *Ibid.*, 630.

Sin embargo, el proceso significante individual se enfrenta con la precedencia de significaciones contingentes: “Todo sería muy simple [...] si perteneciera a un mundo cuyas significaciones se descubrieran simplemente a la luz de mis propios fines”²⁷. El sujeto no se ve arrojado dentro de un mundo virgen de significaciones. Ya existen varias redes de significaciones a las cuales se somete, ya que siempre habita un mundo: “[...] infestado por [su] prójimo”²⁸. Existen, en *El ser y la nada*, tres capas de significaciones que se imponen al sujeto antes que él las haya significado. Por lo tanto, dentro de tres campos, el sujeto está rodeado de sentidos de los cuales no es la fuente: “[...] los utensilios ya significantes (la estación, el semáforo del ferrocarril, la obra de arte, el cartel de movilización), la significación que descubro como ya mía (mi nacionalidad, mi raza, mi aspecto físico) y, por último, el Otro como centro de referencia al que esas significaciones remiten”²⁹.

El primer tipo de significaciones preexistentes es aquel de las significaciones que se dan dentro de un mundo. Al pasearse por la calle, no se descubre un existente bruto que se fenomenaliza a la luz de un proyecto individual; más bien, se percibe una casa, una tienda, un policía o un escritor público. Por todas partes surgen significaciones colectivas que des-individualizan las manifestaciones subjetivas para dirigirse a un hombre colectivo. Las señales de tránsito, ya sean órdenes, prohibiciones o recomendaciones, significan antes que el sujeto las haya significado, y se dirigen a todos, es decir, a un hombre anónimo: “[...] los modos de empleo, las designaciones, las órdenes o las prohibiciones, los carteles indicadores, se dirigen a mí en tanto que soy cualquiera; [...]”³⁰. Hablemos de señales de “entrada”, de “salida” o de “pare”, el sujeto vive rodeado de significaciones colectivas que lo preceden, que le recuerdan el carácter habitado y pre-humanizado de todo mundo: “[...] las indicaciones agregan, al coeficiente de adversidad que hago nacer yo sobre las cosas, un coeficiente de adversidad propiamente humano”³¹.

El segundo tipo de significaciones que se impone al sujeto se da mediante la mirada del otro. No solo el mundo se fenomenaliza mediante significaciones colectivas sino además el mismo sujeto existe de manera significada, dentro del mundo del otro: “[...] el prójimo no es solamente aquel que veo, sino aquel que me ve. Encaro el prójimo en tanto que éste es un sistema conexo de experiencias fuera de alcance, en el cual yo figuro como un objeto entre los otros”³². En la mirada del otro, el sujeto descubre un nuevo mundo de significaciones del cual

27 *Ibid.*, 691.

28 *Ibid.*, 691. Lastimosamente, “infestar” no refleja el carácter fantasmático y espectral del verbo francés “hanter”.

29 *Ibid.*, 691.

30 *Ibid.*, 693.

31 *Ibid.*, 693.

32 *Ibid.*, 322.

forma parte en tanto que objeto significado. De ahí la nueva modalidad de ser del sujeto que Sartre describe como ser-para-otro. Mediante el surgimiento del otro, el sujeto adquiere la consistencia de un ser; en su mundo, el sujeto es francés, burgués o profesor, y el entrecruzamiento de las miradas fundamenta el carácter originariamente conflictivo de las relaciones con el otro. Dada la fuerza objetivadora de la mirada, uno experimenta que existe calificado por y para el otro, como objeto de otra consciencia significativa. De ahí que Sartre pueda afirmar que: “[...] el Infierno son los Otros”³³, o que: “El conflicto es el sentido del ser-para-otro”³⁴.

Finalmente, el tercer tipo de significaciones contingentes es el mismo fenómeno de la alteridad que no aparece dentro de un mundo sino en tanto que mundo. Entre los fenómenos, uno goza de una característica singular, la de ser una fuente de significación: “La percepción del objeto-prójimo remite a un sistema coherente de representaciones, y este sistema no es el mío”³⁵. El otro no aparece como objeto sometido a las reglas de constitución del mundo por el sujeto; se manifiesta, al contrario, como fuente de sentido que desestabiliza el mundo, fenomenalizando nuevas relaciones entre los objetos. Cuando un objeto aparecía como medianamente deseable y cercano, basta con que otra persona lo desee para que aparezca de repente como lejano y muy deseable³⁶. Así, el otro es un ladrón³⁷ del mundo, una caída³⁸, o un hueco dentro del cual se hunden y desaparecen las significaciones: “[...] no se trata de una fuga del mundo hacia la nada o fuera de sí mismo. Más bien, parecería como horadado en medio de su ser por la boca de un vaciadero, por donde perpetuamente se me oscurece”³⁹. El otro es una significación que se impone al sujeto, más allá de su constitución del mundo y a través de una descentración de éste: “La aparición del prójimo en el mundo corresponde, pues, a una descentración del mundo, que socava por debajo la centralización operada por mí al mismo tiempo”⁴⁰. Su fenomenalización redobla entonces el movimiento de significación del sujeto ya que resignifica las significaciones del sujeto. Por este motivo, cuando el sujeto significa el existente

33 SARTRE, Jean-Paul, *A Puerta cerrada, Escena V*. En. SARTRE, Jean-Paul, *La puta respetuosa / A puerta cerrada*. Bogotá: Orbis, 1983, p. 186.

34 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 499.

35 *Ibid.*, 321.

36 VINOLO, Stéphane, « Critique de la raison mimétique – René Girard lecteur de Jean-Paul Sartre ». En RAMOND, Charles (ed.), *René Girard, La théorie mimétique, de l'apprentissage à l'apocalypse*. Paris : PUF, 2010, pp. 59-104.7

37 “Así, de pronto, ha aparecido un objeto que me ha robado el mundo.”, SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 358.

38 “[...] mi caída original es la existencia del otro; [...]”, *Ibid.*, 367.

39 *Ibid.*, 358.

40 *Ibid.*, 358.

bruto, se enfrenta con la significación-significante del fenómeno singular que es el otro: “[...] el otro, como mirada, no es sino eso: mi trascendencia trascendida”⁴¹.

A pesar de la libertad absoluta del sujeto que yace en su capacidad significativa que hace de cualquier obstáculo a su libertad, un obstáculo para su libertad, evidenciándola a la hora que la limita, encontramos en Sartre tres tipos de significaciones que preceden el proceso de significación, lo que se evidencia en la oposición entre el Estado del mundo y la situación: “[...] no hay ya situación como organización de un mundo significativo en torno de la libre elección de mi espontaneidad, sino un estado que me es impuesto”⁴². Ahora bien, el Estado se encuentra de manera paradigmática en el lenguaje. Tanto el vocabulario como la gramática cristalizan la precedencia de las significaciones que se imponen al sujeto: “Yo, por quien las significaciones vienen a las cosas, me encuentro comprometido en un mundo ya significativo, que me refleja significaciones no puestas por mí”⁴³. Cada ser humano, uno podría pensar, utiliza un lenguaje mediante estructuras que le son impuestas, razón por la cual es posible hacer del ejemplo sartreano del lenguaje el paradigma de cualquier significación. Más aún, es legítimo unificar estas dos problemáticas ya que Sartre habla de un “modo fundamental del lenguaje”⁴⁴ que habita toda relación intersubjetiva, y por lo cual todo proceso de significación, en tanto que surge de una libertad, es una forma de lenguaje: “Por el solo hecho de que, por mucho que yo haga, mis actos libremente concebidos y ejecutados, mis pro-yectos hacia mis posibilidades, tienen afuera un sentido que me escapa y que experimento, soy lenguaje”⁴⁵. El lenguaje es entonces el paradigma de la preexistencia de significaciones y de estructuras que obstaculizan la libertad creadora de significaciones del sujeto.

2. INTERPRETACIÓN Y CREACIÓN: SARTRE LECTOR DE PAULHAN

La respuesta de Sartre a la precedencia de significaciones está sintetizada en una sola frase que sustenta su teoría del lenguaje: “[...] la estructura elemental del lenguaje es la oración”⁴⁶. La primacía de la oración se opone a la comprensión común del lenguaje según la cual éste se compone de entidades significantes autónomas: las palabras. Se suele pensar que éstas son el átomo significativo puesto

41 *Ibid.*, 368.

42 *Ibid.*, 694.

43 *Ibid.*, 692.

44 *Ibid.*, 509.

45 *Ibid.*, 510.

46 *Ibid.*, 697.

que cada palabra refleja un sentido y una referencia que se utilizan a la hora de construir frases. Al ser así, existiría un límite a la libertad marcado por el sentido de las palabras que solo se pueden combinar según reglas preestablecidas. Esta concepción común del lenguaje se fundamenta en dos argumentos. Primero, los estudios lingüísticos tratan los fonemas, así como los morfemas, tal como si fuesen entidades autónomas, es decir, cosas: “Las investigaciones de los lingüistas pueden engañar acerca de esto: sus estadísticas sacan a la luz constantes deformaciones fonéticas o semánticas de un tipo dado; permiten reconstruir la evolución de un fonema o de un morfema en un periodo dado, de suerte que parece que la palabra o la regla sintética sea una realidad individual, con su significación y su historia”⁴⁷. Segundo, el poco impacto que tienen los seres humanos sobre la evolución del lenguaje lleva a pensar que sus átomos, las palabras, se comportan tal como si estuviesen fuera del alcance del ser humano: “[...] los individuos parecen tener muy poco influjo sobre la evolución de la lengua. Hechos sociales como las invasiones, las grandes vías de comunicación, las relaciones comerciales, parecen ser las causas esenciales de los cambios lingüísticos”⁴⁸. No obstante, para Sartre, esta concepción atomística del lenguaje se basa en una visión externa de éste, cuando un análisis más concreto de su funcionamiento nos lleva, con los psicólogos, a entender que: “[...] la palabra no es el elemento concreto del lenguaje [...]”⁴⁹.

La palabra es un elemento incompleto del lenguaje que adquiere su sentido dentro de la frase. De por sí, la palabra es una “función proposicional”⁵⁰ que Sartre ha definido como una fórmula del tipo “x es”⁵¹, es decir, una fórmula relacional en espera de relleno para comenzar a significar. Por este motivo: “La palabra no tiene, pues, sino una existencia puramente virtual fuera de las organizaciones complejas y activas en que se integra”⁵². Sólo la frase la da el relleno faltante a la palabra para que ésta pueda significar. Desde una posición estructuralista extranjera a Sartre, se podría pensar que el relleno se recibe desde la gramática, desde la posición de la palabra dentro de la frase. En este caso, la semántica no provendría de la esencia significante de la palabra, sino de la posición de las palabras dentro de la frase. Sin embargo, cuando Sartre afirma que la palabra es una función

47 *Ibid.*, 697.

48 *Ibid.*, 697.

49 *Ibid.*, 697.

50 *Ibid.*, 697.

51 “En cuanto a la Nada, tendría su origen en los juicios negativos; sería un concepto por el cual se establece la unidad trascendente de todos esos juicios, una función proposicional del tipo: “x no es”.”, *Ibid.*, 45. Las dos únicas ocurrencias de la expresión “función proposicional” en *El ser y la nada*, se encuentran en las páginas 45 y 697.

52 *Ibid.*, 697.

proposicional, significa que la palabra adquiere su significado en el acto de decir, mediante el sujeto que, al pronunciar la frase, resignifica cada una de las palabras, y no en su simple posición relacional. De ahí que el sentido de las palabras no se imponga al sujeto ya que, si bien pueden tener un sentido convencional, nadie, al hablar, las utiliza sin resignificarlas según su libre proyecto.

No obstante, la resignificación subjetiva de las palabras plantea el problema de la comunicación. Si cada sujeto resignifica las palabras según su propio proyecto que es, por definición, inaccesible al otro, la misma comunicación se hace imposible ya que el locutor y el receptor no estarían compartiendo el sentido de las palabras.

Es en este contexto que Sartre remite al pensamiento de Jean Paulhan que mostró que no solo las palabras no significan antes de la frase, sino que además pedazos enteros de frases, los “lugares comunes”, tampoco lo hacen: “Paulhan ha mostrado, en *Les fleurs de Tarbes*, que ciertas frases enteras, los “lugares comunes”, exactamente como las palabras, no preexisten al empleo que de ellas se hace”⁵³. Sartre conoce a Paulhan por el papel que ocupó en la publicación de sus primeros textos, entre los cuales *Melancholia*, que sería publicado, en 1938, como *La náusea*⁵⁴. En *Les fleurs de Tarbes*⁵⁵, Paulhan distingue dos puntos de vista a raíz de los cuales se analiza el significado de los lugares comunes. Por un lado, el receptor o lector, ignorante de la resignificación que opera el locutor, recibe el lugar común mediante su significación social, es decir, común: “Lugares comunes si encaradas desde afuera por el lector, que recompone el sentido del párrafo de una oración a otra [...]”⁵⁶. En cambio, los lugares comunes, desde el punto de vista del locutor: “[...] pierden su carácter trivial y convencional [...]”⁵⁷. Así, el sentido convencional o social de las palabras refleja el punto de vista del receptor que no dispone de otra manera de leer la frase puesto que no tiene acceso a la intencionalidad originaria del autor, cuando el autor, en su proyecto individual, las resignifica de manera subjetiva y singular: “[...] el lenguaje, libre proyecto para mí, tiene leyes específicas para el otro”⁵⁸. Esta oposición entre dos puntos de

53 *Ibid.*, 697.

54 En la carta a Simone de Beauvoir del 30 de abril 1937, Sartre relata sus encuentros con Paulhan y su apoyo en la publicación de sus textos en la NRF. Cf. SARTRE, Jean-Paul, *Lettres au castor et à quelques autres, Tome 1, 1926-1939*. Paris : Gallimard, 1983, pp. 113-117. En *La force de l'âge*, Beauvoir relató las dificultades de estos años para Sartre. Cf. BEAUVOIR, Simone (de), *Mémoires, I*. Paris : Gallimard, 2018.

55 PAULHAN, Jean, *Les fleurs de Tarbes ou La Terreur dans les Lettres*. Paris : Gallimard, 1941, 2006.7

56 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, pp. 697-698.

57 *Ibid.*, 698.

58 *Ibid.*, 701.

vista permite conservar un sentido convencional de las palabras pero también su resignificación subjetiva.

Sin duda, Paulhan separó estos dos puntos de vista, a tal punto que, para él, el lector ve el autor al revés⁵⁹. Retomando la filosofía de Bergson –filósofo del Terror⁶⁰, opone el pensamiento y el lenguaje, separándolos según los conceptos de fluidez, creatividad y plasticidad del lado del pensamiento, y con aquellos de lo rígido, predeterminado y común en lo que se refiere al lenguaje: “Nuestra vida interior, según Bergson, no alcanza su expresión sin dejar en camino lo que tiene de más valioso. El espíritu se ve, en cada momento, oprimido por el lenguaje”⁶¹. Por este motivo, el lenguaje obstaculiza la vida y el pensamiento, cristalizado y fijando lo que es en realidad un flujo: “[...] todo hombre debe finalmente romper, al querer acceder a su pensamiento auténtico, una costra de palabras que siempre amenaza con volver a formarse, de la cual los lugares comunes, clichés y convenciones son una de las formas posibles, la más evidente”⁶². Sartre tendría entonces razón en su lectura de Paulhan ya que se debería oponer el lenguaje tal como lo entiende el lector o el crítico que piensan que el lenguaje se impone al autor⁶³ y lo influencia, y el lenguaje que utiliza el autor, creador subjetivo hasta de los lugares comunes: “Por muy banal que sea un lugar común siempre puede haber sido inventado por aquel que lo pronuncia: en este caso, se acompaña, incluso, de un intenso sentimiento de novedad”⁶⁴.

Esta posibilidad de creación de los lugares comunes (y de toda palabra) a la hora de pronunciarlos se evidencia, para Paulhan y contra Bergson, en las declaraciones de amor que, a pesar de darse mediante el lenguaje, el vocabulario y los lugares comunes más banales, siempre significan un amor singular y único⁶⁵. Decirle “te amo” a alguien siempre es decirle algo muy banal y no obstante singular: “Esto es lo que las cartas de amor evidencian. Son infinitamente ricas y conllevan un sentido excepcional para quien las recibe o las escribe, pero permanecen

59 El título del capítulo 5 del libro de Paulhan es « Où le lecteur voit l’auteur à l’envers. <Dónde el lector ve el autor al revés> », PAULHAN, Jean, *Les fleurs de Tarbes ou La Terreur dans les Lettres*. Paris : Gallimard, 1941, 2006, p. 81.

60 « La terreur trouve son philosophe. <El terror encuentra su filósofo> », *Ibid.*, 70.

61 « Notre vie intérieure, si l’on en croit Bergson, ne parvient pas à l’expression sans laisser en route le plus précieux d’elle-même. L’esprit se trouve, à chaque moment, opprimé par le langage. », *Ibid.*, 70.

62 « [...] tout homme enfin doit briser, lorsqu’il veut atteindre à sa pensée authentique, une croûte de mots, trop prompte à se reformer, dont les lieux communs, clichés, conventions, ne sont qu’une forme, la plus évidente. », *Ibid.*, 71.

63 *Ibid.*, 88.

64 « Pour banal que soit un lieu commun, il peut toujours avoir été inventé par qui le prononce : il s’accompagne même, en ce cas, d’un vif sentiment de nouveauté. », *Ibid.*, 95.

65 Cf. VINOLO, Stéphane, *Jean-Luc Marion, Apologie de l’inexistence, Tome II, Une phénoménologie discursive*. Paris : L’Harmattan, 2019, pp. 242-254.

enigmáticas para un extranjero, banales y puros verbalismos”⁶⁶. Por todas partes, encontramos entonces en Paulhan la idea según la cual el lenguaje, incluso los lugares comunes, no es una interpretación personal que se da mediante la articulación individual de sentidos preexistentes, sino una verdadera creación de sentido, lo que justificaría la tesis de Sartre.

No obstante, Moati⁶⁷ mostró que Sartre invierte en parte la tesis de Paulhan. Efectivamente, una vez separados el punto de vista del locutor que percibe el lenguaje a raíz de la creación de sentido, y el punto de vista del receptor que entiende el mensaje según el sentido colectivo y despersonalizado de las palabras, Sartre posibilita la comunicación mediante la co-presencia de los interlocutores dentro del proyecto del emisor. Es posible entendernos porque el receptor puede entrar en las significaciones individuales que el locutor crea al hablar: “Comprender una oración de mi interlocutor es, en efecto, comprender lo que éste quiere decir, o sea adherirme a su movimiento de trascendencia, arrojarme con él hacia posibles, hacia fines, y volver en seguida sobre el conjunto de los medios organizados, para comprenderlos por su función y su objetivo”⁶⁸. Al entrar en el “querer decir”, el receptor puede penetrar la intencionalidad del locutor y, por lo tanto, su mismo proyecto. En Sartre, el encuentro entre locutor y receptor se da entonces sobre el campo del locutor, en la intimidad de lo que resignifica, de lo que quiere decir. De ahí que comprender el discurso de alguien siempre sea “[...] hablar con”⁶⁹ él.

Al contrario, en Paulhan, el encuentro que permite la comprensión se da en el otro polo de la comunicación. A pesar de la fractura entre significaciones colectivas y significaciones individuales, es posible comunicar porque la diferencia entre locutor y receptor es muy frágil, ya que el diálogo entre locutor y receptor se da en la misma intimidad del locutor, en cualquier proceso de reflexión. Cualquier locutor es también lector o receptor de sus propias palabras por lo que, cada uno a su nivel, tanto locutor como receptor son, en última instancia, lectores del discurso del otro o de sus propios discursos: “No existe distancia absoluta entre la conversación común y la conversación secreta que cada uno de nosotros mantiene consigo mismo. Cada autor es también su propio lector, cada hablante su hablado- [...]”⁷⁰. Así, Paulhan posibilita la comunicación haciendo del locutor un

66 « C’est de quoi les lettres d’amour sont l’exemple : infiniment riches et d’un sens exceptionnel pour qui les écrit ou les reçoit, – mais énigmatiques pour un étranger, à force de banalité, et (dira-t-il) de verbalisme. », PAULHAN, Jean, *Les fleurs de Tarbes ou La Terreur dans les Lettres*. Paris : Gallimard, 1941, 2006, p. 100.

67 MOATI, Raoul, *Sartre et le mystère en pleine lumière*. Paris : Cerf, 2019, pp. 86-87.

68 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 698.

69 *Ibid.*, 698.

70 « Il n’est pas d’écart absolu entre l’entretien commun et cet entretien secret, que chacun de nous poursuit avec soi. Chaque auteur est aussi son propre lecteur, chaque parlant son parlé – [...] ». »

receptor cuando Sartre permite que el receptor entre dentro de las significaciones intencionales del locutor.

La filosofía del lenguaje de Sartre es entonces en parte una aplicación de la teoría de Paulhan, con una inversión final en cuanto a sus consecuencias. Cuando Paulhan hace de todo locutor un lector (del otro, así como de sí mismo), Sartre abre la posibilidad que cualquier lector pueda entrar dentro de las significaciones individuales que surgen del proyecto del locutor para resignificar el lenguaje. De esta manera, Sartre supera las significaciones preexistentes que podrían obstaculizar la libertad del sujeto por la afirmación que todo sentido es resignificado por el mismo sujeto y que toda significación, dentro de los tres campos que hemos señalado, opera como una función proposicional en espera de un relleno significativo que solo le puede llegar del libre proyecto de un sujeto.

3. LA VIDA DEL LENGUAJE: SARTRE LECTOR DE PARAIN

La relación entre Sartre y Parain es más compleja dado que evolucionó entre la publicación de *El ser y la nada*, la *Crítica de la razón dialéctica*⁷¹ y, finalmente, *El idiota de la familia*⁷². Dado el límite de este trabajo a la primera filosofía de Sartre, se analizará exclusivamente el primer momento de esta relación⁷³. Tal como para Paulhan, Sartre cita de manera explícita el texto de Brice Parain en *El ser y la nada*. Sin embargo, cuando el nombre de Paulhan aparece dos veces, el de Parain aparece una única vez, en una nota al pie de página⁷⁴.

Existen dos textos de Parain que Sartre hubiera podido utilizar y criticar. El texto citado en *El ser y la nada* es la tesis complementaria de Parain titulada *Essai sur le logos platonicien*⁷⁵, publicada en 1942. El segundo, su tesis doctoral titulada *Recherches sur la nature et la fonction du langage*⁷⁶ y publicada también en el año 1942 presenta una dificultad más compleja. Sartre nunca lo cita en *El ser*

PAULHAN, Jean, *Les fleurs de Tarbes ou La Terreur dans les Lettres*. Paris : Gallimard, 1941, 2006, p. 135.

71 SARTRE, Jean-Paul, *Critique de la raison dialectique*. Paris : Gallimard, *Tome I* (1960), *Tome II* (1985).

72 SARTRE, Jean-Paul, *L'Idiot de la famille*. Paris: Gallimard, *Tome I* (1971), *Tome II* (1971), *Tome III* (1972).

73 El segundo momento de esta relación ha sido analizado por DESCOMBES, Vincent, *Les institutions du sens*. Paris: Éditions de minuit, 1996, pp. 267-291.

74 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 699, nota 1.

75 PARAIN, Brice, *Essai sur le logos platonicien*. Paris : Gallimard, 1942, 1969.

76 PARAIN, Brice, *Recherches sur la nature et la fonction du langage*. Paris : Gallimard, 1942,

y *la nada*; no obstante, en febrero de 1944, publicó un artículo completo acerca de este libro⁷⁷. Se sabe, además, por una carta dirigida a Beauvoir⁷⁸, que su redacción fue finalizada a comienzos de 1944 por lo que, aunque sea muy probable, es imposible afirmar con certeza que Sartre había leído las *Recherches sur la nature et la fonction du langage* antes de la publicación de *El ser y la nada*. Por esta razón, aunque nuestra argumentación se pueda reforzar por la lectura de este último texto de Parain que prefigura las tesis del *Essai sur le logos platonicien*, y que muchos temas o ejemplos que encontramos en *El ser y la nada* parecen responder directamente a ideas desarrolladas en éste, el despliegue de nuestros argumentos se concentrará de manera prioritaria sobre el texto de Parain que Sartre cita.

La crítica de Sartre a Parain aparece en una sola frase: “Recientemente ha podido sostenerse que hay un como orden vivo de las palabras, una vida impersonal del logos; en suma, que el lenguaje es una Naturaleza y que el hombre debe servirla para poder utilizarla sobre algunos puntos, como lo hace con la Naturaleza”⁷⁹. Para Sartre, Parain otorgó una autonomía al lenguaje, haciendo de éste una técnica que el ser humano debe aprender a domar para poder expresarse. De la misma manera que el marinero debe aprender a utilizar el viento para poder navegar de bolina, el locutor debe domar las reglas autónomas del lenguaje para poder expresarse. Encontramos entonces una crítica que hemos visto en el debate con Paulhan, salvo que, cuando el debate con Paulhan se focalizaba sobre la semántica, el diálogo con Parain se concentra en la gramática y en el lenguaje en tanto que técnica, razón por la cual es necesario precisar la tesis de Parain antes de entender en qué medida Sartre se opone a ésta.

El mismo título del libro de Parain evidencia el problema con el cual se va a enfrentar: las posibles traducciones de la palabra Logos. Ésta viajó fuera del mundo heleno, por ejemplo, en el mundo bíblico⁸⁰, y estos desplazamientos exponen al riesgo de perder el sentido originario de las palabras. De hecho, los modernos traducen habitualmente logos por diversas palabras que provienen de dos campos

77 Este texto fue publicado en el número 264 de la revista *Cahiers du Sud*, y fue retomado en el volumen I de las *Situaciones*. Cf. SARTRE, Jean-Paul, “*Ida y vuelta*”. En: SARTRE, Jean-Paul, *El hombre y las cosas*. Buenos Aires: Losada, 1960, pp. 146-188.

78 “Cuestión trabajo, he terminado el artículo Parain y volví gozosamente a la novela.”, SARTRE, Jean-Paul, “*Carta a Simone de Beauvoir, comienzos del año 1944*”. En: SARTRE, Jean-Paul, *Cartas al castor y a algunos otros, II, 1940-1963*. Buenos Aires: Editorial Sudamericana, 1987, p. 216.

79 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 699.

80 « On connaît le rôle que le logos a joué dans la pensée grecque depuis Héraclite jusqu'à la théologie chrétienne, en passant par les Stoïciens. <Sabemos el papel que ocupó el logos en el pensamiento griego desde Heráclito hasta la teología cristiana, pasando por los estoicos> », PARAIN, Brice, *Essai sur le logos platonicien*. Paris : Gallimard, 1942, 1969, pp. 9-10.

semánticos diferentes: la razón y el lenguaje⁸¹, lo que no deja de sorprender pues-
to que se suele distinguir, señala Parain, el pensamiento del lenguaje, dado que el
lenguaje se percibe como un simple instrumento de expresión del pensamiento⁸².
La ambivalencia del término *logos* en el mundo moderno es aún más sorprenden-
te cuando el mismo Platón lo había precisado de manera clara y distinta: “[...] la
combinación de los nombres es la esencia de una explicación (*logou ousian*)”⁸³.
La posición de Platón en referencia al *logos* es clara: la esencia del *logos* es una
simple combinación de nombres, afirmación repetida en el Sofista, cuando el Ex-
tranjero propone una nueva definición del discurso (*logos*) en tanto que relación
entre los elementos de una frase⁸⁴. Así, en Platón, el *logos* se encuentra claramen-
te del lado del lenguaje lo que complica la persistencia de su uso moderno para
designar tanto el pensamiento como el lenguaje.

Se podría pensar, siguiendo los análisis de Léon Robin, que *logos* significa
a la vez el pensamiento y su expresión; no obstante, sería entonces inentendible
la permanencia, en la filosofía griega, de la diferencia entre *logos* y *nous*⁸⁵. Para
resolver este problema, Parain propone una nueva hipótesis: es necesario invertir
la teoría de Robin y afirmar que *logos* es un lenguaje que contiene cierto tipo de
pensamiento⁸⁶. De esta hipótesis surge la interpretación de Sartre según la cual,
para Parain, el lenguaje tiene vida propia. Contra Robin, Parain afirma que, si el
pensamiento es a la vez su propio verbo, su propia expresión, no hay razón por la
cual se hubiera conservado el término *logos* para señalar el pensamiento, ya que

81 « [...] nous avons pris l'habitude de l'interpréter doublement, tantôt par langage et tantôt par
raison. <es costumbre traducirlo de dos maneras: a veces como lenguaje, a veces como razón> », *Ibid.*, 10.

82 « Pour nous le langage est l'instrument docile de la pensée (ou de la raison), et c'est la
pensée (ou la raison) qui nous fournit la connaissance que nous avons des choses. <Para nosotros,
el lenguaje es el instrumento dócil del pensamiento (o de la razón) y es éste (o la razón) que nos
provee de los conocimientos de las cosas> », *Ibid.*, 11-12.

83 PLATÓN, *Teeteto*, 202b. En: PLATÓN, *Diálogos*, V. Madrid: Gredos, 1988, 1992, p.
299.

84 “Extranjero- No obstante, los nombres tomados por sí solos, y mencionados en forma
continuada, no constituyen discurso alguno, ni tampoco los verbos mencionados separadamente de
los nombres.”, Platón, *Sofista*, 262a, *Ibid.*, 465.

85 « Peut-être la considéraient-ils, en effet, selon l'exégèse de M. Léon Robin, comme une
pensée qui serait à la fois la pensée et son verbe. Mais pourquoi lui auraient-ils choisi, dans ce cas, le
nom de *logos* plutôt que celui de *nous* ou de *phronesis*, en désignant ainsi l'exprimé par l'expression,
la réalité par son apparence ? <Tal vez lo consideraban, según la exegesis de M. Léon Robin, como
un pensamiento que fuese a la vez el pensamiento y su verbo. Pero ¿por qué hubiesen escogido,
en este caso, el nombre de *logos* más que el de *nous* o de *phronesis*, designando así lo expresado
por la expresión, y la realidad por su apariencia?> », PARAIN, Brice, *Essai sur le logos platonicien*.
Paris : Gallimard, 1942, 1969, p. 12.

86 « Le *logos* ne serait-il pas plutôt un verbe qui par sa nature contiendrait de la pensée ? <¿No
sería más bien el *logos* un verbo que, por su naturaleza, contendría cierto pensamiento?> », *Ibid.*,
12-13.

logos sería únicamente la expresión del pensamiento, y no el mismo pensamiento. Ahora bien, dado que los griegos mantuvieron la palabra logos como acepción de cierto pensamiento a pesar de haber distinguido claramente lo que los modernos llaman pensamiento y lenguaje, el mundo griego remite a una concepción del pensamiento diferente a la de los modernos (a las operaciones mentales). En esta otra concepción del pensamiento, éste no contiene su verbo sino más bien el lenguaje contiene cierto tipo de pensamiento. Parain invierte entonces la tesis de Robin al postular que el pensamiento contiene cierto tipo de lenguaje, razón por la cual Sartre lo acusó de haber pensado un lenguaje que dispone de una “vida impersonal”⁸⁷.

Esta tesis de Parain ya se encontraba en su tesis doctoral. Oponiéndose a la teoría del expresionismo⁸⁸ según la cual el lenguaje no es más que un instrumento neutro que sirve para expresar un pensamiento que ha sido formado fuera de éste, Parain insiste sobre el lenguaje en tanto que determina al locutor, por la precedencia de sus reglas. Es notable, desde un punto de vista sartreano, que todo el libro de Parain está habitado por el problema de la libertad que constituye su misma razón de ser⁸⁹, y que su conclusión sea, al contrario de la de Sartre, que el lenguaje es un obstáculo radical a la libertad: “El lenguaje, al actuar en contra del orden de la libertad pura es un personaje molesto en nuestra relación con el mundo”⁹⁰. Para Parain, el lenguaje no es el vector neutro del pensamiento. Es una técnica que, tal como cualquier técnica, impone al ser humano ciertas condiciones dado que no se puede hacer cualquier cosa con cualquier herramienta. Al utilizar un cuchillo, uno puede cortar un objeto o, por torpeza, cortarse un dedo. Su habilidad determina el cortar bien o el cortar mal, pero el cortar como tal proviene del cuchillo: “El uso es mío, pero la herramienta proviene de fuera de mí. [...] la herramienta dispone de sus propias leyes que no son las mías”⁹¹. De la misma manera, un hacha impone ciertas reglas que no son las del sujeto. Uno puede hacer muchas cosas con un hacha, pero un hacha no puede ser un mazo⁹². Aplicando

87 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 699.

88 « N'hésitons pas à nous placer dans l'hypothèse expressionniste : [...]. Mais nous serons obligés de l'abandonner bientôt, et d'autant plus vite que nous l'aurons adoptée plus pleinement. <No dudemos en situarnos dentro de la hipótesis expresionista: [...]. Pero estaremos obligados de abandonarla muy pronto, aún más rápido cuando la hayamos adoptado plenamente.> », PARAIN, Brice, *Recherches sur la nature et la fonction du langage*. Paris : Gallimard, 1942, p. 224.

89 « Ce livre est né, comme il s'achève, dans une inquiétude au sujet de la liberté. <Este libro nació tal como se acaba, sobre una inquietud acerca de la libertad.> », *Ibid.*, 235.

90 « Le langage, agissant contre l'ordre de la liberté pure, est un personnage gênant dans notre commerce avec le monde. », *Ibid.*, 238.

91 « L'usage est de mon fait, mais l'outil me vient d'ailleurs. [...] l'outil a ses lois, qui ne sont pas les miennes. », *Ibid.*, 221.

92 *Ibid.*, 228.

esta lógica al lenguaje, para Parain, uno no decide del sentido de las palabras, ni de las reglas según las cuales se pueden articular. Para expresarse, el ser hablante escoja la palabra más apropiada⁹³ a lo que quiere significar, la que más refleja el sentido de lo que quiere decir. Además, al hablar, uno reconoce las reglas de la gramática al someterme a éstas con el objetivo de formar frases: “Hablar es una aceptación, como mínimo tácita, o un reconocimiento, como mínimo tácito, del orden dentro del cual uno penetra al hablar”⁹⁴.

Por este motivo, el lenguaje, al ser una norma semántica y sintáctica externa al locutor, siempre pierde lo singular, lo individual y lo íntimo de éste. Toda persona que habla diluye el carácter único de lo que quiere decir dentro de la impersonalidad o de la universalidad del lenguaje: “Se espera de él [del lenguaje] que exprese lo que el hombre tiene de más íntimamente individual. No es adecuado para esto. Su destino es formular lo que el hombre tiene de más íntimamente impersonal, de más íntimamente idéntico con los otros”⁹⁵. Aparece entonces que desde las *Recherches sur la nature et la fonction du langage*, Parain defendió la tesis de la autonomía de lenguaje, tanto en lo que se refiere al sentido de las palabras como a la regla de su articulación dentro de una frase. Encontramos, en el lenguaje, significaciones y estructuras preexistentes al sujeto que lo determinan desde el exterior, ya que el lenguaje es necesariamente un hecho social⁹⁶.

La respuesta de Sartre a la tesis de Parain es que éste “[...] ha considerado al lenguaje una vez muerto, o sea, una vez que ha sido hablado, [...]”⁹⁷. La vida que Parain ve en el lenguaje es precisamente lo que, para Sartre, le es dado por el para-sí, en el mismo acto de habla. Parain olvidó que el para-sí da vida al lenguaje a la hora de hablar, una vida y unas fuerzas que: “[...] han sido tomadas en préstamo a la libertad personal del para-sí hablante”⁹⁸. Por esta razón, habla de un lenguaje ideal que proviene de la abstracción del lenguaje hablado, retirándole el

93 « Je l’emploie [le mot « faim »] parce que je le trouve approprié à mon état, parce que je le lui attribue. <La utilizo [la palabra “hambre”] porque la pienso adecuada a mi estado, porque se la atribuyo.> », *Ibid.*, 223.

94 « Parler est une acceptation, au moins tacite, une reconnaissance, au moins tacite, de l’ordre dans lequel nous pénétrons en parlant. », *Ibid.*, 225.

95 « On lui demande d’exprimer ce que l’homme a de plus intimement individuel. Il n’y est pas propre. Sa destination est de formuler ce que l’homme a de plus intimement impersonnel, de plus intimement pareil aux autres. », *Ibid.*, 227.

96 « L’étude des langues n’enseigne pas autre chose. Elle enseigne, en effet, que le langage et un fait social, qui dépasse l’individu. <El estudio de las lenguas no enseña otra cosa. Enseña, en efecto, que el lenguaje es un hecho social que sobrepasa el individuo.> », *Ibid.*, 229.

97 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 699.

98 *Ibid.*, 699.

papel fundamental del locutor que lo habita y que le inyecta la vida: “Se ha hecho del lenguaje una lengua que se habla sola”⁹⁹.

