

Daimon

Revista Internacional de Filosofía

Número 84. Septiembre-Diciembre 2021

EXPRESSING HATRED: THE POLITICAL DIMENSION OF EXPRESSIVES

Editores

Eduardo Pérez-Navarro y María J. Frápolli

UNIVERSIDAD DE MURCIA
DEPARTAMENTO DE FILOSOFÍA

Daimon. Revista Internacional de Filosofía, fundada en 1989, es una publicación cuatrimestral del Departamento de Filosofía de la Universidad de Murcia (España). Desde entonces, *Daimon* ha abierto un espacio filosófico de reflexión, análisis y crítica de problemas referidos principalmente al ser humano, en todas las dimensiones de su existencia.

Tiene como objetivo la publicación de investigaciones originales: publica por lo tanto trabajos que abordan, desde una perspectiva filosófica, las múltiples dimensiones o esferas de la existencia humana. *Daimon* es, pues, una revista que se dirige a investigadores, pero también a todo el que se interesa por el pensamiento filosófico en sentido amplio, desde la frontera de la ciencia hasta la de la literatura. En las páginas de *Daimon*, el especialista puede encontrar nuevos enfoques de un determinado problema o autor; el investigador, un espacio en el que publicar, contrastar o confirmar sus ideas; y el lector aficionado, artículos, revisiones críticas y reseñas de libros que pueden alimentar su curiosidad y ampliar su formación.

Daimon combina, en fin, el rigor académico con la originalidad de las investigaciones, sin olvidar la apertura y pluralidad necesarias en una publicación filosófica que quiere interesar también al lector ilustrado en general.

Daimon figura en el *European Reference Index for the Humanities* (ERIH) en la categoría ERIH PLUS (*Philosophy, 2016-01-11*); sus artículos son registrados en las bases filosóficas de datos nacionales e internacionales siguientes: *Base ISOC - Filosofía*. CINDOC (España); *Dialnet* (España); *Francis, Philosophie*. INIST. CNRS (France); *Philosopher's Index* (Bowling Green, OH, USA); *Repertoire Bibliographique de Philosophie* (Louvain, Belgique); *Ulrich's International Periodicals Directory* (New York, USA), *Scopus* (Editora Elsevier, Ámsterdam, Holanda), *Web of Science* (Clarivate Analytics, Estados Unidos).

Daimon obtuvo por vez primera en 2014 el Certificado de Revista Excelente de la Fundación Española para la Ciencia y la Tecnología (FECYT, <http://calidadrevistas.fecyt.es/Paginas/Home.aspx>).



Edición electrónica: www.um.es/daimon

Daimon. Revista Internacional de Filosofía

Daimon. Revista Internacional de Filosofía

Publicación cuatrimestral. Número 84. Septiembre-Diciembre 2021

*Expressing Hatred:
The Political Dimension of Expressives*

Editores

Eduardo Pérez-Navarro y María J. Frápolli

UNIVERSIDAD DE MURCIA
DEPARTAMENTO DE FILOSOFÍA

Daimon. Revista Internacional de Filosofía

Publicación cuatrimestral. Número 84. Septiembre-Diciembre 2021

Directora / Editor: Francisca Pérez Carreño (Universidad de Murcia).

Secretario / Secretary: Emilio Martínez Navarro (Universidad de Murcia).

Consejo Editorial / Editorial Board

Mabel Campagnoli (Universidad de La Plata), Alfonso García Marqués (Universidad de Murcia), Ricardo Gutiérrez Aguilar (Universidad Complutense de Madrid), Manuel Liz Gutiérrez (Universidad de La Laguna), María Teresa López de la Vieja de la Torre (Universidad de Salamanca), Claudia Mársico (Universidad de Buenos Aires), Miriam Molinar Varela (Instituto Tecnológico y de Estudios Superiores de Monterrey, México), Jesús Navarro Reyes (Universidad de Sevilla), Diana Pérez (Universidad de Buenos Aires), Ángel Puyol González (Universidad Autónoma de Barcelona), Luisa Paz Rodríguez Suárez (Universidad de Zaragoza), Salvador Rubio Marco (Universidad de Murcia).

Comité Científico / Scientific Committee

Florencia Dora Abadi (Universidad de Buenos Aires y CONICET), Atocha Aliseda Llera (Universidad Nacional Autónoma de México), Mauricio Amar Díaz (Universidad de Chile), Diego Fernando Barragán Giraldo (Universidad de La Salle, Bogotá), Eduardo Bello Reguera (†), Noelia Billi (Universidad de Buenos Aires), Antonio Campillo Meseguer (Universidad de Murcia), Germán Cano Cuenca (España), Cinta Canterla González (Universidad Pablo de Olavide, Sevilla), Fernando Cardona Suárez (Colombia), Adelino Cardoso (Universidade Nova de Lisboa), Salvador Cayuela Sánchez (Universidad de Murcia), Luz Gloria Cárdenas Mejía (Universidad de Antioquia, Medellín), Pablo Chiuminatto (Chile), Jesús Conill Sancho (Universidad de Valencia), Adela Cortina Orts (Universidad de Valencia), Kamal Cumsille (Universidad de Chile), Juan José Escobar López (Colombia), Ángel Manuel Faerna García-Bermejo (Universidad de Castilla-La Mancha), Hernán Fair (Universidad Nacional de Quilmes y CONICET), María José Frápolli Sanz (Universidad de Granada), Ángela Lorena Fuster (Universidad de Barcelona), Domingo García Marzá (Universitat Jaume I, Castellón), Mariano Gaudio (Universidad de Buenos Aires), Juan Carlos González González (Universidad Autónoma del Estado de Morelos, México), María Antonia González Valerio (Universidad Nacional Autónoma de México), María José Guerra Palmero (Universidad de La Laguna), Valeriano Iranzo Garcia (Universidad de Valencia), Rodrigo Karmy Bolton (Universidad de Chile), Elena Laurenzi (Università del Salento y Universidad de Barcelona), Juan Carlos León Sánchez (Universidad de Murcia), Gerardo López Sastre (Universidad de Castilla-La Mancha), José Lorite Mena (Universidad de Murcia), Alfredo Marcos Martínez (Universidad de Valladolid), António Pedro Mesquita (Universidade de Lisboa), Marina Mestre Zaragoza (ENS de Lyon), Javier Moscoso Sarabia (Instituto de Filosofía, CCHS-CSIC, Madrid), Paula Cristina Mira Bohórquez (Universidad de Antioquia, Medellín), Jose María Nieva (Universidad Nacional de Tucumán), Laura Nuño de la Rosa (KLI, Austria), Patricio Peñalver Gómez (Universidad de Murcia), Angelo Pellegrini (Italia), Francisca Pérez Carreño (Universidad de Murcia), Manuel de Pinedo García (Universidad de Granada), Miguel Ángel Polo Santillán (Universidad Nacional Mayor de San Marcos, Lima), Hilda María Rangel Vázquez (Universidad Pontificia de México), Jacinto Rivera de Rosales Chacón (Universidad Nacional de Educación a Distancia, Madrid), Antonio Rivera García (Universidad Complutense de Madrid), Concha Roldán Panadero (Instituto de Filosofía del CSIC, Madrid), Adriana Rodríguez Barraza (Universidad Veracruzana, México), Miguel Ruiz Stull (Chile), Vicente Sanfélix Vidarte (Universidad de Valencia), Merio Scattola (Università degli Studi di Padova), Francisco Vázquez García (Universidad de Cádiz), José Luis Villacañas Berlanga (Universidad Complutense de Madrid).

© **Daimon. Revista Internacional de Filosofía**, de todos los trabajos. Para su uso impreso o reproducción del material publicado en esta revista se deberá solicitar autorización a la Dirección de la revista. Esta no se hace responsable de las opiniones vertidas por los autores de los trabajos que en ella se publican.

Administración: *Daimon* es una revista cuatrimestral, editada y distribuida por el Servicio de Publicaciones de la Universidad de Murcia. Apartado 4021. 30080 Murcia (España). Tfno.: 868883012. Fax: 868883414.

Redacción e intercambios: ver *Normas de publicación*, al final de la revista.

ISSN de la edición en papel: 1130-0507.

ISSN de la edición digital (disponible en <http://revistas.um.es/daimon>): 1989-4651.

Depósito legal: V 2459-1989.

Maquetación, diseño de cubierta: Compobell, S.L. Murcia.



Daimon. Revista Internacional de Filosofía

Publicación cuatrimestral. Número 84. Septiembre-Diciembre 2021

Monográfico sobre

Expressing Hatred: The Political Dimension of Expressives

(Editores: Eduardo Pérez-Navarro y María J. Frápolli)

Introducción

- Presentación del número monográfico sobre Expressing Hatred: The Political Dimension of Expressives. *Eduardo Pérez-Navarro y María José Frápolli* 5

Artículos

- Exactly why are slurs wrong? *Thaddeus Metz* 13
- The resistant effect of slurs: A nonpropositional, presuppositional account. *Alba Moreno Zurita y Eduardo Pérez-Navarro* 31
- Paving the road to hell: The Spanish word *menas* as a case study. *David Bordonaba-Plou y José Ramón Torices* 47
- Provocative insinuations. *Álvaro Domínguez-Armas y Andrés Soria-Ruiz* 63
- Not all speakers are equal: Harm and conversational standing. *Claudia Picazo* 81
- Exactly, what do you mean? *Manuel Almagro y Nefthalí Villanueva* 97
- I am large, I contain multitudes: Epistemic pragmatism, testimonial injustice and positive intersectionalism. *María J. Frápolli y Llanos Navarro* 115
- Las cualidades secundarias y el valor epistémico de la experiencia estética. *Nieves Acedo* 131
- Una concepción pragmatista sobre las reglas. *Pedro Antonio García Jorge* 147
- Enactivismo y valoración. Cómo superar la querrela entre teorías somáticas y cognitivas de las emociones. *Andrea F. Melamed* 163
- La autoridad de la primera persona como criterio de determinación y reconocimiento de la identidad de género. *Jorge Suárez Muñoz* 183
- Investigación filosófica y emociones: la ontología histórica de Ian Hacking como alternativa. *Juan Manuel Zaragoza Bernal* 199

Reseñas

- BUTLER, Judith (2021). *Excitable Speech. A Politics of the Performative*. Nueva York: Routledge, 212 pp. (*Javier Moscoso Cala*) 219
- BRANDOM, Robert M. (2019). *A Spirit of Trust. A Reading of Hegel's Phenomenology*. Cambridge, Massachusetts: The Belnap Press of Harvard University Press, 768 pp. (*Christian Iber; Agemir Bavaresco; Danilo Vaz Costa*) 223

NÚÑEZ, Amanda (2019), <i>Gilles Deleuze, una estética del espacio para una ontología menor</i> . Madrid, Arena Libros, 232 págs. (Rosa Jiménez Asensio).....	228
CORCHIA, L., MÜLLER-DOOHM, S. und OUTHWAITE, W. (Hg.). <i>Habermas global. Wirkungsgeschichte eines Werks</i> . Berlin: Suhrkamp, 2019, 894 pp. (José Luis López de Lizaga)	232
HEGEL, G. W. F. (2019). <i>Lecciones sobre la filosofía del espíritu subjetivo II. Antropología</i> . Traducción Alberto Ciria y Juan J. Padial, notas y estudio preliminar por Juan J. Padial, Sevilla: Thémata, 540 pp. (Andrés Ortigosa).....	236
CAVALLETTI, A. (2019). <i>Vértigo. La tentación de la identidad</i> . Traducción de María Teresa D'Meza. Buenos Aires: Adriana Hidalgo editora. (Juan Cruz Aponiuk).....	240
GONZÁLEZ BLANCO, Azucena (ed.) (2019). <i>Política y literatura. Nuevas perspectivas teóricas</i> . Berlín: De Gruyter. (Beltrán Jiménez Villar)	243
WITTGENSTEIN, Ludwig (2019). <i>Correspondencia con Rudolf Koder</i> . Traducción de Venancio Andreu Baldó e Isabel Gamero Cabrera. Madrid: Editorial Apeiron (Jeremy Mederos).....	247
LUMBREERAS SANCHO, Sara (2020). <i>Respuestas al Transhumanismo. Cuerpo, autenticidad y sentido</i> . Madrid: Digital Reasons, 202 pp. (Adrián Santamaría Pérez y José María Santamaría García)	251
HUSSERL, Edmund (2019). <i>Textos Breves (1887-1936)</i> . Coord.: Agustín Serrano de Haro y Antonio Ziriñón Quijano. Salamanca: Sígueme, 716 pp. (Jesús Miguel Marcos del Cano).....	255
HINDRICH, Gunnar (2020). <i>La autonomía del sonido. Una filosofía de la música</i> . Traducción de Pedro García-Baró Huarte. Salamanca: Ediciones Sígueme. (Laura Carolina Durán).....	260
RIVERA, A. y MATEO, E. (eds.) (2020). <i>Las narrativas del terrorismo. Cómo contamos, cómo transmitimos, cómo entendemos</i> . Vitoria, Madrid: Fundación Fernando Buesa y Ediciones Catarata. (Isabel Roldán Gómez)	263
PASCAL, Blas (2020). <i>Escritos espirituales</i> . Edición de Alicia Villar. Madrid: Tecnos. 242 pp. (Carlos Gómez)	266
BRENTANO, Franz C. (2019). <i>Vermischte Schriften. Sämtliche veröffentlichte Schriften, Band IX. Mit einer Einleitung von D. FISETTE</i> . Herausgegeben von T. BINDER und A. CHRUDZIMSKI. Berlin/Boston: Walter de Gruyter. (David Torrijos Castrillejo).....	268
ÇRETU, Ana-Maria; MASSIMI, Michela (ed.) (2020). <i>Knowledge from a human point of view</i> , Springer. (Abraham Hernández)	270
ARAGÜÉS, Juan Manuel (2020): <i>De idiotas a koinotas. Para una política de la multitud</i> . Madrid: Arena Libros. (Julien Canavera)	273
ASSMANN, J., CACCIARI, M., DUQUE, F., PONS, J. y VELA, A. (2020). <i>La palabra que falta. El dios indecible de Moisés y Aarón, de Schönberg</i> . Madrid: Ediciones CBA. (Gonzalo Pérez Santonja)	277



Presentación del número monográfico sobre *Expressing Hatred: The Political Dimension of Expressives*

Presentation of the monographic issue on *Expressing Hatred: The Political Dimension of Expressives*

EDUARDO PÉREZ-NAVARRO*

MARÍA J. FRÁPOLLI**

Resumen: Desde hace unos años, se ha producido dentro de la filosofía analítica un movimiento de acercamiento a las prácticas reales y de huida de las idealizaciones no justificadas que pretende poner las herramientas conceptuales desarrolladas durante el último siglo al servicio de la justicia social. En el ámbito de la filosofía del lenguaje, este giro ha pasado por el análisis de expresiones del lenguaje natural que, por no encajar de forma completamente satisfactoria con la concepción del significado como condiciones de verdad, han recibido tradicionalmente poca atención. Sin embargo, estas expresiones juegan un papel fundamental en la comunicación con impacto político. Hablamos de los expresivos, esto es, expresiones que se utilizan para comunicar una cierta actitud. El propósito de este número especial de *Daimon* es ofrecer una panorámica de algunos debates que se están desarrollando en la actualidad en relación con la dimensión política de los expresivos, pero también de discusiones cercanas que en ocasiones se solapan con esta, tanto en filosofía del lenguaje como en ramas de la filosofía afines.

Palabras clave: expresivos, filosofía del lenguaje, giro político de la filosofía analítica, injusticia

Abstract: For some years now, there has been a movement within analytic philosophy to get closer to real practices and to flee from unwarranted idealizations in order to put the conceptual tools developed over the last century at the service of social justice. In the field of philosophy of language, this turn has involved the analysis of natural language expressions that, not fitting in a completely satisfactory way with the conception of meaning as truth-conditions, have traditionally received little attention. However, these expressions play a fundamental role in communication with political impact. We are talking about expressives, that is, expressions that are used to communicate a certain attitude. The purpose of this special issue of *Daimon* is to offer an overview of some of the debates that are currently taking place in relation to the political dimension of expressives, but also of related discussions that sometimes overlap with it, both in philosophy of language and in related branches of philosophy.

Keywords: expressives, philosophy of language, political turn in analytic philosophy, injustice

Recibido: 02/07/2021. Aceptado: 02/07/2021.

* Investigador postdoctoral, Departamento de Filosofía I, Universidad de Granada. Correo electrónico: edperez@ugr.es.

** Catedrática de Lógica y Filosofía de la Ciencia, Departamento de Filosofía I, Universidad de Granada. Correo electrónico: frapolli@ugr.es.

Desde sus orígenes, la filosofía analítica ha pretendido ofrecer un tratamiento universal de los rasgos característicos de fenómenos humanos como el conocimiento o el lenguaje. Por este motivo, ha dejado de lado las particularidades que surgen cuando estas capacidades se ponen en práctica en contextos no ideales y ha asumido, por ejemplo, que, a igualdad de evidencia y capacidades, todas las personas que participan en un intercambio epistémico tienen la misma credibilidad. Desde hace unos años, sin embargo, se ha producido un movimiento de acercamiento a las prácticas reales y de huida de las idealizaciones no justificadas que pretende poner las herramientas conceptuales desarrolladas durante el último siglo al servicio de la justicia social. Un paso dentro de este movimiento es la introducción en la epistemología contemporánea del concepto de *injusticia epistémica* (Fricker, 2007). La variedad de injusticia correspondiente en filosofía del lenguaje es la *injusticia discursiva* (Kukla, 2014) o *injusticia del habla* (Ayala, 2016), ambos conceptos herederos de las discusiones sobre los mecanismos lingüísticos detrás de la pornografía que tuvieron lugar en la década de los noventa (Langton, 1993; Hornsby y Langton, 1998).

En este ámbito, el giro político de la filosofía analítica ha pasado también por el análisis de expresiones del lenguaje natural que, por no encajar de forma completamente satisfactoria con la concepción del significado como condiciones de verdad, han recibido tradicionalmente poca atención desde la filosofía del lenguaje. Sin embargo, estas expresiones juegan un papel fundamental en la comunicación con impacto político. Hablamos de los *expresivos*. Llamamos “expresivos” a aquellas expresiones del lenguaje natural que se utilizan para comunicar una cierta actitud. Si lo que nos interesa de los expresivos es su potencial para el uso político, son especialmente relevantes expresiones como los *slurs* (véanse, por ejemplo, Potts, 2007; Hom, 2008; Camp, 2013; Jeshion, 2013; Cepollaro y Stojanovic, 2016), los *dogwhistles* (Saul, 2018) y los términos étnicos y sociales neutros usados como insultos o TESNI (véase Castroviejo, Fraser y Vicente, 2020), así como las insinuaciones (véanse Fricker, 2012; Camp, 2018). Los *slurs* son insultos que se dirigen específicamente a un determinado grupo social. Como tales, estas expresiones son sumamente dañinas, pero poco sutiles. Los *dogwhistles*, por el contrario, no son generalmente percibidos como herramientas que comuniquen una actitud, lo que los hace potencialmente más peligrosos. Por ejemplo, utilizar expresiones como “PER” puede movilizar ciertas actitudes contra andaluces y extremeños en el auditorio, pero de forma tal que al menos una parte del mismo ignorará que esto está ocurriendo. Los TESNI, por su parte, son términos en principio neutrales que se utilizan en ocasiones para transmitir estereotipos negativos. Las insinuaciones, finalmente, son un tipo particular de inferencia pragmática.

El propósito de este número especial de *Daimon*, resultado de una iniciativa de las personas participantes en el proyecto de investigación de excelencia de la Junta de Andalucía “La identificación inferencial de las proposiciones. Una reconsideración de las dicotomías clásicas en metafísica, semántica y pragmática” (P18-FR-2907), es ofrecer una panorámica de algunos debates que se están desarrollando en la actualidad en relación con la dimensión política de los expresivos, pero también de discusiones cercanas que en ocasiones se solapan con esta, tanto en filosofía del lenguaje como en ramas de la filosofía afines.

Los artículos incluidos en este número especial se han ordenado siguiendo una narrativa que parte de temáticas clásicas dentro de este tipo de discusiones para ir abriendo el espectro hacia debates menos estándar que superan finalmente los límites de la filosofía del lenguaje.

El de Thaddeus Metz (“Exactly why are slurs wrong?”) y el de Alba Moreno y Eduardo Pérez-Navarro (“The resistant effect of slurs: A nonpropositional, presuppositional account”) se centran en viejos conocidos como son los *slurs*, si bien el primero deja de lado la discusión semántica para investigar las razones por las que ciertos usos de estas expresiones son moralmente incorrectos. Los siguientes artículos exploran fenómenos lingüísticos similares que solo recientemente se han empezado a describir, sean los TESNI, como en el de David Bordonaba-Plou y José Ramón Torices (“Paving the road to Hell: The Spanish word *menas* as a case study”), o las insinuaciones provocativas, como en el de Álvaro Domínguez-Armas y Andrés Soria-Ruiz (“Provocative insinuations”). Dentro de la filosofía del lenguaje encontramos también el artículo de Claudia Picazo (“Not all speakers are equal: Harm and conversational standing”) y el de Manuel Almagro y Neftalí Villanueva (“Exactly, what do you mean?”), que exploran la influencia que distintos aspectos del contexto tienen a la hora de determinar la fuerza con la que se implementa una norma de permisibilidad, en el primer caso, y de clasificar un uso del lenguaje como evaluativo, en el segundo. Finalmente, el artículo de María J. Frápolli y Llanos Navarro (“I am large, I contain multitudes: Epistemic pragmatism, testimonial injustice and positive intersectionalism”) permite a este número ampliar su foco para cubrir otra de las protagonistas del giro político en filosofía analítica: la epistemología. El número se completa con una reseña a cargo de Javier Moscoso de *Excitable Speech: A Politics of the Performative*, de Judith Butler, con motivo de su reedición en Routledge Classics. En lo que sigue, recorreremos con un poco más de detalle los contenidos del número.

Como se ha adelantado anteriormente, la pregunta que guía el artículo de Thaddeus Metz es por qué es moralmente incorrecto usar un *slur* en los casos en los que lo es. Metz considera tres posibles explicaciones: la bienestarista, la kantiana y una teoría relacional que toma prestadas ideas de la filosofía sudafricana. Es esta última opción la que se va a defender como más adecuada frente a las desventajas que presentan las otras dos alternativas. El artículo de Alba Moreno y Eduardo Pérez-Navarro, por su parte, tiene un propósito doble. Por un lado, se propone describir de manera unificada el comportamiento que hace de los *slurs* expresiones particulares. La peculiaridad se achaca, en concreto, a su carácter resistente frente a la cancelación, el rechazo y la retractación. Por otro lado, el artículo pretende explicar este comportamiento a través de una teoría presuposicional del significado de los *slurs* que, como otras propuestas recientes, entiende el contenido presupuesto en estos casos como no proposicional.

Si los dos primeros artículos de este número se centran en los *slurs*, los dos siguientes se ocupan de fenómenos relacionados pero mucho menos discutidos. Por un lado, David Bordonaba-Plou y José Ramón Torices argumentan a partir de resultados experimentales que la palabra “mena” exhibe los rasgos característicos de los TESNI y exploran las distintas direcciones en las que el uso de la palabra podría evolucionar. En algunos de estos casos la palabra seguiría siendo un TESNI, mientras que en otros se convertiría en un *slur*. Por otro lado, Andrés Soria-Ruiz y Álvaro Domínguez-Armas tratan de localizar el mecanismo que se halla detrás de ciertas inferencias racistas. Tras descartar que se trate de *slurs*, TESNI, *dogwhistles* o implicaturas conversacionales, sugieren explicar estas inferencias como *insinuaciones provocativas*.

El artículo de Claudia Picazo y el de Manuel Almagro y Neftalí Villanueva se encuadran también dentro de la filosofía del lenguaje, en esta ocasión con el contexto como protago-

nista. En su artículo, Claudia Picazo argumenta que la fuerza de las normas que determinan qué es permisible en un determinado contexto es sensible a la posición conversacional del hablante y que, en casos extremos, esta posición puede ser insuficiente para establecer ningún tipo de norma. Por su parte, el artículo de Manuel Almagro y Nefthalí Villanueva pretende, por un lado, rastrear la relación que la literatura atribuye al contexto y los usos evaluativos del lenguaje y, por otro, argumentar que algunas de las propuestas teóricas que actualmente se encuentran en el mercado son incapaces de explicar los hallazgos experimentales recientes.

El artículo de María J. Frápolli y Llanos Navarro, por último, contrapone dos concepciones del conocimiento: el pragmatismo epistémico de Brandom y la concepción del conocimiento implícita en la obra de Fricker y quienes han seguido su estela. Frápolli y Navarro argumentan que, por sí solas, estas dos concepciones del conocimiento están en tensión, pero que pueden hacerse compatibles si se complementan con un “interseccionalismo positivo”.

Los artículos contenidos en este número especial parten de debates de gran actualidad y contribuyen a ellos con propuestas novedosas que exploran los límites del aparatage teórico disponible hasta el momento. Sea poniendo sobre la mesa preguntas hasta ahora ignoradas sobre fenómenos ampliamente explotados (¿por qué son moralmente incorrectos algunos usos de los *slurs*?), contribuyendo a la incipiente literatura sobre fenómenos de reciente descripción, como los TESNI o las insinuaciones provocativas, aprovechando los avances de la filosofía experimental o poniendo en diálogo tradiciones que se han dado la espalda durante largo tiempo, como el pragmatismo epistémico de Brandom y la ética del conocimiento, estos artículos suponen un paso adelante en la aplicación de la filosofía contemporánea a problemas que deberían preocupar a cualquier persona interesada en cuestiones sociales y políticas. Esperamos, por tanto, que estos trabajos sean la base sobre la que puedan proponerse herramientas de intervención en los ámbitos abarcados por el número.

Referencias

- Ayala, S. (2016), “Speech affordances: A structural take on how much we can do with our words”, *European Journal of Philosophy*, 24(4), 879-891.
- Camp, E. (2013), “Slurring perspectives”, *Analytic Philosophy*, 54(3), 330-349.
- Camp, E. (2018), “Insinuation, common ground, and the conversational record”, en D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 40-66.
- Castroviejo, E., Fraser, K. & Vicente, A. (2020), “More on pejorative language: Insults that go beyond their extension”, *Synthese*.
- Cepollaro, B. & Stojanovic, I. (2016), “Hybrid evaluatives: In defense of a presuppositional account”, *Grazer Philosophische Studien*, 93(3), 458-488.
- Fricker, E. (2012), “Stating and insinuating”, *Proceedings of the Aristotelian Society Supplementary Volume*, 86(1), 61-94.
- Fricker, M. (2007), *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press.
- Hom, C. (2008), “The semantics of racial epithets”, *The Journal of Philosophy*, 105(8), 416-440.