Este error de Parain impacta sobre la concepción de la técnica en Sartre, puesto que el lenguaje, en tanto que lengua muerta, se puede asemejar a una técnica. Frente a cualquier técnica, afirmar la lógica de Parain, el hombre se encuentra en la posición de un piloto de barco que debe adaptarse a la naturaleza y domarla para poder manejarlo¹⁰⁰. Para Sartre, esta concepción de la técnica es ontológicamente errónea ya que cristaliza la finalidad de la herramienta dentro del en-sí, cuando las finalidades o intenciones solo provienen, tal como lo mostró Sartre con sus análisis ontológicos, del para-sí. La finalidad de la técnica solo le puede llegar del exterior. Sartre toma el ejemplo de un hacha y de un martillo, exactamente tal como lo hacía Parain¹⁰¹, pero, a diferencia de éste, afirma que la actividad determina el objeto: “[...] el hachazo revela el hacha, el martillar al martillo”¹⁰². He aquí la diferencia fundamental que establece Sartre entre un acontecimiento natural y una técnica artificial como el lenguaje. El acontecimiento natural evidencia una ley externa que lo rige desde el exterior y que no hace más que ejemplificar, cuando las técnicas artificiales se encarnan y evidencian leyes internas que las rigen¹⁰³ a partir de su uso. Parain confundió la técnica artificial que es el lenguaje con un acontecimiento natural. Para que una frase sea una frase, es necesario que el carácter sintético que le da sentido unificando a sus palabras sea interno a éstas: “[...] para que las palabras mantengan relaciones entre sí, para que se junten –o se rechacen– mutuamente, es menester que estén unidas en una síntesis que no proviene de ellas; suprimamos esta unidad sintética, y el bloque “lenguaje” se desmigaja: [...]”¹⁰⁴. Tal como el hacha se hace hacha con la acción del ser humano que la utiliza, las reglas del lenguaje se actualizan en el acto de habla. Esto no significa que el lenguaje no disponga de ciertas reglas, de

99 *Ibid.*, 699.

100 “Si se hace surgir al hombre en medio de técnicas que se aplican solas, en medio de una lengua que se habla sola, de una ciencia que se hace por sí misma, de una ciudad que se construye de por sí según sus leyes propias; si se fijan las significaciones en en-sí conservándoles a la vez una trascendencia humana, entonces se reducirá el papel del hombre al de un piloto que utiliza las fuerzas determinadas de los vientos, las olas y las mareas para dirigir un navío.”, *Ibid.*, 700.

101 Para asimilar el lenguaje a una técnica, Parain tomaba el ejemplo de un hacha y de un mazo. Estos dos ejemplos son paradigmáticos dentro de la historia de la filosofía. No obstante, el hecho que Sartre retome los dos ejemplos de Parain puede dejar pensar que su texto es una respuesta explícita al texto de Parain que, por lo tanto, habría leído antes de la publicación de *El ser y la nada*.

102 *Ibid.*, 699.

103 “El hecho natural se produce conforme a una ley que él manifiesta, pero que es pura regla exterior de producción, de la cual el hecho considerado no es sino un ejemplo. La “oración” como acaecimiento contiene en sí misma la ley de su organización, y sólo en el interior del libre proyecto de *designar* pueden surgir relaciones legales entre las palabras.”, *Ibid.*, 701.

104 *Ibid.*, 700.

la misma manera que el hecho de esquiar –para retomar el ejemplo de Sartre– obedece a las leyes de la física, y que se podría describir el hecho de esquiar de manera neutra; no obstante, tal como nadie puede simplemente esquiar sino que al esquiar encarna cierta manera singular y única de hacerlo –en este sentido que nadie esquiaba según la definición neutra e idealizada del hecho de esquiar¹⁰⁵, nadie puede simplemente hablar, utilizando de manera neutra el lenguaje, hablando un lenguaje anónimo, sin sobrecargarlo de significaciones y emociones individuales; por esta razón: “[...] hablando, hago la gramática; la libertad es el único fundamento posible de las leyes del idioma”¹⁰⁶.

CONCLUSIÓN

La filosofía del lenguaje de Sartre tal como aparece en *El ser y la nada* responde ante todo al problema de la libertad del sujeto y a sus posibles límites. Frente a los tres campos fenoménicos que amenazan con presentar significaciones que se imponen al sujeto (significaciones colectivas, ser-para-otro y fenómeno del otro), Sartre responde mediante una confrontación con los pensamientos de Paulhan y de Parain. A pesar de invertir el papel del locutor y del receptor en el pensamiento de Paulhan, Sartre retoma de éste la idea de un lenguaje que se actualiza en el acto de habla, lo que implica que toda interpretación semántica o sintética sea, en última instancia, una libre creación del sujeto: “En tal sentido, la oración aparece como libre invención de sus leyes”¹⁰⁷.

La oposición con Parain es más fructífera ya que éste presenta de manera detallada y fundamentada la tesis que Sartre pretende combatir. En sus dos libros de 1942, afirma la existencia de una autonomía del lenguaje que limita la libertad del sujeto, imponiéndole un marco de significaciones y de articulaciones al cual

105 “Será dado descubrir, en una carrera particular, el método francés de esquiar, y, en este método, el arte general del esquí como posibilidad humana. Pero este arte humano jamás es nada por sí solo; no existe *en potencia*, sino que se encarna y manifiesta en el arte *actual* y concreto del esquiador.”, *Ibid.*, 703.

106 *Ibid.*, 700.

107 *Ibid.*, 702.

debe adaptarse¹⁰⁸. El lenguaje (semántica y sintáctica) sería entonces trascendente¹⁰⁹ al sujeto: “Toda palabra existe de manera independiente al contenido de imágenes y de acciones que le atribuimos”¹¹⁰. La dificultad, para Sartre, es que Parain mantiene una concepción metafísica de la trascendencia del lenguaje, al mantenerlo dentro de cierta sustancialidad estable que existe por fuera de su uso subjetivo. Sartre acepta que la significación y el sentido estén vinculados con la trascendencia, pero con la acepción fenomenológica de ésta en tanto que es la misma estructura de la consciencia, la estructura correlacional de significación: “La consciencia es consciencia de algo: esto significa que la trascendencia es estructura constitutiva de la consciencia; es decir, que la consciencia nace conducida sobre un ser que no es ella misma”¹¹¹. Por esta razón, todo lo que se da a la consciencia se da mediante una significación del sujeto. Por mucho que una significación sea colectiva o que preceda la existencia del sujeto, ésta no puede tener impacto sobre él si no es resignificada por la misma consciencia del sujeto. Así, el lenguaje, tal como toda significación, no obstaculiza la libertad del sujeto ya que es una libre creación de éste.

Es posible ahora responder al problema de los tres campos de significaciones que preceden el arrojamiento del sujeto al mundo. Las significaciones colectivas solo significan a medida que el sujeto las resignifique. Es el parar al semáforo rojo que hace de éste una orden. Segundo, el ser-para-otro sólo puede manifestarse de manera resignificada: “Por este libre reconocimiento del prójimo a través de la experiencia de mi alienación, asumo mi ser-para-otro, cualquiera que fuere, y lo asumo precisamente porque es mi nexo concreto con el prójimo”¹¹². Tercero, la asunción del ser-para-otro solo es posible después de que el sujeto haya

108 « Le langage n'est pas un vide devant nous, dans lequel nous serions attirés comme le vent des zones hautes pressions dans les zones basses pressions. Il est la règle de notre pensée et de notre action humaine, extérieure et par conséquent transcendante à nous, parce qu'il est le lieu de l'universel et de la volonté réfléchie. <El lenguaje no es un vacío frente a nosotros, por el cual estuviéramos atraídos tal como el viento de las zonas de altas presiones lo es por las zonas de bajas presiones. Es la regla de nuestro pensamiento y de nuestra acción humana, exterior al sujeto y por consiguiente trascendente, porque es el lugar de lo universal y de la voluntad reflexionada.> », PARAIN, Brice, *Recherches sur la nature et la fonction du langage*. Paris : Gallimard, 1942, p. 229.

109 « Mon dessein, qui s'achève ici, n'a été que de montrer en quoi l'idée que nous pouvons nous faire de la nature du langage d'après ses fonctions nous conduit vers une métaphysique de la transcendance plutôt que vers une métaphysique de l'immanence. <Mi objetivo, que aquí termina, fue el mostrar en qué medida la idea que podemos tener acerca de la naturaleza del lenguaje desde sus funciones nos conduce hacia una metafísica de la trascendencia más que de la inmanencia.> » *Ibid.*, 230.

110 « Tout mot existe indépendamment du contenu d'images et d'actions que nous lui attribuons. », *Ibid.*, 227.

111 SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 31.

112 *Ibid.*, 712.

reconocido, en primera instancia, al otro en tanto que otro, en tanto que transcendencia que lo objetiva: “[...] no puedo captar al prójimo como libertad sino en el libre proyecto de captarlo como tal (en efecto, siempre queda la posibilidad de que capte libremente al otro como objeto), y el libre proyecto de reconocimiento del prójimo no se distingue de la libre asunción de mi ser-para-otro”¹¹³.

Las páginas de *El ser y la nada* en las cuales Sartre desarrolla su filosofía del lenguaje son entonces fundamentales para entender su fenomenología. Insistiendo sobre el papel creador del sujeto en tanto que resignifica todo lo que le aparece, el tema del lenguaje vincula la filosofía sartreana con la problemática de Dios, central en su obra¹¹⁴. En su artículo sobre Descartes¹¹⁵, Sartre lamentaba que Descartes haya puesto en Dios, mediante la teoría de la creación de las verdades eternas¹¹⁶, la libertad creadora que le pertenece al ser humano. Con la primera filosofía del lenguaje de Sartre que impacta sobre toda su concepción de la significación y del sentido, y, por lo tanto, sobre su lugar en la historia de la fenomenología, encontramos nuevamente esta posición divina del sujeto humano. Éste no solo es creador del mundo y de su ser mediante el uso de significaciones y del lenguaje, sino además es creador del mismo lenguaje tanto en su carácter semántico que en su carácter sintáctico, tal como si fuese Dios¹¹⁷.

REFERENCIAS BIBLIOGRÁFICAS

- BEAUVOIR, Simone (de), *Mémoires, I*. Paris : Gallimard, 2018.
 BELAVAL, Yvon, *Les philosophes et leur langage*. Paris : Gallimard, 1952.
 CAMUS, Renaud, *Syntaxe ou l'autre dans la langue*. Paris : POL, 2004.
 DESCARTES, René, *Tres cartas a Marin Mersenne (Primavera de 1630)*. Madrid: Encuentro, 2011.
 DESCOMBES, Vincent, *Les institutions du sens*. Paris: Éditions de minuit, 1996.
 HUSSERL, Edmund, *La idea de fenomenología, cuarta lección*. México: FCE, 1950, 1982.
 JEANSON, Francis, *Sartre devant Dieu*. Paris : Éditions Cécile Default, 2005.

113 *Ibid.*, 712.

114 JEANSON, Francis, *Sartre devant Dieu*. Paris : Éditions Cécile Default, 2005.

115 SARTRE, Jean-Paul, “La libertad cartesiana”. En: SARTRE, Jean-Paul, *El hombre y las cosas*. Buenos Aires: Losada, 1960, pp. 242-258.

116 DESCARTES, René, *Tres cartas a Marin Mersenne (Primavera de 1630)*. MADRID: Encuentro, 2011.

117 “[...] la responsabilidad del para-sí es agobiadora, pues es aquel por quien se hace que haya un mundo; y, [...] es también aquel que se hace ser, [...]”, SARTRE, Jean-Paul, *El ser y la nada*. Buenos Aires: Losada, 2013, p. 747.

- MOATI, Raoul, *Sartre et le mystère en pleine lumière*. Paris : Cerf, 2019.
- PARAIN, Brice, *Essai sur le logos platonicien*. Paris : Gallimard, 1942, 1969.
- , *Recherches sur la nature et la fonction du langage*. Paris : Gallimard, 1942.
- PAULHAN, Jean, *Les fleurs de Tarbes ou La Terreur dans les Lettres*. Paris : Gallimard, 1941, 2006.
- PLATÓN, *Diálogos*, V. Madrid: Gredos, 1988, 1992.
- RAMOND, Charles (ed.), *René Girard, La théorie mimétique, de l'apprentissage à l'apocalypse*. Paris : PUF, 2010.
- RICCEUR, Paul, *Ser, esencia y sustancia en Platón y Aristóteles*. México: Siglo veintiuno editores, 2013.
- SARTRE, Jean-Paul, *El hombre y las cosas*. Buenos Aires: Losada, 1960.
- , *Critique de la raison dialectique*. Paris : Gallimard, Tome I (1960), Tome II (1985).
- , *Bosquejo de una teoría de las emociones*. Madrid: Alianza, 1971, 2015.
- , *L'Idiot de la famille*. Paris: Gallimard, Tome I (1971), Tome II (1971), Tome III (1972).
- , *Lettres au castor et à quelques autres, Tome 1, 1926-1939*. Paris : Gallimard, 1983.
- , *A Puerta cerrada, Escena V, en La puta respetuosa / A puerta cerrada*. Bogotá: Orbis, 1983.
- , *Cartas al castor y a algunos otros, II, 1940-1963*. Buenos Aires: Editorial Sudamericana, 1987.
- , *El ser y la nada*. Buenos Aires: Losada, 2013.
- VERNANT, Jean-Pierre, *Œuvres. Religions, Rationalités, Politique*. Paris : Seuil, Paris, 2007.
- VINOLO, Stéphane, "El sujeto amoroso en Sartre, un puente entre la metafísica y la postmetafísica". *Cuadernos salmantinos de filosofía*, Vol. 45, 2018, pp. 323-341.
- , Jean-Luc Marion, *Apologie de l'inexistence, Tome II, Une phénoménologie discursive*. Paris : L'Harmattan, 2019.
- , "El neorrealismo absoluto en *El ser y la nada*, de Jean-Paul Sartre". *Eidos*, 2021, pp.194-223.

EL DIÁLOGO COMO PROXIMIDAD EN EL PENSAMIENTO DE EMMANUEL LEVINAS

*DIALOGUE AS PROXIMITY IN THE THOUGHT
OF EMMANUEL LEVINAS*

FRANCISCO JAVIER RODRÍGUEZ

Licenciado en filosofía
Becario investigador del Instituto de filosofía
Universidad Católica de Santa Fe (Argentina)
Santa Fe/Argentina
franciscorodriguez@ucsf.edu.ar
ORCID: 0000-0002-0733-3153

Recibido: 19/12/2020

Revisado: 11/04/2021

Aceptado: 6/09/2021

Resumen: Emanuel Levinas ha desarrollado su ética de la alteridad a partir de una visión de la subjetividad como lenguaje y responsabilidad. Este artículo indaga en la concepción levinasiana del diálogo como proximidad y en su noción dialógica de la subjetividad. Al diálogo que busca coincidir en la razón, dando cuenta de toda diversidad y alteridad, antepone el diálogo como proximidad y diaconía. La subjetividad lingüística es enunciada por Levinas como un diálogo original previo a la abstracción de la razón y al lenguaje como trasmisor de contenidos, un diálogo que es proximidad y contacto con la altura del Rostro. El llamado sensible del Otro y la respuesta responsable se articulan en un diálogo original que mantiene la trascendencia de la relación, un diálogo que es praxis ética de fraternidad y hospitalidad. Este artículo es el resultado de una investigación bibliográfica.

Palabras clave: corporeidad, diaconía, diálogo, espacio ético, fraternidad, Levinas, proximidad.

Abstract: Emanuel Levinas has developed his ethics of otherness from a vision of subjectivity as language and responsibility. This article investigates the Levinasian conception of dialogue as proximity and his dialogic notion of subjectivity. To the dialogue that seeks to coincide in reason, accounting for all diversity and alterity, it puts dialogue first as proximity and diakonia. Linguistic subjectivity is enunciated by Levinas as an original dialogue prior to the abstraction of reason and language as a transmitter of content, a dialogue that is proximity and contact with the height of the Face. The sensitive call of the Other and the responsible response are articulated in an original dialogue that maintains the transcendence

of the relationship, a dialogue that is the ethical praxis of fraternity and hospitality. This article is the result of a bibliographic research.

Keywords: corporeity, diakonia, dialogue, ethical space, fraternity, Levinas, proximity.

INTRODUCCIÓN¹

La propuesta de Levinas de una filosofía primera como ética busca romper con la primacía de la ontología y de la representación del pensamiento occidental. Antes de que la conciencia dé inicio a su labor sintetizadora, el sujeto es interpelado y reclamado por el Otro. El llamado a la responsabilidad acontece en el encuentro con el Otro, el sujeto es elegido como el único que puede responder, su unicidad de sujeto consiste en la unicidad de la elección. En la *proximidad* es en donde tiene lugar la elección, es la incesante puesta en cuestión del sujeto, que impide la evasión y la abdicación de la responsabilidad. La subjetividad lingüístico-dialógica que esboza Levinas, como Decir original anterior a todo Dicho², permite delinear lo humano desde la intersubjetividad y la proximidad. El sujeto inquieto y obsesionado por el Otro surge como oposición al sujeto moderno que reposa en la comodidad de su pensamiento inmanente e indiferente. La proximidad no es una simple característica del encuentro intersubjetivo, sino que hace posible la emergencia del sujeto como tal, por medio de la palabra. En este artículo se considerará la proximidad como noción central en el pensamiento levinasiano, especialmente como constitutiva de la subjetividad y del diálogo original.

La relación yo-tú, en la ética de la alteridad levinasiana, es asimétrica, pues la presencia del Otro reclama responsabilidad. La relación intersubjetiva no es *horizontal* sino *vertical*, en el sentido de que el Otro no está al mismo nivel (Moreno Márquez, 1987). El Rostro expresa una dimensión de altura, es autoridad para el Mismo. Autoridad que nace de la desnudez del Rostro, de su debilidad y vulnerabilidad. El Otro es maestro, su enseñanza original consiste en la permanencia de la exterioridad en la relación intersubjetiva y en la interpelación a la responsabilidad.

1 El presente artículo es resultado del Proyecto de investigación “Aspectos performativos del lenguaje. Sentido ético y político de los actos de habla y su correspondiente noción de subjetividad” (Resolución de aprobación Nro.: 10.253) dirigido por el Dr. Federico Viola en el Instituto de Filosofía de la Universidad Católica de Santa Fe. El proyecto fue posible gracias al financiamiento y el apoyo de la nombrada institución.

2 Levinas entiende por *Dicho* la forma estática del lenguaje como conjunto de signos, cuya función es comunicar y referir. Es el lenguaje que refleja el ser, que fija y sistematiza la realidad. El *Decir*, por el contrario, es el lenguaje como proximidad, el cual posee una significancia ética que antecede a la sistematización de la conciencia.

En el pensamiento de Levinas, la subjetividad surge como inspiración a la responsabilidad. Interpelación ética que acontece en la proximidad. Asimismo, el Rostro es maestro que ordena la no-indiferencia y la responsabilidad ineludible. La forma original de concebir lo humano en Levinas plantea algunos interrogantes: ¿Qué importancia tiene la proximidad en la relación intersubjetiva? ¿Qué implica la “curvatura del espacio intersubjetivo”? (Levinas, 2016, p. 327) ¿Es el lugar de la proximidad un espacio ético y lingüístico? ¿Qué lugar tiene la corporeidad en la subjetividad inspirada? La asimetría de la relación intersubjetiva ¿no impide el diálogo y la fraternidad? ¿Cómo entiende Levinas el diálogo y qué relación tiene con la subjetividad? ¿Cómo se integran la verticalidad de la asimetría de la relación ética con la horizontalidad del diálogo?

A la luz de estos interrogantes, y con el objetivo de presentar la concepción levinasiana del diálogo como proximidad, en este artículo se analizará, en los dos primeros apartados, la desformalización del espacio geométrico que genera el encuentro intersubjetivo y el surgimiento de un espacio ético y lingüístico; en el tercer apartado, el papel de la corporeidad en la subjetividad; y, en los dos últimos apartados, el diálogo como diaconía y fraternidad. La noción de proximidad es abordada especialmente en los dos primeros apartados, al describir la espacialidad ética del encuentro intersubjetivo, pero su análisis es transversal a todo el artículo. El estudio de la corporeidad vincula las nociones de proximidad y de diálogo, en tanto que es corporalmente como irrumpe el llamado del otro; irrupción sensible que acontece en la proximidad de la escucha.

1. DESFORMALIZACIÓN DEL ESPACIO GEOMÉTRICO

El llamado ético a la responsabilidad produce una desformalización del tiempo: el tiempo del Decir pre-original, el tiempo de la proximidad, es un tiempo diacrónico, que nunca fue presente ni es posible re-presentarlo por la conciencia. Es un tiempo pasado inmemorial. De la misma manera que el tiempo, en el encuentro con el Otro, es *desformalizado* por la diacronía, así también el espacio intersubjetivo produce la desformalización del espacio geométrico por la proximidad. El espacio ético de la proximidad es *curvo*, lo cual permite que la exterioridad del Rostro se presente en su superioridad. La curvatura del espacio imposibilita la asimilación de la exterioridad del Otro en la interioridad del Mismo, “cambia la distancia en elevación” (Levinas, 2016, p. 327).

Con la metáfora de la curvatura del espacio, Levinas busca superar el dominio de la visión que conduce a la asimilación del Otro y a la violencia de la historia. El Rostro expresa lo infinitamente Otro, en tanto separado y absolutamente exterior. Paradójicamente, es la rectitud del Rostro, como inmediatez del

cara-a-cara, como imposibilidad de conceptualizar al Otro y de comprenderlo por medio de los conceptos, lo que produce la curvatura del espacio. La distancia intersubjetiva es rectitud (inmediatez) que se curva por la altura del Otro, por su Maestría y exterioridad: “La dimensión de altura en que se sitúa el Otro es como la primera curvatura del ser –en ello está el privilegio del Otro-, el desnivel de la trascendencia” (p. 91).

En su obra de juventud, *De la existencia al existente*, Levinas reconoce y rescata del pensamiento de Descartes el carácter personal y concreto que caracteriza a la *res cogitans*:

El *cogito* no desemboca en la posición impersonal: *hay pensamiento*, sino en la primera persona del presente: *soy una cosa que piensa*. La palabra *cosa* es aquí admirablemente precisa. La enseñanza más profunda del *cogito* cartesiano consiste precisamente en descubrir el pensamiento como sustancia, es decir, como algo que se pone. El pensamiento tiene un punto de partida. No se trata de una conciencia de la localización, sino de una localización de la conciencia. (2000a, p. 94)

El discurrir de la conciencia inicia en un sujeto posicionado en el espacio en un *aquí* concreto. Levinas enfatiza el hecho de que el comienzo de la reflexión consciente acontece en un sujeto ya en posición: “es gracias al hecho de apoyarse en la base como el sujeto se establece como sujeto” (p. 97). Plantea un tiempo pre-original, previo al tiempo de la conciencia, en donde el sujeto es afectado éticamente en su carne por la irrupción del Otro. Es decir, en ese tiempo pre-original el sujeto ya está posicionado en un espacio, es un cuerpo que siente la afección del Otro: “El lugar, antes de ser un espacio geométrico, antes de ser el ambiente concreto del mundo heideggeriano, es una base. De ahí que el cuerpo sea el advenimiento mismo de la conciencia” (p.98). Por el hecho de su corporeidad, de su posición en el espacio, el sujeto es afectado por la alteridad. La inspiración del sujeto, su irrupción sensible, acontece en el *aquí* del cuerpo: “Sin posición, sin localización, no es posible la subjetividad en cuanto tal” (Viola, 2016, p. 133).

Levinas compara su noción de corporeidad como *posición* con la relación novedosa entre la escultura y el pedestal en la obra de Rodin. El cuerpo, en el pensamiento levinasiano, no es un simple transmisor o revelador del mundo interior del sujeto, sino que es el mismo acontecimiento de la interioridad en cuanto posición (Levinas, 2000a, p. 99). Asimismo, las estatuas de Rodin llaman la atención de Levinas por su forma de estar cimentadas: “El acontecimiento que llevan a cabo sus estatuas reside mucho más en su relación con la base, en su posición, que en su relación con un alma-saber o pensamiento, que tendrían que expresar” (p. 99-100). La localización de la conciencia conduce a una concepción descosificada del cuerpo y a una subjetividad encarnada: “el *ahí* o la *posición* constituyen la *instancia* original de una subjetividad corporal” (Garrido-Maturano, 2017, p. 235).

La desformalización del espacio es producida por la asimetría de la relación intersubjetiva. La relación con el Otro es inmediata, no hay mediación posible, no hay un tercer término común que fundamente la relación, formando así una colectividad. La mediación y la pertenencia común no es aplicable al espacio ético de la asimetría del Rostro. La exterioridad del Rostro es distinta de la exterioridad de los objetos, es una exterioridad absoluta en cuanto no reducible a concepto ni asimilable como tema. La relación con el otro es la del Eros, en donde la distancia entre el Mismo y el Otro siempre permanece en la proximidad: “La intersubjetividad asimétrica es el lugar de una trascendencia donde el sujeto, aun conservando su estructura de sujeto, tiene la posibilidad de no retornar fatalmente a sí mismo, de ser fecundo” (Levinas, 2000a, p. 130). La exterioridad del Otro y su presencia que es al mismo tiempo ausencia, con la consecuente asimetría de la relación, conducen a la curvatura del espacio ético. Además, el enigma del Rostro que es al mismo tiempo vulnerable y maestro desformaliza el espacio volviéndolo *curvo*: “curvatura que hace del acceso al Otro un movimiento descendente (hacia el Otro en su postración) y de trascendencia (trans-ascendante)” (Moreno Márquez, 1987, p. 346).

Garrido-Maturano (2017) explicita el papel de la espacialidad en la noción de instante propuesta por Levinas. En su análisis, la espacialidad, junto a la alteridad que comporta, son la condición de posibilidad del instante. El surgimiento del tiempo depende de un Otro, no es posible su inicio en la soledad de la inmanencia: “el tiempo no puede surgir en el sujeto solo (...) el tiempo es la relación misma con el otro” (p. 234). La relación con el Otro que posibilita el tiempo es una relación espacial, parte de la posición del sujeto en el aquí concreto. El sujeto en tanto corporal y posicionado es afectado por la proximidad del Otro: “El hecho de que, en cuanto cuerpo, el sujeto nazca en una posición dada comporta en el núcleo mismo de la propia identidad una afección de la alteridad” (p. 235). La irrupción del Otro en el Mismo, la inspiración, no acontece en el mundo desencarnado de los conceptos, sino en la sensibilidad corpóreo-espacial de la carne. Por tanto, el instante en que el sujeto emerge sobreviene como “la intrusión de un estar” (p. 237) que trastorna la continuidad del tiempo y del espacio. El llamado a la responsabilidad que hace emerger al sujeto acontece en un tiempo diacrónico y en un espacio “dia-tópico” (p. 237):

La responsabilidad para con el otro es el lugar en que se coloca el no-lugar de la subjetividad, allí donde se pierde el privilegio de la pregunta *dónde*. Allí es donde el tiempo de lo *dicho* y de la *esencia* deja escuchar el decir pre-original, responde a la trascendencia, a la dia-cronía. (Levinas, 2003, p. 54)

La subjetividad como Decir pre-original es diacronía del tiempo, porque espacialmente es cuerpo afectado por la proximidad.

La corporeidad es la que enlaza y transforma espacio y tiempo en el sujeto, ya que es en el cuerpo en donde el sujeto es afectado por la proximidad que lo interpela a ser responsable. Es la proximidad sentida en la carne la que transforma el espacio en un lugar ético: “Es la intriga de la proximidad donde se instala un polo absoluto de alteridad desde el cual se hace el tiempo como respuesta o responsabilidad por la asignación del otro” (Garrido-Maturano, 2017, p. 243).

2. EL ESPACIO ÉTICO: ESPACIO LINGÜÍSTICO, ASIMÉTRICO Y DESIGUAL

La relación intersubjetiva produce un trastorno en la espacialidad, la distancia geométrica deja paso a una distancia ética o *separación* (Levinas 2016). La interpelación del Rostro, en su asimetría, es llamado a la responsabilidad que acontece corporalmente, es decir, se siente en la carne su afectación ética. La proximidad, en tanto inquietud y desformalización del espacio, expresa una significancia ética que rompe con la reciprocidad: “Ella llega a su punto *superlativo* como *mi* inquietud que no cesa, se convierte en única y desde ese momento uno olvida la reciprocidad como si se tratase de un amor del que no se espera correspondencia” (Levinas, 2003, p. 142). La inquietud generada por la orden del Rostro, en cuanto conlleva una responsabilidad intransferible, es elección e inspiración. Elección que consiste en la unicidad del sujeto y que tiene su fuente en la asimetría de la relación intersubjetiva. La proximidad, por la intensidad de su reclamo, imposibilita la indiferencia y la transferencia de responsabilidad; perturbando así, la tranquilidad del sujeto que reposa en el ser. Así también, la asimetría del Rostro, su altura exclusiva, escapa a la reciprocidad: “La asimetría de tal relación significa que los derechos del Otro son obligaciones para el Mismo, sin que esto implique que los derechos del Mismo deben ser obligaciones del Otro” (Moreno Márquez, 1987, p. 347).

El carácter asimétrico de la relación intersubjetiva, con la consecuente imposibilidad de la reciprocidad, es un tema primordial del pensamiento de Levinas, en el cual insiste a lo largo de toda su obra. Él mismo afirma que es una de las tesis fundamentales de *Totalidad e Infinito* (Levinas, 2000b, p. 82). Asimismo, es lo que separa y distingue la reflexión levinasiana de la filosofía del diálogo de Martin Buber (1982). Levinas asume en gran medida el pensamiento de Buber, pero plantea sus divergencias. Reconoce como su aporte fundamental, la “valoración de la relación dia-logical y de su irreductibilidad fenomenológica, de su aptitud para construir un orden dotado de sentido autónomo y tan legítimo como la tradicional y privilegiada correlación sujeto-objeto en la operación del conocimiento” (Levinas, 2002, p. 56). La divergencia fundamental propuesta por Levinas consiste

en la no-reciprocidad de la relación Mismo-Otro. La asimetría del Otro es la que impide la reciprocidad. La palabra fundamental Yo-Tú propone una relación horizontal y equivalente, en cambio la socialidad original levinasiana, es una relación desigual, el Otro reclama desde una altura. Reclamo que sólo se puede responder éticamente, es decir, siendo responsable hasta la sustitución, gratuidad del eros que no espera ninguna restitución. Diferenciándose de la igualdad inicial de la relación Yo-Tú de Buber, Levinas plantea una desigualdad original: “desigualdad ética: subordinación al otro, diacronía original: la primera persona en acusativo y no en nominativo” (p. 59).

El espacio ético de la relación intersubjetiva es un espacio lingüístico. Levinas propone un tipo de subjetividad lingüística, como Decir pre-original anterior a todo Dicho. Decir como proximidad, que posee un tipo de “significación anárquica” (Levinas, 2003, p. 141), no localizable en el mundo de las esencias. El Decir, que emerge en el espacio abierto por el Otro al resquebrajar la totalidad ontológica, posee un sentido ético. Significancia ética que va más allá de la significación de los seres. El sentido ético irrumpe en el mundo indiferente de la intencionalidad, como no-indiferencia ante el Otro. Enseñanza ética original que pone en cuestión a la conciencia al poseer un sentido ético que la trasciende: “La conciencia es puesta en cuestión en el cara-a-cara precisamente porque la posición de altura desde la cual el humano Otro viene impone un sentido a la relación desde fuera de los límites de la *conciencia de ese sentido*” (Drabinski, 2001, p. 120). El hecho de que la asimetría del Rostro exponga un sentido que escapa a la intencionalidad, desestabiliza al *cogito* autosuficiente. En la proximidad del cara-a-cara, el sujeto aprende algo absolutamente nuevo. Esta novedad absoluta es posible, porque su significancia es ética, proviene desde fuera de los esquemas y categorías de la conciencia.

Uno de los grandes objetivos de Levinas al desarrollar su ética de la alteridad como filosofía primera, es ampliar la noción de sentido (Levinas, 2002, p. 105). Para él la ética aporta “una significación sin referencia al mundo, al ser, al conocimiento, al Mismo y al conocimiento del Mismo” (Levinas, 2005, p. 162). El sentido ético trasciende la rigidez de la intencionalidad y la visión. Es un sentido que se hace sentir sin aparecer a la visión del intelecto, en tanto “sentir que no se reúne todavía en una percepción (...) [es un] aventurarse más acá de la unidad apriorística de la impresión originaria y de la retención, hacia la impresión originaria concebida en su independencia contra toda tematización” (Duhamel, 2011, p. 87). El Rostro expresa un sentido que trasciende la significación del ser y del conocimiento representado. Es un tipo de racionalidad ética, no ontológica ni fenomenológica; un Decir de la proximidad que antecede y cimienta todo Dicho: “El sentido es el rostro del otro, y todo recurso a la palabra se sitúa ya en el interior del cara a cara original del lenguaje” (Levinas, 2016, p. 230). En el espacio

ético de la intersubjetividad, el Rostro enseña un sentido que antecede al fenómeno, dando testimonio así de una forma de racionalidad distinta y primigenia, una racionalidad ética que significa en tanto proximidad. Esta racionalidad no es independiente de la espacialidad concreta en la que se sitúa el individuo (como pretende serlo la racionalidad ontológica), ya que acontece en el encuentro cara-a-cara de la proximidad. La proximidad como significación conduce al sujeto a una relación con la espacialidad desde la pasividad, es decir, el sentido es recibido desde fuera, no es dado por el sujeto, sino que es enseñado por un Otro. El sujeto ya no es aquel que dota de sentido y nombra los seres y los fenómenos, sino aquél que escucha y acoge en la proximidad al Otro. Doble pasividad del sujeto, el cual renuncia a *nominar* a los seres y se deja enseñar por el Otro: “La significación es lo infinito, o sea, el Otro. Lo inteligible no es un concepto sino una inteligencia. La significación precede a la *Sinnggebung* [donación de sentido] e indica el límite del idealismo” (p. 231).

Un ejemplo concreto que da Levinas de cómo la proximidad es significación y transmisión de sentido ético, es la amistad judeocristiana que se desarrolló en la Europa de la postguerra en el grupo llamado *Amistad judeocristiana de París*. La describe como una “búsqueda de una proximidad que dura, incluso cuando el diálogo se hace imposible” (Levinas, 2014, p. 72). Es posible escuchar el murmullo de la significación de la proximidad, cuando se acalla la apologética de la razón obsesionada por justificarlo todo. La reconciliación de aquellos que sufrieron los desastres de la guerra comenzó con la renuncia a encontrar una solución definitiva o un principio absoluto. En las crisis y conflictos profundos la persuasión es obscena y violenta. Sólo ayuda la presencia próxima que asume y respeta la identidad y responsabilidad de cada uno. Como un velar incesante por el Otro en su alteridad: “ir hacia el Otro ahí donde es verdaderamente otro, en la contradicción radical de su alteridad” (p. 72). En este contexto, Levinas refiere a un pasaje de *El Principito* de Saint Exupery, aquél en que el aviador es despertado por el principito y éste le pide que dibuje un cordero. El aviador al principio rehúsa el pedido diciendo que no sabe dibujar, luego accede al mismo, pero tiene que dibujar el cordero varias veces porque no satisface al que lo demandó. Sólo el dibujo de la caja con el cordero durmiendo en su interior conforma y alegra al Principito. Este fragmento literario es recordado por Levinas porque ilumina la significación de la *proximidad* que renuncia a dotar de sentido problemas insolubles, que prescinde de confiar en un principio racional que cierre la pregunta y desvele el misterio de la complejidad de la realidad:

Yo no sé dibujar la solución de los problemas insolubles. Esta duerme aún en el fondo de una caja sobre la que velan, entre tanto, personas que se han acercado. No tengo ninguna idea de la idea que habría que tener: el dibujo abstracto de un

paralelepípedo, cuna de nuestras esperanzas. Tengo la idea de un posible en donde quizá duerma lo imposible. (p.73)

La proximidad del Otro deshace la centralidad del sujeto en cuanto le interpela a la responsabilidad. El Rostro lo pone en cuestión antes de cualquier actividad y discurrir del pensamiento. La posición del sujeto en la proximidad pierde la primacía, así como su privilegio de ser el que juzga, es decir se pasa de la autoafirmación del sí mismo al cuestionamiento por parte del otro que lo interpela y lo obliga a justificar su ser: “es apología, discurso *pro domo*, pero discurso de justificación ante el Otro” (Levinas, 2016, p. 330). El Rostro *traumatiza* al sujeto al presentarle una significación ética que desmorona la exclusividad de la significación ontológica e intencional. Es una irrupción sensible que el sujeto no puede dominar, trasciende el ámbito de la libertad: “la libertad es cuestionada en la invocación moral porque es pasiva en la constitución del sentido ético. El sentido ético es impuesto, no creado o aceptado por el Yo” (Drabinski, 2001, p. 124). Esta pasividad testimonia la unidireccionalidad de la significación ética. La altura del Otro, su desigualdad respecto del Mismo, suscita, en efecto, una relación asimétrica que va más allá de toda reciprocidad. La unidireccionalidad de la donación de sentido ética provoca la unidireccionalidad de la responsabilidad como respuesta. Ante el exceso de sentido del Rostro, ante su infinitud inasimilable, el sujeto sólo puede responder desde la insaciabilidad de la praxis responsable: “La deuda se acrecienta en la misma medida que se cancela (...) la positividad del infinito es la conversión en responsabilidad, en acercamiento al otro” (Levinas, 2003, p. 56).

El lugar de la emergencia de la subjetividad es el espacio ético-lingüístico del Decir pre-original. *Espacio ético* que es abierto por la proximidad del Rostro que instaura la responsabilidad: “la proximidad, como si fuese un abismo, interrumpe la indescargable esencia del ser” (Levinas, 2003, p. 152). *Espacio eminentemente lingüístico* del sujeto como Decir sin Dicho, destronado del reposo en el ser y de la ilusión del *conatus essendi*, llamado y elegido para la inquietud permanente de la salida de sí: “En el Decir el sujeto se aproxima al prójimo ex-presándose en el sentido literal del término; esto es, expulsándose de todo lugar, no morando ya más, sin pisar ningún suelo” (p. 101). *Espacio asimétrico y desigual*, ya que el Otro no está al mismo nivel, habla desde una altura que no permite la reciprocidad ni la comparación, que produce en la espacialidad una *curvatura* que verticaliza la relación intersubjetiva: “En la relación con el otro, el otro se me aparece como aquel a quien debo algo, respecto del cual tengo una responsabilidad” (Levinas, 2014, p. 81).

3. CORPOREIDAD

La desformalización del tiempo y del espacio que la relación intersubjetiva realiza, comporta una subjetividad encarnada, es decir, corporal. El cara-a-cara de la proximidad no refiere al encuentro inmaterial entre espíritus, sino al contacto del Rostro que es sensibilidad y vulnerabilidad de la carne. El análisis antecedente de la proximidad y la espacialidad conduce a los siguientes interrogantes: ¿Qué papel asume el cuerpo en la subjetividad inspirada que propone Levinas? ¿Cómo el acceso al Otro se da sensiblemente en la corporeidad?

El sujeto lingüístico e inspirado emerge en un tiempo previo a la conciencia, el tiempo de la sensibilidad y del Decir de la proximidad, Decir aún no tematizado en un Dicho. Además, la relación intersubjetiva desformaliza el espacio geométrico creando una espacialidad ética. Es en el espacio-tiempo de la proximidad en donde el sujeto nace como sensibilidad y corporeidad abierta a la irrupción sensible del Otro, como pasividad expuesta a la herida: “Bajo las especies de la corporeidad, cuyos movimientos son fatiga y cuya duración envejecimiento, es donde la pasividad de la significación, del uno-para-el-otro, no es acto, sino paciencia, esto es, sensibilidad o inminencia del dolor” (Levinas, 2003, p. 110). En el cuerpo es en donde se unen el esfuerzo de la vida económica, el disfrute y la necesidad, con el llamado a responder que proviene de la alteridad. Levinas busca “descosificar” el cuerpo y enfatizar su carácter de acontecimiento, es decir, “pensar el cuerpo antes de su incorporación en un horizonte y en un mundo” (Gibu, 2014, p. 101). Pensar la corporeidad como un acontecimiento previo a la intencionalidad de la conciencia y a la representación.

Levinas aborda la pregunta por el cuerpo y la sensibilidad del sujeto desde los primeros años de su reflexión filosófica, siendo además un tema recurrente en toda su obra. En un texto titulado *Algunas reflexiones sobre la filosofía del hitlerismo* (2001a), publicado en la revista *Esprit* en el año 1934, desarrolla una noción de corporeidad que difiere de la concepción clásica del cuerpo como un obstáculo a superar por el espíritu. El desprestigio del cuerpo, e incluso su negación, parten del olvido del *sentimiento de identidad* originario (o preoriginario) que posee el cuerpo: “¿Acaso no nos afirmamos en este calor único de nuestro cuerpo con bastante anterioridad a la dilatación del Yo que busca distinguirse del cuerpo? (...) ¿No permanece el espíritu rebelde encerrado en el dolor, ineluctablemente?” (pp. 164-165). El cuerpo no es un ropaje, algo externo al Yo, cuya misión consiste simplemente en ubicarlo como un ser abierto al mundo. El gozo y el dolor testimonian la irreductibilidad del cuerpo y la incapacidad del espíritu de vivir desencarnadamente. Testimonio que manifiesta la centralidad de la corporeidad en la vida del sujeto: “Lo biológico, con toda la fatalidad que conlleva, se convierte en algo más que un objeto de la vida

espiritual, se convierte en su corazón” (p. 165). La ilusión de una libertad absoluta colisiona con el sentimiento radical de la corporeidad. La libertad no sólo es finita por el hecho de haber tenido un inicio, sino también por su condición encarnada, lo que desemboca en la paradoja de una libertad indefectiblemente *encadenada* a un cuerpo.

En el espacio-tiempo ético de la relación intersubjetiva, la proximidad es en sí misma significación. Antes de la significación de la verdad objetiva epistémica de lo Dicho, el sujeto como Decir pre-original es *enseñado* por la significancia ética de la proximidad del Otro: “una significación que sólo es posible como encarnación. La animación, el propio *pneuma* del psiquismo, la alteridad dentro de la identidad es la identidad de un cuerpo que se expone al otro” (Levinas, 2003, p. 126). La inspiración que caracteriza al sujeto levinasiano acaece como irrupción sensible, en tanto que la inquietud que produce el Otro en el Mismo es un trastorno sensible de la corporeidad. La significación de la proximidad, que llama/inspira a la responsabilidad, es recibida y percibida por el sujeto en su cuerpo, en la materialidad de su condición y posición: “mi cuerpo no sólo es un objeto percibido, sino un sujeto que percibe (...) el horizonte desde el cual las cosas se manifiestan sólo puede esclarecerse a partir de mi situación” (Gibu, 2014, p. 105). El sujeto no puede prescindir, en la búsqueda de sentido y de realización, de su condición encarnada. El drama de lo humano consiste en la imposibilidad de evadirse de sí mismo, en la imposibilidad de jugar el juego del pensamiento sin comprometerse sensiblemente con la verdad: “Encadenado a su cuerpo, el hombre se niega a escapar de sí mismo. La verdad ya no será la contemplación de un espectáculo extraño – la verdad consistirá en un drama cuyo actor es el propio hombre” (Levinas, 2001a, p. 166).

En su obra de juventud *El Tiempo y el Otro* (1993), Levinas enlaza la materialidad corporal con la primacía de la responsabilidad. El cuerpo, en lo que tiene de definitivo, encadena al sujeto a sí mismo, lo hace responsable antes de toda libertad: “No existo como espíritu, como una sonrisa o un viento que sopla, no soy sin responsabilidad. Mi ser se duplica en un deber: estoy a cargo de mí mismo. En esto consiste la existencia material” (p. 94). Ya en sus primeros escritos, se evidencia la intención del filósofo lituano de distanciarse del pensamiento racionalista moderno, el cual fundamenta la subjetividad en el ser y la razón, subordinando la corporeidad a la mitología de una autonomía absoluta del espíritu. Por el contrario, la existencia encarnada del sujeto conduce a una subordinación de la libertad a la primigenia responsabilidad, la cual consiste en asumir el deber que conlleva la materialidad corporal: “una unión cuyo gusto trágico y definitivo nada podrá alterar” (Levinas, 2001a, p. 165). No hay razonamiento ni aventura del pensamiento capaz de hacer olvidar al sujeto de su posición corporal en la existencia y de su consecuente responsabilidad.

Asimismo, es en tanto cuerpo que el sujeto es capaz de relacionarse con la alteridad radical del Otro. Es en la proximidad donde el sí mismo aprende la novedad absoluta del Rostro del Otro, su significancia ética que fundamenta toda posterior reflexión y trasmisión de sentido. A partir de la asunción de su corporeidad, el sujeto es capaz de escuchar y sentir en su carne el llamado de la alteridad: “Existir como cuerpo es estar abierto al otro, al sentido de su altura, a su distancia inalcanzable, insuperable (...) el cuerpo es el lugar de la articulación recíproca entre la interioridad y la exterioridad” (Verano Gamboa, 2018, p. 385). Por tanto, el cuerpo no sólo es el lugar de la responsabilidad del sí Mismo por su carne, sino también el lugar del acceso al Otro. La inspiración descentra e inquieta al sujeto, conduciéndolo a la responsabilidad y al drama de la obsesión por el Otro. No hay posibilidad de evasión o indiferencia ante la herida del clamor del prójimo, ya que es en la proximidad y en la materialidad del cuerpo que acontece la irrupción del Otro. Así como “la materia es el lugar propio del para-el-otro”, así también, la responsabilidad es una deuda sensible y corporal, en tanto que el sujeto es “susceptible de dar el pan de su boca o de dar su piel” (Levinas, 2003, p. 136). El *de otro modo que ser* de la subjetividad comienza como irrupción sensible que hiere la propia carne y conduce a la sustitución, exposición total al Otro sin posibilidad de refugio, “des-interés a modo de vida corporal volcada a la expresión y al dar” (p. 104).

La subjetividad en Levinas no sólo es trazada como corporal y sensible, sino también como *vulnerable*. La apertura del sujeto ya no es la salida de sí de la intencionalidad y de la comprensión del ser, sino “lo descarnado de la piel expuesta a la herida y al ultraje (...) vulnerabilidad de una piel ofrecida” (Levinas, 2009, p. 122). La posición del sujeto no consiste en el poderío de quien domina y conceptualiza toda diferencia, sino en la pasividad radical de aquél que está totalmente expuesto a la ofensa. Lo que da inicio a la subjetividad no es la actividad racional, ni siquiera cualquier otra actividad realizada por el sujeto, sino la pasividad de la sensibilidad que es vulnerable a la herida del Otro: “Pasividad más pasiva que toda pasividad (...) El Yo, de pie a cabeza, hasta la medula de los huesos, es vulnerabilidad” (p. 123). Es esta misma pasividad vulnerable el terreno en donde puede brotar la relación intersubjetiva con su sentido ético: “Desde la sensibilidad, el sujeto es para el otro: sustitución, responsabilidad, expiación” (p. 125). En cuanto el sujeto es vulnerable es capaz de escuchar la palabra/mandato original del Otro que llama a la responsabilidad. La subjetividad del sujeto “no es un acto, sino un sentimiento absolutamente pasivo: sentir la responsabilidad por el otro, sin poder escapar” (Grollo, 2014, p. 86).

Una imagen que utiliza Levinas para ilustrar el acercamiento al Otro por fuera de la intencionalidad, es la del *tacto*. De la misma forma que el oído se opone a la visión del pensamiento tematizador, haciendo estallar el mundo de la luz, y

permitiendo que la gloria de la alteridad resuene (Llorente, 2014); así también el tacto refleja la permanencia de la exterioridad de lo sentido. Si bien hay formas del tacto que pueden obedecer al deseo de dominio y posesión, esto sucede en cuanto ya han superado el ámbito de la pura sensibilidad: “La mano comprende la cosa (...) porque ya no es un órgano del sentir, no es puro disfrute, pura sensibilidad, sino señorío, dominación, disposición – aspectos todos que no pertenecen al orden de la sensibilidad” (Levinas, 2016, p. 178). En el tiempo pre-original de la proximidad el sujeto es pura pasividad de la sensibilidad, aún no pretende dominar la realidad por la actividad del intelecto.

Dentro del sentido del tacto, Levinas refiere a la *caricia* como una forma de tocar sin poseer, que trasciende el mundo de lo percibido: “no es una intencionalidad del des-velamiento sino de la búsqueda: un ir hacia lo invisible (...) busca, más allá del consentimiento o la resistencia de una libertad, *lo que aún no es*” (p. 292). Búsqueda que refleja, paradójicamente, la pasividad del tacto, ya que no impone un porvenir o un horizonte de sentido a lo tocado, sino que es un encuentro en la noche del intelecto con lo que ya está ausente:

En la ciega *episteme* de los amantes, en la invidencia del gozo vuelto sobre sí mismo, se es para el Otro, se recibe al Otro, en la caricia y en la ternura el Otro es ya una visita en la noche y el anuncio de una trascendencia imprevisible. (Di Giacomo, 2016, p. 60)

La caricia es sensibilidad, pero a su vez trasciende lo sensible (Levinas, 2016, p. 291). Consiste en no aprehender nada, en buscar lo que no es aún, más allá del porvenir y de lo posible. Convierte el presente en un “presente-futuro” (p.293) incapaz de abarcar; es un contacto previo a la comprensión y al saber: “El tropo *demasiado pronto* protege el contacto de escapar a la cognición” (Joldersma, 2014, p. 79). Permanece como pura proximidad, prescindiendo de la intencionalidad de la conciencia que llega tarde al acontecimiento de la caricia.

La recuperación realizada por Levinas del papel central de la corporeidad en la subjetividad, es correlativa con la búsqueda de un sentido fuera de los esquemas de la mera razón y el entendimiento. En la distinción entre Decir y Dicho se hace manifiesta esta relación. El Decir significa en tanto proximidad, es decir, en el cara-a-cara corporal de la relación intersubjetiva. El Rostro del Otro en la relación de proximidad, expresa un sentido previo al sentido de la ontología, un sentido ético. Esta expresión de sentido es posible y acontece en la materialidad de los cuerpos próximos: “significancia, el-uno-para-el-otro, exposición de sí mismo a otro, inmediatez en la caricia y en el contacto del decir” (Levinas, 2003, p. 145). Por lo tanto, no es una significación conceptual plausible de una transmisión inmaterial, sino que el sentido ético brota de la proximidad corporal. En cambio, la significación ontológica y conceptual es ya un discurso transmisible independientemente

de la proximidad material, se enmarca en la totalización propia de lo Dicho, en donde toda singularidad se universaliza. Por lo tanto, a una subjetividad encarnada corresponde una racionalidad ética de la proximidad, y viceversa: “el nacimiento del sujeto, de un *sí mismo* anterior a la reflexión, a la indagación explícita, es solo posible en la existencia material, corporal” (Verano Gamboa, 2018, p. 384).

Simultáneamente, es la corporeidad la que impide al sujeto el enclaustramiento en la inmanencia de la conciencia abriéndolo a la trascendencia de la exterioridad. El cuerpo posibilita una subjetividad inspirada como Otro en el Mismo, ya que la alteridad irrumpe sensiblemente en la carne del sujeto: “el contacto de la piel es todavía la proximidad del rostro, responsabilidad, obsesión del otro, ser-uno-para-el-otro; se trata del propio nacimiento de la significación más allá del ser” (Levinas, 2003, p. 153). La existencia corporal del sujeto es el lugar en donde acontece el encuentro con el Otro y en donde la noción de sentido se ensancha más allá del ser. Al *de otro modo que ser* de la subjetividad corresponde un *de otro modo que significar*: “Decir que, como responsabilidad, es la misma significación, el uno-para-el-otro, subjetividad del sujeto que se torna signo” (p. 229). El sujeto como Decir de la proximidad se convierte en signo que significa/testimonia éticamente, en cuanto es responsable. La noción lingüística del sujeto como Decir inspirado permite a Levinas enlazar, por medio del lenguaje, subjetividad encarnada, sentido ético y alteridad. Queda aún por ver, qué lugar ocupa el diálogo en dicha noción de sujeto, y si es posible alguna forma del mismo que se inscriba en el lenguaje como Decir y no como Dicho.

4. DIÁLOGO COMO DIACONÍA

En la obra de Levinas, la noción de diálogo adquiere significados diversos. En muchos de sus textos describe el diálogo como intercambio de saberes y, por tanto, como Dicho. Asimismo, en otros escritos, describe un tipo de diálogo como socialidad y proximidad originaria que pertenece al lenguaje como Decir. Si bien Levinas asume la noción de diálogo para referir a la relación ética primigenia de la intersubjetividad (al menos en alguno de sus textos), lo hace marcando cierta distancia de la concepción del diálogo de Martin Buber (1982), y más aún, de Gabriel Marcel (1969).

En primer lugar, sobresalen los textos en donde Levinas separa el diálogo de la subjetividad lingüística como Decir. En muchos lugares de su obra, en donde esboza la relación ética pre-original de la proximidad del Mismo y del Otro, precisa que de lo que está hablando es de una relación previa al diálogo como transferencia de saberes: “La comunicación de las ideas, la reciprocidad del diálogo, ocultan ya la esencia profunda del lenguaje. Reside ésta en la irreversibilidad de la relación

entre Yo y Otro: en la Maestría del maestro” (Levinas, 2016, p. 108). Es decir, la relación de la proximidad es asimétrica y desigual, el Otro no está al mismo nivel que el Mismo, proviene de una altura. La condición de Maestro del Otro impide la reciprocidad que conlleva el diálogo que meramente informa al otro sobre un contenido cognoscitivo. El lenguaje es primigeniamente proximidad y relación ética con el Otro, testimonio y no diálogo: “no creemos que el diálogo sea la primera forma del lenguaje. Éste es antes testimonio” (Levinas, 2005, p. 227). Levinas distingue entre el testimonio, que responde a la exigencia infinita de responsabilidad diciendo *Heme aquí*, del diálogo que nivela e iguala a los interlocutores en el mero comercio de la información. En otro texto también afirma la separación entre la responsabilidad inspirada por el Otro y el diálogo: “Decir: Heme aquí. Hacer algo para otro. Dar (...) Dia-conía antes de todo diálogo” (Levinas, 2000b, p. 81). El sujeto es pura pasividad que recibe la orden del Rostro de ser responsable; interpelación y asignación unidireccional: “Decir sin correlación noemática dentro de la pura obediencia a la gloria que ordena; sin diálogo, dentro de la pasividad sometida de golpe al heme aquí” (Levinas, 2003, p. 222). En todos estos pasajes se evidencia la intención de Levinas de preservar el carácter asimétrico de la relación, la altura del Rostro del Otro que verticaliza la relación intersubjetiva. El diálogo es entendido en estos textos como perteneciente al lenguaje como Dicho, es decir en tanto mero canal de transmisión de saberes intercambiables; rodeo reflexivo que se aleja del llamado ético del Otro.

En uno de los capítulos de su libro *De Dios que viene a la idea*, Levinas hace una distinción entre dos tipos de diálogo, a los que corresponden dos tipos de subjetividad. Primero describe la concepción tradicional de un diálogo basado en el saber y la inmanencia, el cual parte del sujeto entendido como cogito autónomo que comprende el ser e intercambia conocimientos: “En cuanto aprender, el pensamiento comporta una captura, una prensión y una posesión de lo aprendido (...) Pensamiento y psiquismo de la inmanencia: de la suficiencia para sí” (Levinas, 2001b, pp. 184-185). Este sujeto pensante no logra salir del encierro de la inmanencia, toda relación con una alteridad es reducida a concepto, la intencionalidad termina siendo dominio y homogeneización autorreferencial: “Es lo Mismo que se reencuentra en el Otro. La actividad del pensamiento tiene razón de toda alteridad” (p.9). El sujeto cimentado en el saber es incapaz de aprender, de recibir de Otro algo nuevo. De esta forma el diálogo se convierte en alternancia de discursos que, al fin de cuentas, conforman un *monólogo* que tiende a la unificación de los sentidos comunicados en una única forma de racionalidad. La finalidad de este tipo de diálogo es “penetrar en el pensamiento del otro” y “coincidir en la razón” (p.11). La razón es invocada como mediación indispensable para lograr la coincidencia a costa de la diversidad. Toda alteridad es reducida y hecha funcional a un objetivo común: “Es el famoso diálogo llamado a detener la violencia

al conducir a los interlocutores a la razón, al instalar la paz en la unanimidad, al suprimir la proximidad en la coincidencia” (p. 11). Esta forma de diálogo *racional*, que busca sintetizar toda diferencia, es a la que se opone Levinas, ya que no respeta la exterioridad del Otro al dominarlo y reducirlo por medio del pensamiento. La significancia ética del Decir es ajena a esta forma uniformante de diálogo, el Rostro está ya ausente cuando sobreviene el intercambio del diálogo como saber conceptual.

Pero Levinas, siguiendo en gran parte la tradición de la llamada *filosofía del diálogo* contemporánea, propone otra forma de concebir el diálogo que supera el encierro en la inmanencia del *cogito*. Es un diálogo abierto a la trascendencia en el cual la relación intersubjetiva posee un sentido en sí misma, como socialidad original. Es un diálogo en donde el Otro permanece en su exterioridad y alteridad. La distancia entre el Mismo y el Otro no se suprime, sino que se mantiene y se trasciende en la misma relación de proximidad: “absolutamente distinto el uno en relación con el otro, sin medida común ni dominio disponible para una coincidencia cualquiera (...) una manera de acceder al otro distinta del conocer: aproximar al prójimo” (p. 15). Es decir, la razón pierde su exclusividad y primacía en el sujeto y en el lenguaje. El Decir como proximidad y socialidad posee un sentido más allá del ser y del conocer, un sentido primigenio que acontece antes de toda reflexión. Un sentido ético que adviene al sujeto en la misma interpelación del Otro, y que lo conduce a la responsabilidad: “El diálogo es la no-indiferencia del *tú* al *yo*, sentimiento des-inter-esado (...) en el encuentro el *otro cuenta* por sobre todo (...) es el lugar y la circunstancia originales del advenimiento ético” (pp. 17-18). Es justamente esta dimensión ética lo que aporta el planteo levinasiano como novedad a lo desarrollado por los filósofos del diálogo. Es la inmediatez y urgencia del reclamo ético del Otro la que logra romper con los esquemas de la conciencia y del lenguaje como Dicho. Es más, el mismo lenguaje, en lo que tiene de apertura y solicitud por el otro, testimonia ya su dimensión ética: “La apertura misma del diálogo, ¿no es acaso una manera para el yo de descubrirse, de entregarse, una manera para el Yo de ponerse a disposición?” (p. 20). Este tipo de diálogo, en tanto apertura a la trascendencia en la proximidad, es *diaconía* en la forma de responsabilidad y obsesión por el Otro. Diaconía previa a todo diálogo de la inmanencia. Esta distinción entre diálogos radica en la diferencia entre la inmanencia del *maestro interior* que guía por la razón el discurso, y la trascendencia de la maestría del Otro que conduce a responder por la responsabilidad: “Pese a la afinidad de fórmulas, este diálogo ético/existencial se encuentra distante del diálogo mayéutico socrático, precisamente por esta idea de la exterioridad radical del Otro que, en su trascendencia y proximidad, actúa como el verdadero maestro interior” (Cerezo Galán, 1997, p. 101).

La *diaconía* que caracteriza al diálogo de la trascendencia procede de la asimetría de la relación intersubjetiva. La espacialidad del diálogo levinasiano se caracteriza por la curvatura de la horizontalidad de la relación, como consecuencia de la altura y la maestría del Otro. En el diálogo la respuesta del Yo al llamado del Tú es servicio. El Yo deja de ser el sujeto que domina por medio del pensamiento la realidad, dándole así sentido; y pasa a ser un sujeto inspirado e inquieto, que recibe pasivamente el sentido del Rostro y responde éticamente por el servicio: “el Yo como Yo es servidor del Tú en el diálogo” (p. 21). A un diálogo de la inmanencia corresponde un sujeto autónomo y autosuficiente, que tiende a transformar todo intercambio discursivo en puja de poder. En cambio, al diálogo de la trascendencia corresponde un sujeto descentrado y obsesionado por el reclamo del Otro, un sujeto que se deshace en responsabilidad y servicio.

Levinas no deja de reconocer la importancia de la filosofía de Buber en la búsqueda de superar la cosificación del Otro y la lógica cognoscitiva sujeto-objeto en la relación intersubjetiva. Asume en gran medida su planteo y resalta lo fundamental de su filosofía: “que la presencia de un interlocutor ante el Yo no se reduce a la presencia de un objeto que mi mirada determina y sobre el que enuncia juicios predicativos” (Levinas, 2002, p. 30). La propuesta levinasiana recorre el mismo camino al resaltar la irreductibilidad de la relación Yo-Tú a tema o concepto. La proximidad y la significación ética del Rostro no son objetivables ni plausibles de convertirse en Dicho sin ser traicionadas. El diálogo propuesto tanto por Buber como por Levinas, no es el diálogo que universaliza por la razón el habla de los interlocutores, el diálogo de la política, sino “el diálogo que hace entrar en diálogo” (p. 31). Es el diálogo de la proximidad que significa por sí mismo, independientemente de la razón, y que es fraternidad original. Un diálogo que agrieta la totalidad y la omnipresencia de la ontología al plantear el carácter originario y previo de la relación: “Si lo interhumano es lo originario, la relación interhumana se distancia abismalmente de las relaciones sujeto-objeto, donde no puede regir ninguna forma de encuentro” (Riva, 2005, p. 646). El diálogo rompe con el privilegio del ser y del conocer como únicos dadores y garantes de sentido.

En un texto titulado *Martin Buber, Gabriel Marcel y la filosofía*, Levinas (2002) compara a estos dos autores, marcando sus coincidencias y divergencias respecto al diálogo y la socialidad original. Ambos coinciden en destacar “la originalidad de la socialidad respecto de la estructura sujeto-objeto” (p. 36), describiendo la relación intersubjetiva por el binomio Yo-Tú; en confrontación con la objetivación propia de la relación Yo-Ello. Si bien la relación adquiere un lugar fundamental en la subjetividad, según Levinas, estos autores no terminan de abandonar totalmente el ámbito y el lenguaje de la ontología: “Entre-los-dos es *un modo de ser*: la co-presencia, el co-esse. Si seguimos literalmente los textos, el ser, la presencia, siguen siendo la última referencia de sentido” (p. 38). Tanto

Buber como Marcel buscan dejar atrás el ontologismo y el conocimiento como comprensión del ser, pero se diferencian entre sí en cuanto a la radicalidad y profundidad de esta ruptura. Levinas reconoce cómo Buber se opone a establecer un principio o fundamento común a la relación al describirla lingüísticamente, poniendo así el énfasis en la dinámica excéntrica de la palabra: “La palabra es el entre-los-dos por excelencia. El diálogo no funciona como una síntesis de la Relación, sino como su mismo despliegue” (p. 39). Al caracterizar lingüísticamente la relación logra preservarla parcialmente de la reducción de su trascendencia por la síntesis y del retorno a un esquema ontológico. La preservación de la trascendencia de la relación es parcial ya que sigue recurriendo al lenguaje ontológico para describir el encuentro intersubjetivo.

Por el contrario, el diálogo pierde su carácter originario y radical si necesita del amparo del ser para justificarse, si la esfera del lenguaje no es capaz de significar por sí misma. Este es el problema que descubre Levinas en la filosofía del diálogo de Marcel, su desconfianza del lenguaje: “el lenguaje sería inadecuado a la verdad de la vida interior, mientras que el Yo-Tú es vivido como la inmediatez misma de la co-presencia y, por consiguiente, por encima de las palabras, por encima del diálogo” (p. 40). La noción de *encarnación* desarrollada por Marcel resalta la co-presencia de la relación en detrimento de su dimensión lingüística: “La co-presencia no es un diálogo. Nexo intersubjetivo más profundo que el lenguaje” (p. 42). El ser y el conocer prevalecen por sobre el lenguaje en la relación intersubjetiva esbozada por Marcel. En cambio, el diálogo en Buber, al permanecer como relación lingüística fundamental, se aleja con mayor radicalidad del dominio de la ontología y del intelecto: “El acontecimiento incondicionado del encuentro desborda pensamiento y ser. Es un puro diálogo, pura alianza que no envuelve ninguna presencia común pneumática” (p. 44).

En la búsqueda de una subjetividad lingüística que supere la inmanencia del cogito, Levinas asume el planteo dialógico de Buber, y lo radicaliza. La trascendencia de la relación intersubjetiva adquiere en Levinas una dimensión ética. Es la exterioridad inasumible del Otro la que *introduce* la ética y la desigualdad en la relación. El reclamo de responsabilidad del Rostro constituye al sujeto, en tanto único elegido a responder; su unicidad de sujeto consiste en su elección. La desigualdad de la relación, en donde el Otro es Maestro y el Mismo es elegido, logra preservar el diálogo original de la objetivación de lo Dicho: “Es sólo el reconocimiento del único, y no la afirmación de la relacionalidad como tal, la que puede evitar el reproducirse de una estructura de universalidad neutra” (Petrosino, 1996, p. 33). El sujeto levinasiano, inspirado y elegido, profundiza y preserva la trascendencia de la relación intersubjetiva. El carácter ético de la irrupción del Otro rompe con la tendencia de la filosofía del diálogo de volver a un esquema ontológico: “la Relación, ¿no es ésta el des-inter-és mismo, el des-arraigo fuera del

ser –rectitud de impulso sin regreso a sí? Desinteresamiento que no significa indiferencia, sino fidelidad a lo otro” (p. 51). La propuesta levinasiana de un diálogo original y de su significancia ética, busca llevar a término el objetivo fundamental de la filosofía del diálogo: que la relación intersubjetiva se mantenga en su trascendencia, irreductible a una idea racional e inasumible por un intelecto.

5. DIACONÍA: FRATERNIDAD, AMISTAD Y HOSPITALIDAD

La dimensión lingüística y dialógica de la subjetividad, como expresión de la socialidad original, permite a Levinas conservar la trascendencia de la relación intersubjetiva. El diálogo, deja de ser una trasmisión recíproca de pensamientos que supone cierta tematización del interlocutor, y se transforma en praxis ética. El diálogo de la trascendencia es diaconía, apertura y servicio al otro, que mantiene la exterioridad y la alteridad. La separación y la distancia perduran en la relación, sin que esto signifique indiferencia o degradación. Uno de los objetivos de Levinas al plantear la prioridad de la relación, es romper con la tradición plotiniana de la “excelencia de la unidad, respecto a la cual la relación sería ya privación”; y proponer la “socialidad como independiente de la unidad perdida (...) es la excelencia de lo múltiple” (Levinas, 2001c, p. 138). La multiplicidad adquiere una primacía, ya no es la degradación de lo Uno. La relación intersubjetiva que mantiene y expresa la multiplicidad, no es una búsqueda incesante de recuperar una supuesta unidad original perdida. La riqueza de la relación consiste, precisamente por eso, en la permanencia de la multiplicidad y la diferencia. El lenguaje es servicio y cuidado del otro en su separación.

Esta excelencia de la multiplicidad se manifiesta en la importancia que concede Levinas a la dimensión ética del lenguaje, especialmente al describir al diálogo como *diaconía*. Utiliza para ello diversas expresiones que suponen como un valor la multiplicidad: fraternidad, amistad y hospitalidad. El diálogo de la socialidad y proximidad original es “el uno-para-el-otro de la *fraternidad* (...) la relación ética con el otro hombre en forma de responsabilidad para con él” (Levinas, 2002, p. 107). La fraternidad es originaria y constitutiva de la ipseidad (Levinas, 2016, p. 316), es responsabilidad previa a todo compromiso. Es en este contexto, que Levinas refiere el pasaje del asesinato de Abel por parte de su hermano Caín (Gn. 4, 1-16). La pregunta “¿Soy yo acaso el guardián de mi hermano?” (v.9) expresa la “sobria frialdad cainita” que “consiste en pensar la responsabilidad partiendo de la libertad o según un contrato” (Levinas, 2001b, p. 105). Antes de ser libre, el sujeto es responsable del próximo; la fraternidad como responsabilidad precede a la libertad. La dimensión ética del diálogo/diaconía no radica en una elección libre y consciente del sujeto, sino en la inmediatez del Rostro que le asigna una

responsabilidad, que lo elige como guardián de su hermano *antes* de toda decisión: “La proximidad del prójimo es mi responsabilidad para con él: aproximarse es ser guardián de su hermano, y ser guardián de su hermano es ser rehén” (p. 106). Es la proximidad del cara-a-cara la que hace posible la elección del sujeto como único responsable. Proximidad de un Rostro que no deja de ser extranjero: “la fraternidad (...) se caracteriza paradójicamente como la responsabilidad ante un rostro que me observa como extranjero” (Peñalvert, 2000, p. 145). Fraternidad original que trasciende los lazos biológicos e históricos, que emerge en la proximidad manteniendo la trascendencia y la extranjería del Otro.

En un artículo llamado *Paz y proximidad*, Levinas (2014) distingue entre dos tipos de paz. Por un lado, la paz a la que tradicionalmente Occidente ha tendido, una paz a partir de la verdad; la cual es alérgica a la alteridad y a la diversidad, y se sustenta sobre la preeminencia de la unidad. Por otro lado, ante el fracaso experimentado en el siglo XX de esta paz como unidad y uniformidad, se pregunta por una paz que responda a una apelación más urgente y distinta que la de la verdad; el llamado a la responsabilidad expresado en el Rostro: “Hay que preguntarse si la paz, en lugar de atenerse a la absorción o a la desaparición de la alteridad, no consistiría por el contrario en el modo fraternal de una proximidad del otro” (p. 105). La proximidad y la fraternidad son el fundamento de una paz ética que revaloriza la multiplicidad y la alteridad. Una paz que emerge de la inquietud de la responsabilidad provocada por la acusación del rostro: “La paz como despertar a la precariedad del otro” (p. 107). El lenguaje como proximidad además de poseer una significación ética previa a la intencionalidad, es en cuanto tal, una forma originaria de fraternidad anterior a la libertad: “una significación en la cual el *para* del uno-para-el-otro (...) es un *para* de gratuidad total que rompe con el interés; *para* de la fraternidad humana al margen de todo sistema preestablecido” (Levinas, 2003, p. 161).

Otra palabra que utiliza Levinas para describir la dimensión ética del lenguaje es la de *amistad*. En las conclusiones de su obra *Totalidad e infinito*, afirma que uno de sus objetivos es interpretar el ser como bondad, “interpretar la producción del ser como ser para el otro” (Levinas, 2016, p. 344). Es decir, el sujeto no es un ser aislado que adquiere la capacidad de coexistir con otros con posterioridad a su posición en el ser, ni es un momento del proceso dialéctico. El sujeto en sí mismo es bondad y apertura a la trascendencia. Apertura que se da en la proximidad del cara-a-cara y por medio del lenguaje. Por eso llega a afirmar que “la esencia del lenguaje es bondad (...) amistad y hospitalidad” (p. 345). En su obra posterior *De otro modo que ser o más allá de la esencia*, Levinas desarrolla y profundiza la estructura lingüística de la subjetividad, y describe “el modo en que el sujeto se singulariza como sí-mismo a partir de su respuesta al Otro, ante quien está expuesto” (Garrido-Maturano, 1994, p. 68). El sujeto se constituye como tal

lingüísticamente, como respuesta responsable a la acusación del Rostro. Es un diálogo original en donde el lenguaje no es conceptualizado, sino que permanece como proximidad, como *acercamiento* del/al otro: “un diálogo previo mantiene al Yo, más bien que el Yo mantiene una conversación (...) hay en la trascendencia del lenguaje una relación que no es palabra empírica, sino responsabilidad” (Levinas, 2003, p. 190).

Cuando Levinas caracteriza al lenguaje como amistad, lo que busca resaltar es el carácter ético del mismo. El lenguaje original es proximidad que respeta la alteridad, que no sincroniza ni diluye la diferencia, de la misma forma que la amistad reconoce y valora la diversidad del otro. La amistad testimonia un no-poder en el sujeto, el cual radica en la imposibilidad de reducir e igualar el Otro al Mismo, de anular la diversidad del amigo a pesar de la cercanía: “En su ser él mismo no permite de ningún modo que lo alcance por la sincronía de mi capacidad (...) sentido de la amistad que supera todo mi poder sincronizador” (Casper, 2007, p. 170). Este no-poder que caracteriza a la amistad, la cual queda preservada de la lógica del yo dominador, caracteriza al espacio-tiempo del diálogo original. El sujeto permanece como Decir responsable no tematizado en un Dicho; pura pasividad de quien es acusado en la proximidad del Rostro. El no-poder del sujeto lingüístico funda una sociedad en donde la paz es posible como respeto de la diversidad: “La unidad de la pluralidad es la paz, y no la coherencia de los elementos que constituyen la pluralidad” (Levinas, 2016, p. 345). El diálogo original es diaconía, ya que está al servicio y preserva la significación ética de la proximidad, haciendo posible una sociedad plural en el sentido fuerte del término.