- Jeshion, R. (2013), "Slurs and stereotypes", *Analytic Philosophy*, 54(3), 314-329.
- Kukla, R. (2014), "Performative force, convention, and discursive injustice", *Hypatia*, 29(2), 440-457.
- Langton, R. (1993), "Speech acts and unspeakable acts", *Philosophy and Public Affairs*, 22(4), 293-330.
- Hornsby, J. & Langton, R. (1998), "Free speech and illocution", *Legal Theory*, 4(1), 21-37.
- Potts, M. (2007), "The expressive dimension", *Theoretical Linguistics*, 33(2), 165-198.
- Saul, J. (2018), "Dogwhistles, political manipulation, and philosophy of language", en D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 360-384.

ARTÍCULOS



Exactly why are slurs wrong?*

¿Exactamente por qué están mal los *slurs*?

THADDEUS METZ**

Abstract: This article seeks to provide a comprehensive and fundamental account of why racial epithets and similar slurs are immoral, whenever they are. It considers three major theories, roughly according to which they are immoral because they are harmful (welfarism), because they undermine autonomy (Kantianism), or because they are unfriendly (an under-considered, relational approach informed by ideas from the Global South). This article presents new objections to the former two theories, and concludes in favour of the latter rationale. Deeming slurs to be wrong insofar as they are unfriendly is shown to capture the advantages of the other theories, while avoiding their disadvantages.

Keywords: dignity; harm; immorality; racial epithets; relationality; slurs

Resumen: Este artículo busca proporcionar una descripción completa y fundamental de por qué los epítetos raciales y *slurs* similares son inmorales, allá donde lo sean. Considera tres teorías principales, según las cuales, a grandes rasgos, son inmorales porque son dañinos (bienestarismo), porque socavan la autonomía (kantianismo) o porque son hostiles (un enfoque relacional poco discutido informado por ideas del Sur Global). Este artículo presenta nuevas objeciones a las dos primeras teorías y concluye a favor de la última justificación. Se muestra que considerar que los *slurs* son inmorales en la medida en que son hostiles captura las ventajas de las otras teorías evitando sus desventajas.

Palabras clave: dignidad; daño; inmoralidad; insultos raciales; relacionalidad; *slurs*

1. Introduction

Note the variety of racial epithets and similar slurs there are and the various contexts in which it appears morally wrong to use them. Consider a judge referring to black people as the n-word when handing down a sentence, a blogger describing women as ‘bitches’ on a website, a group of neo-Nazis calling Jewish people ‘kikes’ when in private amongst them-

Recibido: 10/04/2021. Aceptado: 17/06/2021.

* For detailed comments on a previous draft, the author thanks Heidi Savage, Adam Croom, and two thoughtful anonymous referees for Daimon.

** Professor of Philosophy at the University of Pretoria in South Africa. Email: th.metz@up.ac.za. Interests: African philosophy; comparative philosophy; ethics. Recent publications include: Metz, T. (2021), *A Relational Moral Theory*, Oxford: Oxford University Press, and Metz, T. (2021), ‘Recent work in African philosophy: Its relevance beyond the continent’, *Mind*, 130(518), 639–660.

selves, and a person labelling someone who she thinks is gay the f-word¹ Given that such insults targeting people's race, gender, ethnicity, and sexuality are wrong, warranting guilt and censure, precisely why are they wrong?

This article seeks a theoretical answer to this question, meaning that it considers whether a single principle can make good sense of why slurs are ethically impermissible and hence of when they are and to what degree. It would be interesting and revealing if the myriad of wrongful slurs were all immoral in virtue of one basic thing. This article critically explores three theories of what makes slurs immoral, two of which appeal to familiar ideas about harm and autonomy, and one of which will be new to many Western readers, roughly maintaining that slurs are wrong to the extent that they put distance between people or are unfriendly. This article defends the latter, relational theory, which is inspired largely by ideas from grossly under-considered philosophical traditions in sub-Saharan Africa and the Global South more generally, as more justified at this stage of enquiry or at least worthy of consideration.

Little of the English-speaking literature on slurs has sought to answer the question of why slurs are immoral, at least not theoretically in a systematic way. Most have not even sought to analyze the nature of the wrongness of slurs, placing their wrongness in the background so as to focus on capturing their linguistic contents and functions (e.g., Hornsby, 2001; Himma, 2002; Boisvert, 2008; Hom, 2008, 2010; Richard, 2008; Croom, 2011, 2013; Anderson and Lepore, 2013a, 2013b; Camp, 2013; Jeshion, 2013a, 2013b; Whiting, 2013). This article does the opposite; in order to make headway on the quest for a normative principle that best accounts for what all slurs have in common insofar as they are wrong, this article sets aside detailed analysis of what they say, how they say it, and related descriptive questions. It presumes that the reader can readily identify slurs of various kinds and their foreseeable results, and considers whether they might share an underlying structure with regard to their wrongness (as opposed to their nature).

To be sure, those who have undertaken descriptive enquiries into the nature of slurs have made suggestions about their wrongness, which this article uses as springboards to dive in for a much closer look. In addition, there is some analysis of why insults are immoral that is relevant (especially Neu, 2008; Archard, 2014). However, for most in the field, slurs are distinct from more general insults (on which see Boisvert, 2008; Neu, 2008; Hay, 2011), and, more importantly, the standard accounts of why insults are wrong ground theories that do not best capture the immorality of slurs, or so this article contends.

This article does not presume that slurs are by definition immoral or even that they are always immoral. In fact, it is reasonable to think that some acts straightforwardly construed to be calling someone slurs are not immoral when used by certain people in certain contexts for certain purposes, for instance when 'appropriated' by an in-group. One advantage of the appeal to relationality to explain the wrongfulness of slurs is its plausible explanation of why using slurs can intuitively be permissible—roughly when they serve to bring people closer together, albeit without pushing anyone away.

1 As one scholar points out, mentioning slurs is difficult to avoid when theorizing about them in academic spaces (Hornsby, 2001, 129), even if, as others note, doing so still risks harming and offending (Himma, 2002; Anderson and Lepore, 2013b, 354). In the following, this author tries to balance these considerations by not directly mentioning the worst slurs.

Note, by the way, that this article engages in the ‘normal’ analytic practice of evaluating contested general principles by considering their ability to account for *less* contested (not necessarily uncontested) particular cases. That is, it freely appeals to intuitions about when, why, and to what degree slurs are wrong that it expects many enquirers to share, and, insofar as many do share them, they count as evidence that may be used to appraise theories that are *more* controversial than the intuitions. There is no suggestion made that the intuitions are irrefutable, but those who seek to cast doubt on them have a lot of explaining to do about why we should revise them.

Section 2 briefly sketches some salient elements of slurs, so that the target is fairly clear. The next few sections appeal to the most widely invoked moral theories from the past two hundred years of Western philosophy in order to test their ability to account for the wrongfulness of slurs, and they are found wanting. Specifically, section 3 considers welfarist theories of why slurs are wrong, according to which they are immoral insofar as they reduce people’s subjective well-being. After demonstrating that welfarism cannot account for certain intuitive instances of wrongful slurs and the degree to which they are wrong, sections 4 and 5 then take up Kantian, autonomy-based accounts, according to which wrongful acts are those that degrade people’s capacity for self-governance, and they are found to be vulnerable to similar objections. Section 6 articulates a novel approach by appeal to globally neglected values pertaining to harmony salient especially in the Southern African philosophy of *ubuntu*, roughly according to which slurs are wrong insofar as they are unfriendly or discordant, and contends that it is the most viable contender. This article briefly concludes in section 7 by noting that, since this relational approach not only obtains the advantages of the other two approaches, but also avoids (and even explains) their disadvantages, the view that slurs are wrong insofar as they are unfriendly at least merits consideration in future work, if not belief at this stage.

2. Slurs: some salient features

This article does not provide a theoretical account of what all slurs have in common *qua* slurs, i.e., identify the essence of a slur. For one, there might not be any shared core to the nature of what we call ‘slurs’. For another, even if there were, this article lacks the space to consider what it might be, since it is instead devoted to determining what makes slurs at bottom immoral. Even so, it is necessary to lay out some recurrent elements of slurs, or at least of the kinds considered here, to facilitate the project of determining why slurring is wrong.

Slurs of the kind addressed here are *characteristically* forms of (a) communication (b) that is linguistic in which (c) other persons are told that they exhibit (d) undesirable properties (e) associated with their perceived group such as their race, ethnicity, gender, or sexual orientation. The qualification ‘characteristically’ is important, as this article does not wish to provide a closed analysis of the nature of a slur, and there are some likely exceptions to this core account (as well as borderline cases that are difficult to specify). So, the point is not that all slurs have these features, but rather that they often do. The rest of this section briefly spells out each of these salient elements and also highlights plausible exceptions, where some slurs seem not to include them.

To start, using slurs is typically communicative, where one agent more or less transmits a symbol to another agent with the intention that she both understand a certain proposition and recognize that the agent transmitting the symbol intends her to understand that proposition by means of the transmission (e.g., Grice, 1957). However, some uses of slurs appear to be expressive actions without being communicative in this sense. If a misogynist who thinks he is alone mutters ‘bitch’ to himself upon thinking of his wife, he is plausibly slurring but not communicating, even if his wife were unexpectedly to overhear him.

When slurs are communicative, they are usually linguistic. However, it seems possible for there to be slurs that are communicative but not linguistic. Imagine, for instance, that a student routinely left pictures of a baboon under an African-American lecturer’s door and posted them on the whiteboard he uses prior to him arriving to teach. This would probably be a case of a non-linguistic and yet communicative slur. The point is not that such behaviour clearly counts as a genuine slur; it is rather that it might count, and that, in any event, whatever makes slurs wrong would probably make this sort of behaviour wrong, too.

When slurs are communicated, they are often directed towards the one who is the object of the slur. The quintessential instance of a slur is when another person is called a derogatory name. However, there are cases in which the disparaging language is about someone X, but directed towards another person Y. Members of an Aryan supremacist organization might, entirely amongst themselves, make anti-Semitic slurs, for instance.

Many times slurs involve expressing the judgement that certain individuals exhibit particular undesirable properties. Typical are characterizations of people as lazy, stupid, sexually rapacious, miserly, overly talkative, or nagging. However, there are words that are derogatory, apparently not by connoting specific negative traits such as stereotypes, but rather more generally and performatively, plausible examples being the n-word, the f-word, and, on some occasions, even ‘woman’ (cf. Feinberg, 1985, 218–26 on ‘pure insults’; and Anderson and Lepore, 2013a). There are those who say that these terms are implicitly associated with negative traits, particularly stereotypes, even if they do not refer to them explicitly (cf. Camp, 2013, 342n16, 346–7). However, it at least seems possible for a given speaker to slur without having such properties in mind, e.g., someone at his very first KKK rally who doesn’t yet know what ‘kike’ means but is happy to shout it out because his peers are. It is not being suggested here that slurs are necessarily negative; if at this rally Jewish people were instead called ‘clever’, perhaps it would count as a sort of positive slur, one that reduces them to a stereotype. However, this article addresses slurs insofar as they are straightforwardly disparaging, mentioning in the conclusion the possibility of extending the analysis of them.

Finally, it is characteristic of slurs, or at least the ones addressed in this article, that the undesirable properties apply to the individual by virtue of her perceived membership in a certain group pertaining to race, ethnicity, gender, sexual orientation, religion, or the like. It is interesting to consider what these groups have in common, with ‘social salience’, i.e., their determining much about people’s past and present interaction, being a plausible candidate (see Lippert-Rasmussen, 2014, 30–6). When a person is called ‘lazy’ because she is of African descent, one has a clear instance of a slur, but plenty of other occasions of calling someone ‘lazy’ need not be slurring. Some might think that one slurs another’s character if one accuses her of, say, having stolen spoons, where this has nothing to do with her

membership in a group, at least not one that has social salience. However, for this author such an accusation would be to ‘cast aspersions’ or perhaps to ‘insult’, but not to slur in a usefully narrow sense.

3. Harm

Having advanced an account of what slurs typically involve, it is time to consider what makes them wrong, whenever they are. A straightforward thing to say about slurs is that they can be hurtful. Sometimes they are designed to hurt; those who use slurs are on occasion aiming to cause psychological pain. However, beyond the intention to cause such harm, an agent who uses a slur is at least likely to cause it.

These are welfarist considerations. Welfarism is roughly the view that an action is morally wrong insofar as it tends to reduce people’s quality of life. When applied to slurs, welfarism entails that, when they are wrong, they are in virtue of causing harm, which is usually psychological. Utilitarianism is the most influential form of welfarism in the field of moral philosophy and disciplines influenced by it (but is not the only form, as there have been rights-based welfarisms). According to that view, an action is wrong insofar as it fails to produce the most benefit and to reduce the most harm in the long run, taking everyone’s interests into account. For utilitarians, benefit is normally construed subjectively, as a matter of either experiencing pleasure or having desires satisfied, and harm is taken to consist of either experiencing pain or having desires frustrated. The classic utilitarian Jeremy Bentham recommends criminalizing certain speech acts because of the ‘unpleasant sensation’ they usually cause to many who hear them (1822, 546–7). In addition, when explicating utilitarianism, it is common for moral philosophers to note its implications for speech acts such as insults, viz., to point out they are normally wrong on utilitarian grounds for causing unnecessary embarrassment, sadness, and the like (e.g., Feldman, 1978, 25). However, there are those who apply welfarism merely to speech acts and not to morality in general. A good instance is Joel Feinberg, who remarks that ‘to be forced to suffer an offence (...) is an unpleasant inconvenience, and hence an evil’ (1985, 49).

This article focuses on the utilitarian version of welfarism when considering the wrongfulness of slurs, where the advantages and disadvantages of utilitarianism are meant to characterize welfarism generally (including rights-based versions of it). Utilitarianism has the advantage of accounting straightforwardly for the respect in which the hurtfulness of slurs often seems relevant to their wrongness. If a child overhears a slur and is tempted to repeat it, a concerned parent will tell her not to because it would hurt the other person’s feelings. In addition, utilitarianism seems able to explain why some slurs are worse than others, namely, they are likely to hurt more than the alternatives. The reason it would be more wrong to call an Asian businessman a ‘chink’ than a ‘suit’ (as per Anderson and Lepore, 2013a, 25), so the utilitarian suggests, is that doing so would be expected to cause him more psychological distress. Finally, the utilitarian can also reasonably explain why some slurs are not wrong; when slurs have been ‘appropriated’, they do not cause pain when used, making it permissible to use them under that condition.

Despite these strong arguments, there are two glaring problems with the utilitarian, and more generally welfarist, account of what makes slurs wrong. First off, there are occasions

when slurs are intuitively wrong, but utilitarianism entails that they are not. For example, return to the case, mentioned above, of white supremacists who make anti-Semitic remarks only amongst themselves and not to people of Jewish heritage. Since the slurs are not directed towards those who are the object of them, no one can reasonably be expected to be harmed thereby. In addition, one can expect the positive feelings of the ‘in-group’ to be strengthened by virtue of the slurs. Hence, in this circumstance, utility would be maximized by uttering slurs, and hence could even count as morally required by the utilitarian. However, this author’s intuition is that it would still be immoral to do so.

Of course, using slurs in a context in which the object of them is not within earshot is probably not as wrong as doing so purposefully towards those who are the object of them. However, they are still plausibly wrong to some degree. Surely few (if any) readers would ever themselves *encourage* such behaviour or believe it should be encouraged. In addition, most readers, upon reflection, would feel guilty, or at least think they ought to, were they to utter anti-Semitic slurs in such a context.

Supposing, then, that it would indeed be immoral to some extent to utter an anti-Semitic slur only amongst anti-Semites, how might the utilitarian try to explain that? She might initially contend that the reason for the sense of wrongdoing is that one can expect the utterance of slurs in these contexts to get transferred to other ones, where they would indeed cause harm. That is, by indicating that one approved of slurs amongst those of one’s in-group, one would make it more likely that members would say them towards ‘the other’. As a prophylactic, then, one should virtually never communicate slurs, so the reply goes.

However, people are generally pretty good at identifying members of their in-group and regulating their behaviour accordingly. In addition, it would still be wrong to utter a slur even if a neo-Nazi routinely prefaced his remarks this way: ‘I’m calling them “kikes” to you, even though I recommend not doing so to their faces or to anyone else who might report back to them; you know how much power and influence they have, and they could really cause us trouble if they found out we speak this way.’

For a second reply on behalf of the utilitarian, one might suggest that the use of slur terms sustains and fosters attitudes towards particular groups that interfere with building a happier society. Therefore such behaviour is wrong, and not merely because of the immediate pain it can cause those who have been slurred. This reply probably contains a kernel of truth. That is, one reason it would often be wrong to utter a slur in the present context is that it would be likely to make interaction between different groups conflictual and hence to cause unhappiness.

However, it is probably not the whole truth. Imagine a situation in which a neo-Nazi’s fellow anti-Semites are hardcore, meaning their dispositions are extremely unlikely to change, regardless of whether he utters slurs or not. Knowing ahead of time that refraining from slurring would make no difference to anyone else’s behaviour, it appears that the utilitarian must recommend slurring as morally right when it would cause harm that would have happened anyway and would at least make members of the in-group happier for hearing them. However, most readers would still tend to feel bad, and presumably reasonably so, for uttering slurs in those circumstances.

What reflection on these utilitarian replies suggests is that there is something wrong ‘in itself’ with using slurs, by which is meant that they are sometimes immoral apart from their

being responsible for (perhaps indirect and long-term) harmful effects, particularly on those who are the objects of the slurs. Here is a second major way to see this point, in the context of what makes one slur worse than another. Return to the difference between ‘chink’ and ‘suit’, where the former is widely considered to be worse than the latter. According to the welfarist, the reason why calling an Asian businessman ‘chink’ is a greater wrongdoing than calling him ‘suit’ is that he is likely to be more upset upon hearing the former. However, let us consider *why he would be more upset*. Think about the explanation he himself would probably give, and how powerful it is: much of the reason the Asian businessman would be more upset is he deems the slur to be worse! That is, the explanatory relationship is plausibly the *reverse* of what the welfarist suggests: instead of ‘chink’ being more wrong than ‘suit’ merely because the hearer would be more hurt by the former, usually the hearer would be more hurt because he deems ‘chink’ to be more wrong than ‘suit’. Hence, it is apt to search for non-welfarist factors that help to explain the wrongness of slurs.²

4. Autonomy I: Reduction in self-esteem

The other major perspective on right and wrong action in the Western philosophical literature, inspired by the ethic of the German Enlightenment philosopher Immanuel Kant, suggests focusing not on the *hurt* that slurs may be expected to cause, but rather on the *indignity* involved. Instead of accounting for the wrongness of slurs fundamentally in terms of *injury*, the Kantian does so at bottom in terms of *insult*. ‘It is no secret that slurs offend’ (Anderson and Lepore, 2013, 350).³

This article does not invoke Kant’s own ethic, but rather how it has been interpreted of late by Kantian ethicists. According to contemporary Kantians, humans characteristically merit respectful treatment in virtue of their personhood, i.e., their ability to govern themselves as self-conscious beings (perhaps in accordance with moral rules), where wrong acts are those that fail to treat them respectfully, or, equivalently, that degrade them. Clear examples in which a(n innocent) being with the capacity for autonomy is treated in a degrading way are (usually) coercion and deception; where someone has not misused his capacity to make decisions for himself, force or fraud treats that capacity as though it does not exist, is unimportant, or exists merely to be used by others. In contrast, to respect a person, an individual with the capacity for self-governance, normally means engaging with him on the basis of offering him good reasons to perform one action rather than another and seeking his free and informed consent to interact with him in a certain way.

Some contemporary Kantians cash out the degradingness of pejorative speech acts in terms of how they tend foreseeably to reduce people’s autonomy, their ability to make choices for themselves. One way to degrade this capacity is to diminish its potency, trading it off for something worth less than it. And people are plausibly less free, the more they are objects of slurs. To spell this latter claim out, consider John Rawls’ influential discussion of the importance of what he calls ‘self-respect’ in an autonomous life (1999, 91–2, 155–8, 297, 348–9, 386–91) and the fact that slurs are likely to impair it. By ‘self-respect’ Rawls

2 For different objections to welfarism as an account of the wrongness of insults, see Archard (2014, 133–4).

3 See also Jeshion (2013), who attempts to capture how it is that slurs offend.

means what many psychologists and laypeople instead call ‘self-esteem’, which amounts to the combination of three key attitudes: high regard for one’s goals; high confidence in one’s ability to achieve them; and high spirits about oneself and one’s accomplishments. A person who lacks these three elements is less able to make choices, all things being equal, than someone who does not lack them. Rawls plausibly calls self-respect (self-esteem) the most important primary good (1999, 155, 348, 386), by which he means that it is essential for being able to carry out a variety of goals. Crudely stated, a depressed or self-hating person lacks the ability to do much with her life, regardless of how much money, liberty, and other resources she might have. She lacks autonomy in a straightforward respect. Now, a reduction of self-respect (self-esteem) is precisely what tends to be caused by slurs. Slurs have the foreseeable effect of making people not only judge themselves to be unworthy and incapable, but also feel bad about themselves. And reducing a person’s autonomy in this way is patently degrading when done out of indifference, let alone with more nefarious aims.

Although Rawls does not thoroughly discuss pejoratives, he does note that for everyday people, ‘self-respect and their confidence in the value of their own system of ends cannot withstand the indifference much less the contempt of others’ (1999, 297). And Jerome Neu, who does analyze pejoratives in depth, remarks, ‘Feeling insulted is not just anger, not just dejection. There is a disruption of the sense of self, of one’s place in the world and one’s place in the heart and esteem of others’ (2008, 33). Just as segregated schools were deemed unconstitutional by the United States Supreme Court in *Brown v Board of Education* for undermining black pupils’ self-esteem (beyond considerations of the quality and quantity of teachers, books, and other resources), so calling people the n-word and similar slurs should be considered morally forbidden for the same reason of reducing individual autonomy. Conversely, calling others a slur would not be wrong if doing so did not threaten to reduce autonomy, as when certain groups ‘appropriate’ slurs.

As appealing as this rationale is, and even though it is philosophically distinct from a welfarist ethic, it falls prey to the same two objections that faced the latter. First, it has difficulty accounting for the wrongness of slurs that are not directed towards those whom they are about. Above was considered a case of cloistered anti-Semites. Since they do not use slurs beyond their in-group, they do not threaten anyone’s self-esteem. If anything, they might be improving their own self-esteem by putting others down. One can multiply such cases. Imagine a situation in which people slur you behind your back—perhaps literally in the form of deftly taping a derogatory name to the back of your shirt and then removing it before you have a chance to discover it. The wrongness lies not merely in the prospect of making you feel bad about yourself or undermining your confidence; for such an act would be wrong even if there were literally no chance of you finding out or of people treating you differently from the way they would have had the action not been performed. There appears to be something wrong with the action apart from the negative consequences with regard to the self-esteem of the one who is the object of the slur.

Furthermore, relying solely on the self-esteem rationale provides an implausible explanation of degrees of wrongdoing. Returning to the Asian businessman, the present rationale would be that ‘chink’ is worse than ‘suit’ because the former is likely to reduce his self-esteem more than the latter. But why is this the case? Why would he be likely to feel worse about himself and to judge himself less capable, upon being called ‘chink’? Part of the

explanation is that he would judge being called ‘chink’ to be a more degrading treatment. This point again suggests that there is something independently wrong with slurs beyond their negative effects upon being received.

5. Autonomy II: Symbolic disrespect

At this point, the Kantian can and should try another manoeuvre, one that more properly takes advantage of the expressive nature of a respect-based theory. One way to degrade people’s autonomy is causally, e.g., to perform actions that are likely to reduce it, for an end that is worth less than it. However, another way to treat people disrespectfully is symbolically, using conventional representations to convey negative judgements or attitudes. It is plausible to maintain that when slurs are immoral, they are so (at least in part) because of the degrading perspectives they communicate. Wrongful slurs characterize a person, a being with a dignity equal to one’s own, as not having such a value. Although David Archard does not focus on slurs, but rather insults, his remarks about the wrongness of the latter bring to mind this perspective; when it comes to the wrongness of expressive acts, he says that ‘essentially it consists in an attempt to denigrate, humiliate, diminish, dishonour, or disrespect’ (2014, 127).

Since slurs can symbolically degrade another who is their object regardless of whether she becomes aware of the slur or not, the present approach can fairly account for the wrongfulness of those slurs uttered by isolated anti-Semites or posted onto an unsuspecting person’s back. In addition, the symbolic tack does a better job of explaining why some slurs are worse than others than the previous two views; slurs are worse not merely because they tend to have more severe effects on people’s feelings or self-esteem, but in the first instance because they are more degrading in virtue of what they symbolically communicate. What makes one slur symbolically worse than another is a function of a variety of factors, such as the unjust contexts in which it has been used in the past, whether it is used only ever in respect of a particular group of people, and whether it connotes utter sub-humanness. Precisely *how* it becomes the case that a slur means something more degrading than another is something primarily for communication scholars and even historians to sort out (consider, e.g., Anderson and Lepore, 2013a). The point is that one natural thing to say about why ‘chink’ is worse than ‘suit’ is *that* saying the former uses conventional representations to convey a judgement or attitude more degrading than the latter. Hence, the appeal to symbols enables the present version of a Kantian ethic to avoid the major objections made to the other theories considered so far.

Now, it seems clear that a merely symbolic account of why slurs are wrong would be inadequate; after all, their foreseeable effects on people’s lives surely matter, too. It is therefore attractive to suggest, on behalf of the Kantian, that the combination of the self-esteem and symbolic rationales is what does the work. Perhaps slurs are wrong just insofar as they are degrading of persons who have a dignity in virtue of their capacity for autonomy, where degradation can take the form of either reducing self-esteem or using conventional representations to convey an attitude of contempt.

This combined version of Kantianism is attractive. However, there are *prima facie* problems with it, such that it is worth exploring an alternative theoretical approach in the following

section. One worry about the Kantian theory is its difficulty accommodating the intuitions that drive welfarism. Slurs can be designed to hurt, and the pain as something bad for a person—and not merely as something that inhibits her ability to make decisions—is plausibly relevant to their wrongfulness. When someone utters a slur (setting aside ‘appropriated’ contexts), she at best is indifferent to the other person’s feelings and at worst is being cruel. The harm that a slur often foreseeably causes is not adequately captured by the claim that its wrongness consists merely of an ‘attempt to denigrate, humiliate, diminish, dishonour, or disrespect’ (Archard, 2014, 127), supposing that it is the capacity for autonomy that is degraded.

In addition, reflection about why slurs seem worse when face-to-face, amongst intimates, or both suggests that the prospect of reduced self-esteem probably is not the entire explanation. Imagine a white man makes a slur about and directly to his black daughter-in-law. If one were to ask her why that is worse than, say, him saying it merely to his friends or posting the same comment about black people on the web, it is unlikely that she would say that it is *merely* because her self-esteem is expected to be more greatly impaired. *Why* would her self-esteem be more greatly impaired? What might she say in this respect? A plausible account would invoke another, independent consideration, probably about having been in relation with the father-in-law. Similar to the way it is intuitively worse to steal from a friend than a stranger as well as worse to mug someone than to steal behind a person’s back, so it is particularly wrong to direct a slur to a family member and to her face.

The following section presents a theory of why slurring is wrong that promises to capture the advantages of both Kantianism and welfarism, while avoiding their disadvantages. The idea that slurs are wrong roughly because they are unfriendly is now advanced as a plausible monist alternative.

6. Unfriendliness

The gist of the last major theoretical approach to the wrongfulness of slurs is that the ways that family members, neighbours, and colleagues characteristically do and should relate to one another ought to be pursued as ends. Roughly speaking, a moral agent normally ought to strive to come closer to others and to eschew distancing herself from them. Although one finds these kinds of ideas in a variety of philosophical traditions, the prizing of relationality is particularly prominent in those in the Global South, where harmony is central to much indigenous thought in sub-Saharan African, East Asian (especially Confucian), and South American ethics (on which see Bell and Mo, 2014; Metz and Miller, 2016; Graness, 2019; Metz, 2020). Few Western philosophers are familiar with these long-standing traditions. The following draws particularly on the Southern African ethic of *ubuntu* to articulate a principle about what we owe to each other that is revealing in respect of slurs.⁴

There is a maxim that many people indigenous to Southern Africa often invoke to sum up how to live ethically, namely, ‘A person is a person through other persons’ (e.g., Tutu, 1999, 35; Dandala, 2009, 160; for discussion in the context of several sub-Saharan peoples, see Nkulu-N’Sengha, 2009). When they say that ‘a person is a person’, what they mean usually includes the idea that someone who is a person, in the sense of a self-conscious,

4 Particularly as it is articulated in Metz (2021a, 2021b).

deliberative agent, ought to strive to become a real or genuine person, that is, someone who exhibits moral virtue. One with the latter has what is somewhat famously called '*ubuntu*', literally humanness in the Nguni languages of Southern Africa such as Zulu and Ndebele. A true or complete person is someone who lives a genuinely human way of life, who displays ethical traits that human beings are in a position to exhibit in a way nothing else in the animal, vegetable, or mineral kingdoms can. Just as one might say that a jalopy is 'not a real car', so indigenous Africans often say that agents who have been immoral 'are not a person' (Gaie, 2007, 32–3; Dandala, 2009, 260–1) or are even 'animals' (Bhengu, 1996, 27; Letseka, 2000, 186). That does not mean that the wicked are literally not human beings and hence no longer subjects of human rights, but instead connotes the metaphorical point that these individuals utterly fail to exhibit human (moral) excellence and have instead actualized their lower, base nature (Gyekye, 2010, sec. 4).

Turning to the second clause, it tells people how to become real persons and to exhibit *ubuntu*, viz., 'through other persons'. Typically this means by entering into community with others or seeking to live harmoniously with them (Tutu, 1999, 34–5; Mnyaka and Motlhabi, 2005, 218). It is well known that African ethics is characteristically communitarian, but this element is often left vague or is construed in a crude manner, such as the collective taking precedence over the individual. However, it can also, and more plausibly, be interpreted in relational terms, which is suggested by the following comments from South African intellectuals about it. Former South African Constitutional Court Justice Yvonne Mokgoro remarks of an *ubuntu* ethic that 'harmony is achieved through close and sympathetic social relations within the group' (1998, 17). Gessler Muxe Nkondo, who has had positions of leadership on South Africa's National Heritage Council, says that '*ubuntu* advocates (...) express commitment to the good of the community in which their identities were formed, and a need to experience their lives as bound up in that of their community' (2007, 91). Nhlanhla Mkhize, an academic psychologist at the University of KwaZulu-Natal, remarks that a 'sense of community exists if people are mutually responsive to one another's needs (...). (O)ne attains the complements associated with full or mature selfhood through participation in a community of similarly constituted selves (...). To be is to belong and to participate' (2008, 39, 40).

These and other common construals about what it is to commune or to live harmoniously with others suggest two recurrent themes that can be distinguished (Metz, 2021a). On the one hand, there is 'identity', a matter of being close, belonging and participating, and experiencing life as bound up with others. On the other hand, one finds reference to being sympathetic, being committed to others' good, and responding to others' needs, here labelled 'solidarity'.

More carefully, it is revealing to understand identifying with another (or being close, belonging, etc.) to be the combination of exhibiting certain psychological attitudes of 'witness' and cooperative behaviour. The psychological attitudes include a tendency to think of oneself as a member of a group or relationship with the other and to refer to oneself as a 'we' (rather than an 'I'), a disposition to feel pride or shame in what the other or one's group does, and, at a higher level of intensity, an emotional appreciation of the other's nature, value, and presence. The cooperative behaviours include being transparent about the terms of interaction, allowing others to make voluntary choices, acting on the basis of trust, adopting common goals, and, at the extreme end, choosing for the reason that 'this is who we are'.

Exhibiting solidarity with another (or acting for others' good, etc.) is similarly aptly construed as the combination of exhibiting certain psychological attitudes and engaging in helpful behaviour. Here, the attitudes are ones positively oriented toward the other's good and include an empathetic awareness of the other's condition and a sympathetic emotional reaction to this awareness. The actions are not merely those likely to be beneficial, that is, to improve the other's state, but also are ones done consequent to certain motives, say, for the sake of making the other better off or even a better person.

Identity and solidarity are plausibly at the core of what is desirable about a family or a collegial academic department. When both are present, people enjoy a common sense of self, cooperate with each other, seek to promote others' good, and do so for one another's sake. Upon living in such a friendly or harmonious way, one thereby exhibits *ubuntu*.

So far, *ubuntu* has been interpreted as a virtue ethic, indeed one sharing features with Greek *eudaemonism*. Although that is the traditional approach to ethics for indigenous Africans, what is needed is an interpretation similar in form to welfarism and Kantianism, which here have been advanced as accounts of what makes it wrong to slur, not what makes those who slur bad. In order to make sense of the wrongfulness of slurs, one more concept needs to be added, namely, that normally we ought to treat people in a friendly or harmonious way because they have a dignity. More specifically, since people have a dignity in virtue of their capacity to relate harmoniously, we usually have good reason to relate harmoniously to them, at least when they are innocent.

According to this ethic, the reason that it is immoral to kidnap, rape, steal, lie, or exploit is that doing so would be extremely unfriendly or discordant, an egregious failure to treat people as special because of their capacity to be friendly and to be befriended. Such actions characteristically involve: thinking of others as separate and inferior (as opposed to bound up with oneself), subordinating others (instead of coordinating in pursuit of compatible aims), impairing others' quality of life (instead of helping them), and acting on attitudes of indifference towards others' interests (instead of ones of sympathy and altruism). When we fail to harmonize with other innocents, and especially when we act discordantly towards them, we are degrading them, treating their capacity to be party to harmonious relationships as either non-existent or unimportant.

It would be of philosophical interest to weigh up this principle of right and wrong action against utilitarianism and Kantianism as competing comprehensive accounts of right and wrong action (undertaken in Metz, 2021b). However, this article compares them only in regard to the immorality of slurs. According to the view inspired by Afro-communal values, a slur is wrong insofar as it degrades people's capacity to befriend and to be befriended. Typically such degradation is expressed when one treats a person in an unfriendly way and not because this person had initially been unfriendly.

Consider what is happening when a blogger labels women 'bitches' on a website. He is treating women, and specifically their capacity to befriend and be befriended, disrespectfully by virtue of being unfriendly towards them. By communicating the idea that traits such as being controlling and unreasonable are associated with a group of people simply because of their gender, one is thereby acting in an unfriendly manner towards that group; for one is expressing the judgement that women are separate and inferior, participating in a practice that keeps women in a subordinate position, doing what is likely to cause them psycholo-

gical pain, and acting in a way consistent with negative attitudes such as cruelty in regard to others' well-being and excellence. To act in such a discordant manner towards people who have not misused their capacity to be party to harmonious relationships is to express disrespect for that dignified capacity.

Someone who utters a slur such as 'bitch' usually has various things in mind, such as that women are controlling and unreasonable and that they are 'less than' because of that. However, it is possible to express disrespect of women's relational dignity with a slurring term even if one is not thinking these things. Since many people have used the word to convey such judgments, the sentence or conventional meaning of the word can connote them in certain contexts, regardless of what the person who uses the word might intend at the time. Hence, one could act in unfriendly ways towards others, and thereby express disrespect of their capacity to be party to friendly relationships, by using a slur despite lacking a full awareness of what it signifies, e.g., by yelling 'kike' at a KKK rally when one is ignorant of what it means.

The claim is not that slurs are inherently unfriendly or discordant, but is instead that, *when* slurs are wrong, they are so in virtue of that. There are times when slurs can in fact be friendly, for instance when one is recognized as part of an in-group (a member of a 'we') that has 'appropriated' the slur. The present theory provides a strong explanation of why slurring in such a context would not be wrong: it would not be unfriendly, and, indeed, could well be something that is to be encouraged for strengthening ties such as a sense of togetherness and warm feelings! Of course, sometimes cohesion amongst an in-group is enhanced by directing slurs towards those in an innocent out-group. However, then there would be discord exhibited towards the latter and hence wrongness. A slur can intuitively be permissible if it is friendly towards some without being unfriendly towards others. In that case, no one's relational dignity is degraded and there is nothing wrong with using the slur.

Let us return to the welfare- and autonomy-based theories, and consider how the relational approach captures what is appealing about them, while avoiding and even explaining the objections to them. Consider, first, welfarism. Recall it was accepted that the fact that slurs can be hurtful, whether intentionally or merely foreseeably, is partially relevant to why they are wrong. Now, if one is to avoid treating innocent parties in unfriendly ways, then one generally should avoid causing them harm and instead should act out of solidarity. So far, so good.

Turning to objections to welfarism, remember that it had difficulty accounting for the wrongfulness of slurs when there is no chance of those who are the object of them becoming aware of them. There is the case of someone having pinned a slur to your back. The friendliness theory entails that such a slur is nonetheless wrong because, despite the unlikelihood of harm to you, a failure to prize friendliness with you remains. In particular, it is the valuable capacity for identity that is flouted (at least to an extent greater than the capacity for solidarity). In communicating to those besides you the idea that you can be identified with negative traits simply because of your membership in a group, the one slurring is acting in an unfriendly, and particularly divisive, way towards you. For one, it is action consequent to distant attitudes on his part, ones that are hardly expressive of, or disposed to foster, a sense of togetherness or identification with you. For another, there is an undermining of your ends and hence a kind of subordination, insofar as the treatment frustrates your aim not to have your body used without your permission and for the purpose of conveying negative sentiments about you.

Let us also return to the problem with welfarism regarding its explanation for degrees of the wrongfulness of slurs. It does so in terms of amount of (expected) distress or other impairment of subjective well-being, whereas recall that there appears to be an independent moral consideration that explains the greater distress upon being called ‘chink’ compared to ‘suit’. The relational theory accounts for the differential degrees of wrongness between the slurs, and hence differential psychological reactions to them, in terms of there being a greater unfriendliness and hence greater disrespect. Normally someone calls another person a ‘chink’ consequent to a stronger judgement of otherness and inferiority than what is meant by ‘suit’. The hostile attitude conveyed by the former is more intense than what is conveyed by the latter, which moral consideration at bottom explains why there tends to be more intense pain felt upon the victim becoming aware of having been called the former.

Turn, now, to the autonomy-based theory of the wrongfulness of slurs. The strong advantage of this view is its ability to account for the ideas that slurs do seem wrong in virtue of both tending to reduce self-esteem and being symbols that contain disrespectful messages. The friendliness theory can claim these advantages, while avoiding Kantianism’s disadvantages. First off, since both Kantianism and the friendliness prescribe respect for human dignity, they comparably capture the relevance of symbolic considerations. Note that the latter view can also comparably explain the relevance of self-esteem. Performing actions that are likely to reduce another’s self-esteem, at least for intuitively trivial reasons, is to fail to prize a friendly relationship *qua* one that involves cooperation. Doing what one foresees is likely to make others less able to trust and to coordinate behaviour in pursuit of compatible ends is divisively unfriendly and hence degrading.

The previous section identified three logically distinct reasons to question the adequacy of the Kantian view. One is that that it is not readily able to account for the pertinence of welfarist considerations with regard to the wrongness of slurs, as morally relevant beyond impairment of the capacity to choose. In contrast, as noted above, part of treating someone as special in virtue of her ability to be party to relationships of identity and solidarity is to exhibit solidarity with her and hence to avoid causing her harm.

A second problem with Kantianism is that it has difficulty explaining the moral relevance of whether a slur is, say, uttered to someone’s face or instead expressed on the internet. Slurring someone directly makes it personal, which does not seem relevant from the standpoint of respecting the capacity to make decisions autonomously. In contrast, the natural thing to say about a personal slur is that it is characteristically more hostile. There is a greater unfriendliness in slurring a person to his face than simply posting the slur on the web, plausibly in virtue of not merely the greater psychological harm caused, but also the greater flouting of the value of enjoying a sense of togetherness. It is behaviour that undermines any sense of ‘we-ness’ by expressing an intense attitude of ‘me versus you’, for which the relational ethic naturally accounts.

Finally, the third reason to doubt Kantianism is that its focus on autonomy appears unable to capture the moral relevance of who is the object of the slur. Recall that it seems *pro tanto* worse for a person to direct a slur at someone who is part of the family as opposed to someone who is not. It is apt, here, to appeal to relational considerations to make sense of the differential degrees of wrongness. It is a greater flouting of the valuable capacity for friendliness to direct a slur towards someone with whom one has already had somewhat

friendly ties, as opposed to someone with whom one has not. Once one begins to share a sense of self with someone and to coordinate behaviour in the context of a family, one incurs a greater obligation to go out of one's way to share a sense of self with that person, support her ends, help her, and do so for her sake. Slurring her is all the worse, given this extra partial duty that a relational ethic explains particularly well.

7. Conclusion

This article has sought to test theoretical hypotheses about why slurs are wrong in the light of intuitions about when they are, why they are, and to what degree. It has considered three major theories, the views that slurs are wrong insofar as they tend to harm people, to degrade their autonomy, and to be unfriendly. This article probably has not presented conclusive evidence in favour of the theory that slurs are wrong insofar as they are unfriendly or, more carefully, fail to prize people in virtue of their capacity for friendly relationships, which has been developed in the light of ideas about harmony from the globally under-explored tradition of Southern African ethical thought. While it might be too soon to jettison the dominant, welfare- and autonomy-based accounts of the wrongfulness of slurs, the analysis in this article entails that those interested in that topic should not disregard the new, friendly-based approach. Relational ideas from the Global South merit consideration in respect of the immorality of slurs, particularly the negative ones that have been the focus of this article, but also any positive ones that might exist as well as other pejoratives.

It would also be worth considering whether the explanation of the wrongfulness of slurs advanced here might be plausibly extended to other kinds of immorality, such as discrimination in the allocation of benefits, e.g., jobs and housing. Perhaps that, too, is wrong because a negative response is directed towards people simply because they are members of certain (perhaps socially salient) groups, which is degradingly unfriendly since they have not misused their capacity to be party to friendly relationships.

References

- Anderson, L. & Lepore, E. (2013a), 'Slurring words', *Noûs*, 47(1), 25–48.
- Anderson, L. & Lepore, E. (2013b), 'What did you call me? Slurs as prohibited words', *Analytic Philosophy*, 54(3), 350–363.
- Archard, D. (2014), 'Insults, free speech and offensiveness', *Journal of Applied Philosophy*, 31(2), 127–141.
- Bell, D. A. & Mo, Y. (2014), 'Harmony in the world 2013', *Social Indicators Research*, 118(2), 797–818.
- Bentham, J. (1822), 'Letters to Count Toreno, on the proposed penal code', in J. Bowring (ed.), *The Works of Jeremy Bentham, Volume 8*, Edinburgh: William Tait, pp. 487–554.
- Bhengu, M. J. (1996), *Ubuntu: The Essence of Democracy*, Cape Town: Novalis Press.
- Boisvert, D. (2008), 'Expressive-assertivism', *Pacific Philosophical Quarterly*, 89(2), 169–203.
- Camp, E. (2013), 'Slurring perspectives', *Analytic Philosophy*, 54(3), 330–349.
- Croom, A. M. (2011), 'Slurs', *Language Sciences*, 33(3), 343–358.

- Croom, A. M. (2013), 'How to do things with slurs', *Language and Communication*, 33(3), 177–204.
- Dandala, M. (2009), 'Cows never die: Embracing African cosmology in the process of economic growth', in M. F. Murove (ed.), *African Ethics*, Pietermaritzburg: KwaZulu-Natal Press, pp. 259–277.
- Feinberg, J. (1985), *Offence to Others*, Oxford: Oxford University Press.
- Feldman, F. (1978), *Introductory Ethics*, Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Gaie, J. (2007), 'The Setswana concept of *botho*', in J. Gaie & S. Mmolai (eds.), *The Concept of Botho and HIV/AIDS in Botswana*, Eldoret: Zapf Chancery, pp. 29–43.
- Graness, A. (2019), '*Ubuntu* and *buen vivir*', in J. Ogude (ed.), *Ubuntu and the Reconstitution of Community*, Bloomington: Indiana University Press, pp. 150–175.
- Grice, P. (1957), 'Meaning', *The Philosophical Review*, 66(3), 377–388.
- Gyekye, K. (2010), 'African ethics', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/fall2010/entries/african-ethics/>.
- Hay, R. (2013), 'Hybrid expressivism and the analogy between pejoratives and moral language', *European Journal of Philosophy*, 21(3), 450–474.
- Himma, K. (2002), 'On the definition of unconscionable racial and sexual slurs', *Journal of Social Philosophy*, 33(3), 512–522.
- Hom, C. (2008), 'The semantics of racial epithets', *Journal of Philosophy*, 105(8), 416–440.
- Hom, C. (2010), 'Pejoratives', *Philosophy Compass*, 5(2), 164–185.
- Hornsby, J. (2001), 'Meaning and uselessness: how to think about derogatory words', *Midwest Studies in Philosophy*, 25(1), 128–141.
- Jeshion, R. (2013a), 'Expressivism and the offensiveness of slurs', *Philosophical Perspectives*, 27(1), 231–259.
- Jeshion, R. (2013b), 'Slurs and stereotypes', *Analytic Philosophy*, 54(3), 314–329.
- Letseka, M. (2000), 'African philosophy and educational discourse', in P. Higgs et al. (eds.), *African Voices in Education*, Cape Town: Juta, pp. 179–193.
- Lippert-Rasmussen, K. (2014), *Born Free and Equal?*, Oxford: Oxford University Press.
- Metz, T. (2020), 'Relational normative economics: An African theory of justice', *Ethical Perspectives*, 27(1), 35–68.
- Metz, T. (2021a), '*Ubuntu*: The good life', revised edition, in F. Maggino (ed.), *Encyclopedia of Quality of Life and Well-Being Research*, 2nd ed., Cham: Springer, pp. 1–5.
- Metz, T. (2021b), *A Relational Moral Theory: African Ethics in and Beyond the Continent*, Oxford: Oxford University Press.
- Metz, T. & Miller, S. C. (2016), 'Relational ethics', in H. LaFollette, (ed.), *The International Encyclopedia of Ethics*, Malden, MA: Blackwell, pp. 1–10.
- Mkhize, N. (2008), '*Ubuntu* and harmony', in R. Nicolson (ed.), *Persons in Community: African Ethics in a Global Culture*, Pietermaritzburg: University of KwaZulu-Natal Press, pp. 35–44.
- Mnyaka, M. & Motlhabi, M. (2005), 'The African concept of *ubuntu/botho* and its socio-moral significance', *Black Theology*, 3(2), 215–237.
- Mokgoro, Y. (1998), '*Ubuntu* and the law in South Africa', *Potchefstroom Electronic Law Journal*, 1(1), 15–26.

- Neu, J. (2008), *Sticks and Stones: The Philosophy of Insults*, Oxford: Oxford University Press.
- Nkondo, G. M. (2007), 'Ubuntu as a public policy in South Africa: A conceptual framework', *International Journal of African Renaissance Studies*, 2(1), 88–100.
- Nkulu-N'Sengha, M. (2009), 'Bumuntu', in M. K. Asante & A. Mazama (eds), *Encyclopedia of African Religion*, Los Angeles: Sage, pp. 142–147.
- Rawls, J. (1999), *A Theory of Justice*, revised edition, Cambridge, MA: Harvard University Press.
- Richard, M. (2008), *When Truth Gives Out*, Oxford: Oxford University Press.
- Tutu, D. (1999), *No Future Without Forgiveness*, New York: Random House.
- Whiting, D. (2013), 'It's not what you said, it's the way you said it: Slurs and conventional implicatures', *Analytic Philosophy*, 54(3), 364–377.



The resistant effect of slurs: A nonpropositional, presuppositional account*

El efecto resistente de los *slurs*: Una propuesta presuposicional no proposicional

ALBA MORENO ZURITA**
EDUARDO PÉREZ-NAVARRO***, ****

Abstract: The aim of this paper is to account for the *resistance to cancelation, rejection, and retraction* exhibited by slurs. The kind of explanation we offer is a presuppositional one. Like the most recent presuppositional accounts, moreover, ours is a *nonpropositional* presuppositional proposal. Our view is that, to be felicitous, utterances of sentences featuring slurs require certain components to be part of the common ground, but these components are not propositions, but *world-orderings*.

Keywords: slurs, resistance, cancelation, rejection, retraction, presupposition

Resumen: El propósito de este artículo es dar cuenta de la *resistencia a la cancelación, el rechazo y la retractación* exhibida por los *slurs*. La explicación que ofrecemos es presuposicional. Como las teorías presuposicionales más recientes, además, la nuestra es una propuesta presuposicional *no proposicional*. Nuestra posición es que, para ser exitosas, las preferencias de oraciones que contienen *slurs* requieren que el *common ground* incluya ciertos componentes, pero estos componentes no son proposiciones, sino *ordenamientos de mundos*.

Palabras clave: *slurs*, resistencia, cancelación, rechazo, retractación, presuposición

Recibido: 28/05/2021. Aceptado: 01/07/2021.

* This paper has been funded by the Spanish Ministry of Science and Innovation under the research project “Disagreement in Attitudes: Normativity, Affective Polarization and Disagreement” (PID2019-109764RB-I00), by the Regional Government of Andalusia under the research projects “Public Disagreements, Affective Polarization and Immigration in Andalusia” (B-HUM-459-UGR18) and “The Inferential Identification of Propositions: A Reconsideration of Classical Dichotomies in Metaphysics, Semantics and Pragmatics” (P18-FR-2907), and by the University of Granada under a “Contrato Puente” fellowship and the excellence unit FiloLab-UGR (UCE. PPP2017.04). The authors would also like to thank Alex Davies, María José Frápolli, Claudia Picazo, Andrés Soria, and two anonymous reviewers for *Daimon* for their helpful comments and suggestions.

** PhD candidate at the Department of Philosophy I, University of Granada. E-mail: almorenozurita@gmail.com. Her interest areas are philosophy of language and social epistemology. A recent publication of hers is Almagro, M. & Moreno, A. (forthcoming), “Affective polarization and testimonial and discursive injustice”, in D. Bordonaba, V. Fernández Castro and J. R. Torices (eds.), *The Political Turn in Analytic Philosophy: Reflections on Social Injustice and Oppressions*, Berlin: De Gruyter.

*** Postdoctoral researcher at the Department of Philosophy I, University of Granada. E-mail: edperez@ugr.es. His interest areas are philosophy of language and philosophy of logic. Two recent publications of his are Pérez-Navarro, E. (2021), “The way things go: Moral relativism and suspension of judgment”, *Philosophical Studies*, DOI: 10.1007/s11098-021-01650-z, and Pérez-Navarro, E. (2021), “No matter who: What makes one a relativist?”, *Theoria: An International Journal for Theory, History and Foundations of Science*, 36(2), 231–242.

**** The alphabetical order of the authors is meant to reflect their equal contribution to this paper.

1. Introduction

Sentences featuring slurs exhibit peculiar behavior. While there are a number of ways in which the effect of other speech acts can be blocked, exporting these strategies to avoid the harmful consequences of using a slur tends to be pointless. In particular, while one can usually *cancel*, *reject*, or *retract* the effect of other speech acts, these moves seem quite more difficult to make when the uttered sentence includes a slur.¹ Compare the following sentences, in which “S” replaces a slur:

- (1) The cat is on the mat.
- (2) A is an S.
- (3) If the cat is on the mat, we should vacuum it before my parents arrive.
- (4) If A is an S, she shouldn’t go to that restaurant.
- (5) No, it’s not.
- (6) No, she’s not.
- (7) I take that back. (Targeting one’s previous utterance of (1).)
- (8) I take that back. (Targeting one’s previous utterance of (2).)

As we will see in due time, while (3) does not commit the speaker to the proposition that the cat is on the mat like (1) would, (4) is as derogatory as (2). This is what it means for the effect of slurs to be difficult to cancel. While (5) serves to reject the proposition that the cat is on the mat, one cannot successfully reject the derogatory effect of (2) by uttering (6). This is what we mean when we say that the effect of slurs is difficult to reject. Finally, while the speaker can successfully remove from the common ground the previously introduced proposition that the cat is on the mat by uttering (7), (8) does not seem as successful when the aim is to go back to a conversational state prior to the utterance of (2). That is, the effect of slurs is difficult to retract.

The aim of this paper is to account, building upon Marques and García-Carpintero (2020), for a feature of slurs that is meant to encompass this behavior—we say that their effect is *resistant*, either to cancelation, rejection, or retraction. This seems like a project worth embarking on, since the resistant character of slurs is part of what makes these words so harmful. Resistance to cancelation has been widely discussed (Jeshion, 2013a; Sennett and Copp, 2015; Cepollaro and Stojanovic, 2016; Camp, 2018; Marques and García-Carpintero, 2020), and resistance to rejection has been addressed to a certain extent in the literature too (Camp, 2013; Jeshion, 2013a, 2013b; Cepollaro and Stojanovic, 2016). Our explanation will be similar to previous ones given for resistance to cancelation, but we will also get inspiration from these strategies to account for resistance to rejection in a novel way. This paper also breaks new ground insofar as it focuses on resistance to retraction,

1 In what follows, we talk about the effect of slurs to refer to their *derogatory* effect. For the time being, we would like to remain neutral as to whether this effect is illocutionary or perlocutionary—our point is just that the derogatory effect of slurs, whether illocutionary or perlocutionary, is difficult to cancel, reject, and retract. Since the present discussion shows that this effect of slurs behaves differently from the illocutionary effect of assertions, one might be tempted to say that it does so because it is perlocutionary. As we will see, however, our proposal in this paper will eventually be that derogation is something that takes place *in* using a slur.

a largely underexplored aspect of the behavior of slurs (see, however, McGowan, 2009; Simpson, 2013). Still, we take the work done here to be a programmatic effort, in need of further development in future contributions.