Además de la amistad, Levinas utiliza el término *hospitalidad* para describir al diálogo original como diaconía. El lenguaje de la proximidad es acogida y recibimiento del Otro, de aquél que irrumpe como extranjero: “acontecimiento metafísico de la trascendencia – la acogida del otro, la hospitalidad; Deseo y lenguaje” (p. 288). La relación con el Otro es un dar-se, un *desinteresarse* de sí para cuidar del otro. Pero el sentido del dar no es la donación consentida de algo externo y superfluo al sujeto, sino “un arrancarse a sí mismo a su pesar (...) arrancarse a la complacencia de sí mismo en el gozo, como arrancar el pan de su boca” (Levinas, 2003, p. 132). Es un abrir las puertas de la propia casa y compartir el mismo pan³: “apertura, mucho más que del portamonedas, de las puertas de su casa” (p. 133). La revalorización del lenguaje como hospitalidad que propone Levinas, contrasta con el miedo al extranjero que caracteriza a la modernidad, la cual considera toda alteridad como barbarie. Este miedo a la alteridad se fundamenta

3 Levinas hace referencia, en este contexto, a pasajes proféticos del AT como el cap. 58 del libro de Isaías, en los cuales se interpela a la compasión y a la misericordia como ayuno agradable a Dios: “compartir tu pan con el hambriento, acoger en tu hogar a los sin techos” (v.7).

en una visión ontológica y unitaria de la realidad, en donde la multiplicidad es considerada peyorativamente. Racionalidad de la unidad como uniformidad que produce, inevitablemente, violencia: “La *civilización*, en su odio y temor por el Otro, amparada bajo la supremacía del Ser y la ipseidad incuestionadas, engendra la violencia y el exterminio” (Rabinovich, 2001, p. 31). El Decir de la proximidad es acogida del Otro y hospitalidad, que rompe con la totalidad de la ontología al ser una relación ética, dando así lugar al pluralismo en la sociedad.

El sujeto es elegido en la proximidad para ser responsable. Elección que lo convierte en único, no pudiendo escapar ni transferir dicha responsabilidad. Es inquietud y puesta en cuestión del reposo en el ser que se transforma, al acoger al Otro, en servicio como des-inter-es: “El Yo es, en su misma posición y de un extremo al otro, responsabilidad o diaconía” (Levinas, 2001d, p. 63). Esta dimensión ética de la subjetividad ya es anunciada por Levinas en el prefacio de su obra *Totalidad e infinito* (2016): “Este libro presentará a la subjetividad como acogiendo al Otro, como hospitalidad” (p. 20). Su nueva concepción ética y lingüística del sujeto, hace posible el surgimiento de una política que no deje de ser ética, que no anule la significación de la proximidad: “La hospitalidad no es solo una categoría ética en el pensamiento levinasiano, sino también el vehículo o, más apropiadamente, el pasaje para ingresar al campo político” (Makris, 2018, p. 82). Levinas, a lo largo de su vida se opone a la dinámica totalizante de los regímenes políticos de su tiempo, dando primacía a la ética frente a la política. En la noción de hospitalidad se vislumbra cierta apertura a una “política ética” (p. 83, referencia a Anya Topolski). Más allá de la intención de Levinas, es cierto que muchos filósofos posteriores se basaron en su ética para elaborar una política de la alteridad y la hospitalidad. Enrique Dussel (2007), así como otros filósofos latinoamericanos, asumen en parte el pensamiento levinasiano y lo aplican a la filosofía de la liberación. Asimismo, es Jacques Derrida el que, especialmente, busca continuar la ética levinasiana, desarrollando una política de la amistad y la hospitalidad (1998; 2001). Para Levinas, la hospitalidad es primera (Derrida, 2001, p.50), lo que implica que es constitutiva de la subjetividad dialógica. La irrupción sensible del Otro, que se da lingüísticamente, constituye al sujeto como rehén; es el carácter inspirado del sujeto, como lo Otro en lo mismo, que hace posible una “política profética” (Philip Harold, 2009).

CONCLUSIÓN

Recapitulando, Levinas propone una subjetividad lingüístico-dialógica que emerge en la proximidad de la relación intersubjetiva y que es corporal. El espacio de la intersubjetividad se desformaliza al curvarse por la asimetría de la relación,

produciendo así un espacio ético, asimétrico y desigual. Espacialidad ética que permanece como tal al fundarse lingüísticamente. El llamado sensible del Otro y la respuesta responsable se articulan como un diálogo original que es diaconía. Diálogo de la proximidad que mantiene la trascendencia de la relación y que se transforma en amistad y hospitalidad. De esta manera, la propuesta levinasiana busca romper con una visión de lo humano a partir del reposo en el ser y del poder, planteando una subjetividad que, al cimentarse en el lenguaje y en la ética, es capaz de respetar radicalmente la diversidad haciendo posible una sociedad verdaderamente pluralista, fraterna y democrática. El sujeto lingüístico resiste éticamente a la totalización y a la funcionalización. Resistencia del cuidado del Otro que, a pesar de la indiferencia y la violencia generalizada, persevera: “El mundo nada puede contra un hombre que canta en la miseria” (Sabato, 2018, p. 149). Es en el diálogo original de la proximidad en donde la resistencia a la totalidad encuentra su raíz y su fuerza, transformándose en canto de liberación, como respuesta responsable al llamado del Otro.

REFERENCIAS BIBLIOGRÁFICAS

- BUBER, M., *Yo y Tú*. Trad. Horacio Crespo. Buenos Aires: Nueva Visión, 1982.
- CASPER, B., *Pensar de cara a otro. Elementos del pensamiento de Emmanuel Levinas*. Trad. Diego Fonti. Córdoba: EDUCC, 2007.
- CEREZO GALÁN, P., *Reivindicación del diálogo*. Madrid: Real academia de ciencias morales y políticas, 1997.
- DERRIDA, J., *Políticas de la amistad*. Trad. Patricio Peñalvert. Madrid: Trotta, 1998.
- , *¡Palabra! Instantáneas filosóficas*. Trad. C. de Peretti y P. Vidarte. Madrid: Trotta, 2001.
- DI GIACOMO, M., “Caricia, alteridad y trascendencia en el pensamiento de Emmanuel Levinas”. *Apuntes filosóficos*, 48/25, 2016, 46-68.
- DRABINSKI, J., *Sensibility and Singularity. The problem of Phenomenology in Levinas*. Albany: State University of New York, 2001.
- DUHAMEL, V., “Dissolution de la temporalité et temporalité de la dissolution”. *Philosophy Today*, 55, 2011, 86-91.
- DUSSEL, E., “Lo político en Levinas (Hacia una filosofía política crítica)”. *Materiales para una filosofía de la liberación*. México D. F.: Plaza y Valdés editores, 2007.
- GARRIDO-MATURANO, A., “Emmanuel Levinas: continuidad de una perspectiva”. *Revista de filosofía*, 43/44, 1994, 63-74.
- , “La intrusión del estar. Análisis hermenéutico de la significación del espacio para la noción levinasiana de instante”. *Eidos*, 27, 2017, 224-250.

- GIBU, R., "Sensibilidad, corporeidad y significación en Levinas". En: XOLOCOTZI, A.; GIBU, R. (Coords.). *Fenomenología del cuerpo y hermenéutica de la corporeidad*. Madrid: Plaza y Valdés editores, 2014.
- GROLLO, S., "La sensibilità di là dal tempo. Passività e affezione nel pensiero di Levinas". *Discipline Filosofiche*, XXIV, 2014, 75-96.
- HAROLD, P., *Prophetic Politics. Emmanuel Levinas and the Sanctification of suffering*. Ohio: Ohio University Press, 2009.
- JOLDERSMA, C., *A Levinasian ethics for education's commonplaces. Between Calling and Inspiration*. New York: Palgrave Macmillan, 2014. DOI: 10.1057/9781137415493.
- LEVINAS, E., *El Tiempo y el Otro*. Trad. José Luis Pardo. Barcelona: Paidós, 1993.
- , *De la existencia al existente*. Trad. Patricio Peñalver. Madrid: Arena Libros, 2000a.
- , *Ética e infinito*. Trad. Jesús María Ayuso Díez. Madrid: La balsa de la medusa, 2000b.
- , "Algunas reflexiones sobre la filosofía del hitlerismo". Trad. Tania Checchi y Gabriel Aranzueque. *Cuaderno Gris, Época III*, 5, 2001a, 161-167.
- , *De Dios que viene a la idea*. Trad. Graciano González y Jesús María Ayuso Díez. Madrid: Caparrós, 2001b.
- , *Entre nosotros. Ensayos para pensar en otro*. Trad. José Luis Pardo. Valencia: Pretextos, 2001c.
- , *La huella del otro*. Trad. Esther Cohen, Silvana Rabinovich y Manrico Montero. México D.F.: Taurus, 2001d.
- , *Fuera del sujeto*. Trad. Roberto Ranz Torrejón y Cristina Jarillot Rodal. Madrid: Caparrós, 2002.
- , *De otro modo que ser o más allá de la esencia*. 4ª ed. Trad. Antonio Pintor. Salamanca: Sigueme, 2003.
- , *Dios, la muerte y el tiempo*. Trad. María Luisa Rodríguez Tapia. Madrid: Cátedra, 2005.
- , *Humanismo del Otro hombre*. Trad. Daniel Enrique Guillot. México: Siglo XXI editores, 2009.
- , *Alteridad y trascendencia*. Trad. Miguel Lancho. Madrid: Arena libros, 2014.
- , *Totalidad e infinito*. 3ª ed. Trad. Miguel García-Baró. Salamanca: Ediciones Sigueme, 2016.
- LLORENTE, J., "De fosforescencias que configuran mundos: Ser, alteridad y lenguaje en las conferencias inéditas de Levinas en el Collège philosophique". *Ápeiron. Estudios de Filosofía*, (1), 2014, 207-241.
- MAKRIS, S., "Emmanuel Levinas on hospitality: ethical and political aspects". *International Journal of Theology, Philosophy and Science*, 2/2, 2018, 79-96. DOI: 10.26520/ijtps.2018.2.2.79-96.
- MARCEL, G., *Diario metafísico (1928-1933)*. Trad. F. del Hoyo. Madrid: Guadarrama, 1969.

- MORENO MÁRQUEZ, C., "The curvature of intersubjective space: sociality and responsibility in the thought of Emmanuel Levinas". *Analecta Husserliana*, XXII, 1987, 343-352.
- PEÑALVERT, P., *Argumento de Alteridad. La hipérbole metafísica de Emmanuel Levinas*. Madrid: Caparrós, 2000.
- PETROSINO, S., "Levinas e l'essenza del linguaggio". *Idee, Rivista di filosofia*, 33, 1996, 15-38.
- RABINOVICH, S., "Levinas: un pensador de la excedencia". LEVINAS, E., *La huella del otro*. México D.F.: Taurus, 2001, 9-44.
- RIVA, F., "Ética como sociabilidad. Buber, Marcel y Levinas". *Anuario filosófico* XXXVIII/2, 2005, 633-655.
- SABATO, E., *La resistencia*. 14ª ed. Buenos Aires: Booket, 2018.
- VERANO GAMBOA, L., "El cuerpo y el Otro en Merleau-Ponty y Levinas". *Rev. Filos. Aurora*, 30, 50, 2018, 379-396. DOI: 10.7213/1980-5934.30.050.AO02
- VIOLA, F., "La reiteración del inicio. Aportes para una nueva concepción del tiempo a partir de la filosofía de Emmanuel Levinas". *Franciscanum*, 165, LVIII, 2016, 119-143.

NOTAS CRÍTICAS

EDICIÓN DE LOS ESCRITOS ZUBIRIANOS

EDITION OF ZUBIRI'S WRITINGS

X. ZUBIRI, *Ciencia y realidad (1945-1946)*. Presentación y edición de E. Vargas Abarzúa. Madrid: Alianza, 2020, XVI+867 pp.

X. ZUBIRI, *Filosofía primera (1952-1953) Volumen I*. Presentación y edición de E. Vargas Abarzúa. Madrid: Alianza, 2021, XI+692 pp.

ANTONIO PINTOR-RAMOS

Doctor en Filosofía
Universidad Pontificia
Salamanca /España
apintorra@hotmail.com
ORCID: 0000-0002-6289-2855

Recibida: 28/04/2021

Aceptada: 6/09/2021

Resumen. Esta nota pretende informar al lector del estado actual de la edición de los escritos de Zubiri, tomando como referencia las últimas publicaciones. Se puede observar que los criterios de edición se han ido perfeccionando y, al mismo tiempo, se han abierto caminos que inicialmente no se habían tenido en cuenta. No obstante, no todo lo publicado tiene el mismo alcance sistemático ni ofrece la misma madurez filosófica.

Palabras clave: Inéditos, publicaciones, Zubiri.

Abstract. This note aims at informing the reader about the current state of the edition of Zubiri's writings, taking the latest publications as a reference. It can be observed that the publishing criteria have been perfected and, at the same time, paths that had not been taken into account initially have been opened. Nevertheless, not all the works that have been published have the same systematic significance nor do they show the same philosophical maturity.

Keywords: unpublished, publications, Zubiri.

Las dos publicaciones que encabezan esta nota, entre las cuales media un breve lapso de tiempo, ofrecen una ocasión para informar al lector de los criterios y los materiales que conforman actualmente este ya abultado *Corpus* zubiriano, que probablemente ofrecerá al principiante más de una incógnita. Los criterios que se han seguido en la publicación de inéditos de Zubiri, desde el primero en 1984 a cargo de I. Ellacuría, quizá mantienen constantes importantes, pero parece claro que se han ido puliendo y adaptando a medida que los conocimientos y el mismo desarrollo de la investigación lo aconsejaban. De hecho, hoy bajo la autoría de Zubiri y en el mismo sello editorial coexisten productos de alcance diverso, sin que ningún distintivo externo advierta de ello, aunque es cierto que las “presentaciones” que inician cada volumen suelen dar cuenta del alcance de cada texto. Me ceñiré aquí fundamentalmente a lo publicado en lo que va de siglo XXI cuando una nueva generación de investigadores ha ido tomando el relevo con otros intereses filosóficos y menos condicionamientos (para bien y para mal) pues ya no conocieron a Zubiri. Al frente de ellos y como Director de Publicaciones, está una persona tan competente como Antonio González.

El lector tendrá una idea de la necesidad de esta clarificación si se percata de que en este momento puede haber dudas razonables sobre el número de volúmenes que componen el legado de Zubiri¹ Declarados ya obsoletos *El problema filosófico de la historia de las religiones* (1993) y *El problema teológico del hombre: Cristianismo* (1997), no porque lo allí expuesto sea falso sino porque su contenido se ha publicado con mayor rigor, refundido íntegramente *El hombre y Dios* (1984), está en una especie de cuarentena el volumen *Sobre el hombre*, elaborado por I. Ellacuría en 1986. Esta “cuarentena” presumiblemente va a ser larga porque se trata de un libro que presenta una dificultad única y no fácil de resolver, al estar compuesto dentro de un esquema con pretensiones sistemáticas por fragmentos de cursos de fechas distintas y sin que sepamos a ciencia cierta el grado de “actualización” (a la altura de 1986, como es obvio) a que fueron sometidos; de hecho, coexisten allí textos de 1953 con el último que Zubiri dejó preparado en 1983 pocos meses antes de su muerte; en teoría ese volumen quedará obsoleto cuando se publiquen completos los cursos de los que están entresacados sus capítulos y, al mismo tiempo, se encuentre su lugar para el texto de 1983 “Génesis de la realidad humana” (actual cap. VIII); ya se ve que la tarea es a muy largo plazo en el mejor de los escenarios y, mientras tanto, con todas las precauciones críticas habrá que seguir acudiendo a ese texto.

En lo que sigue, me voy a ceñir a tres ámbitos en los que creo que existen novedades importantes. Entiendo que el proyecto primero y urgente era revisar

1 En el texto me limitaré a indicar títulos y fechas. La “bibliografía final” recoge las ediciones aludidas

las pocas obras publicadas por Zubiri a veces en ediciones descuidadas ofreciendo textos limpios y seguros (labor concluida, aunque siempre perfectible) y, al mismo tiempo, publicar los cursos posteriores a *Sobre la esencia* porque se presumía que eran los más maduros y a los que el propio Zubiri había aludido como más fácilmente publicables, aunque a primera vista ya aparecían como desiguales y tampoco había suficiente claridad sobre el alcance de la filosofía de Zubiri. Hay que tener en cuenta, sin embargo, que él no se decidió a publicar ninguno de esos cursos, a pesar de que muchos de ellos conservan huellas evidentes de haber sido manejados y corregidos. Esta fue la tarea que, con algún altibajo, se realizó básicamente bajo la guía de Diego Gracia. No son ni pretenden ser ediciones “críticas” en sentido estricto, pero sí textos limpios y fiables, cosa que en general se consiguió; es posible que algún título hoy necesite ser revisado, también que algún texto necesite ser recolocado en un lugar más idóneo.

A partir de aquí, se afianzó un nuevo plantel de investigadores, ya avezados y acostumbrados tanto a los usos zubirianos como a sus difícilísimas anotaciones, entre los que destacan el propio Antonio González, Manuel Mazón y el profesor chileno Esteban Vargas Abarzúa. El primer camino que exploraron fue rehacer volúmenes importantes que en su primera edición no respondían adecuadamente a las necesidades del estudioso ni a las posibilidades abiertas después de que el Archivo de Zubiri y su biblioteca personal fuesen ordenados y catalogados por Germán Marquínez Argote. El contenido de ese Archivo es conocido², es accesible digitalmente a los investigadores, aunque su manejo no siempre es fácil, aspecto que debe tenerse en cuenta para juzgar luego algunas decisiones adoptadas. En esta situación se encontraba *Espacio. Tiempo. Materia* (1996), reelaborado en una nueva edición de 2008 al localizarse un texto original de Zubiri, que substituía con ventaja a la redacción procedente de I. Ellacuría sobre “la materia”. La gran refundición, sin embargo, afecta al conjunto de los que podríamos llamar escritos teológico-religiosos. En primer lugar, se hace una revisión total del difundido texto *El hombre y Dios*, el primer texto póstumo que I. Ellacuría había editado en 1984, y que E. Vargas Abarzúa recompone editando el curso de Roma (1973) y separándolo, a su vez, de lo que llama “Redacción final” (fecha indicativa de 1983), de hecho una revisión a fondo del texto de I. Ellacuría sin las “adiciones” que este había tomado directamente de textos anteriores; el cambio, a simple vista, es muy llamativo porque se pasó de un volumen medio de 385 a otro imponente de 625 páginas, sin que merme su accesibilidad al lector no especialista, que puede limitarse a la redacción publicada en primer lugar. Esto gana, a su vez, una contextualización impagable cuando se puede ver como una ampliación

2 Véase su catálogo, por ejemplo, en la biografía de referencia: J. COROMINAS-J.A. VICENS, *Xavier Zubiri*, pp. 843-855.

monográfica derivada del gran curso de 1971, el último de los cursos “largos” de Zubiri, publicado como *El problema teológico del hombre: Dios, religión, cristianismo* sólo en 2015, también por E. Vargas Abarzúa. A su vez, Antonio González editó los cursos preparativos de 1965 (uno en Madrid y otro en Barcelona) y de 1968 bajo el título general *Sobre la religión* (2017), con alguna sorpresa por lo que respecta a la evolución intelectual de Zubiri y tornando obsoletos los volúmenes mencionados más arriba, que ya no se reimprimirán pero de los que todavía es posible encontrar algunos ejemplares. La temática la completan E. Colomer Espinoza y E. Vargas Abarzúa con la publicación de un libro nuevo y peculiar dentro de la producción de Zubiri: *Reflexiones filosóficas sobre algunos problemas de teología* (2019). Se trata de ediciones en las que se ha extremado el cuidado en la fidelidad a los textos existentes y a su disposición formal.

El segundo objetivo, inicialmente no previsto y presumiblemente ya completo, son los cursos universitarios de Zubiri. En este apartado están no sólo los cuatro nutridos volúmenes titulados *Cursos universitarios* (2007-2014), en total más de dos mil páginas, preparados por M. Mazón, sino también el volumen *Introducción a la filosofía de los griegos* (2017), preparado por J. Corominas y J. A. Vicens y que se presenta como “Lecciones en Barcelona, Ameixenda y Madrid (1941-1943)”, aunque lo fundamental de este importante texto data de la última fecha señalada. Ya me referí en esta misma revista al inicio de esta labor que, con excepciones poco significativas, no comprende textos directos escritos ni hablados por Zubiri, sino que parte de apuntes tomados en sus clases por oyentes, la mayoría de ellos seguidores fieles y entusiastas de Zubiri que no eran alumnos matriculados en sus cursos (la excepción aquí es un joven Julián Marías), hábiles en tomar notas del veloz y denso discurso zubiriano y que, a todas luces, sometieron esas notas a un proceso posterior de reconstrucción y reelaboración personales, como demuestra el hecho de colecciones de apuntes sobre el mismo curso con una estructura similar, pero con peculiaridades características en cada caso, que el editor terminó conociendo a la perfección. Los seguidores y redactores más asiduos son un competente Luis Felipe Vivanco, Carmen Castro que luego será su esposa, y una siempre rigurosa María Araujo. El editor es muy escrupuloso en el manejo de estos apuntes y, tomando siempre como guía el que en cada caso le parece más adecuado, anota las variantes de otros que le parecen de interés; esto es muy útil para el investigador, pero hace menos fluida la lectura por las interrupciones y las repeticiones inevitables. No todos los cursos llegaron en el mismo estado, algunos están incompletos, de otros sólo se conservan fichas escuetas, pero el lector dispone de toda la información pertinente. A lo largo de este proceso de publicación se produjo un hallazgo inesperado: las intervenciones de Zubiri en 1933 y 1934 dentro del marco de Universidad Internacional de verano de Santander. Del estudio de todo este material se deducen algunos resultados:

1) La actividad docente de Zubiri no consistió nunca en la explicación de un curso completo de Historia de la Filosofía, como se denominaba su cátedra; esto se daba por supuesto, se puede sospechar qué manuales recomendaba y, suponiendo esa base de manera muy optimista, cada año se dedicaba atención amplia y monográfica a cuestiones que coincidían con sus intereses intelectuales en ese momento, como lo muestra la insistencia en un grupo de temas típicos de esta etapa. 2) Esos cursos son el laboratorio del pensamiento de Zubiri en una fase de gran efervescencia y cristalizan en una serie de publicaciones importantes, cuya referencia será la publicación muy revisada de la mayoría de ellos en *Naturaleza, Historia, Dios* (1944), que seguramente sigue siendo el libro más leído de Zubiri; no deja de resultar asombroso, por ejemplo, que en una época socialmente turbulenta como 1934-35 el curso básico de Historia de la Filosofía en la Universidad Central sea el estupendo “Helenismo y Cristianismo”, cuyo texto guía sin ninguna concesión es la epístola de San Pablo a los Efesios. 3) A su vez, se abre un interrogante, no resuelto, respecto al tiempo que podía sostenerse esa identidad entre investigación y docencia, teniendo en cuenta que la primera lógicamente va progresando, mientras que en la docencia hay que comenzar cada año desde cero; Zubiri no estuvo tiempo suficiente en la universidad para que tengamos una respuesta segura a esto. Es decir, son textos preciosos para conocer el desarrollo intelectual de Zubiri, que, sin embargo, arrojan incertidumbres respecto a su eficacia pedagógica para sus alumnos naturales, cosa que no parece haber preocupado mucho a Zubiri.

El curso dedicado a la filosofía de los griegos merece una breve reflexión porque puede servir como referencia que condensa algunas de las características de este grupo de publicaciones. Aunque aparezca como volumen independiente dentro de la colección, puede reflejar parte de la docencia de Zubiri en Barcelona, pero de hecho parte de un intenso curso privado al final de esta etapa docente de Zubiri cuando está en pleno dominio de sus capacidades. No se trata de un tema de “Historia de la Filosofía” en sentido habitual, aunque conserve la forma literaria del “relato”; se trataría más bien de un relato de “segundo orden”, marcado y guiado por un interés teórico y sistemático. En el subsuelo pueden apreciarse referencias a problemas textuales y discusiones históricas; los editores no dejan de anotar que en aquellas fechas era casi inevitable que a este tema le alcanzase la polvareda levantada por la publicación del famoso libro de W. Nestle, *Vom Mythos zum Logos* (1940), del que existe un ejemplar en la biblioteca privada de Zubiri; sin embargo, su visión de los presocráticos (en detalle no sobrepasa ese período) no se distingue en nada básico de los trabajos anteriores, del mismo modo que W. Nestle sistematiza y afina la clásica interpretación positivista, que para los presocráticos aparecía ya en la obra de J. Burnet, de la que también tenía un ejemplar en su biblioteca. Zubiri, como tantos filósofos sistemáticos, construye una concepción “orgánica” de la filosofía presocrática (p. 162) en la que los

nombres individuales (conocidos precariamente, como es sabido) están en función de una especie de magno programa filosófico, que en todas las etapas de Zubiri tiene su cima en Aristóteles –“el maestro de los maestros”-, e incluso de forma más específica en lo que conocemos como su *Metafísica*; por la filosofía helenística no mostró interés y el neoplatonismo le interesó en función de su utilización por el Cristianismo. Esto lleva al hecho sorprendente de tratar a Platón como una transición hacia Aristóteles, sin duda alguna muy importante; de hecho Zubiri sólo prestó atención directa al diálogo *Parménides* y nos restan unas notas de un curso breve sobre el *Fedón*; dicho con algo de crudeza, no parece que Platón fuese para Zubiri uno de los “imprescindibles” de la filosofía de todos los tiempos más que como puente necesario hacia Aristóteles, lo cual seguramente remite a contextos generacionales muy determinados. Pues bien, es llamativo que en este curso y para sistematizar los pasos de la primera filosofía griega Zubiri adopte como esquema la “alegoría” (es sabido que Platón no la llamó nunca “mito”) que constituye el que quizá es el texto más citado de la filosofía de todos los tiempos; por supuesto, en esa alegoría de la caverna a Zubiri no le preocupa su función concreta en Platón cuando aparece colocada precisamente en el libro VII de la *República*. Un hecho que podría parecer anecdótico da pie para ponderar el esmero que se ha puesto en el tratamiento de estos textos; Zubiri seguramente llevó el texto en griego y lo fue traduciendo sobre la marcha, pero ninguna de las fuentes recoge esa traducción, aunque sí el largo comentario y la curiosa disección en pasos; sobre ese comentario, Antonio González reelaboró la traducción de la alegoría de la caverna, no buscando exactitud filológica, sino un texto que pueda aproximarse al que Zubiri estableció en aquel momento.

¿Qué valor tienen estos textos indirectos para el pensamiento de Zubiri? Es obvio que ni son libros de Zubiri ni están contruidos a partir de la transcripción de sus textos, pero el hecho de que suela existir más de una fuente de apoyo les da seguridad en sus rasgos generales. Al menos la misma seguridad que en su momento pudo transmitir la edición de los cursos de Hegel en Berlín por parte de sus discípulos, mientras que los libros elaborados a partir de sus cursos sistemáticos se parecerían más a los cursos de Husserl o de Heidegger, aunque con la peculiaridad del importante papel de la “oralidad” en Zubiri, pues la primera plasmación de esos textos fue en cursos orales que se transcribieron a partir de documentos taquigráficos o de grabaciones magnetofónicas. ¿Está completo este apartado, como apuntaba más arriba? Materialmente, no del todo; Andrés Sánchez Pascual descubrió en la Biblioteca de Cataluña un volumen que contiene unos apuntes mecanografiados del primer curso que Zubiri explicó en Madrid y que, desde el 14 de febrero al 12 de mayo de 1928, está dedicado a la *Metafísica* de Aristóteles; no se ha podido identificar al autor de unos apuntes, que también parecen haber sido refundidos para su copia mecanográfica, y quizá hoy la publicación de esos

apuntes no tenga interés para un tema tratado tantas veces en la obra de Zubiri, pero al menos es un dato de interés biográfico.

A lo largo de estos años tuvo lugar un cambio importante. Los libros de Zubiri dejan de contener los índices de autores y de conceptos, siempre trabajosos y a veces muy nutridos, que eran una de sus señas de identidad. A partir de 1910 (*Acerca del mundo*), tales índices desaparecen, substituidos por índices electrónicos de consulta libre en la Fundación (www.zubiri.net) y, más recientemente, estos han sido refundidos en un *Index zubiriano*, que es un instrumento preciso y muy valioso para el investigador a medida que se vaya perfeccionando con su uso.

Queda un tercer camino, iniciado recientemente, que no significa ningún corte drástico con lo anterior, aunque creo que sí un replanteamiento de los criterios de edición. Si había un consenso respecto al hecho de que los primeros cursos que debían publicarse eran los posteriores a *Sobre la esencia*, la obra aceptaba por todos como plenamente “madura” en la trayectoria de Zubiri, es evidente que esa obra no surgió de la nada y es muy probable que en su gestación estén varios ensayos previos fracasados, es también indudable que Zubiri aprovechó muchos desarrollos de trabajos anteriores, pero asimismo es claro que el libro como tal no es resultado directo de ningún curso, como tampoco lo será en otras circunstancias *Inteligencia sentiente*. Es cierto también que ese criterio restrictivo, basado en indicaciones del propio Zubiri, se había entendido en sentido amplio; así, en 1992 Diego Gracia había editado con el título *Sobre el sentimiento y la volición*, tres cursos breves de temática emparentada, de los cuales sólo el último cumple con ese criterio cronológico. Por lo demás, la colocación de algunos textos, de extensión y elaboración muy desigual, a veces tuvo que buscar cobijo circunstancial en publicaciones que quizá no eran su lugar más adecuado; no es un problema grave porque los distintos editores en general dan cuenta de sus criterios de edición y suelen ser muy sobrios ciñéndose a aspectos textuales. Hasta donde a mí se me alcanza, no hay ningún impedimento legal para que se publiquen los inéditos que parezca conveniente pues, fallecida Carmen Castro Madinaveitia en 1997, la propietaria de los derechos legales es la Fundación Xavier Zubiri. Por lo demás, el caso extremo de Bergson en este punto puede ilustrar bien que la “voluntad” de un autor muerto es legalmente poco relevante, cualquiera que sea la “autoridad” que se le conceda.

De hecho, ese acuerdo tácito se desbordó con la publicación a cargo de Antonio González, el Director de Publicaciones de la Fundación, del curso *Acerca del mundo (1960)*, que apareció en 2010. Es evidente que ese curso ofrece problemas respecto a las publicaciones que le precedieron: hay un recurso abundante a teorías físicas y doctrinas cosmológicas que hoy están desfasadas y no son meros aditamentos, sino que, por ejemplo, sirven de base para una insólita reformulación de un argumento cosmológico de la existencia de Dios, entiendo por mi parte que se trata de algo rechazado de raíz una década más tarde.

Aunque el problema de fondo sea el mismo, los otros cursos de los años cuarenta y cincuenta añaden a la dificultad de sus esperables desfases su extensión. Si se observa el catálogo de cursos y su estado de conservación, ofrecido en la “presentación” de *Ciencia y realidad* con todo rigor y detalle (pp. VII-VIII), hay al menos cuatro cuya transcripción en cada caso se aproxima a los dos millares de páginas, algo que no alcanzó ni de lejos ninguna de las obras anteriormente publicadas o inéditas; de ellos, parece a simple vista que el curso largo “El problema de Dios” (1948-1949) estará ahora obsoleto por la reiteración del tema en tratamientos posteriores. La cuestión no es si estos cursos tienen interés para los investigadores porque es evidente que para un investigador todo puede tener interés; como esos cursos ya son accesibles en formato digital (cierto es que no en un texto depurado por un editor tan riguroso como el de estos volúmenes), la cuestión es si resulta razonable su edición en papel y, dando por hecha una respuesta básicamente afirmativa, qué función pueden cumplir en el aprendizaje y la investigación de Zubiri. Lo más probable es que, también en este tema, termine decidiendo el principio de realidad, que en este caso significa cuántas de estas publicaciones puede asumir el mercado del libro en un momento que no es precisamente de gran euforia. Mientras tanto, se siguen *seleccionando* cursos.

Se pueden invocar diversas razones para comenzar esta nueva fase con el curso largo *Ciencia y realidad* (CR). Las hay externas: si, como es sabido, Zubiri señaló que en 1945 comenzó una nueva “etapa” (término al que sobrecargó de significado) en su trayectoria filosófica y esa etapa fue la definitiva de su “madurez”, CR significaría el primer mojón de ese largo camino; además, ya en el título figura el término “realidad”, que será el tema central de su etapa madura (aunque el título coincida exactamente con el de un capítulo de *Naturaleza, Historia, Dios*). Pero, sobre todo, hay razones internas; el tema de la realidad se coloca como central en un esbozo de nueva *metafísica*, que significa ahora el abandono de cualquier forma de *ontología*. Sin embargo, la mayor parte del libro está dedicado al análisis de cosas reales agrupadas en distintos órdenes: físicas, matemáticas, biológicas y humanas, conectándolas con todo el trabajo de las ciencias correspondientes; digamos de paso que aquí se encuentra el tratamiento más extenso de las realidades matemáticas existente hasta ahora en la obra de Zubiri (pp. 237-330), aunque habrá que analizarlo a la luz del “epílogo” dedicado al mismo tema en *Inteligencia y logos*. Después de la crisis de fundamentos de las ciencias y de las grandes revoluciones en el ámbito de la física, el positivismo decimonónico mostraba abultadas grietas por las que se introdujeron teorías epistemológicas logicistas y convencionalistas, que encerraban el antiintuitivo conocimiento científico en una construcción endogámica o autorreferencial. Frente a ello y con sus mismas armas, Zubiri propugna que los desafíos que lanzan las cosas reales a la inteligencia son el origen y también la “prueba” definitiva de todo saber científico.

Aunque las cosas reales, casi por definición, muestran ya una dimensión de alteridad, queda un largo trecho para hacer de “realidad” un neologismo semántico y verla como mera “formalidad” en el darse las cosas a la inteligencia como “de suyo”, frente al carácter estímulo de los signos en el mero sentir. Ya se puede columbrar que el desafío es que se analicen estos escritos sin proyectar sobre ellos elaboraciones muy posteriores, sin dejarse seducir por los cantos de sirena de una terminología permanente, aún no del todo afinada.

Por su parte, *Filosofía primera* (FM) tiene ya su pequeña historia en la investigación zubiriana, en parte facilitada por un velo nebuloso que lo señala como el modelo de los cursos largos de esta etapa. El mismo título, de raigambre aristotélica como es de sobras conocido, había servido como recurso en filósofos de distintas épocas (el último caso fue probablemente Husserl en el curso en Friburgo de 1923-24, que en este momento con toda probabilidad Zubiri desconocía), es entendido ahora como “la justificación intelectual de los principios firmes y absolutos de las cosas” (p. 23) y a estas alturas ofrecía nominalmente la ventaja de un título más neutro para lo mismo que desde Andrónico se llama “metafísica”. De este curso era conocido el título y una vaga idea de su contenido, se habían hecho en él algunas calas movidas por el hecho de que en él aparece ya naturalizada y desarrollada la expresión “inteligencia sentiente”. En este contexto y poco después de la muerte de Zubiri, cuando nos afanábamos por asimilar el desafío intelectual de su última trilogía. D. Gracia en una obra decisiva³ recurrió a esta expresión para diferenciar *metodológicamente* lo que era descripción de lo dado y reservó “metafísica” para lo que era una construcción de la razón a la búsqueda de un fundamento de eso dado; se trataba de una distinción didáctica con alto grado de convencionalidad y sin apoyo textual; no se trataba, no podía tratarse de oponer “filosofía primera” a “filosofías segundas”, sino de discernir conforme a los criterios internos de aquella trilogía las pretensiones de verdad que podía reclamar cada modalidad intelectual. El texto que ahora se publica seguramente hará las delicias de muchos por el largo desarrollo de la inteligencia sentiente (toda la I Parte), pero, sin entrar en análisis que no son de este lugar, es fácil advertir que ese término está apoyado ahora en una antropología dualista y la propia expresión se refiere a una *facultad* híbrida, de composición dual a partir de dos potencias heterogéneas.

Así es fácil comprender que, una vez publicado el curso CR por su carácter pionero en esa nueva etapa, se acometa la publicación de FP porque, además, es quizá el más revisado, anotado y prolongado entre todos los cursos de Zubiri; el resultado de ello es que también resultará el más largo. El editor distribuye el

3 D. GRACIA, *Voluntad de verdad*.

material en tres volúmenes, algo perfectamente creíble porque es una de las personas más expertas en inéditos zubirianos; cuando se publiquen los dos volúmenes que restan y aún previendo que el tercero se reservará para agrupar los distintos complementos, vamos a encontrarnos con cerca de dos mil páginas impresas. Sospecho que el comportamiento del mercado del libro ante esta publicación va a ser decisivo para el destino final de los cursos restantes. Asimismo, en una perspectiva más lejana, es probable que la sucesiva publicación de materiales obligue a un cribado más fino que la comodidad, propiciada por unas expresiones de la ancianidad del autor, de una larga etapa de “madurez” en la obra de Zubiri extendida a lo largo de cuarenta años; ya hay ensayos muy prometedores que apuntan en esta línea.

Estas breves indicaciones dejan claro que ni estas nuevas publicaciones, ni previsiblemente las que puedan seguir las, substituyen las grandes obras publicadas por Zubiri, tampoco ofrecen atajos para esquivar la dureza de algunos de aquellos textos, que seguirán siendo los pilares de una filosofía que, sin esas referencias tantas veces abstrusas y áridas, se diluiría en análisis parciales sin profundidad ni faro orientador.

Quiero concluir esta nota llamando la atención de nuevo sobre el perfeccionamiento de los criterios de edición y el escrupuloso manejo de los textos, sin escatimar esfuerzos ante dificultades en detalles que muchos lectores no apreciarán; la investigación sobre Zubiri se apoyará ahora sobre un suelo más sólido. El trabajo de presentación formal por parte de la editorial es impecable, como es habitual; solamente, como anécdota, habrá que ver cómo resiste el estudio intensivo este tipo de encuadernación tratándose de obras tan voluminosas y que, por lo demás, tienen un precio alto en el mercado.

REFERENCIAS BIBLIOGRÁFICAS

- J. COROMINAS-J. A. VICENS, *Xavier Zubiri. La soledad sonora*. Madrid: Taurus, 2006
- D. GRACIA, *Voluntad de verdad. Para leer a Zubiri*. Barcelona: Labor, 1986 (2 ed., Madrid: Triacastela 2008).
- A. PINTOR-RAMOS, Crítica a X. ZUBIRI, *Cursos Universitarios, I. Cuadernos salmantinos de Filosofía*, 35, 2008, 508-511.
- X. ZUBIRI, *Sobre el hombre*. Madrid: Alianza, 1986.
- , *Sobre el sentimiento y la volición*. Madrid: Alianza, 1992.
- , *El problema filosófico de la historia de las religiones*. Madrid: Alianza, 1993.
- , *El problema teológico del hombre: Cristianismo*. Madrid, Alianza, 1997.
- , *Espacio. Tiempo. Materia*. Nueva ed., Madrid: Alianza, 2006.
- , *El hombre y Dios*. Nueva ed., Madrid: Alianza, 2012.

- , *Cursos universitarios*. Vols. I-IV, Madrid: Alianza, 2007-2014.
- , *Acerca del mundo (1960)*. Madrid: Alianza, 2010.
- , *El problema teológico del hombre: Dios, Religión, Cristianismo*. Madrid: Alianza, 2015.
- , *Sobre la religión*. Madrid: Alianza, 2017.
- , *Introducción a la filosofía de los griegos*. Madrid: Alianza, 2018.
- , *Reflexiones filosóficas sobre algunos problemas de teología*. Madrid: Alianza, 2019.
- , *Ciencia y realidad (1945-1946)*. Madrid: Alianza, 2020.
- , *Filosofía primera (1952-1953). Volumen I*. Madrid: Alianza, 2021.

RESEÑAS

SACRISTÁN, Manuel, Sobre Jean-Paul Sartre. Edición de Salvador López Arnal y José Sarrión Andaluz. Prensas de la Universidad de Zaragoza, 2021, 188 pp.

ARIANE AVIÑO MCCHESENEY

Doctora en Filosofía
arianependue@gmail.com
ORCID: 0000-0003-3575-9274

Recibida: 18/08/2021
Aceptada: 6/09/2021

La presente obra agrupa seis materiales del filósofo Manuel Sacristán sobre Jean-Paul Sartre, en un trabajo conjunto del filósofo y escritor Salvador López Arnal, reconocido experto en el pensamiento de Sacristán, y autor de numerosos ensayos, y el político y profesor de universidad José Sarrión Andaluz, autor de la obra *La noción de ciencia en Manuel Sacristán* (Dykinson, 2017). Los textos aquí reunidos fueron producidos entre los años 1958 y 1980, y abarcan materiales heterogéneos, como son: dos artículos cuyo destino era ser incluidos en una enciclopedia (los textos que se enumeran en el índice como I y III); una nota editorial (IV); el esquema inédito de una conferencia (II); y dos conferencias inéditas (V y VI). Se acompaña cada uno de los materiales de un gran aparato crítico consistente en abundantes notas complementarias. Este trabajo de notas de edición pone de manifiesto el profundo conocimiento que los responsables de este volumen poseen del pensamiento de Sacristán y de su contexto, pero también hace evidente la coincidencia en el compromiso ético y político de transformación social que definió la trayectoria de Sacristán, y que define la de Salvador López Arnal y la de José Sarrión Andaluz. El libro incluye un estudio preliminar a cargo de los dos autores mencionados, así como una introducción escrita por el profesor y experto en Jean-Paul Sartre, Juan Manuel Aragüés. En las últimas páginas podemos encontrar un índice analítico y onomástico.

Vista en conjunto, esta obra nos presenta una serie de análisis, reflexiones y críticas de Manuel Sacristán en torno a la relevante figura de Jean-Paul Sartre, que tenían, en el momento de su realización, la vocación de dar a conocer el pensamiento del filósofo francés en España. Los límites de estas aproximaciones

a Sartre vienen dados por el hecho, que señala Juan Manuel Aragüés, de que las obras póstumas de Sartre representaron toda una dislocación en la concepción del pensamiento sartreano, y Sacristán no tuvo la posibilidad de conocer estas obras, debido a lo prematuro de su fallecimiento en el año 1985. Dejando a un lado estas limitaciones, la obra continúa siendo profundamente relevante.

Advertimos, desde nuestro primer encuentro con esta obra, que no iba a ser posible permanecer cómodamente en su superficie, quizá porque la lectura de Sacristán siempre es una lectura exigente, incluso cuando nos enfrentamos a textos aligerados de aparato crítico, como son algunos de los que aparecen en esta edición, y que tienen un carácter fundamentalmente expositivo. No obstante, lo cierto es que la gravedad, el peso de esta obra, lo sentimos ya antes de enfrentarnos a los propios textos de Sacristán, con la lectura tanto de la introducción de Juan Manuel Aragüés, como del estudio preliminar de los dos responsables de la edición: Salvador López Arnal y José Sarrión Andaluz. La introducción y el estudio preliminar merecen, a nuestro entender, una atención especial por su claridad y por el profundo conocimiento y compromiso con la obra de los autores implicados, que es manifiesta en ambos casos. Los dos textos, la introducción centrada en Sartre y el estudio que se ocupa del propio Sacristán, parecen visitarse mutuamente, mostrando esta obra como lo que es: un caminar de difícil compás de dos autores imprescindibles. Y no nos referimos solamente a la dificultad de acompañar los ritmos de ambos filósofos en un caminar conjunto, sino, principalmente, a la complejidad que ambos autores encontraron y manifestaron, desde una honda honestidad, a la hora de adherirse a algunos de los ritmos de su tiempo. Ambas cuestiones aparecen perfectamente tratadas en la introducción y en el estudio, y nos guían en la lectura, a primera vista fragmentada, de Sacristán sobre Sartre. De hecho, la primera virtud que advertimos en los dos textos introductorios es que consiguen tejer seis materiales heterogéneos elaborados en el espacio de veintidós años, entre los que hay tanto textos escritos con un fin enciclopédico como transcripciones de conferencias. De este modo nos muestran, más allá de los fotogramas, una trama teórica y vital no solamente de Sacristán, sino también del propio Sartre.

El tono de esa trama creemos que no es otro que el desacomodo. Este desacomodo viene dado por la dificultad, como hemos dicho, de encajar las piezas de la teoría y de la praxis, pero también de encontrar una posición no impostada, verdadera, en la problemática cartografía tanto posbélica como posrevolucionaria. La aproximación de Sacristán a Sartre en los textos contenidos en esta obra hace patente que el movimiento telúrico que caracteriza el suelo de Sartre no le es en absoluto ajeno a Sacristán, y que el cielo tormentoso de los conceptos aún envueltos en las nubes del romanticismo anuncia, como en la canción de Dylan, que va a caer una fuerte lluvia.

Gran parte de los conflictos que atraviesa la filosofía desde la II Guerra Mundial, incluida la propia constatación de hasta qué punto la guerra definirá a los intelectuales, aparecen de una u otra forma en los textos que reúne este volumen. El posicionamiento teórico con respecto a Heidegger es uno de esos conflictos, y así lo recoge el primero de los textos de Sacristán, del año 1958, titulado *Corrientes filosóficas de la posguerra, existencialismo y corrientes afines*. El propio Sacristán, tal y como relatan Salvador López Arnal y José Sarrión Andaluz, experimentó en carne propia esa relación problemática con el pensamiento heideggeriano, objeto de estudio de su propia tesis doctoral. Otro de los conflictos característicos de una época y presente en el tercero de textos de este volumen, del año 1968, titulado *Corrientes principales del pensamiento filosófico. Jean-Paul Sartre: la noción del filosofar y la ética existencialista*, es la cuestión de la ideología. Sacristán aborda de forma crítica este conflicto a partir de su lectura particular de la ambición existencialista de Sartre, es decir, desde el difícil encaje del papel mediador del existencialismo con respecto al marxismo. También encontramos en este texto uno de los aspectos más oscuros de la aproximación de Sacristán a Sartre, y que tiene que ver con el problema de la ciencia. En este punto, Salvador López Arnal y José Sarrión Andaluz arrojan luz con sus apreciaciones sobre un posible equívoco conceptual en torno al cientificismo, que ayuda a la lectura de ese aspecto de la crítica de Sacristán a Sartre. Una vez más, constatamos el gran valor de los textos introductorios en este volumen, sin los cuales los materiales de Sacristán presentados probablemente resultarían menos significativos, tanto para el lector de Sartre como para el lector de Sacristán. Cabe destacar también el trabajo de las notas complementarias que acompaña a cada uno de los materiales, y que constituye también uno de los aspectos a destacar de la edición.

Conforme avanzamos en los textos y, consecuentemente, en el tiempo, nos parece advertir que la relación de Sacristán con Sartre ha sufrido una transformación, volviéndose más humana e íntima y menos teórica. Estamos en un escenario que se aleja de la posguerra y en el que ahora se representa la posrevolución, un nuevo lugar desde el que por fin parece posible hacerse una idea más precisa de quiénes fuimos como resultado del trauma bélico. El horizonte nuevo permite a la filosofía pensarse y hacer repaso, y la actualización de cualquier posicionamiento político y ético pasa por esa honestidad autocrítica que, como podemos leer fundamentalmente en el último texto recogido en esta obra, caracteriza al Sartre maduro. No obstante la admiración y el reconocimiento de Sacristán hacia Sartre, en esa relación que se humaniza con el paso del tiempo, hay también, como ocurre en cualquier relación que se vuelve personal, algún que otro reproche. Y es aquí cuando percibimos que el último texto recogido nos reconduce, a través una crítica que ha adquirido ahora un tono más íntimo, a una relectura de la introducción de Juan Manuel Aragüés.

El retrato de un Sartre que no está a la altura del lugar que ocupa como intelectual universal en la conciencia colectiva, incapaz de vibrar con la urgencia del acontecimiento, resulta coincidente en la introducción y en la conferencia inédita recogida al final de este volumen, pronunciada en 1980 y titulada *Sartre desde el final*. Si para Juan Manuel Aragués la urgencia desatendida fue la Guerra Civil Española, Sacristán destaca una urgencia que, a día de hoy, cobra un sentido absolutamente pleno: la cuestión ecologista. Es por ello que creemos que el último de los textos presentados en esta obra e inédito, destaca entre el resto de materiales recogidos, y nos invita a leer o releer al último y valioso Manuel Sacristán.

En conclusión, de esta obra destacamos el trabajo de urdir que queda plasmado en la introducción y en el estudio preliminar, ambos necesarios para una lectura plena y profunda de los textos recogidos. Del mismo modo, la complementación de los textos con las notas constituye una fuente valiosísima de fragmentos que, en muchos casos, actúan como una corriente subterránea que recorre y sostiene los propios textos de Sacristán, convirtiendo la lectura de los mismos en una especie de lectura polifónica. Existe, creemos, cierto desequilibrio en los materiales que se recogen, pues hay materiales que brillan más que otros por motivos diferentes, y son los que hemos ido señalando y destacando. En conjunto, no obstante, creemos que esta obra, en tanto que actualización de problemas fundamentales del pensamiento contemporáneo, y en tanto que retrato de la no resuelta problemática de los intelectuales y su relación con la praxis, constituye una obra valiosa tanto para aquellos que conocen bien a Sacristán y/o a Sartre, como para quienes quieren aproximarse por primera vez a estos pensadores imprescindibles.

HACKER, P. M. S. (Peter Michael Stephan), *The Moral Powers. A Study of Human Nature*. Chichester, West Sussex; Hoboken, NJ, (EE. UU), Wiley-Blackwell, 2021, 424 pp. ISBN: 9781119657774.

JESÚS PADILLA GÁLVEZ

Facultad de Ciencias Jurídicas y Sociales
Universidad de Castilla-La Mancha
Toledo/España
jesus.padilla@uclm.es
ORCID: 0000-0002-2890-3514

Recibida: 23/08/2021
Aceptada: 6/09/2021

El libro que reseñamos estudia la naturaleza humana desde una perspectiva muy amplia al abordar el “potencial moral”. Efectivamente, la expresión inglesa “*moral powers*” es compleja de trasladar al castellano ya que hace referencia a la fuerza moral, aludiendo pues al rendimiento o la potestad moral; así como, la potencia humana y su capacidad moral. Nos encontramos ante un hito que estudia el valor en la vida y el pensamiento humano, escrito por uno de los pensadores contemporáneos más activos e importantes del momento. Estamos pues ante una investigación filosófica sobre las potencialidades y sensibilidades morales de los seres humanos, del significado de la vida humana y el lugar que ocupa la muerte en la vida. Se nos presentan unas reflexiones sobre antropología filosófica, es decir el estudio del marco conceptual mediante el cual pensamos, hablamos e investigamos acerca del ser humano, entendido en su horizonte social y cultural. El volumen examina la diversidad de valores en la vida humana y el lugar del valor moral dentro de las variedades de estos valores. El asunto crucial gira alrededor de la naturaleza del bien y el mal y la propensión humana a la virtud y al vicio. Examina las concepciones tradicionales sobre el valor ético y repasa críticamente los conceptos erróneos que se han ido introduciendo en la filosofía, la psicología o la neurociencia cognitiva.

El autor realiza al inicio de esta apasionante obra una observación conspicua que debe tenerse en cuenta a la hora de leer estas páginas: el libro se adentra en el estudio de la antropología filosófica; sin embargo, no debe ser confundida

con la filosofía de la antropología. Esta distinción es relevante debido a que la antropología filosófica ha estado durante mucho tiempo descuidada en la tradición anglosajona. Todo lo contrario a lo que ha ocurrido en el ámbito continental europeo que reviste una vitalidad significativa. Para Hacker, el campo de estudio de la antropología filosófica se limita al examen del marco conceptual en términos de lo que pensamos, hablamos e investigamos acerca del hombre como animal social y cultural (p. xi). En este contexto, el autor resalta el papel fundamental que juegan los valores en la vida humana. El ser humano es considerado como un agente moral, responsable tanto de sus actos como de sus fechorías. No hay que perder de vista que el hombre persigue el placer y la felicidad, esforzándose por vivir una vida dichosa. No todas las acciones que realizan los hombres son positivas para conseguir dicho fin ya que muchos de sus actos son meras felonías, maldades, infamias en las que el individuo comete faltas, perpetra atrocidades, hace disparates y miente para conseguir unos fines que a la postre son inmorales y muestran la otra cara de la facultad humana. El libro investiga las actitudes específicas, denominadas generalmente como la condición humana.

El Prolegómeno se divide en dos partes: primero, introduce su punto de vista acerca de la antropología filosófica y sobre el papel que juega el valor. Segundo, muestra un bosquejo preliminar antes de estudiar los temas en concreto. Dicho esto, deseo advertir de una dificultad que surge en la lectura de este libro para los lectores europeos. A pesar de reconocer el descuido existente en la filosofía anglosajona con respecto a la antropología filosófica, estas páginas ignoran las propuestas europeas por lo que desiste dialogar con la tradición continental. Es más, para un lector continental a veces le puede resultar extravagante el modo de enfocar los problemas. Esto puede generar cierta fascinación. Como es sabido. I. Kant distinguió en el prefacio de su *Antropología* (1798) entre la *antropología fisiológica*, que tiene como objetivo investigar lo que la naturaleza hace del hombre, y la *antropología pragmática*, que inquiere cómo el hombre actúa, como puede ser libre, o que es lo que puede o debe hacer en un contexto social e histórico concreto. Aunque esta distinción adolezca de cierta imprecisión, en la actualidad sigue teniendo cierta resonancia. Tampoco encontraremos referencia alguna a las investigaciones llevadas a cabo por Max Scheler, Helmuth Plessner o Arnold Gehlen. Podríamos afirmar que esta obra está inscrita en lo que Plessner denominó una “posición excéntrica” del hombre. Sin embargo, desde nuestro punto de vista, el libro adquiere una connotación inconfundible. La posición excéntrica describe una característica que se da entre los seres humanos al relacionarse con su entorno: la persona toma distancia de sí misma, sin exonerarse de sí misma. Esto requiere un distanciamiento entre el ser humano y su entorno que se expresa en la expresión “posición excéntrica”. Esta expresión genera una alternativa a los conceptos cartesiano-dualistas. Descartes hizo una distinción entre *res cogitans*

y *res extensa* que asienta la diferenciación entre cuerpo y mente que se expresa en la división entre las ciencias humanas y las naturales, y que a su vez se refleja en la mayoría de las disciplinas de las ciencias sociales en forma de una preocupación casi exclusiva por la dimensión social de los seres humanos. Este libro no se posiciona claramente ante estas discusiones tratadas en la filosofía alemana. Tampoco argumenta acerca del concepto de *animal symbolicum* de Cassirer, ya que su campo de investigación es muy distinto.

En este cuarto y último volumen de su meritorio proyecto, el autor reflexiona sobre la acción moral, la búsqueda de la felicidad y la mortalidad. El libro se subdivide en cuatro secciones temáticas que exploran nuestra vida moral y su significado que a su vez están subdivididos en un total de doce capítulos y cuatro apéndices. Presentare un breve repaso de su estructura y me centraré en comentar algunos pasajes que me parecen relevantes.

La primera sección es la más amplia y estudia el bien y el mal. El capítulo primero investiga las raíces del bien y sitúa la virtud y el vicio en relación con otras categorías de valor, prestando especial atención a la naturaleza del mal y del perdón. Se adentra en las raíces del valor y la naturaleza de la moral. Por ello ubica el lugar que ocupa el valor en el mundo de los hechos. Seguidamente presenta ciertas variedades de la bondad y la dispone dentro de un marco moral, finalizando con una disquisición general acerca de la moralidad. Hacker estudia la propuesta de von Wright y acepta la proposición disconforme con la tradición acerca de la autonomía conceptual de la moral, por lo que los conceptos de bondad o corrección moral deben ser estudiados en relación con una red de conceptos que se refieren al hombre como una totalidad. Para Hacker, la clave para el desarrollo de una visión perspicua de la axiología se encuentra en la comprensión de que todos los valores surgen de la vida. Los ejemplos que afianzan dicha propuesta se asientan en la felicidad, la salud y el bienestar. Desde esta perspectiva, la conducta individual afectará al bienestar de los humanos. En este primer apartado repasa diferentes tipos de habilidades bondadosas –es decir, todas aquellas que tienen que ver con la bondad médica, instrumental, técnica, utilitaria o hedónica–. Desarrolla y establece el significado de los conceptos de deber, bien o virtud.

El capítulo segundo estudia las raíces de la moralidad y la naturaleza de la bondad moral para, acto seguido, seguir inquiriendo acerca de las raíces del valor moral y el respeto. Pasa a explorar la relativa permanencia de las virtudes y sus constantes en la naturaleza humana.

El capítulo tercero analiza el lado contrario: las raíces del mal: ¡El horror! Se adentra en la gramática del mal, los sentimientos que acompañan a este y la repulsión que son causados por algo terrible o repugnante. Por ello se plantea si existe el mal y si se puede explicar como tal. Hacker indica que la maldad humana concierne a la depravación moral, los vicios, la corrupción y, sobre todo, la crueldad.

El léxico de la maldad está estrechamente vinculado a unos sujetos concretos que realizan los actos malvados: las *acciones* mismas en las que se desatan estos actos malvados (*evil-doing*); las consecuencias o resultados de las fechorías (*evil-done*); y, sus agentes, malhechores o delincuentes (*evil-doers*). Como podemos comprobar, la maldad se analiza en el marco directo de la autoría y la mediata de la coautoría. Por ello, se explyaba en investigar el mal común y el generalizado y omnipresente (*ubiquitous*). En este contexto estudia el mal producido en el holocausto. Si bien, el término “maldad” es considerado un término obsoleto, Hacker es de la opinión que debe ser rescatado de dicho olvido a pesar de que carezca de un carácter explicativo y, por tanto, no se pueda dilucidar. Sin embargo, el término “maldad” hace referencia a una serie de fenómenos que están presentes en nuestra vida cotidiana, como el hecho de que una serie de individuos sean propensos a cometer una maldad cuando se dan ciertos factores que son estudiados de manera sistemática. Entre estos factores encontramos la proliferación de un cierto narcisismo, el embrutecimiento progresivo, la indiferencia, el disfrute por el sufrimiento ajeno, la excitación lujuriosa por el poder –lo que se conoce como el nexo entre poder y sexo–, los celos y la envidia, la codicia, la renuncia a reflexionar, la privación infantil a sentir afecto y amor, la carencia de simpatía o empatía, el exceso de competitividad y el egoísmo. De todos estos fenómenos resultan las fechorías, por lo que el mal se expande en nuestra sociedad.

El capítulo cuarto presenta una variedad de explicaciones acerca del mal y se adentra en presentar las razones y motivos humanos para hacer el mal mediante actos malvados (*evil-doing*). Estos se dividen en cuatro ámbitos diferentes. Por un lado, se llevan a cabo en unas condiciones determinadas como puede ser ciertos contextos sociales, condiciones económicas definidas, entornos políticos concretos o ambientes educativos. El segundo ambiente está estrechamente vinculado a las creencias que incluyen los aspectos religiosos, ideológicos, la intolerancia o el fanatismo y la deshumanización de las víctimas. El tercer aspecto a tener en cuenta son los motivos emocionales entre los que se deben considerar la venganza, la envidia, la ira, el odio, los resentimientos y la codicia, la lujuria y la lascivia, la soberbia, la vanidad, la arrogancia y el hastío, el tedio o fastidio. El cuarto entorno comprende las características personales vinculadas a la crueldad que está íntimamente emparejada a la ambición corrompida, el egoísmo ligado a la codicia, el sadismo, la malevolencia, la insensibilidad, el narcisismo, el ansia de poder y la agresión. Hacker pone en duda que la ética y la estética se consideren lo mismo (en contra de Wittgenstein, TLP, 6.421). Comparto su punto de vista respecto al contenido, pero no respecto de la forma. Desde mi punto de vista, Wittgenstein recalca que las expresiones éticas y estéticas comparten ciertas características formales, si bien el contenido es completamente disímil. Finaliza este apartado indagando si podemos conocer el bien y el mal y repasa las repercusiones en el

ámbito de la psicología experimental acerca de las malas acciones. Objeta que el mal sea meramente banal.

El capítulo quinto discute la relación que existe entre el mal y la muerte del alma. Hacker propone superar el concepto tradicional de “alma” concebida como una sustancia inmaterial, espiritual e inmortal, e introduce una concepción secular, según la cual el alma es entendida como el poder moral de la humanidad que permite hablar sobre la vida moral. Hacker distingue entre mente y alma. La primera se considera un conjunto de habilidades intelectuales adquiridas y voluntades que generalmente vienen a ser ejercitadas por el *homo loquens*. El segundo distingue entre lo bueno y lo malo por lo que se ejercita en el ámbito moral. La posesión de un alma presupone la posesión de la mente pero no *viceversa*. La muerte del alma se lleva a cabo mediante la comisión de un delito. Su tesis se reduce al principio: “To do evil is to destroy one’s soul” (p. 138). Leyendo este pasaje pensé, ¡Cuan falta de almas hay en el mundo! Analizados los acontecimientos de mi alrededor cuantos individuos actúan diariamente destruyendo la dignidad humana. Tantos necios y vagos abandonados a la “*ennui*” se auto-adscriben conocimientos que no han adquirido y por una combinación entre la codicia, la soberbia, la vanidad y la arrogancia han llegado a ser catedráticos, políticos, directores de bancos con el único afán de destruir a aquellos que se han esforzado por aprender, corrompiendo con su sadismo y malevolencia a todo aquel que se acerca a su lado y no da pábulo a su narcisismo. Cuantos, con ansia de poder, carecen de alma y desatan las fuerzas psicológicas desintegradoras. Por eso concluye “Evil-doing is a psychologically desintegrative force” (p. 143) Y esa fuerza desintegradora la vemos diariamente en los medios de comunicación, la experimentamos en el trabajo, la padecemos en las relaciones personales. Hacker nos muestra el mundo tal y como es y denuncia su cara más siniestra en unas páginas repletas de humanidad. Por ello exige que para que se conceda el perdón (*Forgiveness*) a esos individuos sin alma –yo diría, sin corazón–, primero tienen que hacer el esfuerzo de perdonarse a sí mismos (*Self-forgiveness*). Finaliza “triseando” un ángulo nuevo del alma. A lo largo del libro, Hacker recurre a numerosos ejemplos del canon literario occidental para fortalecer sus argumentos. Pero, ante todo, nos encontramos con un libro que, en parte, es un ensayo, en parte, ha de ser incluido entre las “*meditaciones*”.

En la segunda sección, Hacker considera la libertad y la responsabilidad como presupuestos de las buenas y malas acciones, examinando el fatalismo, el determinismo y la lógica de la responsabilidad. El capítulo sexto aborda diferentes asuntos como el destino y la fortuna, el fatalismo, el determinismo nomológico, los defectos en el determinismo reductivo, lo aleatorio y lo determinado. El capítulo séptimo se para a considerar el determinismo neurocientífico, por lo que reflexiona acerca de la libertad y la responsabilidad humana. Muestra las pautas más relevantes en las explicaciones del comportamiento humano y expone los

límites de la explicación neurocientífica. Al final se adentra en las variedades de la responsabilidad y expone mediante ejemplos los impulsos y las tentaciones irresistibles.

La tercera sección relaciona el hedonismo y el eudemonismo con la moralidad. El capítulo octavo inspecciona el placer y el disfrute como variaciones inequívocas de hedonismo. Se adentra en el placer, el goce y la complacencia, así como el dolor y los placeres que producen las sensaciones. Pasa a sondear el disfrute y el placer que generan ciertas actividades como el deseo y la satisfacción. Finalmente, profundiza en los juicios que se emiten acerca del placer en primera persona y las consecuencias de la vida hedónica. En el capítulo noveno estudia la felicidad así como sus condiciones previas. Repasa la epistemología de la felicidad y la compara con la moralidad. El capítulo décimo lleva a cabo un estudio histórico acerca de la felicidad como es considerada en el siglo XVIII. A este respecto escruta el concepto científico de felicidad, desplegando los presupuestos psicológicos y epistemológicos y su valoración cuantitativa.

Debido a la verosímil sospecha de incurrir en la falacia naturalista del ser al deber, la mayoría de las propuestas antropológicas han prescindido de vincular el análisis de la identidad humana y el mundo de la vida (*Lebenswelt*) con las reglas que rigen la acción. Sin embargo, ciertamente, se puede juzgar el estatus pertinente desde el punto de vista de la conveniencia y hacer declaraciones normativas desde este punto de vista. Desde este enfoque, el libro analiza las acciones humanas en función de sus necesidades y experiencias. El ser humano no es un ser individual, sino que siempre depende del otro. De esto se deriva la necesidad de plantear reglas de acción, que se puedan asentarse en una ética eudemónica como una filosofía del arte de vivir.

La cuarta sección pretende aclarar el significado de la muerte. Por esta razón comienza buscando un sentido al término en el capítulo undécimo. Observa que premia la pérdida de significado y la sensación de falta de un significado del término. En dicha carencia se asientan las raíces del sinsentido de la muerte. Hacker invierte la cuestión y plantea la pregunta crucial que primero debemos resolver: ¿tiene sentido la vida? El capítulo duodécimo intenta ubicar la muerte en la vida. Primero pregunta: ¿Qué es la muerte? Seguidamente escruta si hay una vida después de la muerte. De nuevo vuelve a plantearse el valor de la vida y si vivimos eternamente. Estas disquisiciones finalizan planteando el persistente, anormal e injustificado miedo a la muerte, es decir la denominada “tanatofobia”. Esta se muestra en cuatro creencias irracionales: la primera, el miedo a una vida después de la muerte que es peor que la vida terrenal; segundo, miedo a lo desconocido; tercero, miedo a morir dolorosamente; y, cuarto, miedo a la inexistencia. En este contexto reflexiona sobre el papel que juega desde la Antigüedad la relación entre el sueño y la muerte. El primero se considera una interrupción de la vida mientras

que el último equivale a su finalización. El libro termina describiendo el ciclo vital en el que el pleno uso de las facultades físicas y mentales menguan y decae la relevancia social por lo que parece conveniente abandonar el escenario. En este contexto no se debe temer la muerte o *in propria voce*: “At some stage in the natural life cycle of mankind, one’s physical powers decline, one’s mental powers begin to wither, one’s role is over, and it is time to leave the stage. Then death should be nothing to fear.” (p. 360)

El volumen concluye con la publicación de cuatro apéndices. Los apéndices enumerados enriquecen el texto con discusiones críticas sobre el pensamiento y el razonamiento animal. Por ello discute acerca de las creencias animales y la moralidad animal. En el primer apartado se centra en explorar el pensamiento animal, los pensamientos animales y su memoria. Presenta algunos contraargumentos y su refutación. Sigue rastreando el conocimiento animal de la mente de otros animales y finaliza mostrando las emociones animales. El siguiente trabajo lleva a cabo un estudio acerca del diablo –lo que denomina “diabología”– que no ha de confundirse con el satanismo. Estudia estos fenómenos en el pensamiento occidental. Sigue un trabajo publicado anteriormente en castellano sobre Hannah Arendt y la banalidad del mal. Finaliza divulgando un trabajo titulado ‘La representación del placer en el arte occidental posmedieval’.

Actuando como la culminación de cinco décadas de reflexión sobre la filosofía de la mente, la epistemología, la ética y la naturaleza humana, este volumen titulado ‘*The Intellectual Powers: A Study of Human Nature*’ concluye la aclamada tetralogía de la naturaleza humana de Peter Hacker entre los que se encuentran los títulos: ‘*Human Nature: The Categorical Framework*’, y ‘*The Passions: A Study of Human Nature*’. La implacable entrega final sobre la naturaleza humana es una lectura fundamental para filósofos morales, filósofos de la mente, psicólogos, neurocientíficos y lectores interesados en comprender quiénes y qué somos, por qué actuamos deliberadamente de manera virtuosa o cometiendo felonías; recapitulando, haciendo el bien o produciendo el mal.

**MARTÍN GÓMEZ, M., *Diario de una filósofa embarazada*. Madrid:
Tecnos, 179 páginas.**

PEDRO MARTÍN LAGO

Doctor en Filosofía
MEC - UNED
pmal0002@gmail.com

Recibida: 9/04/2021
Aceptada: 6/09/2021

En la Historia de la Filosofía podemos encontrar todo tipo de temáticas y problemáticas. Algunos dicen filosofemas. O sea, el tema como problema: qué es la vida, qué es el alma, qué es lo que está bien o mal, etc., temas que, posteriormente, llevan su desarrollo o explicación. Son las filosofías. Porque la filosofía no es –o no es sólo– un saber académico, sino una manera de entender nuestra propia existencia. Y esto es clave para lo que viene después. No hay forma de vivir una vida como seres humanos sin tener antes una filosofía de vida. (Eso que los alemanes llaman una *Weltanschauung*).

De vida, desde luego, pero también de muerte, como último tramo de la vida y sus tejemanejes (consecuencias) morales o religiosos. Por eso algunos filósofos, padres y escritores que publican libros, se han atrevido a dedicar sus obras a sus hijos, con la intención de exhortarles, para que en su *neotenia* (los educadores llaman *educabilidad*) se dejen llevar por el buen camino de la *eudaimonía* (Aristóteles), la moralidad práctica (Séneca), o más recientemente, entre nosotros, la ética o la política (Fernando Savater escribió a finales del siglo pasado un manual de ética y otro de política dedicados a su hijo Amador) en interacción e interactuación. No en vano Nietzsche escribió, “el niño es el padre del filósofo”.

Resulta asombroso que en estos y en otros casos (San Agustín, Rousseau, Hegel, Freud o Piaget) hayan sido filósofos, quienes, si bien experimentaron su paternidad con hijos naturales o legítimos, no llevaron “en sus entrañas” a estos mismos hijos y tal vez por esto tampoco se les ocurrió reflexionar sobre el momento glorioso de la recreación de una vida. Por esta razón, María Martín, profesora de filosofía en la Universidad de Salamanca, y ahora *madre admirable*, se ha atrevido a romper reglas y tradiciones (traiciones), ofreciéndonos este

original y refrescante ensayo que la autora decidió escribirse a sí misma para explicárselo a ese ser que estaba surgiendo a la existencia. Y de esa escritura, íntima y reflexiva, salieron estas hermosas páginas que componen el *Diario de una Filósofa Embarazada*.

Desde el principio, su autora, nos hace partícipe del objetivo: “Ahora seremos tú y yo, hijo mío, los protagonistas de esta historia”. Una historia que como en toda filosofía (Historia de la Filosofía) empieza con el *thaumazein*, el momento del asombro. Este asombro no es otro que la admiración ante las cosas (Ortega) y que tiempo atrás, veintitantos siglos, ya señalaron los maestros griegos (Platón y Aristóteles) como causa y origen del filosofar, razón por la cual nuestra joven autora “embarazada” (“lo confirmo: no hay más deseo de conocer que cuando se está embarazadas por primera vez”) ha querido preguntar e interpelar en solitario y en solidario. “Lo he hablado con otras mujeres que han estado embarazadas y la mayoría coincidimos”.

El cuerpo, y en especial el cuerpo humano, ha sido considerado en filosofía una parte esencial del *ser personal*. María ha rastreado en los filósofos magnos (Platón y Aristóteles), ha mencionado a los cínicos de toda laya (Cioran, con sus inconvenientes), escépticos, consoladores y terapeutas (Séneca, Boecio), existenciales (Kierkegaard, Schopenhauer) y existencialistas (los Gabrieles, Heidegger y Sartre), perspectivistas (Ortega y Gasset), negacionistas y antinatalistas. A todos mira y a ninguno admira filosóficamente, porque ella, como futura madre, milita en la filosofía de la ilustración: “Jamás me ha parecido mal que alguien decida libremente no tener descendencia, pero, por favor, dejen de juzgar a los que sí lo hemos decidido”.

Abandonando el lenguaje de las corrientes filosóficas, paso ahora a adoptar el de los conceptos límite, por el cual abogo, tal y como los encontramos en el ensayo: un Prólogo, que presenta los fundamentos que van desde el primer hombre (¿Adán?) hasta el moderno giro lingüístico wittgensteiniano, universalizando al hombre y el nombre. (“Elegir nombre es apasionante”). Todos los demás capítulos que se suceden son sólidos y rocosos (Sísifo), empezando por el origen heracliteano (diríamos dialéctico, si la dialéctica no se viese tan mellada hoy por un uso excesivo) y la ensoñación de Segismundo (*La vida es sueño*) que nos conduce directamente a la ensoñación unamuniana (“Tener o no tener”) y las grandes filósofas “no madres” (las dos Simones, Beauvoir y Weil, y Hannah Arendt) o las “madres arrepentidas” de las que habla Orna Donath. Son actos de *in-moralidad* que usted debe experimentar sin clichés (cap. 4), porque, cuando *se tiene*, la sociedad toda con-mueve hacia la eticidad universal: “Fue muy bonito apreciar, en general, el respaldo de la sociedad”.

Es la vida que sale al encuentro (cap. 5) y tal vez por esto mismo, como en el viaje de Megan Foster, debieron enviar a un poeta, porque las emociones

suscitadas son lógica y filosóficamente indescriptibles. Tintinea en el alma un *sermo interior* (¿agustino?), desde un preexistir platónico o de María (Zambrano) que sostiene que en el nacer hay también un olvido. “El alma, que residía en un mundo ideal, debe olvidar todo lo que sabe antes de venir a este mundo”. Poesía y Filosofía, claros en el bosque, imaginarios sociales y colectivos donde el Hado, el Feto y el Fado que lo canta (enfadados) harían las delicias de Fernando Pessoa (María lo cita en dos ocasiones), al no conocer el por qué (cap. 8) o bien sentir el temor y el temblor del filósofo de la angustia. (Soren Kierkegaard).

Todo el libro nos sabe a deliciosa poesía (“Sabré el porqué cuando termine el tiempo”, Dickinson). También a filosofía moral. (“Se debe soportar el dolor”, Gadamer), porque toda gran poesía tiende siempre hacia el pensamiento, en tanto que el pensamiento que piensa lo que merece ser pensado camina por sendas vecinas a la poesía (nuevamente Heidegger) sin caer en la baratija de la autoayuda o cualquier otro texto de psicología práctica. Por eso los capítulos que restan tienen que ver con el no en el decir (y aun decidir), el acompañamiento, la duda, no cartesiana sino existencial, y nuevamente la locura (“bendita locura”) o la *sinrazón* de Rosa Chacel. Todas ellas ensortijadas en la vida (*via et vita*).