The kind of explanation we offer is a presuppositional one. Like the most recent presuppositional accounts, moreover, ours is a *nonpropositional* presuppositional proposal. To be felicitous, utterances of sentences featuring slurs require certain components to be part of the common ground, but these components are not propositions. As advanced above, this is inspired by Marques and García-Carpintero's (2020) work, according to which it is reactive attitudes taken to be appropriate that the common ground needs to be enriched with. We follow Marques and García-Carpintero in using a nonpropositional account to explain why the effect of slurs persists in a range of constructions, as presuppositions do, but also in environments in which presuppositions are not expected to survive. We also use it to account for slurs' resistance to rejection and retraction. However, our explanation only depends on one aspect of Marques and García-Carpintero's proposal—the idea that the presuppositions involved are nonpropositional. We do not need them to concern reactive attitudes, and in fact will offer some reasons to understand them in terms of *world-orderings* instead. By doing so, we expect to provide an account of nonpropositionality that has independent interest.

The structure of the paper is as follows. In section 2, we explore what it means to say that the effect of slurs is resistant. We contrast the effect of sentences featuring slurs with that of other sentences and find that it is much more difficult to cancel, reject, and retract. In section 3, we survey a number of proposals that have relied on presuppositions to account for the meaning of slurs and how they have evolved to explain what makes the presuppositions triggered by slurs special. In our view, however, no account has satisfactorily done this until Marques and García-Carpintero's, which we discuss in section 4 together with our proposed modifications and our reasons to incorporate them. Sections 5–7 are devoted to showing how the proposal resulting from section 4 can account for the resistant effect of slurs. In particular, section 5 explain slurs' resistance to cancelation along Marques and García-Carpintero's terms, while section 6 extends this explanation to flesh out Cepollaro and Stojanovic's (2016) account of slurs' resistance to rejection and section 7 applies it to slurs' resistance to retraction.

2. The resistant effect of slurs

In this section, we describe the peculiar behavior of slurs, for which it is the aim of this paper to account. In particular, we characterize the derogatory effect of slurs as difficult to cancel, reject, and retract. We will summarize these features of slurs by saying that their effect is *resistant*.

There are a number of places where communication can go awry. Suppose I say

(1) The cat is on the mat.

The effect of this assertion is to add to the common ground the proposition that the cat is on the mat. This happens in two steps. In the first place, by uttering (1), I propose to add to the common ground the proposition that the cat is on the mat. In the second place, the proposition is added if my audience accepts my proposal.

Correspondingly, there are three ways in which I could utter the words in (1) without them having their default effect. First, something could go wrong before even the first step takes place. If the words are uttered in a linguistic environment that cancels their effect, the proposal to add to the common ground the proposition that the cat is on the mat will not even be made. This is what happens, for instance, when the words appear in the antecedent of a conditional, as they do in

(3) If the cat is on the mat, we should vacuum it before my parents arrive.

We summarize this by saying that the effect of an assertion is *cancelable*. Second, something could go wrong between the first and the second step. In particular, it could be that my audience does not accept my proposal. This is what happens when they reply to my assertion with something like (5) “No, it’s not”. We summarize this by saying that the effect of an assertion is *rejectable*. Lastly, something could go wrong after the second step. This is what happens when, after the proposition that the cat is on the mat has gone into the common ground, I say something like (7) “I take that back”. If this later assertion is accepted by my audience, the proposition will be taken out of the common ground. We summarize this by saying that the effect of an assertion is *retractable*.

It is worth noting that, in spite of the order in which we have introduced cancellation, rejection, and retraction, the first and the last of these phenomena can be seen as closer to each other than to rejection. After all, they are things that the speaker does, while rejection is done by the hearer. But note too that rejection and retraction target a certain content, while cancellation aims precisely at preventing such content from being conveyed. It is also worth mentioning that Caponetto (2020) advances a different taxonomy of ways of blocking the effect of a speech act. She discusses retraction like we do (Caponetto, 2020, 2407–10) and her *amendment* (Caponetto, 2020, 2410–12) bears a (passing) resemblance with our cancellation, but instead of rejection she considers *annulment* (Caponetto, 2020, 2404–6), which seems a wholly different phenomenon.

The picture above seems to work just fine with ordinary assertions. There are cases, however, in which cancellation, rejection, and retraction seem a bit harder. In particular, the derogatory effect of slurs, however it takes place, seems harder to cancel, reject and retract than the effect of an assertion. It is present even when they occur in the antecedent of a conditional, as in

(4) If A is an S, she shouldn’t go to that restaurant.

And also when the sentence in which they appear is negated or posed in the form of a question:

(9) A is not an S.

(10) Is A an S?

Slurs are also difficult to reject—it is not obvious that replying to a predicative use of a slur with (6) “No, she’s not” directly addresses its derogatory effect, and trying to do so is

commonly felt to interrupt the flow of the conversation (Chilton, 2004, 64; Stanley, 2015, 170). And they are difficult to retract—derogation does not seem able to be undone by saying (8) “I take that back”.²

We summarize this by saying that the derogatory effect of slurs is *resistant*. The first two ways in which it is so—especially its resistance to cancelation—have been considerably explored in the literature (see Jeshion, 2013a; Sennett and Copp, 2015; Cepollaro and Stojanovic, 2016; Camp, 2018; Marques and García-Carpintero, 2020 for cancelation and Camp, 2013; Jeshion, 2013a, 2013b; Cepollaro and Stojanovic, 2016 for rejection). Slurs’ resistance to retraction, however, has gone relatively unacknowledged (see, however, McGowan, 2009; Simpson, 2013). In this paper, we advance an explanation of it, but our proposal is intended to account for slurs’ resistance to cancelation and rejection as well. The proposal is a presuppositional one. In the next section, accordingly, we survey this family of views.

3. Presuppositional accounts of slurs

In the previous section, we saw that the derogatory effect of slurs seems to be resistant to cancelation, rejection, and retraction. In the next section, we propose an account of the meaning of slurs that, as we will devote the rest of the paper to argue, explains this set of features. This is a presuppositional account of the meaning of slurs. There have been a number of proposals that have aimed at accounting for the meaning of slurs in presuppositional terms. Only the most recent of them, however, have tried to combine the presuppositional treatment with an explanation of what makes the derogatory content of slurs *evaluative*. We will follow the latter path, characterizing derogatory content as nonpropositional, presupposed content. Before doing this, however, we will devote this section to surveying how presuppositional accounts have evolved to this point.

The presuppositional account of the meaning of slurs can be traced back to Macià (2002). According to Macià, sentences featuring expressives trigger presuppositions; thus, they determine partial functions from worlds to truth-values, defined only for those worlds in which the sentence’s presuppositions are true. For instance,

(11) Toni Morrison is a Nobel laureate in spite of being African American.

determines a partial function that assigns truth to all worlds in which Toni Morrison is a Nobel laureate and is an African American, but only among the worlds in which African Americans are unlikely to win a Nobel prize. Making the sentence presuppose this is what “in spite of” does. As expressives, slurs will behave in a similar way.

It should be noted that utterances of sentences whose presuppositions are false in the context are infelicitous, but the presupposition can be made true in the context *as the utterance is made*. In cases like this, a process of *accommodation* is said to take place (Lewis, 1979). This means that the context is modified in any way needed for the utterance to be felicitous. In this case, in particular, the presupposition at issue would automatically be added

2 We will describe these features of the behavior of slurs in more detail in sections 5–7, in which we will show how our proposal is able to account for them.

to the context. The fact that this addition is automatic will play a fundamental role when we use presuppositions to account for the resistant character of the derogatory effect of slurs.

It is in Schlenker's (2007) work that, for the first time, we find an account specifically targeting slurs (along with other expressives). Schlenker takes sentences featuring slurs to trigger presuppositions that have the following features (2007, 237). First, they are *semantic* presuppositions, as opposed to pragmatic ones, insofar as it is the meaning of the slur that triggers them. Second, they are *indexical* presuppositions, for they are bound to the context of utterance. Third, they have an *attitudinal* component, i.e., they make reference to the agent's mental state. Finally, the agent whose mental state the proposition refers to does not necessarily have to be the speaker. Unlike in Macià's view, these four features together make the presuppositions triggered by slurs expressive. Thus, if someone utters (2) "A is an S" and it is presupposed in the context that the agent believes that *Ns* are bad, where "*N*" is the neutral counterpart for "*S*", the sentence will be true if *A* is an *N* and false otherwise. If there is no such presupposition in the context, however, (2) will lack truth-value.

Williamson (2009) and Nunberg (2018) find Schlenker's proposal defective for the following reason. If, previously to the utterance of (2), it is not presupposed in the context that the agent believes that *Ns* are bad, the presupposition will be accommodated and become part of the common ground. However, it is problematic to accept an utterance of (2), while there should be no problem with accepting that a certain agent has a certain belief (Williamson, 2009, 151–2; Nunberg, 2018, 284; see also Marques and García-Carpintero, 2020, 142–3).

Objections like this have driven some authors to propose to specify the content of the presupposition triggered by the utterance of a slur in ways other than Schlenker's. Cepollaro and Stojanovic (2016), first, and later Cepollaro (2020) in a more developed form, coincide with Schlenker in deeming these presuppositions different from descriptive ones,³ but render them as the mere presupposition that *Ns* are bad because of being *Ns* (Cepollaro and Stojanovic, 2016, 459). When we utter (2), therefore, we express two kinds of content. On the one hand, we express a descriptive content, which determines the truth-conditions of the sentence. This descriptive content is just like that of "A is an *N*". On the other hand, we express an expressive content that is presupposed and can be rendered as "*Ns* are bad because of being *Ns*". Similarly to Schlenker's proposal, the sentence will be true or false depending on whether *A* is an *N* only if it is presupposed that *Ns* are bad because of being *Ns*. If it is not, it will lack truth-value.

4. A nonpropositional, presuppositional account of slurs

Cepollaro and Stojanovic avoid the problem faced by Schlenker's presuppositions. However, they do so at the cost of leaving the attitudinal component out of the picture. In this way, they leave quite unexplained what makes their presuppositions expressive, other than being able to be made explicit through sentences that include evaluative terms like "bad". Marques and García-Carpintero (2020) propose a presuppositional account that fleshes out what it means for a presupposition to be expressive while avoiding Schlenker's problem.

3 Unlike Schlenker, and later Marques and García-Carpintero, Cepollaro and Stojanovic call these presuppositions "evaluative" instead of "expressive". However, the two terms should be equivalent for our purposes.

To do this, they need to extend the common ground to include not only possible worlds, but also other components. This goes in line with how the common ground has been proposed to incorporate partitions to account for the effect of questions (Stalnaker, 2014) or to-do lists to account for the effect of directives (Portner, 2007).

What Marques and García-Carpintero propose to add to the common ground, in particular, is a set of *reactive attitudes* (Strawson, 1962) that are taken to be appropriate by the participants in the conversation. Reactive attitudes are emotions that, as such, consist of two aspects—the intentional object that is the target of the attitude and the formal object or attitude that is directed toward that object (Kenny, 1963). We could thus have as part of the common ground a set that includes the reactive attitude of contempt toward a given group. Such an attitude will then be taken to be accepted by participants in the conversation. For an utterance of (2) “A is an S” to be felicitous, the set of reactive attitudes in the common ground needs to include contempt toward the Ns. If it does, whether the sentence is true or false will depend on whether A is an N. If it does not, the sentence will lack truth-value.

Marques and García-Carpintero’s view manages to put together two features that we find worth taking into consideration, and which might seem difficult to combine. On the one hand, by situating the derogatory content of slurs at the presuppositional level, they account for its elusive character, which is accentuated by the nonpropositional character of this content. As we will see in the next three sections, this makes it possible to explain the resistance that slurs exhibit to cancelation, rejection, and retraction. On the other hand, the fact that they rely on *semantic* presuppositions allows them not to lower the degree of responsibility that a speaker takes in using a slur. The problematic presuppositions are triggered by the meaning of the word; they are not an unexpected consequence at the end of the causal chain. Thus, it is *in* using a slur that the speaker is harming the target group, and she can be said to be responsible for this. Additionally, Marques and García-Carpintero contribute to the discussion by fleshing out what it means for content to be expressive. By doing so in a nonpropositional way, they account for slurs’ special connection with action—in particular, for their derogatory character, which is what makes them dangerous.

We think, however, that modeling the effect of accommodating a slur in the common ground in the way in which Marques and García-Carpintero do can lead to a psychologistic interpretation. A way of understanding what it means for reactive attitudes to be part of the common ground is to take speakers to share them, at least for the purposes of the conversation. This is what drives us to characterize this proposal as psychologistic, since it makes the derogatory effect of slurs depend on the particular attitudes of speakers. Once the derogatory effect of slurs is characterized in this way, we might be inclined to think that situations of injustice stem from implicit attitudes and, in prioritizing the latter as an explanatory element, make intervention measures depend on them too.

An alternative way of modeling the effect of accommodating a slur in a way that is more akin to *structural* explanations of what makes them dangerous (Haslanger, 2016) is through *world-orderings*. This move is inspired in Stanley’s (2015) account of political propaganda, which relies on the distinction between at-issue content and not-at-issue content. At-issue content is that understood as proposed to be added to the common ground (Potts, 2007, 666), while not-at-issue content is directly added to the common ground, without any need for the audience to accept it (Stanley, 2015, 135). In fact, the latter is added to the common ground

without us having the chance to decide whether we accept it or not. According to Stanley, propagandistic discourse features not-at-issue content whose effect can be modeled along the lines of Starr's (2020) account of imperatives. Starr models the effect that imperatives have on the common ground as that of ordering the possible worlds in it in accordance with their preferability. For instance, if Víctor commands his child Pablo not to eat any more cookies, the worlds in the common ground are ordered so that those in which Pablo stops eating cookies are ranked as preferable to those in which he does not.

Stanley takes the relevant order to have to do with socialization, and says that, when accommodated, slurs rank those worlds in which the speaker socializes less with the target group above those in which she socializes more. This account suffices for the purposes of this paper. Hence, our view is that sentences featuring slurs presuppose that the worlds in the common ground are ranked so that they are more preferable the less contact the speaker has with the target group. If no such ordering is in place, the worlds will be rearranged according to it to accommodate the use of the slur. Presuppositions are a variety of not-at-issue content: when they are accommodated, they are directly added to the common ground, with no proposal involved, so participants in the conversation do not have the chance to accept or reject them. Thus, we do not depart too much from Stanley's account; if anything, we flesh it out a little.

An additional argument for understanding the effect of slurs in terms of world-orderings rather than reactive attitudes is that it results in a more parsimonious ontology. World-orderings also serve to model the effect of directives, while we know of no speech act, other than utterances of sentences featuring slurs, whose effect we can model in terms of reactive attitudes.⁴ By modeling the effect of slurs in terms of world-orderings, we subsume these speech acts under the more general class of directives, even if the associated command is made in a back-handed way. Moreover, enriching the set of operations that utterances can perform on the common ground instead of enriching the common ground itself may be seen as a generally preferable move.

5. Resistance to cancelation

As we saw in section 2, the derogatory effect of slurs seems resistant to cancelation, rejection, and retraction. An account such as the one resulting from the previous section will allow us to explain why they are so. In this section, we begin by using the account to make sense of slurs' resistance to cancelation.

The derogatory effect of slurs seems to project along a variety of contexts, which has driven many authors to defend that it is the result of a presupposition. The phenomenon of projection occurs whenever the sentence resulting from embedding a sentence in a complex context triggers the same presuppositions as the original sentence. Consider, for instance, the following sentences:

4 It could be argued that the effect of thick terms too can be modeled using reactive attitudes, so we would need to enrich the common ground with them anyway. Following Cepollaro and Stojanovic (2016), however, we could take thick terms to belong with slurs to the wider class of *hybrid evaluatives*. Our point would then be that the effect of slurs and thick terms alike can be modeled either using reactive attitudes, and thus complicating our ontology, or using world-orderings, and thus assimilating sentences featuring them to directives. Thanks to María José Frápolli for making us take this into consideration.

- (12) Nerea has stopped going to jiu-jitsu classes.
- (13) If Nerea has stopped going to jiu-jitsu classes, I've bought her this gi for nothing.
- (14) Nerea hasn't stopped going to jiu-jitsu classes.
- (15) Has Nerea stopped going to jiu-jitsu classes?

(12) presupposes that Nerea used to go to jiu-jitsu classes, and so do (13–15). We say that the presupposition that Nerea used to go to jiu-jitsu classes projects when the sentence that triggers it is embedded as the antecedent of a conditional, negated, or posed in the form of a question. The fact that presuppositions behave in this way may lead us to think that a presuppositional account of slurs should be given, as their derogatory character seems to project in similar cases:

- (2) *A* is an *S*.
- (4) If *A* is an *S*, she shouldn't go to that restaurant.
- (9) *A* is not an *S*.
- (10) Is *A* an *S*?

(4), (9), and (10) are just as derogatory as (2) is, just like (13–15) presupposed that Nerea used to go to jiu-jitsu classes just like (12) did. We could thus say that this happens because (2) triggers a certain presupposition that projects in (4), (9), and (10). In particular, we could say that (2) presupposes that *Ns* are bad because of being *Ns* (Cepollaro and Stojanovic, 2016, 459).

However, the projective behavior of slurs does not exactly coincide with that of presuppositions.⁵ Consider these two sentences:

- (16) If Nerea used to go to jiu-jitsu classes, she has stopped doing so.
- (17) If *Ns* are bad because of being *Ns*, *A* is an *S*.

These cases are structurally identical: in both of them, we have placed the sentence that is supposed to trigger a certain presupposition—(12) in (16), (2) in (17)—in the consequent of a conditional whose antecedent is a sentence expressing the presupposition allegedly triggered by the former sentence.⁶ However, they behave differently. (16) does not presuppose that Nerea used to go to jiu-jitsu classes, while (17) is still derogatory. If it is so in virtue of its triggering the presupposition that *Ns* are bad because of being *Ns*, it presupposes that *Ns* are bad because of being *Ns* just like (2) does. This feature has been called the *hyperprojectivity* of slurs (Jeshion, 2013a; Sennett and Copp, 2015; Cepollaro and Stojanovic, 2016; Camp, 2018; Marques and García-Carpintero, 2020). Thus, the derogatory content of slurs does not exactly behave like a presupposition, which speaks against explaining it in this way.

5 This does not only apply to conditional environments, but also to belief reports and quotations. In this paper, however, we stick to conditional environments, and leave a presuppositional explanation of slurs' projective behavior under belief operators and quotation marks for further work. Thanks to Andrés Soria for making us take this into account.

6 For the conditional to behave in the way we need it to, it is enough for the antecedent to *entail* the content presupposed by the consequent (see Cepollaro and Stojanovic, 2016, 478). This includes cases in which the antecedent *states* this content, which are the only ones considered by Marques and García-Carpintero (2020, 149).

Since presuppositions do not project in contexts like (16), we can embed sentences that in some sense imply a certain content in contexts with that structure to know whether the content is presupposed or not. If the content does not project, it will be a presupposition; if it does, it will not. This is what has been called the *binding test* (van der Sandt, 1992).

Proponents of presuppositional accounts of slurs have offered different explanations that try to make these accounts compatible with the hyperprojectivity of slurs. Cepollaro and Stojanovic (2016, 480) rely on the distinction, made among others by Abusch (2002) and Abbott (2006), between *soft* and *hard* triggers, where the former give rise to presuppositions that can be canceled in some contexts, while the presuppositions that result from the latter project in contexts such as (17). For instance, gendered pronouns are hard triggers (Heim, 2008):

(18) If Nerea is a woman, she can't do jiu-jitsu.

(18) presupposes that Nerea is a woman even if it is a conditional whose antecedent expresses exactly that presupposition. The fact that this content projects in (18) seems to constitute no obstacle to understanding the implication that Nerea is a woman as a presupposition. In the same way, Cepollaro and Stojanovic say, the fact that (17) is as derogatory as (2) should not drive us to reject a presuppositional account of slurs.

Cepollaro and Stojanovic's reply to the hyperprojectivity objection, however, seems stipulatory to us. They just introduce the distinction between soft and hard triggers, but they are not clear as to *why* slurs and gendered pronouns belong to the latter category.⁷ We find Marques and García-Carpintero (2020, 149–50), by contrast, closer to offering a really explanatory account. Their point is that, when we place (2) in a context such as (17), we are not really applying the binding test to it. For the binding test to be correctly applied, the antecedent of (17) should express the presupposition in (2) (see McCready, 2010, 9). But, as we said in the previous section, this presupposition is not propositional, while the antecedent of (17) will necessarily have propositional form⁸—otherwise, (17) will be ungrammatical. Thus, no matter what we place in the antecedent of (17), it will not express the presupposition in (2). It should not worry us, therefore, if no sentence with the form of (17) is free of derogatory character. No such sentence will be the result of applying the binding test. Thus, we cannot say that slurs fail the binding test just because the presupposition that *Ns* are bad because of being *Ns* projects in (17).

Despite appearances, though, Cepollaro and Stojanovic's and Marques and García-Carpintero's explanations can be said to be complementary rather than incompatible. We could say that slurs are hard triggers because the presuppositions they trigger are nonpropositional. Moreover, Marques and García-Carpintero's account can also give us a hint as to why gendered pronouns are hard triggers too. Rather than propositional presuppositions, gendered pronouns can be taken to trigger presuppositions concerning the assignment of values to the variables—they pose certain restrictions on this assignment for utterances of sentences containing them to be felicitous. Since we can only place sentences with proposi-

7 They relate soft triggers to pragmatic presuppositions and hard triggers to semantic ones (Cepollaro and Stojanovic, 2016, 480), but we do not think this is enough of an explanation.

8 This is related to Potts' (2007, 176–9) discussion of the *ineffability* of the meaning of expressives.

tional form in the antecedent of a conditional like (17), no conditional will allow us to apply the binding test to a sentence featuring a gendered pronoun.

6. Resistance to rejection

The derogatory effect of slurs is not only hard to cancel. It is also hard to reject. In this section, we flesh out what resistance to rejection amounts to and show how a nonpropositional, presuppositional account of slurs like the one we have advanced can make sense of this feature.

Remember (2):

(2) *A* is an *S*.

An utterance of (2) could be replied by saying

(6) No, she isn't.

However, we cannot be sure that, by doing so, it is the derogatory effect of the utterance that we are rejecting. A natural interpretation of (6) is just that *A* is not an *N*, but this is compatible with our sharing whatever negative attitude toward *Ns* the speaker of (2) holds (see Stanley, 2015, 135–6). Something else is needed if we want to make sure that we are understood as rejecting this attitude.

This fact is smoothly accommodated by a presuppositional account like ours. When we say something like (6), we are targeting the content asserted by the speaker of (2). If the derogatory effect of this sentence is explained through its presuppositions, it is then natural that replies like (6) do not manage to block that effect, as it is not through replies like this that we target presuppositions.

Instead, presuppositions are usually assumed to be identifiable through the “Hey, wait a minute!” test (von Fintel, 2004). Remember (12):

(12) Nerea has stopped going to jiu-jitsu classes.

If the hearer does not share the assumption that Nerea used to go to jiu-jitsu classes, it is natural for her to reply with:

(19) Hey, wait a minute! I didn't know Nerea used to go to jiu-jitsu classes.

This is a standard way in which presuppositions can be rejected. Cepollaro and Stojanovic (2016, 467), however, note that this kind of reply is not natural when it is an utterance of a slur that we are replying to. It would be weird, for instance, to reply to (2) with

(20) ?? Hey, wait a minute! I didn't know *Ns* are bad because of being *Ns*.

or whatever formulation we have chosen for the presupposition triggered by (2). Instead, they say, a rejection is likely to go along the following lines:

(21) Hey, wait a minute! You shouldn't talk about *Ns* like that.⁹

Cepollaro and Stojanovic take this to be “a crucial difference between rejecting descriptive and evaluative presuppositions” (Cepollaro and Stojanovic, 2016, 467). It seems to us, however, that, by saying this, Cepollaro and Stojanovic are again describing the behavior of evaluative presuppositions rather than explaining why they behave in this way. Fortunately, we think we can find an explanation in Marques and García-Carpintero's (2020) view just like we found one for the projective behavior of slurs. Note, first, that “I didn't know that” requires a propositional prejacent. Thus, no matter what we place after “I didn't know that”, it will not express the presupposition in (2), which is not propositional. Replies like (20) will therefore always fail to target this presupposition.

But also, propositional presuppositions are meant to model the information shared by participants in a conversation as to how the world is. They serve to describe the stage at which participants are in a process of inquiry, which is aimed at *knowing* which of the candidate worlds we inhabit. Nonpropositional presuppositions are nothing like this. Suppose we take these presuppositions to concern the participants' reactive attitudes, as Marques and García-Carpintero do. In what kind of process do they play a role? Coordination is a natural candidate: we could say that part of the aim of a conversation is for speakers to align their reactive attitudes (cf. Gibbard, 1990, 110), and expressive presuppositions allow us to characterize the point they are at. However, it is far from settled that this is the role that reactive attitudes play in conversations (Pérez Carballo and Santorio, 2016). At any rate, we can be quite certain that, whatever process we should be talking about, it is not a process of inquiry. Speakers do not aim at *discovering* what the appropriate reactive attitudes are. This explains that utterances of sentences like (2) cannot be challenged with replies like (20). If, instead of reactive attitudes, we use world-orderings to model the effect of slurs, we will reach a similar conclusion.

Replying to (2) with (21) is better, but it also comes at a cost. Insofar as (21) concerns how we should talk instead of what the speaker of (2) actually said, its utterance is likely to be interpreted as an attempt at changing the subject of the conversation and, as such, a sign of lack of cooperativity (Chilton, 2004, 64; Stanley, 2015, 170).

We take these considerations to explain why it is so hard to reject the derogatory effect of a slur. Direct negation, as in (6), does not target the presupposition, but standard ways of rejecting presuppositions, as in (20), do not seem to work either. The fact that the derogatory effect of slurs is explained through presuppositions and that these presuppositions are nonpropositional in character is what is needed to account for slurs' resistance to rejection.

7. Resistance to retraction

The last and least discussed feature of slurs for which we should want our proposal to account is their resistance to retraction. Retraction is the phenomenon whereby a speaker says something like “I retract that” or “I take that back” (MacFarlane, 2014, 108). By doing

9 This reply is meant to encompass the different *metalinguistic negations* considered by Cepollaro and Stojanovic (2016, 466–7).

so, according to MacFarlane, the speaker cancels the normative effects of the speech act to which “that” refers, which is the *target* of the speech act of retraction. For example, by retracting an offer we withdraw a permission we had previously extended. In the case of utterances featuring slurs, however, the normative effects associated with such words seem hard to cancel in this way. In other words, you cannot “unring a bell” once it has been rung (see McGowan, 2009, 403; Simpson, 2013, 570).

A presuppositional account like the one we have advanced in this paper can explain this fact along the lines drawn by our explanation of slurs’ resistance to rejection. When I say “I take that back”, what I am taking back is the asserted content, not the presupposed content.¹⁰ The expression “I take that back” is naturally read as elliptical for “I take what I said back”, not for “I take what I presupposed back”—while we can find actual occurrences of “I take what I said back” in ordinary conversations, “I take what I presupposed back” is a philosophical construct. If we say “I take that back” trying to retract a presupposition, all we will do is retract the assertion that triggered that proposition. Now, if the derogatory content of slurs is presupposed, this would explain why we cannot undo the damage associated with the utterance of a slur.