Porque la vida es vía y camino (“En el camino”, cap. 13), categorías muy presentes en la poesía (camino de perfección en Teresa, sendas luisianas, caminante de Machado...) y también en la filosofía medieval (*itinerarium mentis in Deum*), o en la metafísica moderna como alegoría del pensamiento (los caminos de Heidegger). María hace de lo pensado en el camino (caminado) la vía para hacer patente lo no pensado (no caminado) porque no de otro modo es la vida: “La vida como camino, con multitud de alternativas, atajos y direcciones” (p. 142), y que ella imagina peregrina (no olvidemos su origen vigués) en el *camino* de Compostela, con sus compañeros de Facultad, que “lo saben bien”; y tal vez por todo esto, calla, como el marinero del Conde Arnaldos: “Yo no digo mi canción sino a quien conmigo va”.

Los dos últimos capítulos 15 y 16 están comprometidos. Son de bienvenida a Bruno (nombre de origen germano y también nombre del filósofo y poeta Giordano Bruno) y la presentación de Cecilia (cristiana, noble y patrona de los músicos). “De ellos pienso hablarlos cuando seáis mayores”. Tal vez ahí se encuentren otras claves.

SNELLER, Rico, *Perspectives on Synchronicity, Inspiration and the Soul*. Newcastle: Cambridge Scholars Publishing, 2019, 389 pages

GERMÁN ULISES BULA CARABALLO

Doctor en Educación
Mg. en Filosofía
Facultad de Filosofía y Humanidades
Universidad de la Salle
Bogotá/Colombia
gbulalo@unisalle.edu.co
ORCID: 0000-0002-1296-0610

Recibida: 5/05/2021
Aceptada: 6/09/2021

La lectura de libros especializados es, casi siempre, un ejercicio que implica un cierto grado de redundancia. Para el filósofo de profesión, la lectura de un libro académico suele implicar el repaso de una buena cantidad de ideas ya conocidas (a las que se añaden unas cuantas, que constituyen el aporte específico del libro en cuestión). En este aspecto, el libro de Rico Sneller es una sorprendente excepción: aborda temas bien poco comunes, haciendo uso de un arsenal de referentes académicos inusuales aunque de muy buen calibre: desde la artista brasilera Lygia Clark hasta la biología vitalista de Hans Driesch y la *Lebensphilosophie* de Ludwig Klages, pasando por los sesudos estudios de Max Picard sobre la fisiognomía.

El carácter inusual del libro se debe a su temática, y a la manera intrépida de abordarla: Sneller aborda el fenómeno del alma en cuanto fenómeno espiritual (que no psicológico, biológico, antropológico o cultural). De este modo, Sneller rechaza de plano los prejuicios cientificistas de nuestro tiempo; prejuicios que, en nombre del empirismo, rechazan partes cruciales de la experiencia humana, tales como la inspiración divina, la sincronicidad, o lo que pueden enseñar las experiencias de estar cercano a la muerte. Sneller (2020, 5-6) articula su epistemología alternativa a partir de la distinción que hace Hermann Friedmann entre la orientación háptica, táctil, y la orientación óptica hacia el mundo. El cientificismo es háptico: toma por existente e importante sólo lo que es tangible, manipulable, dominable: lo que puede aprehender entre dedos y pulgar. Sneller lo caracteriza con la figura de Polifemo: pobre de vista, palpando a sus prisioneros en la oscuridad.

Esta orientación contrasta con una orientación óptica, dispuesta a contemplar: capaz de manejar una cantidad mayor de incertidumbre, de dejarse afectar por los fenómenos tal como se presentan, de incluir en su ontología todo lo que dura en el tiempo, no sólo lo que ocupa el cartesiano ancho, alto y largo (Sneller, 2020, 36).

Desde esta perspectiva, el alma aparece como un fenómeno de bordes indefinidos: los límites entre lo consciente y lo inconsciente, lo externo y lo interno, lo propio de mi alma personal y del alma del mundo, se hacen difusos; y lo que es más, navegables. No se trata aquí, exclusivamente, de fenómenos místicos: más bien, Sneller trata de recuperar el carácter misterioso de la existencia cotidiana (Sneller, 2020, 42). Nuestra alma es siempre un misterio para nosotros mismos, pero no por ello es objeto de conocimiento, en el sentido usual del término. La exploración del alma excluye la intencionalidad (Sneller, 2020, 130). A medida que se gana lucidez en el alma, se pierde el yo, de modo que esta lucidez, este conocimiento, no se comprende a través de ninguna epistemología que trate de sujetos que tienen conocimiento de objetos; ni tampoco una especie de observación de segundo orden, que hace de las observaciones de primer orden un objeto de consideración. Más bien, la lucidez respecto a la propia alma es una forma de existir.

En contra de la antropología que considera al ser humano como un agente racional autointeresado, dotado de una computadora de carne (digamos, la visión Hobbesiana del ser humano, o la que resulta de un evolucionismo darwiniano unilateral); Sneller considera que somos movidos por una tensión entre la autoprotección y el autoabandono del alma, un jugar con los límites difusos de la misma, un deseo incómodo de exteriorizar lo más íntimo, que siempre está permeado por la vergüenza (Sneller, 2020, 91-93). Discutiendo la obra de Lygia Clark, Sneller muestra de manera magistral cómo la tensión entre exhibicionismo y vergüenza permea la actividad artística. La inspiración artística es, bajo esta perspectiva, equivalente a la lucidez respecto a nuestra propia alma (Sneller, 2020, 278).

Es sólo desde el dualismo háptico y cientificista que este tipo de exploraciones se entienden como alejadas del mundo material. Por el contrario, el vernos como almas cuyos bordes con el exterior son difusos arroja una luz nueva e intrigante sobre algunos fenómenos materiales: tales como el aún misterioso proceso por el que la vida reproduce sus formas (que explora Hans Driesch), o aquello poderoso que puede decirnos la contemplación atenta a un rostro (que explora Picard). En este punto creo necesario el uso del vocativo: no pienses, oh lector, en lo que sabes, sino en lo que vives, en lo que has podido ver en el rostro de alguien que amas, o en hermandad que a veces descubres en el rostro de alguien que recién conoces. El rostro contemplado (no mirado para el uso y la clasificación, hápticamente; Sneller, 2020, 239), es, por así decirlo, un rayo del sol infinito, una concreción específica y particular del alma universal, un marcador de la eternidad (*Platzhalter der Ewigkeit*, (cfr. Sneller, 2020, 244).

Leer a Sneller es entrenarse en una manera de pensar inusual para nuestro tiempo, en particular para un académico (pues en la academia se suele confundir el rigor con el enfoque háptico). En lugar de centrarse en el espacio y sus entes cartesianos, Sneller se centra en el tiempo y sus sincronicidades, en las alegres y misteriosas confluencias de eventos que producen el sentido en la geometría del tiempo (Sneller, 2020, 213-254); en aquellos eventos externos que parecen responder a lo que ocurre en nuestra alma, que interesaran a C.G Jung. Sneller cita el caso de un terapeuta en un edificio de Manhattan quien tiene que interrumpir su sesión con un paciente, quien le relata su miedo a las serpientes, porque una serpiente ha entrado a la oficina (Sneller, 2020, 250). Quizás es el aspecto abisal de estos fenómenos lo que nos tienta a considerarlos meras coincidencias (como si las serpientes fueran muy comunes en Manhattan): en general, el deseo de facilidad, de certidumbre nos ancla en lo háptico. Abandonar la mirada de Polifemo es lanzarse al mar, que (al contrario de la orilla) es tridimensional y ofrece más posibilidades de perderse. Es peligroso, ajeno e impredecible; pero guarda insospechadas riquezas.

HAN, Byung. Ausencia. Acerca de la cultura y la filosofía del Lejano Oriente (Trad. Graciela Calderón). Buenos Aires: Caja Negra, 2019, 133 p.

JUAN DAVID ALMEYDA SARMIENTO

Profesor de Filosofía
Escuela de Filosofía
Universidad Industrial de Santander
Bucaramanga/Colombia
Juanalmeyda96@gmail.com
ORCID: 0000-0002-6463-6388

Recibida: 29/06/2021
Aceptada: 6/09/2021

Existe dentro del trabajo filosófico de Byung-Chul Han un ejercicio *amable* por intentar conciliar dos tradiciones en una hibridación teórico-práctica, sépase, poner en comparación (en una conversación o una dialéctica) el pensamiento oriental y el occidental. La lectura apresurada de los *haikus ensayísticos* de Han llevan a encasillarlo meramente en las bases de la tradición de Occidente. Los trabajos sobre el neoliberalismo, el cansancio, la transparencia y lo digital, parten de bases filosóficas que son comunes a los habitantes de este lado del mundo, sin embargo, las investigaciones que tiene este mismo autor sobre la tradición oriental pasan sin pena ni gloria. *Ausencia* pertenece a esta segunda categoría, es una obra que parte de una metodología expuesta en el pasado por Han: “el presente estudio se desarrolla mediante comparaciones (...) Empleamos aquí la comparación como método que saca a la luz el sentido”¹. No obstante, el centro de atención está principalmente enfocado en el trabajo sobre la cultura y el pensamiento de Oriente, pero siempre contrastada en pie de página desde una postura de Occidente que saca a la luz ese sentido del que habla Han. Este último, en este libro, tiene por objetivo poner sobre la mesa lo que para este lado del globo todavía es visto como *extraño*. De este modo, lo que Han quiere es dar lugar a

1 HAN, B. *Filosofía del budismo zen* (Trad. Raúl Gabás). Barcelona: Herder, 2015a, p. 11.

un *espacio distinto* que deje de lado el prejuicio para generar una conversación: “Sería una expresión de amabilidad que también haría posible *volverse distinto*”².

De manera general, la obra está constituida por siete capítulos que intentan abarcar las distintas dimensiones, estos son: i) esencia y ausencia; ii) cerrado y abierto; iii) luz y sombra; iv) conocimiento y timidez; v) tierra y mar; vi) hacer y acontecer; vii) salud y reverencia. Cada momento intenta explicar la complejidad detrás de la cultura del *Lejano Oriente*, la cual responde a todo un bagaje de autores orientales que llenan de contenido las páginas de Han: Zhuang zi, Eihei Dōgen, Lao zi, Engo Kokugon, entre otros, se encuentran con Heidegger, Benjamin, Hegel y Leibniz, en una *amable conversación simpatética* que termina por abrir lugares distintos frente a la relación del ser humano con el mundo, específicamente, en la era contemporánea.

Este libro se enfoca en el concepto de *vacío*, tanto en el uso zen como en el taoísta (principalmente, puesto que no se limita a estas dos interpretaciones), en contraste con la aprehensiva tradición de la *esencia*, con la que Han identifica a toda la historia de Occidente. Para este pensador coreano, la base de la cultura asiática tiene por piedra angular *la amabilidad*³ con el mundo por medio del vacío antiesencialista que implica la práctica del camino (道). El primer capítulo es la puerta que da sustento argumentativo al resto de la obra, ya que el contraste entre la dicotomía: esencia (Occidente)⁴ y vacío (Oriente), se retoma en toda la obra desde distintas latitudes. De esta manera: “En la figura de la esencia está prefigurado el poder. Debido a esta prefiguración la cultura o el pensamiento que se orienta a la esencia desarrolla necesariamente una firmeza *respecto de sí* que se expresa en deseo de poder”⁵. A esta esencia (*wesen*) se le contrapone la *no esencia* (無物), la cual se asemeja al ave que carece de nido, por lo que siempre vuela sin dejar rastro alguno por donde va: “la ausencia se sustrae a toda fijación sustancial. En consecuencia, la «no esencia» está asociada al caminar, al no habitar. El sabio

2 Ibid. p. 11.

3 La amabilidad no es definida concretamente dentro de la obra, pero en otro de sus trabajos existe una explicación más precisa de qué es dentro del pensamiento de Han: “La amabilidad es una «participación» completa, es decir, una forma intensiva del *espíritu*”. HAN, B. *Hegel y el poder. Un ensayo sobre la amabilidad* (Trad. Miguel Alberti). Barcelona: Herder, 2019a, p. 154.

4 Principalmente, el concepto de esencia occidental que el autor toma como paradigmático es el de Heidegger: “en el combate esencial, los elementos en lucha se elevan mutuamente en la autoafirmación de su esencia. La autoafirmación de la esencia no consiste nunca en afirmarse en un estado casual, sino en aband-narse eri el oculto estado originario de la procedencia del propio ser” HEIDEGGER, M. “El origen de la obra de arte (1935/36)”. En: *Caminos de bosque* (Trad. Helena Cortés y Arturo Leyte). Madrid: Alianza, 2010, p. 35.

5 HAN, B. (2019). *Ausencia. Acerca de la cultura y la filosofía del Lejano Oriente* (Trad. Graciela Calderón). Buenos Aires: Caja Negra, 2019, p. 15.

se mueve hacia donde no hay «puertas ni aposentos»⁶⁷. En este orden de ideas, el trabajo de Han en *Ausencia* gira en torno a una conversación entre estos dos conceptos (*vacío* y *esencia*), pasando por su espacialidad, su temporalidad, su cultura, su sociabilidad, etc.

En este sentido, esta obra forma parte de las piezas de Han por constituir una *filosofía de la amabilidad*. Ciertamente es que el trabajo de este pensador no está constituido por hacer esto explícito, pero los rastros filosóficos que sus investigaciones sobre Oriente y sus abordajes a las problemáticas contemporáneas dejan detrás, hacen posible sentar las bases de un proyecto más amplio de lo que los haikus hanianos permiten. En este marco se posiciona *Ausencia*, abre lugar a pensar las estructuras axiomático-conceptuales que componen la armonía teórica que busca Han en su *filosofía de la amabilidad*:

La ética de la amabilidad no abandona solamente el nominativo, sino también el acusativo (...) Abandona también el espacio interior dialógico y se dirige al espacio de la ausencia, al entre vacío que no está ocupado ni por el yo ni por el otro⁸.

Este proyecto implica una lucha por parte de Han de despsicologizar al yo de los sujetos, esto es, pasar del *habitar* al *caminar*, en un amable intento por reconciliar (*Versöhnung*) la negatividad y la positividad que compone al mundo, una relación que ha sido desequilibrada. Dentro de la filosofía hasta el momento trabajada por Han, el pensamiento oriental cuenta con la facilidad de entenderse en una *reverencia* indiferenciada con el mundo, la cual se vuelve fundamental para concebir una existencia del ser humano desde una interpretación narrativa de lo que son las acciones humanas más allá de la idea aprehensiva de yo:

El ser se sustancializa haciéndose camino. También Heidegger se vale a menudo de la figura del camino. Pero el suyo es un camino que se distingue claramente del camino taoísta, puesto que no trans-curre, sino que pro-fundiza. Los conocidos “Caminos de bosque” de Heidegger son “caminos [Wege], por lo general medio ocultos por la maleza, que cesan bruscamente en lo no hollado”. En cambio, el camino chino se extiende en la planicie, cambia de recorrido constantemente, sin cesar⁹.

6 Ibid. 17.

7 Las bases de esto vienen del pensamiento de Lao Tse y Zhuang Zi. En conjunto, estos dos proveen la *no esencia* de un contenido filosófico que contrasta con la tradición de Occidente, consultar: Tse, L. *Tao Te Ching. Libros del Tao* (Trad. Iñaki Preciado). Madrid: Trotta, 2012 y Zi, Z. *Maestro Chuang Tsé* (Trad. Iñaki Preciado). Barcelona: Kairós. 1996.

8 HAN, *Ausencia...*, p. 127.

9 HAN, B. *Shanzhai: el arte de la falsificación y la deconstrucción en chino* (Trad. Paula Kuffer). Buenos Aires: Caja Negra, 2017, p. 13.

En este sentido, el proyecto de *Ausencia* tiene una originalidad inmanente, es un intento por abarcar una tierra firme, en lo correspondiente a un marco teórico de autores que provee de argumentos sus ideas, y, al mismo tiempo, intenta buscar con qué regar las semillas filosóficas de su proyecto sobre la amabilidad. La tierra, filosóficamente hablando, representa para Han un ejercicio de cuidado y respeto, los conceptos y argumentos que el coreano intenta poner dentro de su obra poseen la *fragancia* que el autor en el pasado tanto ha recalcado¹. La jardinería teórica² que paulatinamente se realiza en *Ausencia* representa una muestra de las alternativas que Han ha venido señalando en sus críticas a la sociedad del rendimiento, pero las cuales se veían opacadas dentro de los variados temas que el coreano profundiza en sus obras.

Así, *Ausencia* es una obra de *jardinería* de la cual brota un carácter *aromático* en la medida que se desenvuelve teóricamente y a la vez que postula su filosofía práctica. El trabajo de Han con esta investigación comparativa es la de proponer fundamentos para una revitalización de la *vita contemplativa* desde la *no esencia* oriental, además del enfoque en los andamiajes conceptuales de la filosofía asiática. En el proceso, reafirma implícitamente su proyecto de una *filosofía de la amabilidad* y de un principio de armonía con el mundo desde una despsicologización del ser humano. Es por esto que la obra es una lectura obligada para quienes intentan profundizar tanto en este autor como en la investigación sobre las alternativas a los dispositivos de poder suave que actualmente implementa el sistema hegemónico neoliberal. Este libro representa un interesante contraste entre dos culturas, llegando a puntos de sintetizar ambas en una sola amalgama indiferenciada, la cual permite cultivar conceptos y apuestas prácticas atractivas para el lector acostumbrado al lugar común del pensamiento occidental y que busca nuevos paisajes para aproximarse a las problemáticas contemporáneas.

1 Algo que se enmarca en la propuesta haniana de revitalización de la contemplación sobre la acción: “Las llamadas estrategias de desaceleración no son capaces de acabar con la crisis temporal contemporánea (...) es necesaria una revitalización de la *vita contemplativa*”. HAN, B. *Aroma del tiempo. Un ensayo filosófico sobre el arte de demorarse* (Trad. Paula Kuffer). Barcelona: Herder, 2015b, p. 11.

2 Esto debe señalarse con énfasis, ya que para Han el ejercicio de respeto y amabilidad que surge del ejercicio del cuidado de un jardín provee al ser humano de un *aroma temporal de lo diferente*: “El trabajo de la jardinería ha sido para mí una meditación silenciosa, un demorarme en el *silencio*. Ese trabajo hacía que el tiempo *se detuviera y se volviera fragante*” HAN, B. *Loa a la tierra. Un viaje al jardín* (Trad. Aberto Ciria). Barcelona: Herder, 2019b, p. 11.

TRADUCCIONES

MI PROPÓSITO PRINCIPAL¹

Gabriel Marcel

Traducción e Introducción de Yefrey Antonio Ramírez

YEFREY ANTONIO RAMÍREZ

Docente investigador
Universidad Agustiniana
Bogotá/Colombia
yefrey12@gmail.com
ORCID: 0000-0001-7972-5095

Recibido: 6/07/2020
Aceptado: 6/09/2021

Resumen: En esta intervención en la Universidad de Gante, Gabriel Marcel intentó dar cuenta, como el mismo título lo expresa, del propósito principal que persigue su filosofía. Se encuentran entonces aquí tanto la dirección de su pensamiento, como algunos de los fundamentos de su gnoseología. Debido a ello, su contenido es una puerta útil de entrada a su obra.

Palabras claves: Gabriel Marcel, ser, verdad, inquietud, intersubjetividad, teatro.

Abstract: As the title suggests, Gabriel Marcel intends to account of the main purpose pursued by his philosophy with this conference at the University of Ghent. It shows the direction of his thought and the foundations of his gnoseology. Its content is therefore a useful gateway to his work.

Keywords: Gabriel Marcel, being, truth, restlessness, intersubjectivity, theater.

“Mi propósito principal”, hasta ahora inédita en español, es una conferencia que Gabriel Marcel impartió en 1937. Fue publicada en la obra *Présence et immortalité* en 1959 por la editorial Flammarion, París, y reeditada en la misma

1 Título original: *Mon propos fondamental*.

ciudad, en 2001, por l'Association Présence de Gabriel Marcel. *Présence et immortalité* es el tercer *Diario metafísico* de este autor francés.

Marcel comienza destacando uno de sus aspectos más propios y que marca definitivamente su tipo de pensamiento: su dramaturgia está íntimamente ligada a su filosofía. Aspecto que explica por qué en su teatro se encuentran varias de las conclusiones filosóficas a las que llegó posteriormente². En efecto, los diálogos teatrales le sirvieron para interrogarse por las situaciones concretas de la vida que le parecían conllevar una importancia ontológica.

Las cuestiones a través de las cuales gira la filosofía de Marcel están, pues, vinculadas con la vida particular. Sus preguntas no versan sobre objetos, sino sobre sujetos en situación. Su teatro es la evidencia de esas situaciones. Por este motivo, para Marcel, hablar de “su filosofía” le resulta difícil porque no entiende por ella un sistema de proposiciones abstractas, sino una profundización en la experiencia concreta del yo. Esta singular posición tiene una explicación gnoseológica: esta experiencia es inabarcable pues depende del sujeto, por eso no se pueden emitir juicios fijos que la determinen completamente y definitivamente³.

El carácter inagotable de las cuestiones que ocupan esta filosofía, causa la distinción de una esfera de inteligibilidad que se aparta de la epistemología conocida por Marcel. Esa esfera es la de la presencia y la reflexión segunda, por oposición a la de los objetos y la reflexión primera. Pensar la situación particular de una persona no puede hacerse con los mismos parámetros gnoseológicos que se utilizan para pensar un objeto, si no se le estaría objetivando. Para Marcel, el hombre no se puede objetivar sin traicionarse como ser.

Estas distinciones tienen un impacto en sus reflexiones sobre la verdad: la verdad buscada aquí no depende de una técnica ni de su carácter impersonal, sino que, al contrario, es profundamente personal. Esto quiere decir que, entre otras cosas, esta verdad no puede ser tratada como un tener, sino como dependiente del ser en el plano de la intersubjetividad⁴. De esta manera, la verdad en esa esfera

2 Este aspecto permite comprender por qué Marcel prefirió calificar su filosofía como “socrática” en lugar de “existencialista”: su pensamiento estuvo marcado por el movimiento y el cambio que comportan las experiencias y los diálogos al estilo de Sócrates.

3 Estas son las cuestiones llamadas por él “misterios”, por oposición a los “problemas” que se pueden agotar en un enunciado y anulan al sujeto que las piensa.

4 La filosofía de Marcel más que del “ser”, trata de “mi ser”. Hay ser allí donde el hombre encuentra plenitud. Por eso no se trata de un concepto abstracto y estático que se opone a la libertad, sino de un reconocimiento de las experiencias que llenan ontológicamente la vida. Desde esta perspectiva, para este autor francés, ser en plenitud es imposible aisladamente: es indispensable el otro. La búsqueda del ser es, pues, intersubjetiva. Lo verdadero, en este marco, es aquello que permite el reconocimiento y la vivencia de sí mismo como un ser intersubjetivo.

de inteligibilidad se distingue de la verdad objetiva y subjetiva, para ser comprendida como intersubjetiva: el yo no se puede comprender sin el tú.

La tarea del filósofo para acceder al ser y a la verdad es ilustrada por Marcel a través del acto de leer: el lector se enfrenta a un papel con letras y a un significado que se da en diferentes grados según la apertura de quien busca descifrarlo. La tarea del filósofo es, igualmente, un descifrar, en este caso, el ser que se le presenta también gradualmente. Sin embargo, el ejemplo que más agrada a Marcel para la explicación de este vínculo es el de una partitura musical: el intérprete cuando está delante de ella, debe abrirse mediante una actitud creadora a la verdad expresada en el pentagrama. Cuando logra hacerlo, entra en comunicación con el creador de la partitura.

Lo que se quiere resaltar con ambos ejemplos es la existencia de un orden situado más allá de lo que se considera válido y objetivo según los parámetros científicos. Ese orden es del tipo de la comprensión de un texto literario o musical, en el que se sobrepasa lo objetivo para abrirle paso a una comunicación entre seres. Este orden es, por consiguiente, intersubjetivo y se revela particularmente en las relaciones personales que se dan en situaciones concretas, las cuales son analizadas a través de la reflexión fenomenológica y con ayuda de los personajes de su teatro. Este es el horizonte de acción de la filosofía marceliana.

MI PROPÓSITO PRINCIPAL⁵

Gabriel Marcel

En vista de una conferencia que debo dar en Gante en los próximos meses, me vi llevado a reconocer, con desconcierto, la dificultad que me provoca determinar *ex cathedra* algunos temas esenciales de mi pensamiento filosófico. Esto se debe a múltiples razones sobre las cuales me veré obligado a insistir posteriormente, pues creo que debo mostrar ciertos obstáculos a los que se enfrenta, necesariamente, cualquiera que afirme dar una idea global de mi obra.

La primera dificultad, relativamente superficial, cuyas raíces profundas tendré que descubrir, consiste en que mi obra filosófica no se deja disociar realmente de mi teatro. Éste no constituye, en absoluto, una ilustración o una traducción de un cierto cuerpo de pensamientos preexistentes y que habrían sido antes expuestos de forma abstracta. En realidad, como lo han visto los mejores intérpretes de mi pensamiento, hay que otorgar al drama una verdadera primacía en mi obra. ¡Cuántas veces anticipé en mi teatro conclusiones a las cuales debí llegar más tarde en el plano filosófico! Nada más característico que la escena final de *L'Iconoclaste*, en el que se reconoce el valor positivo del misterio.

Pero me dirían: “poco importa, nada le impide aislar esas conclusiones del contexto dramático en el que se encontraban primitivamente y formularlas abstractamente”. Aquí, sin embargo, hay que tener cuidado: mis obras más importantes y significativas, *Le quatuor*, *L'iconoclaste*, *Le chemin de Crête*, y sobre todo *Les coeurs avides*, terminan en un acuerdo complejo, en una armonía a la que el espectador –o el lector– está invitado a participar, pero que no se deja reducir en formulas sin alterarse en cualquier grado. Como lo dijo un personaje de *Le quatuor*: “¿Va a morir en las palabras este pensamiento en el que fuimos uno solo?”. Corresponde al drama, como a la música, desarrollar más allá del conocimiento discursivo una consciencia superior en la que nuestro ser se presente en su integridad y trascienda los enunciados abstractos en los que la inteligencia pura está obligada a contentarse.

Hacia un conocimiento de este orden es que mi pensamiento filosófico está orientado. Es polifónico por esencia, y se opone radicalmente a todas las ideologías más o menos derivadas del pensamiento filosófico francés del siglo XVIII.

5 Agradecemos la amable autorización de *L'Association Présence de Gabriel Marcel* para publicar esta conferencia en español. Data de 1937.

Pero hay más, me pongo en graves dificultades cuando intento proceder por relación a mis propios escritos como si se tratara de rendir cuenta de la filosofía de otro, aunque esto parezca un proceso no sólo lícito, sino recomendable.

Cuando me remonto a mis escritos antiguos, reconozco claramente los puntos esenciales de mi pensamiento que he intentado expresar desde hace diez años. Sin embargo, ese lenguaje es otro y no puedo evitar preguntarme si ahora confiero abusivamente otro sentido a esos escritos del que presentaban en la época en que los escribí. Esto contribuye, además, a que me vuelva bastante escéptico del valor de muchos trabajos de historia de la filosofía que implican, justamente, reinterpretaciones de las que no podemos decir en realidad si son o no legítimas.

Yo podría, en verdad, decir borrón y cuenta nueva de esos escritos añejos. Pero debo confesar que mi malestar es aún más fundamental. Pienso que siempre me será imposible presentar algo que se parezca a una exposición o a un boceto de mi filosofía a manera de una teoría más o menos conectada hábilmente con teorías vecinas. En realidad, no me parece que la investigación filosófica conlleve, en el mismo sentido de la científica, un *acervo* que el espíritu no ponga en cuestión y de donde parta para ampliar sus conquistas.

Tiendo a creer que es propio de la esencia de un pensamiento filosófico vivo poner siempre en cuestión, a cualquier grado, las conclusiones a las que ha llegado gradualmente. Esto pelagra de parecer, a primera vista, no solamente desalentador, sino desesperante, y debo hacer hincapié en ello.

A menudo he concentrado mi atención sobre estas dos palabras: *mi sistema* o *mi filosofía*, y he sentido una especie de horror en presencia de su carácter irrisorio. Esto se traduce en la molestia insoportable que siento cuando me piden – por lo general las mujeres de clase alta– definir en pocas palabras “en qué consiste mi filosofía”. Esta petición es suficiente para terminar con la consciencia que puedo tener de la misma. Me parece que siempre he sentido, sin quizás formulármelo con esta nitidez, que la filosofía por definición no es y no puede ser cualquier cosa que uno posee, que uno tiene. Pienso que ningún filósofo digno de este nombre ha podido entender su propia doctrina como algo que uno posee o tiene, salvo en el curso de una polémica, donde la cuestionan y donde ha tenido que defenderla como se defiende la piel. Pero habría que preguntarse si a partir del momento en el que es tratada como tener, una filosofía no se degrada, no se convierte en su propio cadáver. Cuestión difícil que sólo me limito aquí a plantear.

Pero, me objetarán: ¿No es de la esencia de una búsqueda filosófica tender hacia la constitución de un conjunto de proposiciones coherentes y fuertemente relacionadas? ¿Renunciar a este ideal, no es reducir esta investigación a un juego formal, un ejercicio estéril? La filosofía no es nada si no es una búsqueda, pero buscar es esperar encontrar, es tender hacia lo definitivo.

Hay que penetrar aquí mucho más profundo y descartar malentendidos.

Sin duda la filosofía es esencialmente una búsqueda, y esta búsqueda es la de la Verdad. Pero esta palabra presenta una profunda ambigüedad que debemos detectar. La Verdad, en el sentido filosófico de la palabra, es inconmensurable con respecto a las verdades particulares que obtiene el científico al término de sus pacientes investigaciones. Es importante ver por qué esto es así.

Lo propio de una verdad particular, del orden que sea, es que se pueda formular de manera estricta, corriendo, incluso, el riesgo de confundirse con el enunciado en el cual se resume; o más exactamente, de no poner ninguna resistencia a esta confusión que es tal vez una corrupción. En la medida en que ella es tomada en sí, es decir, independiente de la investigación de la que es el resultado, tiende a aparecer como exterior al sujeto. Allí está la raíz del cientificismo, entendido como degradación de la verdadera ciencia. El rol de la reflexión filosófica será hacer aparecer el carácter falaz de toda verdad particular si esta es reducida a un elemento aislable del saber.

Se podrá mostrar más profundamente que el conocimiento, si es concebido como un modo de posesión, tiende a convertirse en no-saber; y una degradación de este tipo es, al contrario, impensable donde no se trata tanto de saber, sino de comprender.

Conviene aún observar –y todas estas notas son convergentes– que una verdad particular es una verdad impersonal; su valor de verdad y su carácter de impersonalidad son inseparables. Si afirmo que la conclusión a la cual llegué es verdadera, quiero decir que cualquiera en mi lugar, procediendo correctamente, es decir, en conformidad con una cierta lógica impersonal, llegaría a la misma conclusión; yo no intervengo, mi personalidad no se tiene en cuenta. Puede ser que me haya beneficiado de una intuición privilegiada en el camino, pero esta intuición me ha descubierto únicamente una ruta reconocible sobre la cual cualquiera podrá comprometerse a seguirme. En definitiva, las condiciones accidentales en las cuales he descubierto ese camino no tienen importancia, no son susceptibles de interesar más que al sicólogo o al historiador. Eso es lo que expreso al decir que las verdades particulares provienen *del pensamiento en general* (en el sentido kantiano).

Uno se ve primero tentado a afirmar que esto es más verdadero para las llamadas verdades generales. Pero es aquí dónde conviene tener cuidado.

No hay duda que si por verdades generales entendemos ciertas relaciones que serían sacadas por abstracción de unas verdades particulares, a las que nos conducirían las diferentes disciplinas científicas –es decir, lo que podemos llamar generalmente principios, por ejemplo en el sentido cartesiano–, es cierto que esos principios provienen también del pensamiento en general. Pero se trata de saber

si la filosofía en su esencia se propone establecer principios de este tipo. En realidad se trata aquí de una cuestión de definición. La filosofía, tal como siempre la he entendido, al igual que muchos de mis contemporáneos, está orientada hacia un fin muy diferente.

En primer lugar, diría en una palabra que la *verdad* a la cual aspira la investigación filosófica, por esencia, no se puede poseer. De ninguna manera puede ser considerada o tratada como un tener.

Esto fue lo que quise decir entre 1910-1914, cuando sostuve que la metafísica es ante todo una filosofía de la participación, y estuve muy sorprendido de comprobar en el Congreso de Filosofía de París (1937) que Louis Lavelle utilizaba ese mismo término de participación en una acepción sin duda idéntica –y por consecuencia bien diferente de la platónica. De lo que se trata aquí es de la participación del pensamiento en el ser. Sí, en el momento en el que se introduce la palabra participación, estamos conducidos a sustituir el término ser por el término verdad. Sin embargo, en este punto hay que evitar todo malentendido. Antes, afirmaba a la ligera lo que llamaba la trascendencia del ser en relación a la verdad. Era una manera de expresarme peligrosa e inaceptable. Declararía hoy que el ser y la verdad son idénticos, pero a condición de anotar, como lo hice anteriormente, la inconmensurabilidad de la Verdad por relación a las verdades finitas a las que la ciencia nos da acceso, es decir, de reconocer explícitamente que los métodos de verificación por relación a los cuales una verdad se define, son aquí inutilizables. La trascendencia del ser es la de la verdad, por relación a esos métodos y a esas verdades parciales.

Quisiera ahora profundizar un poco en esta afirmación que, tomada en su desnudez abstracta, peligra de parecer casi ininteligible.

Si esta idea tan vaga, tan amplia de conocimiento, la sustituimos por la *lectura* comprenderemos muy bien cómo, bajo la materialidad objetiva de un texto que está ahí para todo el mundo, pueden esconderse sentidos en jerarquía que se descubren de manera sucesiva al lector cuando éste está dotado de un poder de penetración suficiente. Pero todavía es poco decir: lo que interviene aquí, es una voluntad de interpretación que no se satisface con una significación superficial y que pretende ir más allá. Lo propio de un significado es no revelarse más que a una consciencia que se abre para recibirla; es, de cualquier manera, una respuesta a una espera activa y perseverante o, más exactamente, a una exigencia. La jerarquía de sentidos está en función de esta exigencia. De esta manera, lo que es verdad en la lectura de un texto, se aplica casi exactamente a la investigación filosófica tal como yo la entiendo. Ella es, en realidad, un descifrar. Mantengamos esta palabra en memoria: el descifrar musical nos ofrece aquí un término de comparación mucho mejor. He aquí una partitura, hay en ella unos signos cuyo valor es rigurosamente definido por quien conoce el sistema común de anotación

musical; esta lectura corresponde a eso que llamo el pensamiento en general. Pero para el músico se tratará, sea o no el intérprete, de entender esos signos, de descubrir el sentido musical de eso que es a primera vista un criptograma, a saber, de proceder a una verdadera creación a la cual él está expresamente invitado o llamado por la existencia misma de la partitura. Recreación mucho más que creación. Me deberán observar: esto es la vez verdadero y falso. Para el compositor que posee una consciencia perfectamente distinta no sólo de lo que ha querido, sino de lo que ha oído, se trata tanto de recreación como de reproducción. Pero el lector o el intérprete por definición desprovisto de este conocimiento previo, no puede confiar más que en sí mismo, en el poder de su propia intuición. Deberá darse, abrirse a ese misterio del que tiene por delante sólo un delgado rastro sensible y objetivo. Esta interpretación creadora a la cual tiende es una participación efectiva en la inspiración misma del compositor.

Yo le concedo la más grande importancia a esta comparación, ya que ella permite reconocer directamente que existe un orden situado más allá de lo que sería solamente válido, pero así mismo más allá de lo que no sería solamente una subjetividad entendida en un sentido psicológico y privativo. Por ejemplo, es muy claro que hay una cierta realidad de Schumann o de Chopin que el intérprete tiene que encarnar, a la que debe someterse activamente; pero esta realidad no se deja traducir en ninguna fórmula del tipo de las que traducen las verdades particulares. Lo propio de esta realidad es poder ser ignorada o traicionada; sí, es de su esencia el no imponerse al espíritu con la fuerza obligatoria de un texto o de una ley física. Sin embargo, cometeríamos el más grave error buscar reducirla a un simple estado subjetivo, a una disposición fugaz; eso sería ignorar lo que da a una obra su peso, su valor, su virtud.

Reconocer que una comprensión puede ser irregularmente profunda, es implícitamente establecer una jerarquía en la aprehensión, que se ordena por relación al ser y no por relación a lo válido. Y sin duda debería, para hacerme entender mejor, llevar más lejos el análisis de la idea de validez: ella no es separable de su uso. Mostraríamos sin pena que el pragmatismo no ha podido introducirse más que a favor de una confusión sistemática entre verdad y validez. Hay, yo creo, en nuestro propósito, todo el interés de poner a la luz el aspecto práctico de la noción de validez, tomando, por ejemplo, como punto de referencia, la validez de un billete de ferrocarril o de teatro. El tiquete da derecho a tal trayecto o tal representación en unas condiciones determinadas. Esto es aplicable también a una fórmula científica, si bien en un sentido diferente y en unas condiciones netamente específicas. Validez implica siempre especificación. Por ello tendemos a representarnos la fórmula actuando como un clic automático que se produce cuando se aprieta un botón. He dicho explícitamente y he insistido en que eso no es más que un conocimiento degradado; pero en ese registro, dicha degradación

tiende a producirse casi inevitablemente, y sólo el ejercicio de la reflexión, es decir del espíritu en tanto que libre, puede contrapesar eficazmente eso que hay que ver como un proceso de pura y simple esclerosis espiritual.

Concentremos nuestra atención sobre estas palabras de libertad y de reflexión.

En primer lugar libertad. Una filosofía digna de ese nombre, no puede ser más que una filosofía de la libertad, y esto en un doble sentido.

De una parte, la filosofía es el pensamiento centrado sobre la Libertad, el pensamiento al que la Libertad misma se da por contenido.

De otra parte, la libertad no puede ser pensada más que por la libertad, ella misma se crea o se constituye pensándose. Hay ahí una suerte de círculo que la reflexión debe reconocer y que no tiene nada de vicioso. La noción de una libertad que se demostraría ella misma o que se extraería de un determinismo dialéctico es una pseudoidea cuya reflexión muestra que es contradictoria.

Quisiera añadir que, desde mi punto de vista, la oposición que se ha querido establecer entre la filosofía del ser y la filosofía de la libertad no puede ser mantenida. Ella no se justificaría más que sí, “percatándose” imaginativamente el ser, se le identificara a la sustancia tal como se la han representado los epígonos de la filosofía medieval o sus adversarios sistemáticos. Una auténtica filosofía del ser no es una filosofía de la cosa, como lo ha creído Renouvier. Ella no es tampoco como la pensó Bergson, una filosofía de lo estático. Es manifiesto a mis ojos que el ser, tal como lo han concebido los grandes metafísicos, trasciende la oposición entre lo estático y lo dinámico.

El término reflexión marca mejor por dónde mi pensamiento se separa del bergsonismo o por lo menos de la interpretación que sobre éste comúnmente se da. Pienso que el método filosófico es reflexivo por excelencia. Pero se trata de una reflexión de segundo grado que se ejerce menos sobre la conciencia inmediata que sobre las primeras mediaciones por las cuales la misma se constituye como experiencia.

¿Cuál será entonces el punto de partida de la búsqueda así concebida? Será la consideración de la situación fundamental en la que me encuentro situado en tanto que ser humano –no diría aún en tanto que creatura–, por mí misma condición de hombre. Aparece aquí en plena luz la oposición de una filosofía de tipo existencial con una filosofía de tipo cartesiana –a pesar que hubiera, tal vez, en Descartes, pero no en sus sucesores, indicaciones que podrían ser explotadas en un sentido no cartesiano–. Sin embargo el cogito, como el pensamiento en general de Kant, está exento de todo indicio antropológico. La situación o la condición humana no es, para este pensamiento impersonal, más que un objeto de consideración como cualquier otro. Es tratada por él como no afectándolo; pero en esta medida

se suprime como situación o como condición. Supresión ficticia; abstracción por la cual el pensamiento se prohíbe el acceso al ser. Aquí, no podría mostrar demasiado mi acuerdo con Heidegger y con Jaspers –con Jaspers sobretodo que ha puesto un acento tan fuerte sobre las situaciones límites, estas no podrían ser más que iluminadas, pero no justificadas o explicadas, si por explicación entendemos una deducción cualquiera a partir de otra cosa.

Creo que debo citar aquí una nota sobre la inquietud que es ya antigua y que considero que puede aclarar el sentido profundo de mi esfuerzo.

Inquietud metafísica. Me parece que una metafísica no es nada si ella no es el acto por el cual una inquietud se define, y parcialmente –misteriosamente– se suprime o por lo menos se transpone. Inquietud: pero, ¿cuál inquietud? En primer lugar me refiero a cualquier cosa que no sea una curiosidad. Ser curioso, es partir de un centro inmóvil, extenderse para tomar, para abrazar un objeto, que no sería más que una representación confusa y esquemática. Diría, en este sentido, que toda curiosidad tiene la mirada puesta hacia la periferia. Al contrario, estar inquieto, es no estar en su centro, es buscar su centro, su equilibrio. Ello es verdadero en todo el sentido de la palabra. Si yo estoy inquieto a propósito de la salud de uno de mis familiares, eso quiere decir que las dudas o los temores que mantengo, tiende a destruir mi estabilidad interior. En mi caso, una curiosidad tiende a convertirse cada vez más en una inquietud a medida que el objeto sobre el que se refiere, forma más parte de mí y está más estrechamente ligado a un confort moral. (Esa palabra confort, diría hoy, es insuficiente. Se trata de un cierto orden, en cuyo error no caigo más). De otra parte, una inquietud es más metafísica cuando se refiere directamente a eso que no puede ser separado de mí mismo y, en un sentido fuerte, alienado sin que yo mismo me aniquile. Hay un sentido en el cual es verdadero decir que el único problema metafísico es: ¿Qué soy yo?, y que a éste se reducen los otros, incluso el problema de la existencia de los demás. Algo poderoso y secreto me asegura que *si los otros no son, yo no soy tampoco*; que no puedo atribuirme una existencia que los otros no posean. Y aquí *no puedo* no significa que no tenga el derecho, sino más bien, *esto me es imposible*. Si los otros se me escapan, me escapo a mí mismo, ya que mi sustancia se hace de ellos.

Esta inquietud metafísica, ¿puedo decir que la siento como un estado actualmente dado? Diría simplemente que en una circunstancia determinada podría tomar conciencia de una inquietud que, en la reflexión, me aparecería como desbordando infinitamente esa circunstancia misma. Ella presenta un carácter de permanencia, no está unida a tal o cual *ahora*; mucho más, tan pronto como ella se formula, se extiende a todo los seres que puedo ver como participantes en la misma experiencia que yo. Yo no podría no mirarlos como sintiendo esta misma inquietud que es la mía.

Pero, me podrían objetar, ¿la inquietud no es considerada tradicionalmente como un estado del espíritu antifilosófico, por oposición a una serenidad, a una ataraxia que se define, además, basada en diversas escuelas filosóficas?

Yo respondería, en efecto, que si la filosofía tiende universalmente a la instauración de una cierta paz interior, de una cierta armonía, no es menos verdadero que esta no podría estar dada desde el principio. Lo que es dado es la inspiración hacia ella, y esta inspiración no puede dejar de ser sentida como nostalgia, es decir como inquietud. La única cuestión es saber si se considera deber o no poner el acento sobre este problema inicial. Cuánto más una filosofía sea conceptual, más ella creará arrojar un velo sobre eso que ve como un simple suceso subjetivo del que se trata de extraer el contenido inteligible. Esta simple cuestión de acento se revela capital a la reflexión, ya que de lo que se trata aquí es del valor, positivo o no, que será conferido al indicio personal. Una filosofía de tipo hegeliano que desarrolla hasta el final la idea de una dialéctica interna del concepto, planteará o tenderá a plantear en principio que la persona es el simple teatro accidental sobre el cual se realiza esta dialéctica en su pureza. Hará así caso omiso de la repercusión que esto puede tener sobre la vida del sujeto empírico. Esto será, por lo menos, para ella un ideal –pero que una filosofía como la que vengo delineando reniega en cuanto tal.

Estar inquieto, he dicho, es buscar su centro. Esas palabras señalan suficientemente que la filosofía tal como yo la concibo, es un desarrollo que no se continúa solamente en el interior del sujeto, considerado éste como organismo espiritual, sino incluso en la realidad de la que ese sujeto es, de cualquier manera, el fin –estaría tentado de decir el desafío, pues ese desarrollo me parece asimilable a una partida o a un drama.

Pero aún es muy poco decir. El término sujeto es bastante ambiguo; el de persona sería preferible si no lo hubiésemos abusado desgraciadamente de él en el transcurso de estos últimos años. Sería preferible en la medida en que se opone al de individuo. El pensamiento filosófico, el más auténtico, me parece situarse en la articulación de sí mismo con el otro. Algo poderoso y secreto me asegura, lo repito, que si los otros no son, yo no soy tampoco. Ciertamente, hay allí una evidencia que no ha sido universalmente reconocida y que un cierto idealismo en particular ha tendido a rechazar; queda saber el valor de los postulados sobre los cuales ella se funda. “¿Cabe la posibilidad, una vez me pregunté (*Etre et Avoir* p. 156), de plantear por derecho la prioridad del acto por el que el yo se constituye como sí mismo, por relación a aquel por el cual plantea la realidad de los otros? Esta prioridad no puede, pienso, ser afirmada más que en virtud de una confusión que, de hecho, Kant ya había detectado.

Anotaré, por cierto, que ningún texto filosófico me ha impresionado tan fuertemente como aquel en el cual la filosofía americana de W. E. Hocking, en su libro

The Meaning of God in Human Experience, ha establecido que nosotros no podemos, en realidad, concebir una comprensión del otro que no sea justamente ésa que es la nuestra y que confiere a nuestra experiencia su peso humano.

Cuando hablo de articulación de sí con el otro, hay que abstenerse de toda espacialización indebida: no podemos admitir un solo instante que exista una suerte de demarcación o de sutura entre un dominio que sería el de sí mismo y otro más que sería el del otro. Lo que nos enseñaban no solamente en la historia de la filosofía, sino en el estudio de las obras literarias, en tanto que contribuciones al conocimiento del hombre, es que el mundo de los otros se ilumina de una luz cada vez más intensa a medida que el yo dilucida más y más heroicamente sus propias tinieblas.

Una potente tradición filosófica representada en nuestros días, por ejemplo, por Léon Brunschvicg, tiende a instituir una distinción fundamental entre lo que es propio de cada uno de nosotros y lo que nos es común a todos y presenta, por lo mismo, un carácter racional. Yo no pretendo que esta distinción sea impracticable, pero me temo que es infecunda y que no afronta lo que llamaría gustosamente una sicología plana que ignora la dimensión que es por excelencia la de la vida espiritual: la profundidad.

Se encontraría aquí lo que he dicho anteriormente sobre los sentidos y las exigencias. La singularidad en nosotros o en los otros puede ser siempre etiquetada como bizarra y relegada a cualquier trastero; pero ella puede presentársenos como un llamado a una comprensión más íntima, a una comunicación más estrecha. Ella podría invitarnos, de manera insidiosa, a una renovación, a una reforma de nuestras categorías, en lugar de los caracteres comunes o generales, por el hecho mismo de que se dejan reconocer, clasificar, catalogar, y de que no hacen más que desencadenar las operaciones más habituales y, en consecuencia, las menos enriquecedoras de nuestra inteligencia. Observaría, por cierto, que este comentario, que está seguramente en la línea de las opiniones más críticas y profundas de Gide, es la base misma de todo mi teatro.

De este conjunto de reflexiones se pueden sacar las siguientes conclusiones: entre más nos esforcemos por comunicarnos con nosotros mismos, entiendo por ello eso que nos parece en principio lo más refractario a una cierta penetración inteligente, más nos liberaremos del automatismo que no es más que un anquilosamiento del juicio. De otra parte, en la medida en que nos abandonemos a este automatismo, el otro deja de ser sentido como otro, no es más que un conjunto de abstracciones con el que no podemos mantener una comunicación viva. Pero esta no-comunicación conlleva para nosotros mismos el más pesado precio que se traduce en una pérdida, a menudo, inconsciente y que se refiere a una cierta cualidad irradiante cuya presencia o ausencia es inmediatamente comprensible por intuición.

Sin duda se me opondría el hecho de que algunos seres muy espontáneos e incapaces de penetrar sus propias tinieblas interiores –suponiendo que existan– son precisamente los más aptos para comunicarse con los otros. Estoy de acuerdo. Pero esta facilidad de comunicación es atribuible en ellos a que están más o menos exentos de lo que consideramos muchas veces una superioridad: la conciencia de sí, entendida en el sentido peyorativo de la palabra inglesa *self-conscious*. La conciencia de sí está ligada a la incapacidad de comunicar o, más exactamente, al hecho de que ésta sigue estando contraída y como crispada. Se podría también distinguir dos niveles donde la comunicación se vuelve real: uno de ellos es la ingenuidad o lo que podríamos llamar el espíritu de la infancia que precede a la conciencia de sí; el otro, infinitamente más elevado, donde ella se ha triunfado a sí misma.

El filósofo tiene, por definición anterior a él, un paraíso perdido. (Este estado de espontaneidad pura es la de un niño, del niño que entra sin esfuerzo a la orilla del otro). No puede imaginar retroceder por debajo de la conciencia de sí; es, al contrario, por un esfuerzo de superación que él puede esperar restablecer unas vías de comunicación de cierta forma obturadas. Sólo falta que emprenda esta exploración con el deseo, mucho mejor, con la intención detenida de desembocar en un mundo de pensamiento concreto donde se reencontrará en el mismo plano con el prójimo, donde le verá en una luz purificada que el día pálido de unas relaciones cotidianas apenas permiten presentir.

En mi opinión, el compromiso filosófico, del que aún no se ha hablado explícitamente, se sitúa precisamente allí. Por oposición a una cierta voluntad hermética de la que el pensador contemporáneo no está, quizás, exento y que es sobre todo discernible en muchos poetas, diría que el filósofo tal como lo concibo debe ponerse en el límite de la voluntad de comunicación. Pero esto no puede ejercerse eficazmente más que por un desvío. Filosóficamente, el camino que conduce de mí al otro pasa por mis propias profundidades. Aquí no quiero decir que éstas sean accesibles en la introspección entregada a sus solas necesidades. Es, incluso, permitido pensar que en ese ámbito la parte que corresponde a la introspección es relativamente reducida. Mucho mejor es considerar la experiencia en sentido amplio, es decir, de la prueba de sí mismo al contacto de la vida; prueba que puede afectar las modalidades más diversas, algunas veces las más desconcertantes. En lo que me concierne, tengo que reconocer, como lo he dicho al comienzo, que esta prueba ha tomado, con frecuencia, la forma de creación dramática; como si mis personajes fueran los reactivos esenciales conjurados por una cierta imaginación supra-consiente para permitir a mi pensamiento, el más íntimo, proceder a una cristalización a la que no habría sido capaz de llegar solo, quiero decir entregado a los solos recursos del *ego*. Y por ello se precisa el vínculo que establecí al principio entre mi obra dramática y mi investigación propiamente filosófica.

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca - Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857

E-ISSN: 2660-9509

INFORME ESTADÍSTICO ANUAL DEL PROCESO EDITORIAL DE LA REVISTA *Cuadernos salmantinos de filosofía* STATISTICS

Vol. 48, 2021

Cuadernos salmantinos de filosofía publica un número al año (noviembre), normalmente de carácter misceláneo, si bien en los últimos volúmenes viene programando secciones monográficas. Los artículos recibidos son en su mayor parte fruto de la iniciativa de sus autores, y están sujetos a un proceso de evaluación por pares. El Consejo de Redacción puede solicitar puntualmente trabajos concretos a determinadas personas atendiendo a su prestigio personal y excelencia académica. Estos trabajos no están sometidos a evaluación. Además, puede publicar números especiales con motivo de homenajes y reuniones científicas de especial relevancia.

RECIBIDOS	29*
SOLICITADOS	1 (0,29)%
RECHAZADOS	7 (24,13%)
ACEPTADOS SIN MODIFICACIONES	0 (0%)
ACEPTADOS CON MODIFICACIONES	22 (75,86%)

*Entre ellos se incluye una Nota crítica recibida, pues las Notas críticas están sometidas a revisión.

RESEÑAS: 5

TRADUCCIONES: 1

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca - Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857

E-ISSN: 2660-9509

INSTRUCCIONES PARA AUTORES VOL. 49, 2022 ISSN: 0210-4857 / E-ISSN: 2660-9509

PRESENTACIÓN Y ALCANCE

Cuadernos salmantinos de filosofía, publicación editada por la Universidad Pontificia de Salamanca, es una Revista científica de investigación filosófica a cargo de la Facultad de Filosofía de la mencionada Universidad. Se fundó en 1974, y se viene editando desde entonces, de forma ininterrumpida, con una periodicidad anual. Pensada para el ámbito académico, publica preferentemente a profesores y doctores en filosofía, sin excluir a jóvenes investigadores que presenten un trabajo de calidad notable. Consta de tres secciones: **Estudios originales** (Artículos), **Notas Críticas** y **Reseñas**.

CONDICIONES ÉTICAS Y LEGALES

1. El manuscrito debe ser inédito, no habrá sido publicado ni en español ni en ninguna otra lengua, ni en parte ni al completo, ni habrá sido propuesto para su publicación en ningún otro medio simultáneamente.
Excepcionalmente, si así lo acuerda el Consejo de Redacción, se podrán publicar trabajos no inéditos, traducciones, etc., atendiendo al alto grado de interés académico que puedan revestir. En tal caso, será necesario el correspondiente permiso y se incluirá la referencia a la publicación original en el documento.
2. El Autor garantiza ser la única persona implicada en la creación del manuscrito, y que el mismo no infringe ni vulnera derechos morales o derechos de autor de terceros, comprometiéndose a mantener indemne al editor en caso contrario. Si el artículo está firmado por varios autores, reconocen que todos ellos han contribuido intelectualmente en su elaboración, y que todos ellos han leído y aprobado el manuscrito remitido.
3. El autor cederá en exclusiva y a título gratuito la totalidad de los derechos patrimoniales (reproducción, traducción, distribución, comunicación pública y transformación) a la Revista Cuadernos Salmantinos de Filosofía sin limitación, para todas las ediciones actuales y futuras de la misma y tanto en formato papel como en formato electrónico, para un ámbito mundial y durante todo el periodo de vigencia de los derechos de autor. El autor se compromete a no hacer uso ni en parte ni al completo, ni en español ni en cualquier otra lengua, del manuscrito publicado sin previo consentimiento por escrito del Director de la Revista, y para su uso se compromete a citar la fuente original.

4. El autor acepta las posibles revisiones del texto que hubieran de realizarse para su publicación, derivadas del proceso de revisión por pares que la revista tiene establecido, en el plazo previsto por la misma.
5. De conformidad con la Ley Orgánica 3/2018, de Protección de Datos Personales y garantía de los derechos digitales y del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, el Autor autoriza a la revista a la publicación, junto con su artículo, de los datos personales necesarios que identifican al mismo (nombre, apellidos, institución de filiación, ciudad, país, correo electrónico, número ORCID, y breve C-V, cuando proceda).

POLÍTICA DE AUTOARCHIVO Y ACCESO ABIERTO

1. **Versión auto-archivo:** Color Romeo: verde. Pre-print y Post-print versión editorial.
 - El autor está autorizado a reproducir su artículo completo, incluyendo siempre un enlace al artículo en SUMMA (Repositorio Documental UPSA).
 - Podrá archivar en web personal, repositorio institucional y repositorio temático.
 - La fuente debe estar claramente reconocida.
 - Cualquier cambio o excepción a la política de la Revista podrá ser negociado por el autor.
2. El contenido de esta Revista, excepto que se establezca de otra forma, cuenta con una **Licencia Creative Commons Reconocimiento-No Comercial-Sin obras derivadas 3.0 España**.

ARTÍCULOS

RECEPCIÓN DE MANUSCRITOS

1. Los artículos se enviarán en formato Word a la siguiente dirección electrónica: cuadernossalmantinos@upsa.es
 - Fecha límite para la recepción de artículos: **1 de mayo**.
2. Es preciso enviar tres archivos:
 - 2.1. Un archivo que incluya:
 - En la primera página: título y traducción al inglés, nombre y apellidos del autor, grado académico, cargo actual, institución a la que pertenece, direcciones electrónica y postal.
 - En la segunda página: 1. Resumen y Palabras clave. 2. Traducción al inglés (Abstract y Keywords).
 - Tercera página: cuerpo del trabajo.

- 2.2.** Otro archivo sólo contendrá el título, los resúmenes y palabras clave en el idioma del artículo y en inglés, y el cuerpo del trabajo. Se evitará cualquier tipo de indicio que de alguna manera revele la autoría del manuscrito a los evaluadores. Las citas referentes al propio autor deben aparecer anonimizadas, del modo que sigue: AUTOR y año de publicación de la obra.
- 2.3.** Además, se adjuntará, escaneada, una **carta de presentación** solicitando que el manuscrito sea tenido en consideración, y responsabilizándose de su autoría. Se señalará brevemente en qué consiste la aportación intelectual del trabajo remitido y se declarará el no envío simultáneo del manuscrito a otra revista. Firmado, lugar y fecha.

PRESENTACIÓN DE MANUSCRITOS

- 1. Idioma:** La Revista publica principalmente en español, pero acepta trabajos en inglés, francés, alemán, italiano y portugués.
- 2. Extensión:** Los artículos no deben sobrepasar las 12.000 palabras, incluyendo el resumen, las notas y la lista de referencias bibliográficas.
- 3. Estructura:**
 - 3.1.** La primera página debe contener, por este orden y según la grafía señalada:
 - **TÍTULO:** En el idioma del artículo (centrado, en mayúscula y en negrita). Separar título y subtítulo con dos puntos.
 - **TÍTULO en inglés:** (Centrado, en mayúscula y en cursiva). Irá en castellano si el idioma original es en inglés.
 - **Nombre(s) y apellido(s) del autor(es):** (en versalitas). En caso de que haya varios autores, indicar el responsable de la correspondencia. El editor respetará el nombre de pluma elegido por el autor, con el fin de evitar los problemas de recuperación en las bases de datos en lengua inglesa.
 - **Grado académico:** (Escrito de forma íntegra y en minúscula. Ej.: Doctor en, Licenciado en...)
 - **Cargo actual o puesto de responsabilidad ocupado:** (Escrito de forma íntegra y en minúscula. Ej.: Catedrático, Profesor Asociado, Director...)
 - **Institución a la que pertenece y Lugar de trabajo:** (Nombre del Centro, Departamento, Institución. Ciudad y País)
 - **Correo electrónico.**
 - **ORCID**
 - **Dirección postal.**
 - 3.2.** Segunda página:
 - **Resumen:** entre 100 y 175 palabras, en el idioma del artículo y en inglés (Abstract). Si el idioma del artículo es el inglés, entonces el resumen va también en castellano.
 - **Palabras clave:** entre 4 y 8 (por orden alfabético), en el idioma original del artículo y en inglés (Keywords). Si el idioma del artículo es en inglés, entonces las palabras clave van también en castellano. Empleo, preferentemente, de tesoro o lista de encabezamientos de materias ampliamente conocida y usada en la especialidad, por ejemplo, el *Tesoro de la Unesco*, *The Philosopher's Index Thesaurus*...
 - 3.3.** Tercera página:

- **Cuerpo del trabajo:** texto.
 - **Lista de referencias bibliográficas**, si procede.
 - En el caso de que se incluyan **agradecimientos** y/o reconocimiento de **apoyo económico y material** (especificar su naturaleza), se abrirá una llamada con asterisco al inicio del texto, mencionándose dichos datos en una nota a pie de página.
4. - Los originales se presentarán con interlineado de 1.5, con tipo de letra Times New Roman 12 en el cuerpo de texto, y 10 en las notas a pie de página.
- Las notas se numeran de modo continuo, en correspondencia con las llamadas en el texto.
 - Los apartados se numerarán del modo **1., 2., 3., ...**, y los subapartados: **1.1., 1.2., 1.3., ... , 1.1.1., 1.1.2. ...** y así sucesivamente: en versalitas sin negrita.
5. Todos los títulos (incluido “Referencias bibliográficas”) van en versalitas sin negrita.
6. **Referencias bibliográficas:** Sistema de estilo humanístico (cita-nota). Podrá combinarse con el sistema *Harvard* (autor-fecha). Puede emplearse también el estilo APA vigente.
- **Libros:**
 APELLIDO(S), Nombre o Inicial(es), *Título del libro*. Edición (siempre que no se trate de la primera). (Responsabilidad secundaria: Trad. / Pról. / Intr. de Nombre y Apellido(s)). Lugar de publicación: Editorial, año, páginas.
 Ej. CHISHOLM, Roderick M., *Theory of Knowledge*. 3ª ed. Englewood Cliffs, New Jersey: Prentice Hall, 1989, 51-52.
 - **Capítulos de libro:**
 APELLIDO(S), Nombre o Inicial(es), “Título del capítulo”. En: APELLIDO(S), Nombre o Inicial(es) (Ed.), *Título del libro*. Edición (siempre que no se trate de la primera). Lugar de publicación: Editorial, año, páginas.
 Ej. RESCHER, Nicholas, “Chisholm’s Ontology of Things”. En: HAHN, Lewis E. (Ed.), *The Philosophy of Roderick M. Chisholm*. Chicago / La Salle (Illinois): Open Court, 1997, 187-199.
 - **Revistas:**
 APELLIDO(S), Nombre o Inicial(es), “Título del artículo”. *Nombre de la revista*, volumen, número, año, páginas.
 Ej. CHISHOLM, Roderick M., “The Place of Epistemic Justification”. *Philosophical Topics*, 14, 1986, 85-92.
 - **Números monográficos de revista:**
 APELLIDO(S), Nombre o Inicial(es), “Título del artículo”. En: APELLIDO(S), Nombre o Inicial(es) (Ed.), *Título del volumen monográfico*. Número Monográfico: *Nombre de la Revista*, volumen, número, año, páginas.
 Ej. MULLIGAN, K.; SMITH, B., “A Husserlian Theory of Indexicality”. En: DAVID, M.; STUBENBERG, L. (Eds.), *Philosophische Aufsätze zu Ehren von Roderick M. Chisholm*. Número monográfico: *Grazer Philosophische Studien*, 28, 1986, 133-164.

– **Lista de referencias bibliográficas:**

En esta lista sólo se incluyen los trabajos citados en el texto. Las referencias deben ser escritas en orden alfabético por el apellido del autor. Las referencias múltiples del mismo autor se ordenarán por año de publicación, empezando por la más antigua. Si el año de la publicación también es el mismo, se han de diferenciar escribiendo una letra *a*, *b*, *c*, etc., después del año.

Ej. CHISHOLM, Roderick M., *Person and Object: A Metaphysical Study*. Chicago / La Salle (Illinois): Open Court, 1976.

—, “Epistemic Principles”. *Dialectica*, 35, 1981a, 341-345.

—, “The Paradox of Analysis: A Solution”. *Metaphilosophy*, 12, 1981b, 1-7.

—, “Epistemic Reasoning”. *Dialectica*, 36, 1982a, 169-77.

—, *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982b.

—, “On the Nature of the Psychological”. *Philosophical Studies*, 43, 1982c, 155-164.

7. Recursos electrónicos:

– **Monografías electrónicas:**

Responsabilidad principal, *Título*. [Tipo de soporte]. Edición. Lugar de publicación: Editorial, fecha de publicación (fecha de actualización / revisión, si la hubiera). <Disponibilidad y acceso, obligatorio en los documentos en línea; opcional en los demás casos>. [Fecha de consulta, obligatorio en los documentos en línea; opcional en los demás casos]. Número normalizado.

Ej. GOLDMAN, Alvin, *Knowledge in a Social World*. [en línea]. Oxford: Oxford University Press, 2012-08-05. <<http://ebookbrowse.com/alvin-i-goldman-knowledge-in-a-social-world-pdf-d377480980>>. [Consulta: 28 feb. 2013]. ISBN: 0-19-823777-4.

– **Contribuciones en monografías electrónicas:**

Responsabilidad principal, “Título de la contribución”. En: Responsabilidad principal del documento fuente, *Título del documento fuente*. [Tipo de soporte]. Edición. Lugar de publicación: Editorial, fecha de publicación (fecha de actualización / revisión, si la hubiera). <Disponibilidad y acceso, obligatorio en los documentos en línea; opcional en los demás casos>. [Fecha de consulta, obligatorio en los documentos en línea; opcional en los demás casos]. Número normalizado.

Ej. GRECO, John, “Virtue Epistemology”. En: ZALTA, Edward N., *Stanford Encyclopedia of Philosophy*. [en línea]. Stanford (Ca.): Stanford University, 1999 (actualización 12 en. 2012). <<http://plato.stanford.edu/archives/win2011/entries/epistemology-virtue/>>. [Consulta: 28 feb. 2013]. ISSN: 1095-5054.

– **Artículos en revistas electrónicas:**

Responsabilidad principal, “Título del artículo”. *Título de la revista*. [Tipo de soporte], volumen, número, año, páginas correspondientes si las hubiera. <Disponibilidad y acceso, obligatorio en documentos en línea; opcional en los demás casos>. [Fecha de consulta, obligatorio en los documentos en línea; opcional en los demás casos]. Número normalizado.

Ej. CORDERO, Ronald, "Kant and the Expression of Imperatives". *Sorites*. [en línea], 20, 2012-03-20, 117-124. <http://www.sorites.org/issue_20/cordero.htm>. [Consulta: 28 feb. 2013]. ISSN: 1135-1349.

- **DOI:** Identificar en caso de que el documento lo posea.

Ej. BATTALY, Heather, "Sosa's Reflective Knowledge: How damaging is Epistemic Circularity?". *Synthese*, 188, 2, sept. 2012, 289-308. DOI: 10/1007/s11229-011-9928-2.

8. Citas textuales:

- Todo dato o idea tomado de otro autor debe ser escrupulosamente tratado, y citado de forma conveniente.
- Cuando se citen textos en un idioma distinto al del manuscrito en el que se trabaja, deberá hacerse en el idioma original del texto. Podrá ofrecerse una traducción del mismo, bien a continuación del original entre paréntesis, bien en nota.
- Todas las citas textuales directas incluidas en el cuerpo del texto irán entre "comillas voladas", y se compondrán en letra redonda.
- Si el texto de la cita es largo (más de tres-cuatro líneas), se situará en un párrafo aparte en cuerpo menor (Times New Roman 10) con interlineado sencillo y sangría a la izquierda de 1 cm, y sin comillas.

9. En caso de que los manuscritos contengan textos en griego, caracteres de otras lenguas o fórmulas lógicas, se empleará una fuente Arial Unicode.

10. Algunas abreviaturas:

Capítulo(s): cap. (caps.) / Confróntese: cfr. / Colaboradores: Cols. / Compilador(es): Comp. (Comps.) / Coordinador(es): Coord. (Coords.) / Editor(es): Ed. (Eds.) / Segunda edición: 2ª ed. / Edición revisada: Ed. Rev. / En el mismo lugar: *ibid.* / et altera: et al. / Introducción: Intr. / Obra citada: *op. cit.* / Página(s): p. (pp.) / Parágrafo: § / Parte: Pt. / Prólogo: Pról. / Reporte técnico: Rep. Téc. / Sin fecha: s.f. / Suplemento: Supl. / Traducción de: Trad. / Véase: v. / Volumen(es): Vol. (Vols.)

PROCESO DE EVALUACIÓN POR PARES Y PUBLICACIÓN

1. Habrá acuse de recibo del original por parte de la Revista.
2. El editor hará una primera valoración del manuscrito, evaluando su pertinencia y calidad. Se verificará que cumpla las normas para la presentación de originales que figuran en las instrucciones a autores, y se procederá al cotejo de la documentación requerida. Una vez conforme lo anterior, se remitirá el artículo a los evaluadores.
3.
 - Los manuscritos, tanto de artículos como de notas críticas, serán sometidos a una evaluación externa y anónima, doble ciego, por parte de dos especialistas en la materia objeto de estudio. En caso de discrepancia absoluta entre los dos evaluadores, se recurrirá a un tercer evaluador. Se acepta que el autor recomiende o rescuse el nombre de potenciales revisores, eludiendo a aquel que de alguna manera pueda hacer una crítica sesgada ya sea a favor o en contra.
 - La guía e instrucciones remitidas a los revisores aparecerán publicadas tanto en el volumen impreso como en la página web de la revista.
 - Los artículos solicitados expresamente por la Redacción no estarán sujetos a revisión.

4. Se evaluará el contenido, la calidad metodológica y el estilo. Los temas deberán ajustarse estrictamente a materias de carácter filosófico.
5. Se informará al autor sobre los resultados de la evaluación de los expertos.
6. A tenor de los informes recibidos, el Consejo de Redacción determinará la conveniencia de su aceptación definitiva. La evaluación positiva por parte de los pares no compromete a la Revista a la publicación del texto. También decidirá si son seleccionados para la publicación del volumen del año en curso o de uno ulterior. Esto último dependerá de la fecha de recepción del original, del volumen de trabajos recibidos, y de motivos tales como la publicación de ediciones monográficas, volúmenes de homenaje... La decisión del Consejo de Redacción es inapelable.
7. Una vez aceptado el artículo o la nota crítica se le comunicará al autor, y se le facilitará un documento de cesión de derechos.
8. El autor recibirá las galeras del texto con el fin de que corrija errores tipográficos y de carácter menor. No podrá haber modificación sustancial alguna. El texto revisado deberá ser devuelto en un plazo no superior a diez días desde su recepción.
9. El volumen se publica en noviembre.
10. El Consejo de Redacción se reúne al menos una vez al año.

PERÍODOS DE CARENIA

Se establece un período de carencia o espera de un año para los siguientes supuestos:

- La publicación de un artículo en un libro u otra publicación
- La publicación de un artículo por un mismo autor.

NOTAS CRÍTICAS

1. Encabezamiento: Título relativo al contenido y traducción al inglés, Resumen / Abstract (no superior a 100 palabras) y Palabras clave / Key Words (no más de 5).
2. Descripción de la obra objeto de crítica:
APELLIDO(S), Nombre o Inicial(es), Título del libro. (Responsabilidad secundaria).
Lugar de publicación: Editorial, año, número de páginas.
Ej. CHISHOLM, Roderick M., *On Metaphysics*. Minneapolis: University of Minnesota Press, 1989, 173 pp.
3. Al inicio del manuscrito indicar el nombre y apellidos del autor, grado académico, cargo actual, institución a la que pertenece, direcciones electrónica y postal. ORCID
4. Se enviarán en formato Word a la siguiente dirección electrónica: cuadernosalmantinos@upsa.es
5. Fecha límite para la recepción: **1 de mayo**.
6. **Idioma:** Español, inglés, francés, alemán, italiano y portugués.
7. Extensión: Oscilarán entre un mínimo de 2.500 y un máximo de 5000 palabras.
8. No se aceptarán notas críticas de libros con más de cuatro años de antigüedad.

9. Se aceptan notas a pie de página. Seguir las indicaciones formales del apartado “Referencias bibliográficas”.
10. Interlineado 1.5, letra Times New Roman 12 en el cuerpo del trabajo, y 10 en las notas a pie de página. Las notas se numeran de modo continuo, en correspondencia con las llamadas en el texto.

RESEÑAS

1. Encabezamiento: Título relativo al contenido y traducción al inglés.
2. Descripción de la obra reseñada. Seguirá el mismo modelo que el indicado en el apartado anterior sobre las Notas críticas
3. Al inicio del manuscrito indicar el nombre y apellidos del autor, grado académico, cargo actual, institución a la que pertenece, direcciones electrónica y postal. ORCID
4. Se enviarán en formato Word a la siguiente dirección electrónica: cuadernosalmantinos@upsa.es
5. Fecha límite para la recepción: **1 de septiembre**.
6. **Idioma:** Español, inglés, francés, alemán, italiano y portugués.
7. **Extensión:** Oscilarán entre un mínimo de 1.000 y un máximo de 2.500 palabras.
8. No se aceptarán reseñas de libros con más de cuatro años de antigüedad.
9. Se aceptan notas a pie de página.
10. Listado de referencias bibliográficas: Seguir las indicaciones formales del apartado “Referencias bibliográficas”.
11. Interlineado 1.5, letra Times New Roman 12 en el cuerpo del trabajo, y 10 en las notas a pie de página. Las notas se numeran de modo continuo, en correspondencia con las llamadas en el texto.

TRADUCCIONES

Solo se reciben traducciones al español de textos escritos originalmente en otros idiomas. Es importante que el traductor se asegure de que los derechos de autor del texto que se traduce permiten su traducción y publicación en la Revista. Esta información debe ser anexada en la propuesta para pasar la primera revisión por parte del Comité Editorial.

El manuscrito se presentará del siguiente modo:

1. **Título en español del texto traducido.** El título original del texto traducido se indicará en una nota a pie de página al final del título en español.
2. Fecha de la primera edición del texto traducido o, en su defecto, fecha aproximada de composición precedida por "circa".
3. **Nombre completo del autor del texto traducido.**
4. Debajo se escribirá: “**Traducción e introducción** (si es el caso) **de**” seguido del **nombre completo del traductor**, grado académico más alto obtenido, afiliación institucional, dirección postal y de correo electrónico, número ORCID.
5. Si la traducción va precedida de una Introducción, esta última deberá incluir un

Resumen y Palabras Clave.