Note, however, that a sufficiently rich context might allow an utterance of “I take that back” to target the presupposition triggered by a previous utterance instead of the content asserted through it. Consider the following dialogue:

- (12) Nerea has stopped going to jiu-jitsu classes.
- (22) In fact, she signed up for them but didn’t get to attend a single class.
- (23) Oh, is that so? I take that back then.

Given (22), it is natural to think that (23) does not target the content asserted by (12), but the presupposition that Nerea used to go to jiu-jitsu classes. By uttering (23), the speaker is not introducing back into the common ground worlds in which Nerea has kept going to jiu-jitsu classes, but worlds in which she has never even gone to them. These worlds had been removed from the common ground by the speaker’s utterance of (12) followed by a process of accommodation.

It is not clear, though, that this mechanism can be used to retract a nonpropositional presupposition like the ones we are interested in. Consider the following dialogue, built to be structurally parallel to the one presented above:

- (2) A is an S.
- (24) In fact, Ns are valuable members of our society.
- (25) ?? Oh, is that so? I take that back then.

(25) feels weird if intended to target the content asserted by (2)—why would the fact that Ns are valuable members of our society speak against the claim that A is an N? But,

10 We could also say that the retraction targets the speech act itself. In this case, however, the speech act would be the assertion, and not the presupposition. It is in fact awkward to talk about presuppositions as acts that the speaker carries out. Thanks to Claudia Picazo for suggesting this alternative to us.

if (25) is intended to target the presupposition triggered by (2), it is hard to see how it could achieve its purpose. *Ns* will have been derogated regardless of one's efforts to "go back in time".

Up to this point, we could be accused of what, in the previous two sections, we accused Cepollaro and Stojanovic of. We have *described* how slurs behave with respect to retraction, but we have not *explained* why they do so. As with rejection, however, we think that the fact that the presuppositions triggered by slurs are nonpropositional allows us to account for the way in which they differ from other presuppositions when it comes to retraction. What makes the context of the first dialogue allow (23) to target the presupposition of (12) is that, in (22), the interlocutor provides the speaker of (12) with certain information. Cooperative principles then obligate the speaker to reintroduce into the common ground the worlds excluded through presupposition accommodation. There is no information that the interlocutor of the second dialogue could offer, though, that automatically forced the speaker of (2) to cancel the nonpropositional presupposition triggered by her utterance. Anything the interlocutor says is strictly compatible with the speaker of (2)'s negative evaluation of *Ns*. Of course, knowing certain facts may drive the speaker to reconsider her position. But, unlike in the first dialogue, refusing to do so does not make her uncooperative.

What if, instead of (24), the interlocutor replies denying precisely the presupposition triggered by (2)? This would require this presupposition to be made explicit in propositional form. But, as we saw when discussing slurs' resistance to cancelation, this is not possible, as the presupposition triggered by (2) is not propositional. Thus, presupposition and nonpropositionality together allow us to account for the difficulty of retracting a slur.

8. Conclusion

Following Cepollaro and Stojanovic (2016) and Marques and García-Carpintero (2020), we have offered an account of the meaning of slurs that, by relying on presuppositions, explains much of their behavior, and explains the behavior that ordinary presuppositions leave unexplained by understanding the presuppositions triggered by slurs as nonpropositional. As we said in section 4, we believe that a nonpropositional, presuppositional account of slurs allows us to make sense of two features of this kind of term that are in apparent tension. On the one hand, their effect is elusive—it resists being pointed out, a previous move that is needed whenever we want to block it. On the other hand, however, this effect is triggered by the *meaning* of the slur. It is what slurs *do*, so speakers can be said to be responsible for this effect. Inasmuch as nonpropositional presuppositions are triggered by what slurs mean but impossible to make explicit using declarative sentences, we think that they have all the features that whatever lies behind slurs' derogatory effect should have. Of course, future work should offer a much more detailed account of how a proposal like ours would deal with each particular case in which the presuppositions triggered by slurs behave in nonstandard ways.

References

- Abbott, B. (2006), “Where have some of the presuppositions gone”, in B. Birner & G. Ward (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Philadelphia: John Benjamins, pp. 1–20.
- Abusch, D. (2002), “Lexical alternatives as a source of pragmatic presuppositions”, in B. Jackson (ed.), *Proceedings of SALT XII*, Ithaca, NY: CLC Publications, pp. 1–20.
- Camp, E. (2013), “Slurring perspectives”, *Analytic Philosophy*, 54(3), 330–349.
- Camp, E. (2018), “A dual act analysis of slurs”, in D. Sosa (ed.), *Bad Words: Philosophical Perspectives on Slurs*, Oxford: Oxford University Press, pp. 29–59.
- Caponetto, L. (2020), “Undoing things with words”, *Synthese*, 197(6), 2399–2414.
- Cepollaro, B. (2020), *Slurs and Thick Terms: When Language Encodes Values*, Washington, DC: Rowman & Littlefield.
- Cepollaro, B. & Stojanovic, I. (2016), “Hybrid evaluatives: In defense of a presuppositional account”, *Grazer Philosophische Studien*, 93(3), 458–488.
- Chilton, P. (2004), *Analysing Political Discourse: Theory and Practice*, London: Routledge.
- von Fintel, K. (2004), “Would you believe it? The King of France is back! Presuppositions and truth-value intuitions”, in M. Reimer & A. Bezuidenhout (eds.), *Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions and Other Related Phenomena*, Oxford: Oxford University Press, pp. 315–341.
- Gibbard, A. (1990), *Wise Choices, Apt Feelings: A Theory of Normative Judgment*, Cambridge, MA: Harvard University Press.
- Haslanger, S. (2016), “What is a (social) structural explanation?”, *Philosophical Studies*, 173(1), 113–130.
- Jeshion, R. (2013a), “Slurs and stereotypes”, *Analytic Philosophy*, 54(3), 314–329.
- Jeshion, R. (2013b), “Expressivism and the offensiveness of slurs”, *Philosophical Perspectives*, 27(1), 231–259.
- Kenny, A. (1963), *Action, Emotion and Will*, London: Routledge.
- Lewis, D. (1979), “Scorekeeping in a language game”, *Journal of Philosophical Logic*, 8(1), 339–359.
- MacFarlane, J. (2014), *Assessment Sensitivity: Relative Truth and Its Applications*, Oxford: Oxford University Press.
- Macià, J. (2002), “Presuposición y significado expresivo”, *Theoria: An International Journal for Theory, History and Foundations of Science*, 3(45), 499–513.
- Marques, T. & García-Carpintero, M. (2020), “Really expressive presuppositions and how to block them”, *Grazer Philosophische Studien*, 97(1), 138–158.
- McCready, E. (2010), “Varieties of conventional implicature”, *Semantics and Pragmatics*, 3, article 8.
- McGowan, M. K. (2009), “Oppressive speech”, *Australasian Journal of Philosophy*, 87(3), 389–407.
- Nunberg, G. (2018), “Speech acts in discourse contexts”, in D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 237–295.
- Pérez Carballo, A. & Santorio, P. (2016), “Communication for expressivists”, *Ethics*, 126(3), 607–635.

- Portner, P. (2007), “Imperatives and modals”, *Natural Language Semantics*, 15(4), 351–383.
- Potts, C. (2007), “The expressive dimension”, *Theoretical Linguistics*, 33(2), 165–198.
- van der Sandt, R. A. (1992), “Presupposition projection as anaphora resolution”, *Journal of Semantics*, 9(4), 333–377.
- Schlenker, P. (2007), “Expressive presuppositions”, *Theoretical Linguistics*, 33(2), 237–245.
- Sennet, A. & Copp, D. (2015), “What kind of a mistake is it to use a slur?”, *Philosophical Studies*, 172(4), 1079–1104.
- Simpson, R. M. (2013), “Un-ringing the bell: McGowan on oppressive speech and the asymmetric pliability of conversations”, *Australasian Journal of Philosophy*, 91(3), 555–575.
- Stalnaker, R. (2014), *Context*, Oxford: Oxford University Press.
- Stanley, J. (2015), *How Propaganda Works*, Princeton: Princeton University Press.
- Starr, W. B. (2020), “A preference semantics for imperatives”, *Semantics and Pragmatics*, 13, article 6.
- Strawson, P. F. (1962), “Freedom and resentment”, *Proceedings of the British Academy*, 48, 1–25.
- Williamson, T. (2009), “Reference, inference, and the semantics of pejoratives”, in J. Almog & P. Leonardi (eds.), *The Philosophy of David Kaplan*, Oxford: Oxford University Press, pp. 137–158.



Paving the road to hell: The Spanish word *menas* as a case study*

Allanando el camino al infierno: la palabra española *menas* como caso de estudio

DAVID BORDONABA-PLOU**
JOSÉ RAMÓN TORICES***,****

Abstract: *Menas* is a term that has attracted a great deal of attention on the political scene in Spain at present. Although the term had a neutral usage originally, being an acronym for unaccompanied foreign minors, it has recently evolved into a term with clear negative connotations. This article explores what kind of term *menas* is today. Specifically, we will examine whether *menas* is a slur or an ESTI, an ethnic/social term used as an insult. First, we point out the most defining characteristics of both types of terms. Then, using analyses on linguistic corpora, we show that *menas* exhibits the most defining characteristics of ESTIs. We end by discussing the possible evo-

Resumen: *Menas* es un término que ha ganado notoriedad en la actual escena política española. Aunque el término tenía un uso neutro en su origen, ya que es un acrónimo de “menores extranjeros no acompañados”, recientemente ha evolucionado a un término con claras connotaciones negativas. Este artículo explora qué tipo de término es *menas* actualmente. Específicamente, examinaremos si *menas* es un *slur* o un TESNI, es decir, un término étnico/social neutro usado como insulto. Primero, señalamos las características más definitorias de ambos tipos de términos. Después, por medio de análisis sobre *corpus* lingüísticos, mostramos que *menas* exhibe las características más definitorias

Recibido: 31/05/2021. Aceptado: 02/07/2021.

* This publication resulted (in part) from research supported by CONICYT/FONDECYT/POSTDOCTORADO/ N° Proyecto 3180096, by MINECO under the grant FPI/BcS-2014-067584 and the research project PID2019-109764RB-I00, by the Regional Government of Andalusia under the research projects B-HUM-459-UGR18 and P18-FR-2907, and by the University of Granada under a “Contrato Puento” fellowship and the excellence unit FiloLab-UGR UCE.PPP2017.04. We are grateful to the editors for making this special issue possible, and to Agustín Vicente and the two anonymous referees for their fruitful comments.

** David Bordonaba-Plou is a FONDECYT postdoctoral researcher at the Universidad de Valparaíso, Chile. Contact: davidbordonaba@gmail.com. His most recent publications are Bordonaba, D. (2021) “Metalinguistic negotiations and two senses of taste”, *Diametros*, 18(67), 1–20, and Bordonaba, D. (2021), “Cognitive penetration and taste predicates: Making an exception to the rule”, *Filosofía Unisinos* 22(1), 12–20.

*** José Ramón Torices is a postdoctoral researcher belonging to the excellence unit FiloLab-UGR and the research group Filosofía y Análisis (HUM-975). Contact: jrtorices@gmail.com. His most recent publications are Torices, J. R. (2021), “Understanding dogwhistle politics”, *Theoria: An International Journal for Theory, History and Foundations of Science*, DOI: 10.1387/theoria.22510, and D. Bordonaba, V. Fernández Castro and J. R. Torices (forthcoming), “The political turn: Analytic philosophy as philosophical activism”, in D. Bordonaba, V. Fernández Castro and J. R. Torices (eds.), *The Political Turn in Analytic Philosophy: Reflections on Social Injustice and Oppression*, Berlin: De Gruyter.

**** The names of the two authors are in alphabetical order. Both authors have contributed equally to the paper.

lution of the term, pointing out that, although the term presents the features related to ESTIs, there are two possible scenarios. On the one hand, the term may retain its neutral uses and thus remain an ESTI. On the other hand, the neutral uses may disappear, and thus the term may become a slur.

Keywords: *menas, slurs, ESTIs, corpus linguistics, collocations, concordance*

de los TESNI. Acabamos discutiendo la posible evolución del término, señalando que, aunque el término tiene los rasgos de los TESNI, existen dos posibles escenarios. Por un lado, el término puede conservar sus usos neutros y de este modo seguir siendo un TESNI. Por otro lado, los usos neutros pueden desaparecer y así el término acabar convirtiéndose en un *slur*.

Palabras clave: *menas, slurs, TESNIs, lingüística de corpus, colocaciones, concordancia*

1. Introduction

Not long ago, most Spaniards did not know what the word *mena*¹ meant. It was not an expression we heard in everyday conversations, nor was it used by the media. Most of us did not even know that such a word existed! *MENA*, a technical term that first appeared in the *Boletín Oficial del Estado* (BOE, “Official State Gazette”) in 2009, is an acronym for *Menores Extranjeros No Acompañados*, “Unaccompanied Foreign Minors”, that is, boys and girls under 18, non-European migrants, who are separated from their parents and are not in the care of any other adult in the host country. It was an expression neutrally used in legal and migration contexts to draw attention to the special vulnerability of these children. Nowadays, however, it is not easy to find a single Spaniard who does not know what the word means. A significant part of the racist discourse of Vox, the Spanish alt-right party, has revolved around this group. Vox’s insistence on publicly vilifying them through fake news, hoaxes or hate speech, both in the media and on social networks, has succeeded in putting them at the centre of the political debate and thus popularising the term.²

The word *mena* has practically evolved from an unknown neutral acronym to a popular noun with negative connotations. Vox winning seats in the Andalusian parliament after the 2019 regional elections prompted several NGOs (for instance, Save the Children, SOS-Children’s Village International or UNICEF) to warn on their websites and social networks about the dangers of using the expression *mena* to criminalise migrants children and teenagers. The term is currently used to associate being a Maghrebi male with being a criminal (thief, rapist, violent...).

Throughout this paper, we seek to understand this evolution and reflect on what kind of expression the word *mena* is. Specifically, we will analyse empirically whether we are dealing with a slur or an ESTI, i.e., an ethnic/social term used as an insult (see Castroviejo, Fraser and Vicente, 2020). Then, we will explain the conversational dynamics that have led to this shift from neutral to offensive usage. Our aim is thus twofold: first, to understand what kind of expression the term *mena* has become, and second, to draw some generalisations about how an expression evolves from having only a neutral use to an offensive or simultaneously

1 We follow the extended practice in cross-linguistic studies of italicizing the original expression and using quotation marks for the translation of that expression.

2 See “La vida visible de las palabras: un análisis del concepto “mena” y su uso en redes” (2019), by Isabel Martín Piñeiro for *El Salto Diario*: <https://www.elsaltdiario.com/comunicacion/Analisis-palabra-mena-racismo>.

neutral and offensive use. Concerning the first aim, our thesis is that the expression *mena* is currently used both neutrally and offensively, with a tendency towards the disappearance of the first type of use. The word *mena* is, thus, an ESTI that may end up becoming a slur. As for the second aim, our thesis is that the systematic appearance of a neutral expression (with *descriptive meaning*) together with negative evaluative expressions ends up turning the neutral expression into an expression with a negative use (with *expressive meaning*) which may lose its neutral use, becoming a slur, or keep it together with the offensive one, becoming an ESTI.

To fulfill these two aims we will compile a corpus of Spanish tweets containing the term *menas*. Then, we will perform some quantitative and qualitative analysis to show to what extent *menas* is an ESTI. Specifically, we will perform a collocation analysis, i.e., we will determine which words most often co-occur with *menas*, and a concordance analysis, i.e., we will inspect the term *menas* in context. Through these two analyses, we will obtain the empirical evidence necessary to determine the extent to which *menas* presents the features that are characteristically associated with ESTIs.

The paper is divided as follows. Section 2 characterises slurs, emphasizing the most salient features of this set of terms. Section 3 similarly characterises ESTIs. Section 4 explains the methods and materials used in the study. Section 5 presents the collocation and the concordance analyses. Finally, Section 6 assesses the results of the analyses, discussing the status of the term and its possible future evolution and drawing some generalisations to shed some light on the process by which a neutral term can end up becoming an offensive one.

2. Words That Hurt I: Characterising Slurs

A slur is an evaluative expression used to derogate members of a particular social group on the grounds of race, gender, social status, sexual orientation or religion, among other social identities (see, e.g., Hom, 2008; Anderson and Lepore, 2013; Camp, 2013; Croom, 2013, 2015; Jeshion, 2013; Blakemore, 2015; Cepollaro, 2015). They are a particularly abusive kind of expression because they have the potential to “harm and degrade their targets, making them feel humiliated, dehumanised, disempowered, and silenced” (Popa-Wyatt and Wyatt, 2018, 2880). It is widely agreed that there is a non-pejorative or neutral counterpart used for picking out the same set of individuals for each slur. According to the mainstream view (see, e.g., Croom, 2015 and Falbo, 2021 for exceptions), the slur “dago”, for instance, has roughly the same descriptive content, the same truth conditions, as the expressions “Spaniard”, “Italian” or “Portuguese,” but “dago” also implies that they are, *qua* Spaniard, Italian or Portuguese, bad people. Some characteristic features of slurs, following Bolinger (2017), include the following:³

- *Offensive Autonomy*—slurs are offensive even when the speaker does not intend the use to be derogatory.
- *Embedding Failure*—the derogatoriness of slurs projects out of various forms of embedding, including indirect reports, negations, and mentions.
- *Perspective Dependence*—use of a slur is taken to indicate that the speaker holds derogatory attitudes.

3 These features are also true for slurs in (European) Spanish.

- *Offensive Variation*—not all slurs, even if co-referential, appear to be equally offensive.
- *Insulation*—despite all the above, slurring terms can occasionally occur inoffensively, which is true even of particularly strong ones. (Bolinger, 2017, 439)

Let's say a little more about these properties of slurs. First, the power of degradation, humiliation and offence towards a slur's target is bigger than that of an individual insult. According to Camp (2013, 330), they "are among the most rhetorically powerful and insidious expressions in a language", and their harmfulness does not depend on the speaker's attitudes or intentions. In other words, slurs are *conventionally* derogative (Cepollaro, 2015, 37). The speaker's not being a racist does not remove the harmful and racist ingredient from her racist slurring utterance. In fact, it is quite the opposite. Except on rare occasions, the use of a racist slurring word makes the speaker a racist.

Second, unlike other offensive expressions such as "jerk", "bastard" or "asshole", slurs' offensiveness preserves when embedded under negation, questions, conditional antecedents, or modals (see, e.g., Hom, 2008, 435; Cepollaro, 2015, 37). Let's illustrate this idea with some examples:

- (1) José is an *asshole*.
- (2) José is not an *asshole*.
- (3) José is a *dago*.
- (4) José is not a *dago*.

By negating (1), the speaker rejects that the property *being an asshole* can be predicated of José. "Asshole" applies to him by virtue of his attributes. Thus, when the speaker denies that José is an asshole, she denies that José possesses those attributes. By negating (3), however, the speaker rejects that the property *being a dago* can be predicated of José, but, unlike with "asshole", "dago" is still an offensive way of referring to a certain social group even when José does not belong to it.

Third, the use of a slur is not only an unfair way of presenting a specific social group; it is also, following Nunberg (2018), a way of exhibiting and building camaraderie among the users of such expressions. The display of camaraderie is crucial in explaining one of the social functions of slurs because it explains why many of the users of these expressions employ them, especially in contexts where there are non-targeted members present. As Bolinger (2017, 447) notes, when a speaker chooses from among two expressions α and β , the one that is associated with a certain negative attitude ϕ towards a particular social group rather than choosing the one that is not, that choice is a signal that the speaker endorses ϕ .

Fourth, the derogatory potential of a slur varies from one to another. While all slurs infringe upon the dignity of the social group they attack, they do not all humiliate or degrade in the same way. As Hom (2008) notes, some are more damaging than others. The meaning and harmful potential of a slur may vary over time along with the social norms and values that underpin that meaning and harmful potential. Expressions that were once offensive may no longer be so today.

Finally, using a slur can have a high social cost for the speaker. There are very few contexts if any (see Anderson and Lepore, 2013 on this point), in which a slur can be used without being perceived as inappropriate and derogative. As Hom points out, “[u]ses of epithets are subject to strict social constraints, if not outright forbidden” (Hom, 2008, 427). The use of slurs is, thus, practically a taboo for members of the non-targeted group beyond a few special contexts (direct quotations, dictionary entries, or memos advising against the use of offensive terms).

3. Words That Hurt II: Characterising ESTIs

Slurs, as we have seen, are expressions whose use is offensive by default. In Spanish, as in other languages, there are plenty of slurs.⁴ Some examples of Spanish slurring words are *sudaca* (for Latin Americans), *guiiri* (for white tourists from Northern Europe and America), or *gabacho* (for Frenchs). However, slurs are not the only type of expression used to derogate certain social groups. Some expressions can be used descriptively to refer to certain social groups and can also be used offensively. *Gitano* (for Romani) or *moro* (for people from North Africa) seem to fit into this category. Castroviejo, Fraser and Vicente (2020) have called such expressions “ethnic/social terms used as insults” (ESTIs, for short). In Spanish, ESTIs can be preceded by an indefinite (i-ESTI, for short) or not (w-ESTI, for short). An i-ESTI is primarily offensive, while an w-ESTI can be used neutrally (n-ESTI, for short) or offensively (o-ESTI, for short).

The most prominent features of ESTIs, largely following Castroviejo, Fraser and Vicente (2020), are:

- *Context-dependent Offensiveness*—w-ESTIs’ offensiveness is highly dependent on contextual singularities, while when it is an i-ESTI its offensiveness is autonomous.
- *Stereotype-based Offensiveness*—i-ESTIs’ application hinges on the individual insulted meeting the stereotypes associated with the target group, not on that individual belonging to it.
- *Mixed Embedding Behaviour*—i-ESTIs project out of various forms of embedding. w-ESTIs do not project when embedded.
- *Perspective Dependence*—generally speaking, using an i-ESTI or an o-ESTI signals that the speaker holds derogatory attitudes towards the target group.
- *Offensive Variation*—ESTIs, when used offensively (whether an i-ESTI or an o-ESTI) do not appear to be equally offensive. Their offensiveness depends on how negative the stereotypes associated with the target group are.

Unlike slurs, ESTIs are often used both neutrally to refer to a group based on their ethnicity or background (n-ESTI) and as an insult (i-ESTI) to refer to people who do not belong to the targeted group. Although one of the most defining features of slurs is that they derogate individuals by group membership, several authors have argued that slurs can be used neutrally and refer offensively to out-group people. First, Zeman (2021) argues that the Romanian slur *tigan* can be used neutrally to identify as being of Roma ethnicity. However, this case would not constitute a counterexample against the thesis that ESTIs, unlike slurs, can be used neutrally

4 In what follows, we focus on expressions used in (European) Spanish.

by members of the non-targeted group. According to Zeman (2021, 5), the neutral use of *tigan* is only allowed for the Romanian Romani population. ESTIs, as mentioned above, can be used neutrally by members of the non-targeted group and in everyday contexts, not only in some exceptional contexts, as is the case with slurs. Second, according to Jeshion (2013, 324), “slurs’ extended uses” involve the application of slurs to people who are not members of the targeted group. However, although possible, we believe that these cases represent a small fraction. On the contrary, ESTIs are often applied offensively to people outside the group, which may be just as common as cases where they are not, which vindicates the asymmetry noted.

Another crucial difference is that, unlike slurs, whether an ESTI is used offensively depends entirely on the speaker exploiting the associated stereotype.⁵ As Castroviejo, Fraser and Vicente note, ESTIs are strongly linked to stereotypes. This link is most clear in their application to individuals who do not belong to the target group. The fact that an ESTI, when used derogatorily, can be applied to people who do not belong to the target group shows, in our view, that users of an ESTI must be aware of the stereotypes that characterise it in a way that they do not have to be when using a slur. In other words, in the case of ESTIs, the most crucial factor in determining their *conditions of application* is not group membership, as in the case of slurs, but the compliance with the stereotypes attributed to the target group.

ESTIs have both neutral and offensive uses. At least in (European) Spanish, i-ESTIs are usually offensive. But, as we have seen, there are also o-ESTIs. This has consequences for the behaviour of ESTIs when they are embedded. Appearances of an i-ESTI scope out just like slurs. w-ESTIs do not project even when offensive. Let’s look at some examples:

- (5) *Juan es gitano.* (“Juan is Romani” = *n-ESTI/o-ESTI*)
- (6) *Juan no es gitano.* (“Juan is not Romani” = *n-ESTI/o-ESTI*)
- (7) *¿Juan es gitano?* (“Is Juan Romani?” = *n-ESTI/o-ESTI*)
- (8) *Juan es un gitano.* (“Juan is (a) Romani” = *i-ESTI*)
- (9) *Juan no es un gitano.* (“Juan is not (a) Romani” = *i-ESTI*)
- (10) *¿Juan es un gitano?* (“Is Juan (a) Romani?” = *i-ESTI*)

(5–7) are sentences in which *gitano* can be *n-ESTI* or *o-ESTI* but knowing this is necessary to know some particularities of the context. In any case, it seems that (5–7) in neither of their two possible interpretations, *n-ESTI* or *o-ESTI*, projects out. *o-ESTIs* neutralises the stereotypes when embedded. That it does not project out when the content is neutral is to be expected, that it does not project out when the content is offensive distances *o-ESTIs* from slurs. In (8–10), however, the offensive content does project out. When the ESTI is preceded by the indefinite, the speaker still exploits the stereotypes attached to the *i-ESTI* even though it is embedded under the interrogative. *i-ESTIs* scope the stereotype out, even if it is embedded.

ESTIs, when used to insult, constitute a signal of the speaker’s negative attitude towards the target group. However, it should be noted that an ESTI can also be used to express positive attitudes towards the target group and not only by members thereof (see Castro-

5 According to Jeshion (2013, 324), in “slurs’ extended uses”, the racial (or gender) component still plays a role. However, Tirrell (1999) defends that the stereotype is crucial for the slur to be offensive in these types of uses. It is not our aim to offer arguments for or against stereotype semantics of slurs, but to point out that the offensiveness of an ESTI depends more directly on whether the speaker exploits the stereotypes associated.

viejo, Fraser and Vicente, 2020, 11, n. 14). Some people may associate positive stereotypes with the target group. For these people, saying something like *X es un gitano*, “X is (a) Roma”, means that X is someone who has certain positive qualities (loyalty, freedom-loving...) precisely because X belongs to that group. Therefore, unlike slurs, the very use of the term by members of the non-targeted group is not by default a sign of the speaker’s negative attitude. In addition, it is valuable to note that sometimes the same stereotype can be perceived as positive or negative by members of the non-targeted group. For example, saying that the Roma have music “in their blood” can mean something positive (they are skilled musicians) or something negative (they prefer to be idle rather than work hard). In any case, whether it is used to insult or to praise, the use is not neutral but evaluative, and hence perspective-dependent.

Finally, how offensive an ESTI is will depend on the associated stereotypes. The associated stereotypes are not equally harmful. For example, the stereotype that Germans are humourless is negative but less negative than the stereotype that Roma are fraudsters. Both stereotypes are negative, but to different degrees: it is more harmful to be perceived as a fraudster than as someone lacking a sense of humour.

4. Methods and Materials

In the last years, several authors have advocated the use of linguistic corpora to investigate issues in the philosophy of language (Bluhm, 2013, 2016; Hansen, 2015; Hansen & Chemla, 2015; Hansen et al., 2019; Caton, 2020; Hinton, 2020; Tallant & Andow, 2020). Although it is not the most widespread methodology, it has certain advantages that make it desirable over other options. First, linguistic corpora are unbiased data sets. As Bluhm (2013, p. 12) puts it, corpora “data are, by and large, unfiltered”.

Second, when using linguistic corpora, we analyze the frequencies of terms or the frequencies of the expressions that typically co-occur with them. However, corpora also provide the broader contexts in which the words appear. A detailed analysis of this broader context may discover new features about the phenomenon we are studying.

Third, computer processing allows for automatic and more in-depth analyses. RStudio (see Desagulier, 2017; Gries, 2017) and #Lancsbox⁶ are available options for the analyses. However, since the corpus language is Spanish, we chose Sketch Engine⁷ because of the ease with which it processes non-English languages.

Fourth, the amount of evidence considered is high. Given the specificity of the subject studied in this work, we have not used general corpora, but a “specialized or special corpora” (Pearson, 1998, 51; Bowker & Pearson, 2002, 48). The size of these corpora is significantly smaller. Even so, our corpus reached a size more than sufficient to carry out a representative and in-depth analysis on the subject studied. Specifically, the corpus has 631.418 words and 24.817 occurrences of the term *menas*.

We compiled our special corpus mining Spanish tweets (and their answers) in which the word *menas* appears. The tweets compounding the corpus were mined from three different

6 <http://corpora.lancs.ac.uk/lancsbox/>.

7 <https://www.sketchengine.eu/>.

times: i) September 14, 2020 to September 28, 2020; ii) October 17, 2020 to November 7, 2020; and iii) March 17, 2021 to April 3, 2021. First, we focus on two specific moments, the declaration of Díaz Ayuso, president of the Community of Madrid, Spain, on September 21, 2020, and the riots that occurred in different Spanish cities due to the COVID measures on October 31, 2020.⁸ Specifically, we extracted the tweets from the week before and after both events, i.e., two periods of fifteen days each. We chose the former incident because Díaz Ayuso used the term *menas* explicitly in her statement.⁹ Therefore, we expected that there would be many tweets mentioning the term. The latter event is not related to *menas*. However, being a time of high social tension, we anticipated much movement on Twitter. As we expected, the term appeared in many tweets. Then, we mined tweets from a time (from March 17, 2021, to April 03, 2021) when no significant political or social event occurred to make the corpus more balanced.

To assess the status of the term *menas*, we carry out two analyses: a collocation and a concordance analysis. First, a collocation is “the tendency of words to be biased in the way they co-occur” (Hunston, 2002, 68). That is, collocates are words “that occur frequently within the neighborhood of another word, normally more often than we would expect the two words to appear together because of chance” (Baker, Gabrielatos and McEnery, 2013, 36). The collocation analysis will allow us to see which terms (neutral and pejorative) usually accompany *menas*, and how often each of them appears typically. Second, a concordance analysis consists in obtaining a “list of all of the occurrences of a particular search term in a corpus, presented within the context that they occur in” (Baker, 2006, 71). In a nutshell, a KWIC (for key words in context) analysis serves to visualize all occurrences of a search term in a corpus with its context. In this work, the KWIC analysis will help us find examples to test whether *menas* presents the features that are characteristically associated with ESTIs.

Before presenting the analyses mentioned above, it should be noted that the analyses carried out in the present study are radically different from the linguistic tests reported in Castroviejo, Fraser and Vicente (2020). According to the authors, these linguistic tests make it possible to differentiate which terms are ESTIs. First, we should begin by saying that the corpus evidence allows us to conclude that the term *menas* passes three of the five tests: i) “the emphatic Spanish indefinite” (see Castroviejo, Fraser and Vicente, 2020, 7–10), in examples like *Lacoste y iPhone. Estos no roban en mercadillos, faltaría más. Son unos menas muy monos*, “Lacoste and iPhone. They don’t rob at flea markets, of course. They are cute little *menas*”; ii) “degree modification” (see Castroviejo, Fraser, and Vicente, 2020, 11–12), in examples like *Son unos MENAS de cojones!*, “They are *MENAS* as shit!”; and iii) “the *auténtico* environment” (see Castroviejo, Fraser and Vicente, 2020, 12–13), since we find examples like *No hay que leer nada, solo ver las imágenes. Auténticos menas*, “No need to read anything, just look at the images. Real *menas*”.

Concerning the other two tests, *menas* may pass one of them, “verbless exclamatives” (see Castroviejo, Fraser, and Vicente, 2020, 7),¹⁰ while it is impossible to know whether it

8 <https://www.elmundo.es/espana/2020/10/31/5f9d33bb21efa099058b4638.html>.

9 <https://www.publico.es/tremending/2020/09/21/isabel-diaz-le-pen-ayuso-cuela-un-mensaje-relacionando-la-covid-con-delincuencia-okupas-y-menores-extranjeros/>.

10 Consider the following example: *Por qué todos los pibes de 13 años se visten de menas*, “Why do all 13-year-old boys dress like *menas*”. Given this tweet, we believe that the following example sounds OK: *Todos los pibes*

passes the other, “N1 of an N2 construction” (see Castroviejo, Fraser, and Vicente, 2020, 6), because the test refers to features more directly related to spoken language, something impossible to check in our case.

Even leaving this aside, there are reasons not to apply such linguistic tests in the analyses. First, the intuitions we have regarding some of the tests do not coincide with those of the authors.¹¹ Second, since ESTIs are terms where context plays an important role and where associated stereotypes play a differential role, we prefer to use corpus methods such as concordance and collocation analysis to explore these two issues. The former allows us to observe the term in context, while the latter gives us a set of terms associated with *menas*, allowing us to extract the associated stereotypes.

5. Analyses

As we have seen in Section 3, ESTIs have five distinctive features. In this section, we will expose the outcomes of the analyses to see if the term *menas* exhibits these five features and thus if it is more like an ESTI than a slur. First, *context-dependent offensiveness* refers to the idea that ESTIs, unlike slurs, are frequently used both neutrally and offensively. Specifically, the offensiveness of w-ESTIs depends on the context, while the offensiveness of i-ESTIs, when applied to people outside the target group, does not. Consider the following examples:

- (11) *Criminalizar y señalar a MENAS, por ejemplo desde grupos políticos es violencia y racismo.* “Criminalizing and pointing to *MENAS*, for example from political groups, is violence and racism.”
- (12) *Y sí, estamos hasta la polla de los inmigrantes ilegales y menas que sólo crean problemas.* “And yes, we are sick to death of illegal immigrants and *menas* who only create problems.”
- (13) *Pues los que destrozan no destrozan porque le vaya a faltar el pan, destrozan porque son unos menas hijos de la gran puta.* “Well, those who vandalize do not vandalize because they are going to lack bread, they vandalize because they are *menas* sons of bitches.”
- (14) *Estos no roban en mercadillos, faltaría más. Son unos menas muy monos.* “They don’t steal from flea markets, of course. They are cute little *menas*.”

As shown by the examples, cases where *menas* is used as an i-ESTI are pejorative, whereas when it is used as a w-ESTI can be offensive or neutral. A close examination of the cases where *menas* is an i-ESTI reveals that the contexts are predominantly negative. Of the 19 copular construction cases (i.e., copula + indefinite article + *menas*), there are 17 cases where the context is negative. On the contrary, when *menas* is used as a w-ESTI, we find negative contexts, but also a lot of neutral ones where the term is related with the idea that *menas* are children.

de 13 años se visten de menas. Mira qué se ha puesto hoy Manolito. ¡Todo un mena, tu primo! “All 13-year-old boys dress like *menas*. Look what Manolito is wearing today. A total *mena*, your cousin!”

11 In our view, the N₁ of N₂ construction, for instance, works well with the ESTI *gitano*, but does not sound very idiomatic with the ESTI *portera*. Similarly, with some slurs, the construction N₁ of N₂ sounds idiomatic to us: *la puta de tu hermana* “the bitch of your sister”, or *el maricón de tu hermano* “the fagot of your brother”.

Second, *stereotype-based offensiveness* alludes to the idea that the application of an ESTI (whether an i-ESTI or a w-ESTI) depends on the insulted person meeting the stereotypes usually associated with the target group, and not on the fact that the person belongs to the group. Consider the following cases:

- (15) *Ya no hay jóvenes en este país, solo menas. Mi sobrino de 14 años, lleva ese pelado, ese look, ... será un mena?* “There are no young people in this country anymore, only *menas*. My 14-year-old nephew wears that haircut, that look, ... is he a *mena*?”
- (16) *Parecen todos fotocopias. ¿Solo existe ese corte de pelo? Cuidao que mañana les llaman menas.* “They all look like photocopies. Is there only one haircut? Be careful, tomorrow people will call them *menas*.”
- (17) *Son menas pagados, el fenotipo es magrebí.* “They are paid *menas*, the phenotype is Maghrebi.”

As can be seen, applying the concept in all the examples does not depend on whether the individual in question is an unaccompanied foreign minor, but on whether he or she is wearing a particular haircut or is a Maghrebi. Ultimately, the offensiveness of *menas* depends on whether the individual to whom the concept is applied meets the stereotypes generally associated with the term.

Third, *mixed embedding behaviour* says that w-ESTIs do not project out of various forms of embedding, while i-ESTIs do project. Consider the following examples:

- (18) *Es lo que tiene el racismo, que veis lo que queréis... Como sabéis que son menas?* “That’s the thing about racism, you see what you want to see... How do you know they are *menas*?”
- (19) *Veo a muchas chicas blancas en los videos, esas no son menas, esta claro que algo hemos hecho muy mal con esta juventud.* “I see a lot of white girls in the videos, those are not *menas*, it is clear that we have done something very wrong with this youth.”
- (20) *los troyanos que nuestras queridas élites han introducido en nuestras sociedades para fomentar el caos y el miedo, MENAS. ¿Tenéis pruebas para demostrar que son unos menas?* “the Trojans that our beloved elites have introduced into our societies to foment chaos and fear, *MENAS*. Do you have evidence to prove that they are *menas*?”
- (21) *Esperemos que no hayan sido unos menas, que entonces se te cae todo el argumento.* “Let’s hope they weren’t *menas*, then the whole argument falls apart.”

As the examples show, although in the two examples where *menas* is used as an i-ESTI it does seem that the term projects out, the intuitions regarding the examples where *menas* is used as a w-ESTI are not clear. It could be that the term does not project out in the first example, but projects in the second, but it is not entirely clear for us.

Fourth, *perspective dependence* conveys the idea that ESTIs signal that the speaker holds derogatory attitudes towards the target group. Examining the different collocates of *menas*,

we can find both neutral and offensive uses. Among the most frequent modifiers of *menas* (see Table 1), we find the neutral collocate *menores* “minors”, with a Log Dice¹² of 8.94. However, we also find pejorative collocates. For example: i) *putos*, “fucking” (Log Dice: 10.57); ii) *malditos*, “damned” (Log Dice: 8.66); iii) *violadores*, “rapist” (Log Dice: 8.51); and iv) *violentos*, “violent” (Log Dice: 8.43).

Collocate	Log Dice
<i>putos</i> (“fucking”)	10.57
<i>marroquíes</i> (“Moroccan”)	9.41
<i>pobres</i> (“poor”)	9.31
<i>nazis</i> (“nazis”)	9.25
<i>menores</i> (“minors”)	8.94
<i>morenos</i> (“dark”)	8.85
<i>malditos</i> (“damned”)	8.66
<i>violadores</i> (“rapist”)	8.51
<i>musulmanes</i> (“muslim”)	8.46
<i>violentos</i> (“violent”)	8.43

Table 1. Top ten modifiers of *menas*

Regarding the construction *menas son*, “*menas are*”, we find neutral and pejorative collocates (see Table 2). Among the neutral collocates, we find: i) *menores*, “minors” (LogDice: 11.41); ii) *niños*, “children” (Log Dice: 10.71); and iii) *personas*, “people” (Log Dice: 10.45). However, we also find the pejorative collocate *problema*, “problem” (Log Dice: 11.43).

Collocate	Log Dice
<i>problema</i> (“problem”)	11.43
<i>menores</i> (“minors”)	11.41
<i>niños</i> (“children”)	10.71
<i>ultraderecha</i> (“far right”)	10.53
<i>personas</i> (“people”)	10.45

Table 2. Most frequent collocates of *menas son*, “*menas are*”

Regarding the verbs to which *menas* appear as subject, we also find neutral and pejorative collocates (see Table 3). Among the neutral ones, the following can be noted: i) *nacer*, “be born” (Log Dice: 10.3); ii) *tener*, “have” (Log Dice: 9.7); iii) *votar*, “vote” (Log Dice: 8.98); and iv) *decir*, “say” (Log Dice: 8.89). However, we also find a lot of pejorative collocates among the most frequent ones: i) *robar*, “steal” (Log Dice: 10.98); ii) *saquear*, “loot”

12 Log Dice is a statistical measure that expresses “the tendency of two words to co-occur relative to the frequency of these words in the corpus”. It is a “standardized measure operating on a scale with a fixed maximum value of 14” (Gablasova, Brezina and McEnery, 2017, 164).

(Log Dice: 10.77); iii) *apuñalar*, “stab” (Log Dice: 10.41); iv) *delinquir*, “commit crime” (Log Dice: 9.53); and v) *asaltar*, “rob” (Log Dice: 9.25).

Collocate	Log Dice
<i>robar</i> (“steal”)	10.98
<i>saquear</i> (“loot”)	10.77
<i>apuñalar</i> (“stab”)	10.41
<i>nacer</i> (“be born”)	10.3
<i>tener</i> (“have”)	9.7
<i>delinquir</i> (“commit crime”)	9.53
<i>aprovechar</i> (“take advantage of”)	9.36
<i>asaltar</i> (“rob”)	9.25
<i>votar</i> (“votar”)	8.98
<i>decir</i> (“say”)	8.89

Table 3. Top ten verb collocates of *menas* as subject

In the light of the most frequent collocates of *menas*, it can be said that their use indicates that the speaker maintains derogatory attitudes. Although there are neutral collocates, among the most frequent modifiers and verbs a greater number of pejorative collocates can be observed.

Fifth, *offensive variation* highlights that, when used offensively, ESTIs do not appear to be equally offensive. In other words, we can rank ESTIs according to a scale of offensiveness. Considering *menas* collocations, we can say that the term’s offensiveness is high. We only must consider collocations such as *violadores*, “rapist”, *putos*, “fucking”, *robar*, “steal”, o *apuñalar*, “stab”, to see the high offensiveness associated with the term. To serve as a contrast, consider the most frequent collocations of an ESTI as *portera*, “doorwoman”. When searched as an i-ESTI in the esTenTen Corpus (Kilgarriff and Renau, 2013) in Sketch Engine,¹³ we only find one negative term among the top 200 most frequent collocations: *cotilla*, “gossiper” (Log Dice: 4.36). Moreover, since the offensiveness of ESTIs depends on the stereotypes associated with each term, the more negative the stereotypes, the higher the offensiveness. As can be seen, the stereotype of being someone who is a gossiper, someone who cannot keep a secret, is negative, but it is far from being as negative as being a person who rapes, steals, and stabs.

In summary, the combined evidence from the collocation and concordance analysis concludes that *menas* exhibit four of the five defining characteristics of ESTIs. Regarding mixed embedding behavior, the evidence is not conclusive, as there is no clarity in the intuitions elicited by the examples extracted from the corpus. In the next section, we will discuss some of the reasons why we believe this to be the case. We will also discuss the results to analyse the status of the term and its dynamics and future evolution.

¹³ We conducted the search on May 27, 2021.

6. Discussion

The results of the analyses presented in the previous section conclude that *menas* is an ESTI. First, there are key features of ESTIs that seem to fit well with how the term *menas* is used. On the one hand, there are neutral and offensive uses of the term; on the other hand, their offensiveness depends on the insulted person meeting the stereotypes usually associated with the target group. Besides, as shown by the analyses, *menas* exhibits four of the five defining features of ESTIs, being mixed embedding behaviour doubtful. The concordance analysis allowed us to show the context-dependent offensiveness of the term. In other words, we find that *menas*, when used as an i-ESTI, is almost always offensive, and when used as a w-ESTI, there are offensive and neutral cases. Through the concordance analysis, we also proved the stereotype-based offensiveness of the term. We found examples where the application of the term does not depend on the individual belonging to the group but rather the individual meets the negative stereotypes associated with the term.

The collocation analysis allowed us to show both the perspective dependence and the offensive variation of the term. Concerning the former, we found neutral collocates, for example, *menores*, “minors”, *niños*, “children”, or *nacer*, “to be born”, but also a significantly higher number of pejorative collocates, especially concerning modifiers, for example, *putos*, “fucking”, *malditos*, “damned”, *violadores*, “rapist”, and *violentos*, “violent”, and verbs to which *menas* appears as subject, for example, *robar*, “steal”, *saquear*, “loot”, *apuñalar*, “stab”, *delinquir*, “commit crime”, and *asaltar*, “rob”. Concerning the latter, the analysis allowed us to compare the offensiveness of *menas* and another ESTI such as *portera*, “doorwoman”, showing that *menas* is a much more offensive ESTI.

Regarding the last characteristic, mixed embedding behaviour, the corpus evidence is not clear. We believe that the variation in intuitions about the embedding behavior of *menas*, when used as a w-ESTI, is partly due to the examples being considered out of context. In this way, we believe there will be cases where *menas*, used as a w-ESTI, projects out when embedded in a question or a denial and others where not. We do not know this with total certainty, but we can assume that an excellent guide to know it will be the collocation and concordance analyses. It is likely not to project if the term is surrounded by neutral terms such as *niños*, “children”, or *menores*, “minors”. However, if negative terms like *violentos*, “violent”, or *robar*, “steal”, accompany the term, it is likely to project even if it is embedded in a question or a negation. Consider the following example of the corpus:

(22) *¿Seguro que no son menas?* “Are you sure they are not *menas*?”

Considered in isolation, the term does not project. The speaker asks a question because s/he is not sure, and therefore it cannot be said that s/he maintains derogatory attitudes. However, if we consider the example with more context, the intuition changes:

(23) *¿Seguro que no son menas? La España que madruga y se ducha se dedica a construir y producir. Nunca a destruir.* “Are you sure they are not *menas*? The Spain that gets up early and takes a shower is dedicated to building and producing. Never to destroying.”

In sum, it seems that *menas* projects out when used as an i-ESTI, but more evidence is needed to determine if *menas* projects out or not when used as a w-ESTI. An exhaustive analysis of the embedding behavior of the term would entail a close reading of the context of each case.

In sum, it appears that the term *menas* is an ESTI, since it exhibits almost all the features typical of this type of term. However, we believe that the term may be in the process of becoming a slur. In fact, we believe that two scenarios are possible. On the one hand, on the assumption that the pejorative use of *menas* is relatively recent, it seems clear that the features of the word are not fully defined. Besides, to the extent that there are neutral uses, it may make sense to continue to argue that it is an ESTI. If people who know the origin of the term continue to use it neutrally and fight to prevent the increasing pejorative use of the term, perhaps they can maintain this double use.

On the other hand, the neutral use could be lost over time and *menas* may become a slur. If, in addition to the people that uses the term pejoratively in a systematic manner, people who know its origin prefer not to use it anymore or to use other expression instead—for example, “unaccompanied foreign minors” (as some associations seem to propose, arguing that *menas* makes invisible the fact that they are minors)—, then *menas* may remain exclusively a slur.

Therefore, it seems that we are in a halfway point in which both things are possible. How the term evolves will depend on how speakers use it from now on. What is clear is that there are two tendencies of usage that are in tension with each other. Several examples in the corpus show that the debate on the application of the term is already underway. Consider the following examples:

- (24) *Vosotros sois más de llamar MENAS a cualquiera que haga un delito sin saber su procedencia y si nacieron aquí o están con sus padres.* “You just call *MENAS* anyone who commits a crime without knowing where they come from and whether they were born here or are with their parents.”
- (25) *Y por último, llamar MENAS a chicos que no sabemos si quiera si son españoles o tienen familia solo para criminalizar un colectivo es ser basura.* “Finally, calling *MENAS* kids that we don’t even know if are Spanish or have a family just to criminalize a collective is being trash.”

Besides, the analyses allow us to get some general clues about how to pave the way to hell, that is, the process of conversion of a neutral word into a pejorative one. In particular, the systematic appearance of an expression, initially neutral, together with explicitly pejorative expressions is at least one of the mechanisms through which a descriptive expression ends up acquiring an expressive meaning. This expressive meaning may coexist with the descriptive meaning, turning a descriptive term into an ESTI, or it may end up phagocytizing the descriptive meaning altogether, becoming a slur. What happens depends on what speakers do. For example, in the case of the expression under consideration, there is a perceived tendency not to use the expression by those who consider that the expression has acquired a pejorative meaning. If they cease to use it in a neutral way, the word will become part of the dialect of racists and become a slur. If they struggle to maintain the neutral use of the term, the word will remain an ESTI.

7. Conclusion

In this paper, we have explored the status of the term *menas*, a term that has attracted considerable attention in Spanish politics in the last few years. The main aim was to determine if *menas* is an ESTI or rather a slur. First, we defined both kinds of terms, presenting their most defining characteristics. Second, we compiled a corpus consisting of tweets involving the term. Third, we ran two different types of analysis, a collocation and a concordance analysis, to determine the status of the term. The results of the analyses show that *menas* exhibits the most defining characteristics of ESTIs. However, it seems that neutral uses of the term are on the decline. The corpus yields evidence in this sense since, for example, the term co-occurs more frequently with pejorative terms than with neutral ones, as the collocation analysis shows. For this reason, we believe that the term, currently, is in the process of becoming a slur. To account for this possible dynamic, we have pointed out two possible scenarios. On the one hand, the term could maintain its neutral uses and thus continue to be an ESTI. However, it is also possible that the neutral uses disappear over time, and only the pejorative ones remain. If this were to happen, the term would become a slur.

References

- Anderson, L. & Lepore, E. (2013), “Slurring words”, *Noûs*, 47(1), pp. 25–48.
- Baker, P. (2006), *Using Corpora in Discourse Analysis*, New York: Continuum.
- Baker P., Gabrielatos C. & McEnery T. (2013), *Discourse Analysis and Media Studies: The Representation of Islam in the British Press*, Cambridge: Cambridge University Press.
- Biber et al. (1999), *Grammar of Spoken and Written English Language*, London: Longman.
- Blakemore, D. (2015), “Slurs and expletives: A case against a general account of expressive meaning”, *Language Sciences*, 52, 22–35.
- Bluhm, R. (2013), “Don’t ask, look! Linguistic corpora as a tool for conceptual analysis”, in M. Hoeltje, T. Spitzley, and W. Spohn (eds.): *Was dürfen wir glauben? Was sollen wir tun? Sektionsbeiträge des achten internationalen Kongresses der Gesellschaft für Analytische Philosophie, e.V. Duisburg-Essen*: DuEPublico, pp. 7–15.
- Bluhm, R. (2016), “Corpus analysis in philosophy”, in M. Hinton (ed.), *Evidence, Experiment and Argument in Linguistics and Philosophy of Language*, Oxford: Peter Lang, pp. 91–109.
- Bolinger, R. J. (2017), “The pragmatics of slurs”, *Noûs*, 51(3), 439–462.
- Bowker, L. & Pearson, J. (2002), *Working with Specialised Language: A Practical Guide to Using Corpora*, London: Routledge.
- Camp, E. (2013), “Slurring perspectives”, *Analytic Philosophy*, 54(3), 330–349.
- Castroviejo, E., Fraser, K. & Vicente, A. (2020), “More on pejorative language: Insults that go beyond their extension”, *Synthese*, DOI: 10.1007/s11229-020-02624-0.
- Caton, J. N. (2020), “Using linguistic corpora as a philosophical tool”, *Metaphilosophy*, 51(1), 51–70.
- Cepollaro, B. (2015), “In defense of a presuppositional account of slurs”, *Language Sciences*, 52, 36–45.

- Croom, A. M. (2013), “How to do things with slurs: Studies in the way of derogatory words”, *Language and Communication*, 33(3), 177-204.
- Croom, A. M. (2015), “The semantics of slurs: A refutation of coreferentialism”. *Ampersand*, 2, 30–38.
- Desagulier G. (2017), *Corpus Linguistics and Statistics with R: Introduction to Quantitative Methods in Linguistics*, Cham: Springer.
- Falbo, A. (2021), “Slurs, neutral counterparts, and what you could have said”, *Analytic Philosophy*, DOI: 10.1111/phib.12217.
- Gablasova, D., Brezina, V. & McEnery, T. (2017), “Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence”, *Language Learning*, 67(S1), 155–179.
- Gries S. T. (2017), *Quantitative Corpus Linguistics with R: A Practical Introduction*, London: Routledge.
- Hansen, N. & Chemla, E. (2015), “Linguistic experiments and ordinary language philosophy”, *Ratio*, 28(4), 422–445.
- Hansen, N. (2015), “Experimental philosophy of language”, *Oxford Handbooks Online*, DOI: 10.1093/oxfordhb/9780199935314.013.53.
- Hansen et al. (2019), “A corpus study of “know”: On the verification of philosophers’ frequency claims about language”, *Episteme*, 18(2), 242–268.
- Hinton M. (2020), “Corpus linguistics methods in the study of (meta)argumentation”, *Argumentation*, DOI: 10.1007/s10503-020-09533-z.
- Hom, C. (2008), “The semantics of racial epithets”, *The Journal of Philosophy* 105(8), 416–440.
- Hunston S. (2002), *Corpora in Applied Linguistics*, Cambridge: Cambridge University Press.
- Jeshion, R. (2013), “Slurs and stereotypes”, *Analytic Philosophy*, 54(3), 314–329.
- Kilgariff, A. & Renau, I. (2013), “esTenTen, a vast web corpus of Peninsular and American Spanish”, *Procedia: Social and Behavioral Sciences*, 95, 12–19.
- McEnery T., & Hardie A. (2012), *Corpus Linguistics: Method, Theory and Practice*, Cambridge: Cambridge University Press.
- Nunberg, G. (2018), “The social life of slurs”, in D. Fogal, D. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 237–295.
- Pearson, J. (1998), *Terms in Context*, Philadelphia: John Benjamins Publishing Company.
- Popa-Wyatt, M. & Wyatt, J. L. (2018), “Slurs, roles and power”, *Philosophical Studies*, 175(11), 2879–2906.
- Tallant, J. & Andow, J. (2020), “English language and philosophy”, in S. Adolphs & D. Knight (eds.), *The Routledge Handbook of English Language and Digital Humanities*, London: Routledge, pp. 440–455.
- Tirrell, L. (1999), “Racism, sexism, and the inferential role theory of meaning”, in C. Hendricks & K. Oliver (eds.), *Language and Liberation: Feminism, Philosophy, and Language*, New York: State University of New York Press, pp. 41–79.
- Zeman, D. (2021), “A rich-lexicon theory of slurs and their uses”, *Inquiry*, DOI: 10.1080/0020174X.2021.1903552.