- 5.1.** Resumen: El resumen debe tener entre 150 y 200 palabras en un solo párrafo. El resumen no debe contener abreviaturas indefinidas o referencias no especificadas.
- 5.2.** Palabras clave: una lista de cinco a siete descriptores que se utilizan con fines de indexación.

Formato del texto:

- 1.** El archivo debe enviarse en MS Word 2003-2010 o .docx.
- 2.** Fuente: Times New Roman 12 puntos en el cuerpo del texto; Times New Roman 10 puntos en notas al pie.
- 3.** Espaciado: 1.5.
- 4.** Las notas al pie del texto estarán numeradas consecutivamente.

INSTRUCTIONS FOR AUTHORS

ETHICAL CONDITIONS

- 1.** The manuscript must be unpublished, it will not have been published in Spanish or in any other language, neither in part nor in its entirety, nor will it have been proposed for publication in any other medium simultaneously.
Exceptionally, if so agreed by the Editorial Board, unpublished works, translations, etc. may be published, taking into account the high degree of academic interest that they may have. In this case, the corresponding permission will be necessary and the reference to the original publication will be included in the document.
- 2.** The Author guarantees to be the only person involved in the creation of the manuscript, and that it does not infringe or violate moral rights or copyrights of third parties, promising to hold the publisher harmless otherwise. If the article is signed by several authors, they acknowledge that all of them have contributed intellectually to its preparation, and that all of them have read and approved the submitted manuscript.
- 3.** The author will assign exclusively and free of charge all the patrimonial rights (reproduction, translation, distribution, public communication and transformation) to the Journal Cuadernos Salmantinos de Filosofía without limitation, for all current and future editions of the same and both in paper format as in electronic format, for a global scope and during the entire period of validity of copyright. The author agrees not to make use of the published manuscript in part or in full, in Spanish or in any other language, without the prior written consent of the Director of the Journal, and for its use he / she agrees to cite the original source.
- 4.** The author accepts the possible revisions of the text that should be carried out for its publication, derived from the peer review process that the journal has established, within the term established by the journal.
- 5.** In accordance with Organic Law 3/2018, on the Protection of Personal Data and guarantee of digital rights and Regulation (EU) 2016/679 of the European Parliament and of the Council, of April 27, 2016, regarding the protection of natural persons with regard to the processing of personal data

and the free circulation of these data, the Author authorizes the journal to publish, together with its article, the necessary personal data that identify it (name, surname, Institution of affiliation, city, country, email, ORCID number, and brief CV, when applicable).

SELF-ARCHIVING AND OPEN ACCESS POLICY

1. Self-archiving version: Dulcinea. ROMEO colour: green. Pre-print and post-print publishing version.
 - The author is authorized to reproduce his or her article in full; he or she must always include a link to the article in SUMMA (the document repository of the Pontifical University of Salamanca).
 - The file can be stored on a personal website, institutional repository or subject repository.
 - The source must be clearly recognized.
 - Any change or exception to the journal's policy may be negotiated by the author.
2. Unless established in another form, the journal's content is covered by a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Spain licence.

ARTICLES

RECEIPT OF MANUSCRIPTS

1. Articles are to be sent in Word format to the following email address: cuadernosalmantinos@upsa.es
 - Deadline for the receipt of articles: **1 May**.
2. It is necessary to send three files:
 - 2.1. One file that includes:
 - On the first page: title and translation into English, first name and last name of the author, academic qualification level, current post, institutional affiliation and email and postal addresses.
 - On the second page: 1. Abstract and key words. 2. English translation of the Abstract and Keywords.
 - Third page: main body of the work.
 - 2.2. Another file will contain only the title, the abstracts and key words in the original language and in English, along with the main body of the

work. Any kind of indication that reveals authorship of the manuscript by any means to the reviewers must be avoided. Citations relating to the author must be anonymized in the following manner: AUTHOR and publication year of the work.

- 2.3. Furthermore, a scanned **cover letter** requesting that the manuscript be considered and asserting authorship must be included. This will briefly describe the intellectual contribution of the work submitted and declare that the manuscript has not been simultaneously submitted to another journal. The letter will be signed with place and date of signature.

PRESENTATION OF MANUSCRIPTS

1. **Language:** The journal is published principally in Spanish, but it accepts works in English, French, German, Italian and Portuguese.
2. **Manuscript length:** Articles must not exceed 12,000 words, including the abstract, notes and list of bibliographic references.
3. **Structure:**
 - 3.1. The first page should contain, in this order and according to the specified formatting:
 - **TITLE:** In the language of the article (centred, in capital letters and bold). Separate the title and subtitle with a colon.
 - **TITLE** in English or Spanish if the paper is writing in English: (centred, in capital letters and italics).
 - First name(s) and last name(s) of the author(s): (centred and in block capitals). Where there are several authors, indicate the corresponding author. The editor will respect the pen name chosen by the author, in order to avoid problems in locating the work in English-language databases.
 - Academic qualification level (Written in full and in lower case. For example: Doctor of; Graduate in)
 - Current post or position of responsibility occupied: (Written in full and in lower case. For example: Professor; Associate Professor; Director)
 - Institutional affiliation and place of work: (Name of the department, faculty and institution. City and Country)
 - ORCID
 - E-mail address
 - Postal address

3.2. Second Page

- **Abstract:** between 100 and 175 words, in the original language and in English. If the language of the article is English, then the abstract is also in Spanish.
- **Keywords:** between 4 and 8 (in alphabetical order), in the original language of the article and in English. If the language of the article is English, then the keywords are also in Spanish. Preferably use a thesaurus or list of subject headings that are widely known and used in the specialism—for example, the *Unesco Thesaurus*; *The Philosopher's Index Thesaurus*.

3.3. Third page:

- **Main body of the work:** text.
 - **List of bibliographic references,** if applicable.
 - In the event that **acknowledgements** or recognition of **financial and material support** are included (the nature of which must be specified), an asterisk will be placed at the end of the title, linking to a footnote with the details.
4. Manuscript originals are to be presented with a line spacing of 1.5 and Times New Roman font (size 12 for the main body of text, and size 10 for the footnotes).
- Footnotes are to be numbered continuously and correspond to the numbering placed in the text.
 - Sections should be numbered as **1., 2., 3., etc.,** and subsections as **1.1., 1.2., 1.3., etc.;** **1.1.1., 1.1.2. etc.** and so on, with block capitals and without bold . They will be in capital letters and bold.
5. All titles (including “Bibliographic References”) are in small capitals without boldface.
6. **Bibliographic References:** The journal uses the notes and bibliography style. This can be combined with the Harvard system(author-date). The current APA style can also be used.
- **Books:**
 LAST NAME(S), First Name or Initial(s), *Title of the Book*. Edition (if it is not the first). (Secondary authorship: Trans. / Prol. / Intr. by First Name and Last Name (s)). Place of publication: Publisher, year, pages.
 Example: CHISHOLM, Roderick M., *Theory of Knowledge*. 3rd ed. Englewood Cliffs, New Jersey: Prentice Hall, 1989, 51-52.
 - **Books chapters:**
 LAST NAME(S), First Name or Initial(s), “Title of the chapter”. In: LAST NAME(S), First Name or Initial(s) (Ed.), *Title of the Book*. Edition (if it is not the first). Place of publication: Publisher, year, pages.

Example: RESCHER, Nicholas, "Chisholm's Ontology of Things". In: HAHN, Lewis E. (Ed.), *The Philosophy of Roderick M. Chisholm*. Chicago / La Salle (Illinois): Open Court, 1997, 187-199.

– **Journals:**

LAST NAME(S), First Name or Initial(s), "Title of the article". *Name of the journal*, volume, issue number, year, pages.

Example: CHISHOLM, Roderick M., "The Place of Epistemic Justification". *Philosophical Topics*, 14, 1986, 85-92.

– **Special issues of journals:**

LAST NAME(S), First Name or Initial(s), "Title of the article". In: LAST NAME(S), First Name or Initial(s) (Ed.), *Title of special issue*. Special issue: *Name of the journal*, volume, issue number, year, pages.

Example: MULLIGAN, K.; SMITH, B., "A Husserlian Theory of Indexicality". In: DAVID, M.; STUBENBERG, L. (Eds.), *Philosophische Aufsätze zu Ehren von Roderick M. Chisholm*. Special issue: *Grazer Philosophische Studien*, 28, 1986, 133-164.

– **List of bibliographic references:**

This list should only include the works cited in the text. References should be written in alphabetical order according to author last name. Multiple references from the same author are sorted by year of publication, starting with the oldest. If the year of publication is also the same, works should be differentiated by writing a letter *a*, *b*, *c*, etc., after the year.

Example:

CHISHOLM, Roderick M., *Person and Object: A Metaphysical Study*. Chicago / La Salle (Illinois): Open Court, 1976.

- , "Epistemic Principles". *Dialectica*, 35, 1981a, 341-345.
- , "The Paradox of Analysis: A Solution". *Metaphilosophy*, 12, 1981b, 1-7.
- , "Epistemic Reasoning". *Dialectica*, 36, 1982a, 169-77.
- , *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982b.
- , "On the Nature of the Psychological". *Philosophical Studies*, 43, 1982c, 155-164.

7. Electronic resources:

– **Electronic monographs:**

Primary author, *Title*. [Format type] Edition. Place of publication: Publisher, date of publication (date of update/revision, if any). <Availability and access; obligatory for online documents and optional in other cases>.

[Dated consulted; obligatory for online documents and optional in other cases]. ISBN.

Example: GOLDMAN, Alvin, *Knowledge in a Social World*. [online]. Oxford: Oxford University Press, 2012-08-05. <<http://ebookbrowse.com/alvin-i-goldman-knowledge-in-a-social-world-pdf-d377480980>>. [Consulted: 28 Feb. 2013]. ISBN: 0-19-823777-4.

– **Contributions in electronic monographs:**

Primary author, “title of the contribution”. In: Primary author/editor of the source document, *Title of the source document*. [Format type]. Edition. Place of publication: Publisher, date of publication (date of update/revision, if any).

<Availability and access; obligatory for online documents and optional in other cases>. [Dated consulted; obligatory for online documents and optional in other cases]. ISBN.

Example: GRECO, John, “Virtue Epistemology”. In: ZALTA, Edward N., *Stanford Encyclopedia of Philosophy*. [online]. Stanford (Ca.): Stanford University, 1999 (updated on 12 Jan. 2012). <<http://plato.stanford.edu/archives/win2011/entries/epistemology-virtue/>>. [Consulted: 28 Feb. 2013]. ISSN: 1095-5054.

– **Articles in electronic journals:**

Primary author, “Title of the article”. *Title of the journal*. [Format type], volume, issue number, year, corresponding page numbers, if there are any. <Availability and access; obligatory for online documents and optional in other cases>. [Date consulted; obligatory for online documents and optional in other cases]. ISBN.

Example: CORDERO, Ronald, “Kant and the Expression of Imperatives”. *Sorites*. [online], 20, 2012-03-20, 117-124. <http://www.sorites.org/issue_20/cordero.htm>. [Consulted: 28 Feb. 2013]. ISSN: 1135-1349.

– **DOI:** Include this if the document has one.

Example: BATTALY, Heather, “Sosa’s Reflective Knowledge: How damaging is Epistemic Circularity?”. *Synthese*, 188, 2, Sept. 2012, 289-308. DOI: 10/1007/s11229-011-9928-2.

8. Quotations:

- Any information or idea taken from another author must be handled carefully and cited in a suitable form.
- When quoting texts in a language other than that of the submitted manuscript, the quotation should be in the original language of the quoted text. A translation of it may be provided either after the original in parentheses or in a note.

- All direct quotations (less than three lines) included in the text will appear between double quotation marks and be set in roman typeface.
 - If the text of the quotation is long (more than three lines), it is presented as a separate paragraph with a smaller font size (Times New Roman 10), with single line spacing, indent left 1 cm and without quotation marks.
9. Where manuscripts contain texts in Greek, characters from other languages or logical formulas, an Arial Unicode font will be used.

10. Some abbreviations:

Chapter(s): chap. (chaps.) / Compare: cf. / Collaborators: Collabs. / Compiler(s): Comp. (Comps.) Coordinator(s): Coord. (Coords.) / Editor(s): Ed. (Eds.) / Second edition: 2nd ed. / Revised Edition Rev. Ed. / In the same place: *ibid.* / et altera: et al. / Introduction: Intr. / In the work cited: *op. cit.* / Page(s): p. (pp.) / Paragraph: § / Part: Pt. / Prologue: Prpl. / Technical report: Tec. Rep. / No date: n.d. / Supplement: Suppl. / Translation by: Trans. / See: See / Volume(s): Vol. (Vols.)

PEER REVIEW PROCESS AND PUBLICATION

1. Acknowledgement of receipt of the original will be made by the journal.
2. The editor will make an initial assessment of the manuscript, evaluating its relevance and quality. The editor will verify that it meets the standards for the presentation of originals that are contained in the instructions for authors and will then collate the required documentation. Once these requirements have been met, the article will be sent to the reviewers.
3. The manuscripts of both articles and critical notes will undergo an external and anonymous double-blind review by two specialists in the field under study. It is acceptable for the author to recommend or object to the naming of potential reviewers, avoiding any who may make an evaluation that is either biased in favour or prejudiced against.
 - The guide and instructions sent to the reviewers will be published both in the printed volume and on the website of the magazine.
 - Articles requested expressly by the Editorial Board will not be subject to review.
4. Content, methodological quality and style will be evaluated. Subjects must strictly conform to matters of a philosophical nature.
5. The author will be informed about the results of the reviewers' evaluations.
6. Based on the reports received, the Editorial Board will determine the appropriateness of definitively accepting the article. A positive evaluation by the peer reviewers does not commit the journal to publishing the text. A decision will also be made as to whether texts will be selected for publication

in the current year's volume or in the subsequent volume. This decision will depend on the date of receipt of the original, on the volume of works received, and on factors such as the publication of special issues, tribute volumes and so forth. Decisions made by the Editorial Board cannot be appealed.

7. Once an article or critical note has been accepted, the author will be notified and sent a document for rights transferral.
8. The author will receive the proofs of the text in order to correct minor and typographical errors. No substantial modification may be made. The revised text should be returned in no more than two weeks from the date of its receipt.
9. Volumes are published in November. Authors of articles will receive a link to download a free copy of his or her contribution to the journal volume.
10. The Editorial Board meets at least once a year.

WAITING PERIOD

A grace or waiting period of one year is established for the following cases:

- The publication of an article in a book or other publication
- The publication of an article by the same author.

CRITICAL NOTES

1. Header Title and translation into English, Abstract (this will consist of the bibliographic reference of the book under study) and Key Words (no more than 5).
2. Documents will be sent in Word format to the following email address: cuadernosalmantinos@upsa.es
3. Deadline for receipt: **1 May**.
4. **Language:** Spanish, English, French, German, Italian and Portuguese.
5. **Manuscript length:** *Critical notes* range in length between a minimum of 2,500 and a maximum of 5,000 words.
6. Critical notes of books that are more than four years old will not be accepted.
7. Footnotes are acceptable. Follow the formatting instructions in the of the “Bibliographic references” section.
8. Use 1.5 line-spacing and Times New Roman (size 12 for the body of the work and size 10 for the footnotes). Footnotes are to be numbered continuously and correspond to the numbering placed in the text.

9. Description of the work the under study:
LAST NAME(S), First Name or Initial(s), *Title of the Book*. (Secondary authorship). Place of publication: Publisher, year, number of pages.
Example: CHISHOLM, Roderick M., *On Metaphysics*. Minneapolis: University of Minnesota Press, 1989, 173 pp.
10. At the end of the manuscript, indicate the first name and last name of the author, academic qualification level, current position, institutional affiliation and email and postal addresses.
11. Volumes are published in November. Authors of articles will receive a link to download a free copy of his or her contribution to the journal volume.

REVIEWS

1. Heading: Title related to the content and English translation.
2. Documents will be sent in Word format to the following email address: cuadernosalmantinos@upsa.es
3. Deadline for receipt: **1 September**.
4. **Language:** Spanish, English, French, German, Italian and Portuguese.
5. **Manuscript length:** *Reviews* range in length between a minimum of 1,000 and a maximum of 2,500 words.
6. Reviews of books that are more than four years old will not be accepted.
7. Footnotes are acceptable. Follow the formatting instructions in the of the “Bibliographic references” section.
8. Use 1.5 line-spacing and Times New Roman (size 12 for the body of the work and size 10 for the footnotes). Footnotes are to be numbered continuously and correspond to the numbering placed in the text.
The description of the reviewed work will follow the same model as shown in the previous section on critical notes.
9. At the end of the manuscript, indicate the first name and last name of the author, academic qualification level, current position, institutional affiliation and email and postal addresses.
10. Volumes are published in November. Authors of articles will receive a link to download a free copy of his or her contribution to the journal volume.

TRANSLATIONS

Only translations into Spanish of texts originally written in other languages are received. It is important to be clear that the copyright of the text being translated allows its translation and publication in the Journal. This information must be annexed in the proposal to pass the first review by the Editorial Board.

The **manuscript** should be presented so as to include:

1. **Title of the translated text in Spanish.** A footnote should be inserted at the end of the title in Spanish. The original title of the translated text should be included in the footnote.
2. Date of the first edition of the translated text, or, in the absence of this, approximate date of composition preceded by 'circa'.
3. **Full name of the author of the translated text.**
4. Below it will be written: "**Translation (or Introduction) of**" followed by the **full name** of the translator, highest academic degree attained, institutional affiliation, postal and email address, ORCID number.
5. If the translation is preceded by an Introduction, the latter should include an Abstract and Keywords.
 - 5.1. Abstract: The abstract should have between 150 and 200 words in a single paragraph. The abstract should not contain any undefined abbreviations or unspecified references.
 - 5.2. Keywords: A 5-7 list of descriptors which are used for indexing purposes.

TEXT LAYOUT

1. The file must be sent in MS Word 2003-2010 or .docx.
2. Font: Times New Roman 12 points in the body of the text; Times New Roman 10 points in footnotes.
3. Spacing: 1.5.
4. Footnotes to the text are numbered consecutively.

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca - Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857

E-ISSN: 2660-9509

INSTRUCCIONES PARA LOS REVISORES VOL. 49, 2022

Estimado revisor:

Con el fin de garantizar la calidad científica de nuestra Revista, solicitamos su cooperación con el Consejo de Redacción. Estimamos su juicio académico como una aportación imprescindible para fundamentar la decisión final respecto a la publicación del manuscrito remitido.

Los revisores son elegidos de acuerdo con los siguientes criterios: (i) reconocida solvencia en el área o tema de estudio del artículo; (ii) que el revisor no sea el objeto de la crítica del autor; (iii) imparcialidad respecto al tema, metodología y contenido del artículo; (iv) que el revisor no sea colega del autor o que exista algún tipo de conflicto de intereses que pueda comprometer la neutralidad de su valoración.

El sistema de arbitraje empleado por *Cuadernos salmantinos de filosofía* es el de doble ciego, en el que la identidad de autor y revisor es desconocida por ambas partes.

Tanto si acepta como si rehúsa evaluar el artículo sometido a su consideración ha de comunicarlo al editor en el plazo de una semana. La dirección electrónica para comunicarse con el editor de la revista es la siguiente: cuadernossalmantinos@upsa.es

Si acepta revisar el artículo que le hemos enviado, por favor complete y envíe a la dirección anterior los dos documentos que siguen: “Compromisos adquiridos por el revisor” y “Guía para la revisión”.

En nombre de la Dirección de *Cuadernos salmantinos de filosofía*, agradecemos profundamente su tiempo y su esfuerzo.

El Consejo de Redacción

CUADERNOS SALMANTINOS DE FILOSOFÍA

Revista anual de la Facultad de Filosofía

Universidad Pontificia de Salamanca - Salamanca / España

Vol. 48, 2021 ISSN: 0210-4857

E-ISSN: 2660-9509

COMPROMISOS ADQUIRIDOS POR EL REVISOR VOL. 49, 2022

DON/DOÑA _____,
con D.N.I. nº _____ y domicilio en _____,
correo electrónico _____, tfno. _____
en calidad de revisor del artículo “ _____

_____”, incluido en el proceso editorial del volumen _____ de la Revista *Cuadernos salmantinos de filosofía* de la Facultad de Filosofía de la Universidad Pontificia de Salamanca, con ISSN: 0210-4857, Depósito Legal: S. 643-1974, hago constar que:

Me comprometo a realizar la evaluación del manuscrito de forma crítica, honesta y constructiva, de acuerdo con las siguientes normas:

- Los revisores realizarán su informe en el plazo máximo de un mes y con el formato del formulario de revisión que sigue a esta página. Junto con el formulario de revisión enviarán firmado este documento de compromisos adquiridos.
- Cualquier circunstancia que le impidiera cumplir con el citado plazo será comunicada al editor de forma inmediata.
- Los revisores asumen un compromiso de confidencialidad de forma que no podrán divulgar ni durante todo el proceso de evaluación ni con posterioridad al mismo información alguna del manuscrito objeto de revisión.
- En caso de que considere que pudiera existir algún conflicto de intereses por su excesiva proximidad académica, familiar, etc... con el autor del manuscrito deberá ponerlo en conocimiento del editor.
- Si sospechara que el artículo pudiera haber incurrido en plagio deberá igualmente ponerlo en conocimiento del editor a los efectos de su comprobación con los medios técnicos de los que dispone la revista.

Política de privacidad de la Universidad en cumplimiento del Reglamento UE 2016/679 de Protección de Datos:

Responsable del Tratamiento: Universidad Pontificia de Salamanca. c/ Compañía nº 5. 37002-Salamanca

Finalidad: Gestión de mi participación como evaluador de publicaciones

Legitimación y conservación: Ejecución del contrato o relación jurídica y consentimiento del interesado. Los datos proporcionados se conservarán durante los años necesarios para cumplir con las obligaciones legales o hasta que el interesado solicite su supresión.

Destinatarios de los datos y transferencia internacional: No se cederán datos a terceros, salvo obligación legal.

Derechos de los interesados: Acceder, rectificar y suprimir los datos, así como otros derechos relacionados en la información adicional.

Información adicional: Puedes consultar la información adicional y detallada sobre Protección de Datos y nuestra Política de Privacidad en nuestra página web <https://www.upsa.es/protecciondatos>

Delegado de Protección de Datos: Puedes contactar con el DPO de la Universidad Pontificia de Salamanca en la siguiente dirección: dpo@upsa.es.

Fecha y firma _____

GUÍA PARA LA REVISIÓN

Datos:

Título del artículo:

Fecha del dictamen:

Nombre y apellidos del revisor:

Institución a la que pertenece:

Dirección postal:

e-mail:

DICTAMEN: Aceptado Aceptado, sujeto a revisión Rechazado

FORMULARIO

Puntúe de 0 a 10 los siguientes ítems. Si lo considera conveniente, puede hacer un comentario después de cada ítem. Además, tendremos especialmente en cuenta el juicio que usted emita en el apartado “Valoración y comentarios para el autor”.

1. El título representa adecuadamente el contenido al que se refiere el artículo:
2. El título incluye el mayor número posible de palabras clave para facilitar su posterior indización:
3. El resumen representa correctamente la información cualitativa contenida en el documento:
4. La lista de palabras clave enumera adecuadamente las cuestiones tratadas:
5. El trabajo es original. Queda claro en qué se diferencia de estudios previos y qué aporta a lo ya publicado:
6. El trabajo es significativo y relevante:
7. El trabajo presenta claramente sus objetivos:
8. Los contenidos del trabajo se adecúan a los objetivos enunciados:
9. El artículo desarrolla un hilo argumental continuo, claro y ordenado:
10. Las conclusiones del artículo se encuentran adecuadamente argumentadas y justificadas:
11. El autor muestra dominio del tema, y lo acredita con una discusión directa con la bibliografía relevante:
12. Las referencias bibliográficas se encuentran debidamente actualizadas:
13. Las ediciones utilizadas son de reconocida solvencia académica:
14. Se reconocen de manera clara ante el lector las ideas o perspectivas que se han tomado de otras fuentes:
15. La pulcritud formal y la claridad prosística del manuscrito son adecuadas:

VALORACIÓN Y COMENTARIOS PARA EL AUTOR

CUADERNOS SALMANTINOS DE FILOSOFÍA

Annual journal of the Faculty of Philosophy
Universidad Pontificia de Salamanca - Salamanca / Spain
Vol. 48, 2021 ISSN: 0210-4857
E-ISSN: 2660-9509

GUIDE AND INSTRUCTIONS FOR REVIEWERS

Dear reviewer,

To ensure the scholarly quality of our journal, we request your cooperation with the Editorial Board. We view your academic judgement as making an essential contribution to the basis of the final decision on the publication of the submitted manuscript.

Reviewers are selected in accordance with the following criteria:

(i) recognized competence with regard to the article's area or subject of study; (ii) the reviewer is not the focus of the author's criticism; (iii) impartiality with regard to the subject, methodology and content of the article;

(iv) the reviewer is not a colleague of the author and there is no type of conflict of interests that may compromise the neutrality of the review. Sight should not be lost of the fact that reports must be clear, provide grounds for the judgement issued and offer respectful and constructive comments that are useful for the author. If you believe that there is any reason that would prevent you from properly evaluating the work, please advise the journal that you will be unable to participate in the review.

Cuadernosalmantinos de filosofia uses a system of double-blind peer review, in which the author's and the reviewer's identities are unknown to one another. Please do not forget the need for confidentiality. The article submitted for your consideration is the private property of its author. Prior to the manuscript's publication, do not share or make use of the ideas contained in the original article, and do not make any public assessment of it.

If you deem it appropriate to make a comment that is only for the attention of the Editorial Board, please let us know.

Whether you accept or reject the opportunity to review this article, please send your response to the invitation to the following email address: cuadernosalmantinos@upsa.es. We request that you send your report within a maximum of one month. If this is not possible, and with the goal of facilitating our procedures, we would be grateful if you could let us know as soon as possible.

On behalf of the management of *Cuadernos salmantinos de filosofia*, we greatly appreciate your time and effort.

The Editorial Board

GUIDE FOR REVIEWERS

FORM

Give a score of between 0 and 10 to the following items. You may make a comment after each item if you consider it appropriate to do so. Please also pay special attention to the judgement that you make in the section entitled “assessment and comments for the author.”

1. The title adequately represents the article’s content:
2. The title includes the greatest possible number of keywords to facilitate indexing at a later stage:
3. The abstract accurately represents the qualitative information contained in the document:
4. The list of keywords adequately describes the issues covered:
5. The article is original. How it is different to previous studies and what it contributes to what has already been published can be clearly ascertained:
6. The article is significant and relevant:
7. The article clearly sets out its objectives:
8. The content of the article corresponds to its objectives:
9. The article provides a continuous, clear and organized line of argument:
10. The article’s conclusions are properly argued and justified:
11. The author exhibits mastery of the subject, and this is accredited by a direct discussion with relevant bibliography:
12. The references are sufficiently up to date:
13. The editions used have a recognized academic reputation:
14. Ideas or perspectives that have been taken from other sources are clearly acknowledged to the reader:
15. In terms of clarity of prose and formal aspects the manuscript is sufficient:

ASSESSMENT AND COMMENTS FOR THE AUTHOR

.....

.....

.....

.....

.....

RECOMMENDATION:

- Accepted
 Accepted, subject to modifications
 Rejected

ÍNDICE
TABLE OF CONTENTS

Editorial <i>Editorial note</i> ANA MARÍA ANDALUZ ROMANILLOS	9-12
--	------

ESTUDIOS
ARTICLES

**SECCIÓN MONOGRÁFICA: FILOSOFÍA E
INTELIGENCIA ARTIFICIAL**

Las Humanidades digitales en y para la historia de la filosofía. Una Introducción Digital humanities in and for the history of philosophy. An Introduc- tion MANUEL LÁZARO PULIDO	17-48
Alan Turing: Bases, forma y críticas a la inteligencia artificial. <i>Alan Turing: Base, form and criticism to artificial intelligence</i> GUSTAVO ESPARZA	49-74
Algoritmos, <i>Big Data</i> e inteligencia artificial: ¿Un nihilismo anuncia- do? <i>Algorithms, Big Data and Artificial intelligence: A foretold nihil- ism?</i> DUNE VALLE JIMENEZ Y DIEGO GARCÍA RAMÍREZ	75-103
De las emociones naturales a la emocionalidad artificial <i>From natural emotions to artificial emotionality</i> JOSÉ MIGUEL BISCAIA FERNÁNDEZ	105-139

Deus ex machina: La inteligencia artificial en el futuro del ser humano Deus ex machina: Artificial intelligence in the future of the human being DAVID GONZÁLEZ CUENCA	141-156
Polarización social y control político: Algunas consecuencias de la inteligencia artificial y las redes sociales para la razón práctica <i>Social polarization and political control: Some consequences of artificial intelligence and social networks for practical reason</i> CARLOS RODRÍGUEZ-GORDO Y JOSÉ SARRIÓN-ANDALUZ	157-185
Gobernanza global de la IA <i>Global governance of AI</i> ANÍBAL MONASTERIO ASTOBIZA:	187-211
La UE & la gobernanza ética de la inteligencia artificial: Inteligencia artificial & Diplomacia <i>EU & ethics governance of the artificial intelligence: Artificial intelligence & Diplomacy</i> MARIO TORRES JARRÍN.....	213-234
La Smartificación de la Administración Pública: Apostillas críticas a partir del caso europeo <i>The Smartification of Public Administration: Critical Apostilles from the European case</i> JONATHAN PIEDRA ALEGRÍA.....	235-250
La encrucijada de la filosofía humanística de la tecnología en el siglo XXI. Entre la nueva ilustración europea y la filosofía posteuropea china <i>The humanistic philosophy of technology in the 21st century at the crossroads. Between the New European Enlightenment and Chinese Post-European Philosophy.</i> MONTSERRAT CRESPIÑ PERALES	251-279

MISCELÁNEA

Transhumanismo, el último suspiro de Dios: Análisis del proyecto transhumanista desde la filosofía posthumanista posmoderna <i>Transhumanism, the last breath of God: Analysis of transhumanism project from postmodernism philosophy</i> LUIS LAVINA FAUSTMANN.....	283-312
--	---------

Elogio trágico del reflejo. Sobre la nostalgia antropológica de la identidad <i>Tragic praise of the reflection. On the anthropological nostalgia of identity</i> JOAQUÍN ESTEBAN ORTEGA	313-336
Ciencia y relato verosímil: ¿Es posible hacer filosofía de la naturaleza? <i>Science and plausible stories: ¿Is it possible to do philosophy of nature?</i> MANUEL C. ORTIZ DE LANDÁZURI.....	337-352
<i>Conceptual studies of the historical and modern-day influence of climate on forms of government, political and socio-economic processes: A view from Russia</i> Estudios conceptuales sobre la influencia histórica y actual del clima en las formas de gobierno y en los procesos políticos y socioeconómicos: Una visión desde Rusia ALEKSEY ANISIMOV, M. LEONIDOVNA DAVIDOVA and E. ALIBEKOVICH AHVERDIEV	353-372
Hacia la heteronomía de la experiencia estética: Una aproximación alternativa a la estética de David Hume <i>Towards the heteronomy of the aesthetic experience: An alternative approach to David Hume's aesthetics</i> JORGE LÓPEZ LLORET	373-397
Del paganismo a la gracia: Una interpretación filosófica de la existencia religiosa en el pensamiento de S. Kierkegaard <i>From paganism to grace: A philosophical interpretation of religious existence in the thought of S. Kierkegaard</i> ÁNGEL ENRIQUE GARRIDO MATURANO	399-419
Escribir filosofía: Kierkegaard y su <i>Forord</i> <i>Writing philosophy: Kierkegaard and his Forord</i> ÁNGEL VIÑAS VERA	421-441
Cultura y esclavitud en el pensamiento de Nietzsche. ¿Un aristocratismo meramente espiritual? <i>Culture and slavery in the thought of Nietzsche. A merely spiritual aristocratism?</i> JORGE POLO BLANCO.....	443-470

Simone Weil: La atención y la acción como reconocimiento de la existencia <i>Simone Weil: Attention and action as an acknowledgement of existence</i> MARÍA DEL SOL ROMANO	471-489
El carácter utópico de la felicidad. Una lectura desde Ortega y Gasset <i>The utopian character of happiness. A reading from Ortega y Gasset</i> MORA PERPERE VIÑUALES.....	491-505
Significación y lenguaje en <i>El Ser y la Nada</i> <i>Meaning and language in Being and Nothing</i> STÉPHANE VINOLO	507-528
El diálogo como proximidad en el pensamiento de Emmanuel Levinas <i>Dialogue as proximity in the thought of Emmanuel Levinas</i> FRANCISCO JAVIER RODRÍGUEZ	529-553

NOTAS CRÍTICAS

Edición de los escritos zubirianos Edition of Zubiri's writings: X. ZUBIRI, <i>Ciencia y realidad (1945-1946)</i> . Presentación y edición de E. Vargas Abarzúa. Madrid: Alianza, 2020, XVI+867 pp. X. ZUBIRI, <i>Filosofía primera (1952-1953)</i> Volumen I. Presentación y edición de E. Vargas Abarzúa. Madrid: Alianza, 2021, XI+692 pp. ANTONIO PINTOR-RAMOS	557-567
--	---------

RESEÑAS

SACRISTÁN, Manuel, <i>Sobre Jean-Paul Sartre</i> . Edición de Salvador López Arnal y José Sarrión Andaluz. Zaragoza: Prensas de la Universidad de Zaragoza, 2021, 188 pp. ARIANE AVIÑÓ MCCHESEY	571-574-
--	----------

HACKER, Peter Michael Stephan, <i>The Moral Powers. A Study of Human Nature</i> . Chichester, West Sussex; Hoboken, NJ, (EE. UU.), Wiley-Blackwell, 2021, 424 pp. ISBN: 9781119657774. JESÚS PADILLA GÁLVEZ	575-581-
MARTÍN GÓMEZ, María, <i>Diario de una filósofa embarazada</i> . Madrid: Tecnos, 179 páginas. PEDRO MARTÍN LAGO	583-585
SNELLER, Rico, <i>Perspectives on Synchronicity, Inspiration and the Soul</i> . Newcastle: Cambridge Scholars Publishing, 2019, 389 pages. GERMÁN ULISES BULA CARABALLO	587-589
HAN, Byung. <i>Ausencia. Acerca de la cultura y la filosofía del Lejano Oriente</i> (Trad. Graciela Calderón). Buenos Aires: Caja Negra, 2019, 133 p. JUAN DAVID ALMEYDA SARMIENTO	591-594

TRADUCCIONES

<i>Mi propósito principal</i> Traducción e Introducción de Yefrey Antonio Ramírez GABRIEL MARCEL.....	597-609
---	---------



UNIVERSIDAD PONTIFICIA DE SALAMANCA



Servicio de Publicaciones

OTRAS PUBLICACIONES PERIÓDICAS



Revista Helmántica

Facultad de Filología
Clásica y Hebrea

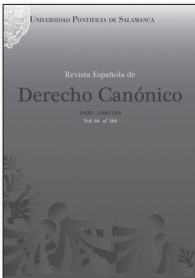
ISSN: 0018-0114
Semestral
Suscripción España: 46,00 €
Número suelto: 18,00 €



Revista Diálogo Ecuménico

Centro de Estudios Orientales y Ecuménicos
“Juan XXIII”

ISSN: 0210-2870
Cuatrimestral
Suscripción España: 39,00 €
Número suelto: 16,00 €



Revista Española de Derecho Canónico

Facultad de Derecho
Canónico

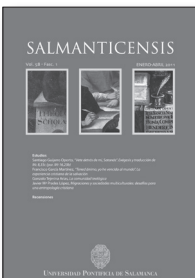
ISSN: 0034-9372
Semestral
Suscripción España: 60,00 €
Número suelto: 32,00 €



Revista Papeles Salmantinos de Educación

Facultad de CC. de la Educación

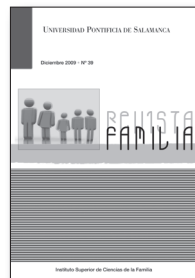
ISSN: 2340-1508
Anual | CD
Suscripción España: 15,00 €



Revista Salmanticensis

Facultad de Teología

ISSN: 0036-3537
Cuatrimestral
Suscripción España: 46,00 €
Número suelto: 18,00 €



Revista Familia

Instituto Superior de Ciencias de la Familia

ISSN: 1138-8893
Semestral
Suscripción España: 27,00 €
Número suelto: 18,00 €

Universidad Pontificia de Salamanca – Servicio de Publicaciones

C/ Compañía, 5 – 37002 Salamanca – Teléfono: 923 277 128 – www.publicaciones.upsa.es – publicaciones@upsa.es