Provocative insinuations*

Insinuaciones provocativas

ÁLVARO DOMÍNGUEZ-ARMAS**

ANDRÉS SORIA-RUIZ***

Abstract: In this paper we analyse utterances that, without explicitly constituting hate speech, nevertheless convey a hateful message. For example, in the headline “Iraqi Refugee is convicted in Germany of raping and murdering teenage girl”, the presence of “Iraqi refugee” does not seem arbitrary. To the contrary, it is responsible for inviting a racist inference against Iraqi refugees. We defend that these inferences cannot be described as slurs, ethnic or social terms used as insults, dogwhistles or conversational implicatures. Rather, we propose that these inferences are insinuations, specifically *provocative* insinuations, as no conversational response seems effective at blocking them.

Keywords: Conversational implicatures, dogwhistles, inferences, insinuations, hate speech.

Resumen: En este artículo analizamos oraciones que, sin constituir explícitamente discurso de odio, no obstante transmiten un mensaje de odio. Por ejemplo, en el titular “Refugiado iraquí es condenado en Alemania por violar y asesinar una chica adolescente”, la presencia de “refugiado iraquí” no parece arbitraria. Al contrario, invita una inferencia racista en contra de los refugiados iraquíes. Argumentamos que estas inferencias no pueden ser descritas como slurs, términos neutros utilizados como insultos, dogwhistles o implicaturas conversacionales. En cambio, proponemos caracterizar estas inferencias como insinuaciones, concretamente insinuaciones *provocativas*, debido a que ninguna respuesta parece efectiva a la hora de bloquearlas.

Palabras clave: Implicatura conversacional, dogwhistles, inferencias, insinuaciones, discurso de odio.

Recibido: 31/05/2021. Aceptado: 02/07/2021.

* Thanks to Pedro Abreu, Marcin Lewiński, Mora Maldonado, Dima Mohammed, Erich Rast, José Ramón Torices, and audiences at IFILNOVA (Portugal), UNICAMP (Brazil), and the University of Granada (Spain). This work has been funded by Portuguese national funds through FCT-Fundação para a Ciência e a Tecnologia, I.P., within PhD scholarship 2020.09683.BD, and project UIDB/00183/2020, as well as by COST Action CA17132, funded by the Horizon 2020 Framework Programme of the European Union.

** PhD student at IFILNOVA, Lisbon, Portugal (a.dgueza@campus.fcsh.unl.pt). Lines of research: philosophy of language, argumentation theory, political philosophy.

*** Postdoctoral researcher at IFILNOVA, Lisbon, Portugal (aruiz@fcsh.unl.pt). Lines of research: philosophy of language, formal linguistics, metaethics. Recent publications: Soria-Ruiz, A., & Faroldi, F. L. (2020), “Moral adjectives, judge-dependency and holistic multidimensionality”, *Inquiry*, DOI: 10.1080/0020174X.2020.1855241, and Soria-Ruiz, A., & Stojanovic, I. (2019), “On linguistic evidence for expressivism”, in M. J. Frápolli (ed.), *Expressivism, Knowledge and Truth*, Cambridge: Cambridge University Press, 86, pp. 155–180.

1 Introduction

Hate speech has the potential to undermine the space we all inhabit by damaging the dignity of members of vulnerable groups through the association of hateful ideas with their identity (McGowan, 2012, 2019; Waldron, 2012). In this paper we are concerned with utterances that, without explicitly constituting hate speech, nevertheless convey a hateful message. Our paradigmatic examples come from headlines published, e.g., in the *New York Times* (NYT) (2019) or the Spanish tabloid *MediterráneoDigital* (2021):

- (1) Iraqi Refugee is convicted in Germany of raping and murdering teenage girl.
 ~ Iraqi refugees are despicable/dangerous/malicious.
- (2) Un marroquí, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
 A Moroccan, arrested for hiding 187kg of hashish under his son's cradle.
 ~ Moroccan people are despicable/dangerous/malicious.

Although these newspapers have little in common,¹ both headlines have arguably a similar effect: they invite an inference (marked with “~”) against Iraqi refugees and Moroccan people, respectively. For lack of a better term, we call this a *racist inference* (RI). The general phenomenon is that the appearance of social, ethnic or national group terms together with actions or events that pose a risk for society suggests that such actions are typically produced by members of such groups, thereby promoting the view that those groups are despicable, dangerous or malicious.

Let us start by pointing out two aspects of this phenomenon. First, RIs do not arise with *any* predicate in subject position. This can be seen by noting that, if we substitute “marroquí” for another predicate, e.g., “cartero” (“postman”) in the subject position of a sentence like (2), the hateful inference disappears:

- (3) Un cartero, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
 A postman, arrested for hiding 187kg of hashish under his son's cradle.
 ~/ postmen are despicable/dangerous/malicious.

Secondly, RIs do not arise simply as a consequence of associating negative actions or events with a group. An inference arises even in cases where the overall message is positive:

- (4) Cuatro vecinos gitanos evitan que una mujer sufra una violación en plena calle.
 Four Romani neighbours save a woman from being raped in the middle of the street.
 ~ Contrary to expectations, these four Romani people are not despicable/dangerous/malicious.

In this case, the headline seems to highlight the surprising character of the situation, where the saviours, and not the attackers, are Romani. Thereby, a racist inference against the Romani community still goes through, even if it is weaker or qualified.

¹ While the *New York Times* is an internationally well-renowned journal, *MediterráneoDigital* is a tabloid known in Spain for spreading racist, sexist and overall extreme right views.

Our research question is: how can RIs be described? In this paper, we consider the possibility of describing them as *slurs* (§2), *ethnic and social terms used as insults* (ESTIs, §3), *dogwhistles* (§4) or *conversational implicatures* (§5). We argue that RIs differ in crucial ways from each of those phenomena. We propose that RIs are a type of *insinuation* (Camp, 2018), in particular, a *provocative* insinuation (§6), whose effects are achieved *via* the mechanism of *conversational eliciture* (Cohen & Kehler, 2021). In §7, we discuss ways of confronting provocative insinuations, aimed at blocking their hateful effects. §8 concludes.

2. RIs as slurs

The presence of “Iraqi refugee”/“Morrocan” in (1–2) contributes to conveying a derogatory conceptualization of these groups via the inference that they are despicable/dangerous/malicious. Slurs give rise to similar derogatory inferences:²

- (5) Bill is a N*
 ~ black people are despicable/dangerous/malicious.

Slurs are lexical items that convey a derogatory conceptualization of a group. Slurs are co-extensional with non-pejorative terms—so-called neutral counterparts (Falbo, 2021; Hess, 2020a, 2020b; Nunberg, 2018).³ The slur “N*” is a conventionally derogatory way to refer to black people. “N*” can be substituted by another non-derogatory term without triggering the derogatory inference—e.g., “Bill is black”.

Given that (1–2) convey a derogatory conceptualization of certain groups, one might be inclined to characterize “Iraqi refugee”/“Morrocan” as slurs (Hess, 2020a, 2020b) or their particular use in sentences like (1–2) as a type of *slurring*⁴ (Camp, 2013; Kukla, 2018). Moreover, slurs acquire and lose their derogatory meaning over time, so it could be the case that “Iraqi refugee” or “marroquí” (in Spanish), which on their face are not as aggressively pejorative as other paradigmatic examples, might be in the process of becoming a slur.⁵

However, there are several differences between our phenomenon and *slurs*. First, the RIs triggered by (1–2) are not limited to a single lexical item. Substituting “Iraqi refugee”/“marroquí” for (i.e.) “Iraqi asylum seeker”/“magrebi” would produce similar inferences.

Secondly, the predicates “Iraqi refugee” or “marroquí” are not conventionally derogative, as shown by the fact that they can appear in sentences that invite no derogatory inferences:

2 The mere mention of a slur is often perceived as offensive, so we abstain from reproducing the word (see Cepolaro, Sulpizio and Bianchi, 2019).

3 In discussions about slurs, the existence of neutral counterparts is generally taken for granted. This has been questioned recently (Bolinger, 2017; Falbo, 2021).

4 Slurring—as the activity of using slurs—is conceived as a device through which power relations, unjust institutions and ideologies are constituted (see Corredor, 2014; Kukla, 2018; McGowan, 2019; Popa-Wyatt & Wyatt, 2018).

5 In support of this possibility, note that Al Jazeera decided in 2015 to stop using “migrant” to refer to people crossing the Mediterranean, owing to alleged pejorative connotations of the term (<https://www.aljazeera.com/features/2015/8/20/why-al-jazeera-will-not-say-mediterranean-migrants>).

- (6) The first Iraqi refugees arrived in Jordan in the aftermath of the 1991 Gulf War.
- (7) Haz cuscús como un marroquí.
Make couscous like a Moroccan.

Moreover, “marroquí” is the neutral counterpart of a Spanish slur (“m*ro/a”).

Thirdly, the derogatory content of slurs and RIs project in different ways, as illustrated by the contrast between (8) and (9). As it’s well established, the derogatory content of slurs projects out of non-assertive environments such as negation, questions or modals (Bolinger, 2017; Camp, 2013; Cepollaro and Stojanovic, 2016; Croom, 2011; DiFranco, 2014; Jeshion, 2013a, 2013b):

- (8)
 - (a) Bill is not a N*.
 - (b) Is Bill a N*?
 - (c) Bill might be a N*.
 ↗ black people are despicable/dangerous/malicious.

However, the RIs of (1–2) do not arise as easily in similar environments. As shown by (9), RIs are substantially weakened in non-assertive contexts:

- (9)
 - (a) Iraqi refugee is not convicted of raping and murdering a teenage girl.
 - (b) Is an Iraqi refugee convicted of raping and murdering a teenage girl?
 - (c) Iraqi refugee might be convicted of raping and murdering a teenage girl.
 ↗ Iraqi refugees are despicable/dangerous/malicious.

Finally, another point of contrast between slurs and RIs is that slurs can be used as vocatives, while the predicates that trigger RIs cannot. Falbo (2021, 15) describes these “weapon uses” as the most toxic and offensive way that a slur can be used. She exemplifies this with a communicative situation where, e.g., a gay couple is on the street holding hands when an unprovoked stranger yells out “Hey F*!”. The thing to note here is that, if the speaker had yelled “Hey gay men!”—the *neutral counterpart* of “F*”—such an utterance would be very unnatural. The same thing is true of the predicates responsible for the inferences in (1–2). If a speaker yells out “Hey Iraqi refugee!”, this is likely to sound strange and unnatural.

In sum, although RIs are similar to slurs insofar as they derogate certain groups, RIs and slurs do not produce this effect via the same linguistic mechanism.

3. RIs as ESTIs

A more promising hypothesis might be to assimilate RIs to terms that, in contrast with *slurs*, do not have a demeaning effect on their targets as a matter of their lexical semantics. Such is the case of *ethnic and social terms used as insults* (henceforth ESTIs, Castroviejo, Fraser, and Vicente, 2020). ESTIs are terms that are generally neutral, but which, in some environments, convey a pejorative meaning built on negative stereotypes

associated with the relevant group. ESTIs include ethnic, professional or even age terms like the following:

- (10)
- (a) Juan es un gitano.
Juan is a roma.
 - (b) Antonio es una portera.
Antonio is a doorwoman.
 - (c) Javier es un viejo.
Javier is an old-man.

Roughly speaking, (10a) attributes a set of stereotyped—negative—properties of Romani people to Juan, e.g., being a thief. Similarly, (10b) says that Antonio is a gossipy person. And (10c) describes Javier as, e.g., grumpy. The sentences in (10) convey such negative stereotypes relying on arguably neutral terms (“gitano”, “portera”, “viejo”).

In our examples (1–2) (repeated here) we also find ethnic or social terms (“Iraqi refugee”, “Moroccan”) triggering negative stereotypes. Therefore, describing RIs as ESTIs could be intuitively promising.

- (1) Iraqi Refugee is convicted in Germany of raping and murdering teenage girl.
- (2) Un marroquí, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
A Moroccan, arrested for hiding 187kg of hashish under his son’s cradle.

Nonetheless, there is a crucial contrast between RIs and ESTIs: while (1–2) entail that the person referred to is an Iraqi refugee/Moroccan, the sentences in (10) do not entail that the subject belongs to the relevant group. The demeaning effect of ESTIs is achieved through the ascription of negative stereotypes to an individual, but without requiring that the individual *actually* belongs to the group in question. Castroviejo, Fraser and Vicente highlight a contrast between (10a) and the following sentence (without the indefinite article):

- (11) Juan es gitano.
Juan is roma.

While (10a) does not entail that Juan is roma, (11) entails that Juan belongs to this group. And while (10a) attributes a set of negative stereotypes to Juan, (11) does not; it simply states that Juan is a member of this group.

In this respect, (1–2) are more similar to (11) than to (10a). Even though (1–2) convey certain stereotyped beliefs about a group, (1–2) entail that the subject is an Iraqi refugee/Moroccan. So even though (1–2) suggest that Iraqi refugees/Moroccans typically cause these dangerous actions, this is not the result of mere stereotyping, since the people alluded to in (1–2) really *do* belong to the relevant groups.⁶

6 Castroviejo, Fraser and Vicente (2020) argue that the use of “portera” as an insult amounts to a conventional way of referring to features that form part of the (negative) stereotype—it is a conventional way of saying, e.g., “gossipy”. Thus, ESTIs rely on stereotypes and not on actual group membership. In contrast, “Iraqi refugee”,

We take this contrast between ESTIs and RIs to show decisively that RIs are a different phenomenon.

4. RIs as dogwhistles

Another possibility is to describe RIs as a kind of *dogwhistle*. A dogwhistle is a type of message designed to be taken up in different ways depending on the audience (Saul, 2018; Torices, 2021). Dogwhistles are usually conveyed by the presence of a word or idiom that, in addition to its normal interpretation, carries a *coded* interpretation targeted for part of the audience (call it the “target” audience). Dogwhistles can shift the boundaries of permissible discourse about a group (Maitra and McGowan, 2012; McGowan, 2019) without resorting to explicitly racist or discriminatory language.

For example, George W. Bush’s State of the Union speech from 2003 contained the following sentence (Saul, 2018, 362):

- (12) Yet there’s power, *wonder-working power*, in the goodness and idealism and faith of the American people.

The phrase “wonder-working power” is a Biblical reference, and thus (12) has been described as an overt dogwhistle: a message with two plausible interpretations, one for the general audience and one for the target audience, namely Christians⁷—who must recognize the intention of the speaker to convey the private message.

Another example is the expression “inner-city”, used in the United States as a coded way of meaning “black” (Saul, 2018, 367):

- (13) Some people want to increase spending for new prisons to lock up violent *inner-city* criminals.
 (13) is a *covert* dogwhistle, because the target audience need not recognize that a coded message is sent to them.

(1–2) might be described as overt dogwhistles if they are conceived as coded messages targeted to a private audience. (1–2) could be inducing in the audience certain beliefs about the dangerousness of Iraqi refugees or Moroccan people which, in turn, might lead to the implementation of policies against those groups. In 2019 there was public concern in Germany regarding the integration of refugees.⁸ One might say that the RI associated with (1) is a coded message from the NYT, aimed against the integration of refugees in Germany.

as used in (1), does not conventionally mean “despicable/dangerous/malicious”. And moreover, the individual picked out by “Iraqi refugee” really belongs to the group.

7 For a non-fundamentalist, (12) is an ordinary piece of political speech. However, Christian fundamentalists can infer that Bush used the expression “wonder-working power” because he supports fundamentalist Christians.

8 After Angela Merkel’s decision to open Germany’s borders in 2015, support for the right-wing party Alternative for Germany and other anti-EU populists was fueled. Although asylum applications in the country peaked in 2016 (722.370 applications), the increasing popularity of conservative and xenophobic views caused much public debate on refugee integration throughout 2018–2019.

The general audience would conceive (1) as a non-remarkable piece of news—after all, there were 720 murders and 7.724 rapes in Germany in 2019 (Statista Research Department, 2021)—but xenophobic subjects would understand (1) as an example in support of their views. If this was the case, (1–2) might be considered as a type of *overt* dogwhistle. But even if a xenophobic subject failed to recognize that a message is being sent to them, a defender of this view might want to treat (1–2) as *covert* dogwhistles—much like “inner city”—so that target audience recognition is not required.

However, there are two crucial differences between RIs and dogwhistles: first, dogwhistles are intended to be perceived only by the “target” audience, since they rely on the presence of expressions whose coded meaning is only understood by some hearers (“wonder-working power”, “inner city”). The inferences associated with (1–2), even if triggered by the presence of certain predicates (“Iraqi refugee”, “Moroccan”), do not rely on those predicates having a coded or hidden meaning: these inferences can be perceived by anyone. Secondly, dogwhistles lose part of their power when they are “uncovered” (Saul, 2018, 373; see also Torices, 2021). In particular, overt dogwhistles are more powerful insofar as the general audience does not recognize them, and covert dogwhistles *only* succeed if the general audience does not recognize the manipulative intention that underlies them. By contrast, the effects of RIs persist even if the audience recognizes the hateful inference.

We will discuss the effects of RIs in Section 7. For now, we take these observations to suggest that RIs cannot be described as dogwhistles.⁹

5. RIs as conversational implicature

Might RIs be a kind of *conversational implicature* (CI, Grice, 1989)?¹⁰ In a CI, the hearer makes an inference based on an utterance that violates one or more conversational maxims,¹¹ while assuming that the speaker is cooperative.

Implicature calculation follows roughly these steps (Grice, 1989; Cohen and Kehler, 2021). First, the hearer recovers the literal semantic content p of the sentence uttered by the speaker. Secondly, the hearer notices that p violates a conversational maxim. Thirdly,

-
- 9 Saul describes a third type of dogwhistle—*unintentional* dogwhistle—as an “unwitting use of words (...) that, used intentionally, constitute[s] an intentional dogwhistle, where this use has the same effect as an intentional dogwhistle” (2018, 368). This class of dogwhistle might come closer to the inferences in (1–2), given that the speaker could be using “Iraqi refugee”/“Moroccan” unwittingly—without the intention to trigger a racist inference. However, as with intentional dogwhistles, the pernicious effects of unintentional dogwhistles are predicted to disappear when they are recognized. We take this to show that RIs cannot be successfully described as *unintentional* dogwhistles either.
 - 10 We set aside the possibility that RIs are described as *conventional* implicatures, for similar reasons to those adduced in Section 2: first, conventional implicatures are usually associated with particular lexical items. Secondly, in many cases they are expected to project out of entailment-cancelling environments. As we saw, RIs lack both of these properties.
 - 11 Speakers might violate the Maxim of (1) Quantity—make the contribution exactly as informative as required—, (2) Quality—try to make the contribution one that is true—, (3) Relation—be relevant—or (4) Manner—avoid obscurity of expression/be perspicuous. We think that our examples (1)-(2) might most plausibly be described as violating the Maxim of Relevance, but perhaps it could be argued that they stem from a violation of Quantity, as they are *more* informative than necessary. In what follows, we discuss and reject the former hypothesis, but our arguments would go against the latter possibility as well.

the hearer—while assuming that the speaker is aware of such maxims—infer that the sentence must have the secondary purpose of conveying an extra-semantic content q . Finally, the hearer treats the utterance as conveying the extra-semantic content q in addition to the literal semantic content p .

Consider the following standard example, based on a violation of the Maxim of Relation (*Be relevant!*, Grice, 1989, 32): Imagine a subject a , whose car is immobilised, approached by another subject b . The following exchange takes place:

- (14)
a: I am out of petrol.
b: There is a garage round the corner.

In that situation, a first recovers the literal semantic content expressed by b as the proposition that there is a garage nearby. Secondly, a notices that this proposition is irrelevant *vis-à-vis* a 's initial utterance. That is, b 's utterance violates the Maxim of Relation. Thirdly, a assumes that b is a cooperative speaker. Thus, there must be a secondary purpose in pointing out the nearby garage, (e.g.) that the garage has petrol to sell. Fourthly, a infers from b 's utterance that there is a garage round the corner, and additionally, that it has petrol to sell.

One of the essential characteristics of conversational implicatures is that they are *cancellable*: the speaker can deny the extra-semantic content without contradiction. In the previous example, speaker b can follow-up their utterance with "... but in fact they do not sell petrol there" without contradicting themselves.

The RIs of (1–2) (repeated here) could be described as conversational implicatures—based on the idea that the literal content of (1–2) is a violation of a conversational maxim.

- (1) Iraqi Refugee is convicted in Germany of raping and murdering teenage girl.
 ↪ Iraqi refugees are despicable/dangerous/malicious.
- (2) Un marroquí, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
 A Moroccan, arrested for hiding 187kg of hashish under his son's cradle.
 ↪ Moroccan people are despicable/dangerous/malicious.

Specifically, it could be argued that mentioning the person's nationality or international status is a violation of the Maxim of Relation and that this generates the inference.¹² This hypothesis is immediately supported by the fact that RIs are cancellable: the speaker of (1), for example, can follow up their utterance with "...but in fact, Iraqi refugees are great people" without contradiction.

The line of reasoning of the audience of (1) might be as follows: first, the audience recovers the literal semantic content of (1), namely that the person convicted in Germany is an Iraqi refugee. Secondly, the audience notes that the literal content of (1) is in violation

12 Multiple journal stylebooks discourage mentioning the nationality or legal/international status of an individual unless it is strictly necessary (Consejo Audiovisual de Andalucía, 2020). The main reasons offered are that mentioning nationality/legal statuses could convey stereotypes against certain groups (see Carrera, 2017; Colegio de Periodistas de Cataluña, 2020) and that connecting terrorism, dangerous actions, etc. with particular groups might fuel alarmism against them (Pellicer, 2015).

of the Maxim of Relation, since the nationality or international status of the perpetrator is irrelevant. Thirdly, in order to reconcile the literal content of (1) with the assumption that the speaker is being cooperative, the audience infers that the nationality of the perpetrator must be necessary to make sense of the story. Fourthly, a possible interpretation reached by the audience would be that these actions (murder, rape) are typically caused by Iraqi refugees and therefore, that Iraqi refugees are despicable/dangerous/malicious (a similar line of reasoning might be followed by the hearer of (2)).

However, there are salient differences between RIs and CIs, which speaks against subsuming the former under the latter. First, it is not clear that the extra-semantic content of (1–2) arises as a result of a threat of communicative failure. Above, we described the RIs of (1–2) as inferences triggered by a violation of the Maxim of Relation. But this account makes a bad prediction, and is independently questionable. First, if violating Relation was determinant for this phenomenon, any irrelevant predicate should invite similar inferences to those triggered by (1–2). But this is not the case, as illustrated by (3) (“A postman, arrested for hiding 187kg of hashish under his son’s cradle”). After all, the perpetrator’s occupation is no more relevant than their nationality. A defender of the implicature view would be hard-pressed to account for the difference between equally irrelevant predicates. As we argue in Section 6, we think that these inferences rely largely on world knowledge, and not just on the status of these different predicates *vis-à-vis* the Gricean maxims.

Moreover, from the point of view of a racist, the idea that mentioning the nationality or origin of the perpetrators in (1–2) violates the Maxim of Relation is questionable. If you are a racist, you will likely think that the nationality or origin of the perpetrator is essential to understand what happened. Put differently, thinking that (1–2) violate the Maxim of Relation requires the antecedent (presumably anti-racist) belief that nationality/origin plays no explanatory role. Without such belief, the premise that the literal contents of (1–2) are conversationally devious is missing, and the implicature calculation cannot take off. Presumably however, the RIs triggered by (1–2) do not require that the audience is anti-racist; such inferences can be drawn by any audience, racist or anti-racist.¹³

Secondly, there is an important contrast between CIs and RIs when they are cancelled in a subsequent utterance. When CIs are cancelled, the speaker appears *uncooperative*. Let us return to Grice’s example (14). Imagine that *a* wants to confirm that the garage sells gasoline. She asks, “Do they sell petrol there?” and *b* answers “No, they don’t”. Clearly, *b* will appear uncooperative. The reason for this is that, without the implicature, *b*’s original contribution seems completely irrelevant to the conversation.

This does not happen with (1–2). Imagine that, after publishing (1), the NYT is accused of racism—e.g., on Twitter. Such accusations of racism might rely on the RI triggered by (1) about Iraqi refugees. The editor of NYT could post on Twitter “The NYT didn’t say/mean that”/“Iraqi refugees are wonderful people”.¹⁴ Setting aside the awkwardness of such

13 One way to test this would be to ask: for a racist who thinks that mentioning nationality/origin is crucial in all cases, would sentences like (1–2) cease to invite a racist inference? We think not.

14 Note that there is an important difference between cancelling an implicature and disavowing it. To cancel an implicature is to deny its content; to disavow it is to deny having intended to communicate it. While cancelling an implicature generally makes the speaker appear uncooperative, disavowing an implicature is acceptable in some contexts. For example, in situations where the speaker is ignorant about the implicated content, disavow-

an apology, these responses would cancel the RI of (1). And crucially, they would not make the editors of NYT look uncooperative for having published (1). However, if (1) was to be described as a CI, one would expect that, when the implicature is cancelled, the speaker should appear uncooperative and the original contribution should seem conversationally defective. Importantly, that does not mean that, if the editors of NYT cancelled such an inference, they would not come across as *hypocrites*. They probably would (especially given the existence of style-books that discourage these practices, see n. 12). But being a hypocrite is not the same as being conversationally uncooperative.¹⁵

In sum: when their inference is subsequently cancelled, (1–2) do not become conversationally defective, and speakers do not seem uncooperative. This suggests that (1–2) do not violate any conversational maxim, and thus that the RIs are not rescuing an otherwise uncooperative contribution. These observations suggest that cases such as (1–2) cannot be described as conversational implicatures.¹⁶

6. RIs as insinuations

Up to now, we have discarded an account of the RIs of (1–2) as a type of slurring, ESTIs, dogwhistles and conversational implicature. In this section, we defend that RIs can be characterized as *insinuations* (Camp, 2018; Fricker, 2012), specifically triggered by the mechanism of *conversational eliciture* described by Cohen and Kehler (2021).

Insinuations are “off-record” pragmatic inferences triggered by certain sentences. For example, suppose that a real estate agent says the following to potential buyers from a different racial background than the local majority (Camp, 2018, 43):

(15) Perhaps you would feel more comfortable locating in a more... transitional neighbourhood, like Ashwood?

Where the buyers perceive the explicit suggestion to seek a house elsewhere, the implicit insinuation is that they should feel uncomfortable in this neighbourhood by virtue of not belonging to the majority group. Another example arises from a situation where a driver

ing it is perfectly fine. This is very clear in scalar implicatures: if a speaker doesn't know whether Mary ate all of the cookies, they can say “Mary ate some of the cookies, but I don't mean to imply that she didn't eat all of them. I don't know that”. In Section 6 we discuss further the contrast between cancelling and disavowing an inference.

- 15 To press this point: one can be a hypocrite without committing any sort of linguistic infelicity (by virtue of saying things one does not mean); and one can say all sorts of conversationally devious things without being a hypocrite (Grice's petrol case is a good example). We think that, in our cases, denying or disavowing these inferences might be hypocritical, but not conversationally devious. We thank an anonymous reviewer for pressing us to clarify this.
- 16 Saul (2002) proposes the notion of an *audience-implicature*, which are similar to CIs except for the fact that it is the audience, and not the speaker, who has authority over what is implicated. Perhaps the inferences triggered by (1–2) could be described as audience-implicatures, specifically drawn by part of the audience—the part of the audience that is not racist. That part of the audience could perceive (1–2) as violating Relation, independently of whether the speaker intended to convey the implicature. We leave this possibility open for future consideration.

is stopped for speeding. They might utter (16) in order to suggest a bribe (Lee and Pinker, 2010; Pinker, Nowak and Lee, 2008, cited in Camp, 2018):

(16) I'm in a bit of a hurry. Is there any way we can settle this right now?

Again, whereas the explicit message is an innocent question, the implicit insinuation is an invitation to accept a bribe. For a final example, if someone utters

(17) You know that Obama's middle name is Hussein. I'm just saying.
the hearer might infer that, e.g., that Obama supports radical Islamism (Camp, 2018, 43).

An insinuation therefore consists of two messages. On the one hand, a literal "on-record" content—e.g., that the buyers would be more comfortable elsewhere—that is unobjectionable. On the other hand, these sentences carry an "off-record" message, e.g., that the buyers would be uncomfortable because they belong to a different racial group. The "off-record" content is such that the speaker cannot be blamed for having conveyed it. As a result, the literal content of these sentences is apparently innocuous, but they manage to subserve a controversial conversational point.

According to Camp, the main property that characterizes insinuations is that speakers can convey them without being held accountable for it. More specifically, Fricker (2012) claims that a distinctive feature of insinuations is that they are *cancellable* (Fricker says "deniable") and *disavowable*: the speaker can cancel their content as well as deny having had the intention to convey that content without appearing uncooperative.¹⁷ For instance, if the speaker of (17) is confronted with "Are you suggesting that Barack Obama is a radical Islamist?", they can answer "Of course not, he's not a radical Islamist!"—cancellation—or "No, I didn't say/mean that!"—disavowal. In this respect, insinuations behave differently from conversational implicature.¹⁸

We submit that (1–2) can be described as a type of insinuation. On the one hand, (1–2) convey the "on-record" content that an individual of a certain background committed a crime. On the other hand, (1–2) convey the "off-record" content that individuals from that background are despicable/dangerous/malicious.

An immediate advantage of this hypothesis is that it bears out the prediction that RIs can be cancelled and/or disavowed by the speaker without appearing uncooperative. Consider the following dialogues as continuations of (1–2):

¹⁷ Torices (2021) shows that dogwhistles also have this property.

¹⁸ It bears pointing out that one of Camp's chief examples of insinuation is also a classic example of Gricean implicature: a letter of recommendation stating "Mr. X's command of English is excellent, and his attendance at tutorials has been regular", which invites the inference that Mr. X is not a promising philosopher. Camp is not particularly invested in making a sharp distinction between implicature and insinuation, but we wish to distinguish them by the property that cancellation implies lack of cooperation. *Per* this test, Grice's example would count as an implicature, since cancelling the inference would certainly leave the addressee of the letter wondering why the writer didn't say that Mr. X was a promising philosopher in the first place.

(1)

a. (An) Iraqi refugee was convicted in Germany because of raping and murdering teenage girl.

b. Are you suggesting that Iraqi refugees are despicable/dangerous/malicious?

a. No, Iraqi refugees are great people/ I didn't say/mean that!

(2)

a. (A) Moroccan was arrested for hiding 187 kg of hashish under his son's cradle.

b. Are you suggesting that Moroccans are despicable/dangerous/malicious?

a. No, Moroccans are great people/ I didn't say/mean that!

Thus, we consider (1–2) as insinuations which trigger in the audience certain associations of hateful ideas about a group. However, as illustrated by the previous examples, insinuations are a varied lot. We still need to point out the specific mechanism by which these inferences arise. The mechanism that underlies such associations is, we think, a form of what Cohen and Kehler (2021) call *conversational eliciture*. A conversational eliciture arises when, by choosing a particular predicate (among others), a speaker elicits inferences on the part of the audience that would not otherwise be drawn. This phenomenon can be illustrated by the following examples (Cohen and Kehler, 2021, 2-3):

(18)

a. The company fired the *manager who was embezzling money*. (Rohde et al. 2011)

b. *The drug-addled undergrad* fell off the Torrey Pines cliffs.

c. *A jogger* was hit by a car in Fresno last night.

On top of conveying their literal content, these sentences invite the inference that the embezzlement was the *reason* for the firing in (18a), that drugs *caused* the undergrad to fall off the cliffs (18b) and that the victim of the accident was jogging *when* the accident happened (18c). Note that the inferences in (18a–c) are cancellable. For instance, (18a) can have the following continuation:

(18a)

a. The company fired the manager who was embezzling money.

b. So she was fired because she was embezzling money?

a. No, she was fired because she's always late!

We think that the inferences associated with (1–2) can be described as insinuations that rely upon this very same mechanism: by linking a group with dangerous actions, they strongly invite the hateful inference that Iraqi refugees/Moroccans typically cause these actions and, thereby, that they are despicable/dangerous/malicious.

This hypothesis is consistent with the properties of RIs that we have observed throughout this paper: first, we observed that the RIs of (1–2) do not project out of entailment-cancelling environments in the same way as the derogatory inferences of slurs do. Conversational

elicitures do not have the projective profile of presuppositions, and therefore, they are not expected to project out of entailment-cancelling environments.

Secondly, in contrast to ESTIs, we observed that the sentences that trigger RIs entail that the predicate applies to the individual in subject position: (1) entails that the murderer is an Iraqi refugee; (2) entails that the subject is Moroccan. Conversational elicitures also have this property: (18a) entails that the manager was embezzling money; (18b) entails that the undergrad was on drugs; (18c) entails that the victim was a jogger.

Thirdly, we saw that, in contrast with dogwhistles, the RIs of (1–2) persist despite being recognized by the audience: if the speaker of (1) is accused of inviting the inference that Iraqi refugees are despicable/dangerous/malicious, the inference will not thereby disappear or become weaker. Conversational elicitures also persist after being recognized by the audience: if a hearer accuses the speaker of (16a) of suggesting that the embezzlement caused the firing, that inference would not vanish.

Fourthly, as we have just seen, the RIs triggered by (1–2) are cancellable and disavowable, and so are conversational elicitures. Moreover, conversational elicitures do not rely on the violation of any conversational maxim, nor are they meant to rescue an otherwise uncooperative contribution.

We still need to account for examples like (3) and (4) (repeated here).

- (3) Un cartero, detenido por esconder 187 kilos de hachís bajo la cuna de su hijo.
A postman, arrested for hiding 187kg of hashish under his son's cradle.
- (4) Cuatro vecinos gitanos evitan que una mujer sufra una violación en plena calle. Four
Romani neighbours saved a woman from being raped in the middle of the street.

As we mentioned in the introduction, (3) does not invite an inference against postmen, and (4) invites the hateful inference that these four people surprisingly prevented, rather than committed, a crime.

Cohen and Kehler (2021) observe that there are sentences with similar structure to (18) that do not trigger the same inferences. For example, consider the following two alternatives to (18a):

- (19)
 - a. The company fired the manager *who was hired in 2002*.
 - b. The company fired the manager *who has a long history of corporate awards*.

(19a) does not offer an explanation for the firing, and therefore does not invite the conversational eliciture that being hired in 2002 caused the firing. By contrast, (19b) gives a positive description of the manager, inviting the inference that the firing is counter to expectations. The reasons for these effects are extralinguistic. Embezzling money is an illegal activity and, as such, it can of course cause someone to be fired from their job. No such thing is true of the property of *being hired in 2002* (although of course we can cook up an appropriate context: suppose that the company is firing its most veteran employees). Similarly, having a long history of corporate awards is a good thing, and thus, it goes against expectations that a company should fire someone like that.

We think that a similar explanation can capture, respectively, examples (3) and (4). First, consider (3). The nationalities and backgrounds of individuals are connected to racist attitudes, prejudices and stereotypes. By contrast, their occupation does not, in general, have this property. Thereby, “postman” in (3) does not offer an explanation for the arrest. Importantly, this is an extra-linguistic consideration. We can of course imagine a context in which there exists a prejudice against postmen. In that context, (3) would invite a similar inference to (1–2). Thus, we submit that sentences like (3) are like (19a) in the following sense: given world knowledge, in most contexts they do not invite an inference that the choice of predicate is guided by an explanatory purpose.

Similarly, due to prevalent racist prejudice against the Romani community, the hateful inference of (4) is that, contrary to expectations, Romani people were the saviours and not the attackers. This inference is, again, due to world knowledge: racist attitudes against the Romani community abound, and in that context, for a speaker to report a situation in which members of this community prevented rather than committed a crime, invites the inference that this is surprising. The situation is parallel to (19b).

7. How to respond? The provocative dimension of RIs

We finish this paper by arguing that the insinuations triggered by (1–2) are *provocative*, because they place the audience in a frustrating position. (1–2) provoke a variety of audience responses, none of which are ultimately successful in blocking their effects. The audience is thereby forced to *succumb* to the provocative dimension of RIs. Audience responses¹⁹ to these inferences can be categorized as follows:

- 1) *Ignore the RI*. Faced with (1–2), the audience might simply ignore the hateful inference. This strategy, however, has a risk: if it is not confronted, the hateful inference gets incorporated into the *common ground* of the conversation (Lewis, 1979; Stalnaker, 2002). If this happens, the common ground will contain racist attitudes about Iraqi refugees/Moroccans.²⁰
- 2) *Confront the RI*. Alternatively, the implicit content of (1–2) might be confronted or challenged²¹ explicitly by the audience. The audience could accuse the speaker of having insinuated the implicit content, for example saying “Are you suggesting that Iraqi refugees are despicable/dangerous/malicious?”.

However, the speaker can reply to such a confrontation by different means. Regarding (1), for example, we saw that the speaker can cancel or disavow the inference (“No, Iraqi refugees are great people!”/“I didn’t say that!”/“I’m just stating facts!”). But the speaker can also reply “You said it, not me! But now that you mention it...” (Camp, 2018, 46). This last type of responses can make the RI turn into the topic of the conversation. This would

19 Camp (2018) proposes that the hearer might respond to insinuations through *pedantry*, whereby the hearer refuses to pick up the implicit content despite having recognized it.

20 Moreover, from the point of view of argumentation theory (Lewiński and Mohammed, 2016), if these inferences get incorporated into the common ground, they might be used as arguments from example in support of racists views. They have argumentative potential (Mohammed, 2019a, 2019b) for a racist speaker. Such speaker can use (1–2) to defend their standpoint saying things like “Those refugees are dangerous, don’t you see what happened in Germany?”.

21 On different types of challenging responses, see Lance and Kukla (2013).

shift the agenda of the conversation. Participants abandon the flow of the dialogue to discuss whether or not there is a hidden inference under the utterance, or whether or not the inference is true. Even further, the content of the inference might *gain credibility* simply by coming under discussion.

In sum, any response that targets the RI explicitly seems *inefficient* at blocking the inference and, moreover, risks turning the conversation into a racist one, thereby serving the purposes of the insinuator.

We want to propose, in closing, that a more effective response might be to “counter-attack” the RI with similar inferences. Consider the following replies to (1):

- (20) a. And another Iraqi refugee was selected to play with the Portuguese national team. So what?
 b. And a German was convicted of the same crime a week ago. So what?

(20a–b) are counter-attacks to the provocative insinuation in (1). These responses also invite *conversational elicitures*: (20a) invites the inference that Iraqi refugees are good football players; (20b) invites the inference that Germans are despicable/dangerous/malicious. Put differently, (20a–b) are attempts to undermine the RI of (1) using the same mechanism. If the inference of (20a) is accepted, or even if it becomes the topic of conversation and speakers start to discuss whether Iraqi refugees are good at football, the hateful effect of the original RI will vanish, or at least it will be “diluted” in combination with other, similar (but not hateful) inferences. Alternatively, if the inference from (20b) is accepted, or if it becomes the topic of the conversation, the hateful effect of (1) would be “compensated” by the hateful effect of (20b). The former strategy is perhaps more desirable, since it does not involve *more* hate speech, but both strategies might seem effective at neutralizing the original inference. On the other hand, note that (20a) is structurally similar to (4) (repeated here), which we observed gives rise to a hateful inference as well.

- (4) Cuatro vecinos gitanos evitan que una mujer sufra una violación en plena calle.
 Four Romani neighbours save a woman from being raped in the middle of the street.

This seems a consideration against saying things like (20a). However, there is a crucial difference between uttering sentences like (20a) or (4) *out of the blue* and uttering them as responses to sentences like (1): in the former case, they may contribute to hate discourse “from scratch”, by triggering the “counter-to-expectations” inferences that we observed; but in the latter case they may succeed at diluting the effects of a previous hateful utterance.

Before wrapping up, one might object to our counter-attack strategy as follows: failing to make explicit the harmful inferences associated with (1–2) can have a normalizing effect, effectively reinforcing the racist prejudices and biased ideology at play in these examples. We think that this depends, to some extent, on whether the conversation involves only two (speaker and addressee) or more than two parties. In the former case, we think that the risks considered above overcome possible benefits of confronting the RIs, especially if the speaker is a racist. But we acknowledge that, in the latter case, making explicit the RIs of

sentences like (1–2) to a general audience might indeed be more beneficial than harmful. This can help other members of the audience spot a type of hate speech that would otherwise remain hidden.²²

This is a mere sketch of different strategies that a speaker who wants to counter hateful inferences of this kind can deploy, and which to our knowledge have not been explored in any great detail in the literature. But this merits more careful attention than we have space for here. We leave this discussion for future work.

Conclusion

In this paper we have defended that headlines like (1–2) do not feature “Iraqi refugee”/“Moroccan” arbitrarily. Rather, they invite racist inferences against the relevant groups. We considered and rejected various possible accounts of this phenomenon, namely the hypothesis that these inferences are a type of *slurring*, that these inferences fall under ethnic or social terms used as insults (ESTIs), that they are understood as dogwhistles, and that they are a kind of conversational implicature.

In contrast to those accounts, we defended that the inferences of (1–2) can be described as insinuations, because they appear to be “off-record”, and thus cancellable and disavowable. Moreover, we argued that the RIs of (1–2) are produced via the mechanism of *conversational eliciture*—by choosing a particular predicate over others, the speaker elicits the inference that the property denoted by the predicate has an explanatory role. By choosing to mention that the perpetrator is an Iraqi refugee/Moroccan, the speaker invites the inference that *being an Iraqi refugee/being Moroccan* explains why they committed those crimes.

Finally, we pointed out that these insinuations have a *provocative* character, because they place the hearer into a frustrating position: any type of response that targets the RI seems inefficient at blocking its hateful effects. We tentatively proposed that the audience may seek a response that *counter-attacks* the RI with similar inferences, so that the hateful effects of the original utterance are diluted or somehow compensated.

References

- Bolinger, R. J. (2017), “The pragmatics of slurs”, *Noûs*, 51(3), 439–462.
- Camp, E. (2013), “Slurring perspectives”. *Analytic Philosophy*, 54(3), 330–349.
- Camp, E. (2018), “Insinuation, common ground, and the conversational record”, in D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 40–66.
- Carrera, M. (2017), “Enfocando a los “digitales de odio””, https://www.eldiario.es/catalunya/opinions/enfocando-digitales-odio_132_3518066.html.
- Castroviejo, E., Fraser, K. & Vicente, A. (2020), “More on pejorative language: Insults that go beyond their extension”, *Synthese*, DOI: 10.1007/s11229-020-02624-0.
- Cepollaro, B. & Stojanovic, I. (2016), “Hybrid evaluatives: In defense of a presuppositional account”, *Grazer Philosophische Studien*, 93(3), 458–488.

²² We thank an anonymous reviewer for pointing this out.

- Cepollaro, B., Sulpizio, S. & Bianchi, C. (2019), “How bad is it to report a slur? An empirical investigation”, *Journal of Pragmatics*, 146, 32–42.
- Cohen, J. & Kehler, A. (2021), “Conversational eliciture”, *Philosopher's Imprint*, 21(12), 1–26.
- Colegio de Periodistas de Cataluña (2020), *Manual de estilo sobre minorías étnicas*, <https://aulaintercultural.org/wp-content/uploads/2014/04/manualdeestilo.pdf>.
- Consejo Audiovisual de Andalucía (2020), “Recomendaciones del Consejo Audiovisual de Andalucía sobre el tratamiento de la Inmigración en los medios audiovisuales”, http://www.consejoaudiovisualdeandalucia.es/sites/default/files/recomendaciones/Recomendaciones_2006_02_Tratamiento_inmigracion.pdf.
- Corredor, C. (2014), “Pejoratives and social interaction”, in P. Stalmaszczyk (ed.), *Issues in Philosophy of Language and Linguistics*, Łódź: Wydawnictwa Uniwersytetu Łódzkiego, pp. 39–55.
- Croom, A. M. (2011), “Slurs”, *Language Sciences*, 33(3), 343–358.
- DiFranco, R. (2014), “Pejorative language”, *Internet Encyclopedia of Philosophy*.
- Falbo, A. (2021), “Slurs, neutral counterparts, and what you could have said”, *Analytic Philosophy*, DOI: 10.1111/phib.12217.
- Fricker, E. (2012), “Stating and insinuating”, *Proceedings of the Aristotelian Society Supplementary Volume*, 86(1), 61–94.
- Grice, P. (1989), *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.
- Hess, L. (2020a), “Practices of slur use”, *Grazer Philosophische Studien*, 97(1), 86–105.
- Hess, L. (2020b), “Slurs and expressive commitments”, *Acta Analytica*, 36(2), 263–290.
- Jeshion, R. (2013a), “Expressivism and the offensiveness of slurs”, *Philosophical Perspectives*, 27(1), 231–259.
- Jeshion, R. (2013b), “Slurs and stereotypes”, *Analytic Philosophy*, 54(3), 314–329.
- Kukla, R. (2018), “Slurs, interpellation, and ideology”, *Southern Journal of Philosophy*, 56, 7–32.
- Lance, M. & Kukla, R. (2013), “Leave the gun; take the cannoli! The pragmatic topography of second-person calls”, *Ethics*, 123(3), 456–478.
- Lee, J. J. & Pinker, S. (2010), “Rationales for indirect speech: The theory of the strategic speaker”, *Psychological Review*, 117(3), 785–807.
- Lewiński, M. & Mohammed, D. (2016), “Argumentation theory”, *The International Encyclopedia of Communication Theory and Philosophy*, DOI: 10.1002/9781118766804.wbiect198.
- Lewis, D. (1979), “Scorekeeping in a language game”, *Journal of Philosophical Logic*, 8(1), 339–359.
- Maitra, I. & McGowan, M. K. (2012), “Introduction and overview”, in I. Maitra & M. K. McGowan (eds.), *Speech and Harm: Controversies over Free Speech*, Oxford: Oxford University Press, pp 1–23.
- McGowan, M. K. (2012), ““Whites only” signs and racist hate speech: Verbal acts of racial discrimination”, in I. Maitra & M. K. McGowan (eds.), *Speech and Harm: Controversies over Free Speech*, Oxford: Oxford University Press, pp. 121–147.
- McGowan, M. K. (2019), *Just Words: On Speech and Hidden Harm*, Oxford: Oxford University Press.

- Mohammed, D. (2019a), “Managing argumentative potential in the networked public sphere: The anti-#MeToo manifesto as a case in point”, in B. Garssen, D. Godden, G. R. Mitchell & J. H. M. Wagemans (eds.), *Proceedings of the 9th Conference of the International Society for the Study of Argumentation*, Amsterdam: SIC SAT, pp. 813–822.
- Mohammed, D. (2019b), “Standing standpoints and argumentative associates: What is at stake in a public political argument?”, *Argumentation*, 33(3), 307–322.
- Nunberg, G. (2018), “The social life of slurs”, in D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 237–295.
- Pellicer, M. (2015), “Periodismo e inmigración: cómo deben informar los medios”, <https://miquelpellicer.com/2015/08/periodismo-inmigracion-como-deben-informar-los-medios/>.
- Pinker, S., Nowak, M. A. & Lee, J. J. (2008), “The logic of indirect speech”, *Proceedings of the National Academy of Sciences of the United States of America*, 105(3), 833–838.
- Popa-Wyatt, M. & Wyatt, J. L. (2018), “Slurs, roles and power”, *Philosophical Studies*, 175(11), 2879–2906.
- Saul, J. (2002), “Speaker meaning, what is said, and what is implicated”, *Noûs*, 36(2), 228–248.
- Saul, J. (2018), “Dogwhistles, political manipulation, and philosophy of language”, in D. Fogal, D. W. Harris & M. Moss (eds.), *New Work on Speech Acts*, Oxford: Oxford University Press, pp. 360–384.
- Stalnaker, R. (2002), “Common ground”, *Linguistics and Philosophy*, 25(5–6), 701–721.
- Statista Research Department (2021), “Number of German male suspects of rape and sexual assault Germany, by age group”, <https://www.statista.com/statistics/1101516/male-suspects-german-rape-sexual-assault-age-group-germany/>.
- Torices, J. R. (2021), “Understanding dogwhistle politics”, *Theoria: An International Journal for Theory, History and Foundations of Science*, DOI: 10.1387/theoria.22510.
- Waldron, J. (2012), *The Harm in Hate Speech*, Cambridge, MA: Harvard University Press.



Not all speakers are equal: harm and conversational standing*

No todos los hablantes son iguales: daño y posición conversacional

CLAUDIA PICAZO**

Abstract: McGowan has provided a linguistic mechanism that explains how speech can constitute harm. Her idea is that utterances routinely enact s-norms about what is permissible in a given context. My aim is to argue that these s-norms are sensitive to the conversational standing of the speaker. In particular, I claim that the strength of the norm enacted depends on the standing of the speaker. In some cases, the speaker might even lack the standing required to enact new s-norms.

Keywords: hate speech, conversational exercise, conversational standing, silencing, discursive injustice, sexism

Resumen: McGowan ha proporcionado un mecanismo lingüístico que explica cómo el lenguaje puede constituir daño. Su idea es que las preferencias de manera rutinaria establecen normas sobre qué está permitido en un contexto dado. Mi objetivo es argumentar que estas normas son sensibles a la posición conversacional del hablante. En concreto, sostengo que la fuerza de la norma establecida depende de la posición del hablante. En algunos casos, el hablante puede incluso carecer de la posición requerida para establecer nuevas normas.

Palabras clave: lenguaje de odio, ejercicio conversacional, posición conversacional, silenciamiento, injusticia discursiva, sexismo

Recibido: 31/05/2021. Aceptado: 01/07/2021.

* I am very grateful to María José Frápolli, Alba Moreno, Eduardo Pérez Navarro and Andrés Soria for discussions on McGowan's framework and their feedback on the ideas presented here. I also thank Carlos Jaén for comments on a previous draft.

** Juan de la Cierva postdoctoral fellow, Department of Philosophy I, Universidad de Granada. E-mail: claudia.picazo@gmail.com. Lines of research: communication, context-dependence, silencing. Publications: Picazo Jaque, C. (2019), "Are mental representations underdeterminacy-free?", *Synthese*, 196(2), 633–654, and Picazo, C. (2021), "Homophonic reports and gradual communication", *Pacific Philosophical Quarterly*, DOI: 10.1111/papq.12368.

1. Introduction

During the campaign for the Madrid elections in 2021, a Vox (a current Spanish far-right party) poster was displayed in a metro station. The poster read: “Un MENA 4.700 euros al mes. Tu abuela 426 euros de pensión al mes.”¹ Many people, as well as some political parties, took the message to spread racist ideas and to be a piece of hate speech. The poster was thought to be not only disturbing, but also dangerous. McGowan has provided a linguistic mechanism that explains how this kind of speech can constitute harm (McGowan, 2019). According to her, assertions and other speech acts are not only assertions and so on, but also conversational exercitives that routinely enact norms about what is permissible. Vox’s poster suggests that the government is spending too much money on unaccompanied migrant children, thus making it permissible or appropriate to complain about migration laws. Moreover, it might make it permissible to use de-humanizing terms to refer to migrants. It might also make it permissible to treat migrant children as second-class citizens, less worthy of respect than our grandmothers, at least in the conversation to which the poster is a contribution. McGowan’s idea is that these permissibility facts follow from the rules of conversation and the state of the conversational score. If the utterance were not rejected (which it was), it would make certain harmful behaviours permissible (in a given domain).²

My aim is to argue that the norms enacted by a conversational exercitive are sensitive to the conversational standing of the speaker. McGowan holds that conversational exercitives are not authoritative speech acts. Although I agree that they do not require the kind of official authority necessary to perform paradigmatic standard exercitives, I claim that the standing of the speaker can affect the strength of the norm enacted. Some speakers have more power than others over the normative landscape. In some cases, the speaker might even lack the standing required to enact new norms. I argue that this is what happens in some cases of silencing, and perhaps also in other ordinary exchanges. To see that the identity of the speaker, and thus its conversational standing, might affect her ability to enact permissibility facts, let us compare Vox’s action with what would happen if it was not a political party but an anonymous citizen who uttered those words. Imagine that this anonymous citizen reaches an audience as wide as that reached by the poster. The anonymous citizen is interviewed by a television program with the aim of having a sample of the political ideas of random people before the election, and uses the words quoted above. I take it that, even if the reporter did not reject the anonymous citizen’s utterance, this utterance would fall short of enacting at least some of the norms enacted by Vox’s action. Perhaps an anonymous citizen simply lacks the standing to make it permissible to treat migrant children as second-class citizens. Or, as I argue, she is only able to enact a weak norm. Thus, whether a strong or a weak norm is enacted does not only depend on the words uttered and the rules of conversation. It also matters who the speaker is.

1 Translation: “A MENA [Menor Extranjero No Acompañado – Unaccompanied Migrant Minor] 4,700 euros per month. Your grandmother 426 euros of pension per month.” MENA can perhaps be considered a slur. My impression is that it works here as a means to de-humanize the target. McGowan’s theory is not specifically about slurs and similar terms. It can be applied to racist or sexist remarks free of slurs.

2 Rejecting an utterance means rejecting as well the permissibility facts it enacts.

As I see it, the difference between authority and standing is a matter of degree. Consequently, it is not clear that standard and conversational exercitives are two separate categories. All exercitives enact norms about what is permissible or appropriate. Depending on the authority or standing of the speaker, these norms might survive the linguistic exchange or be of limited duration, and be stronger or weaker. Once it is admitted that conversational exercitives are sensitive to the standing of the speaker, they can be seen as a weaker variety of exercitives. Exercitives can be ordered along a scale, with standard exercitives at the top and weak conversational exercitives at the bottom.

The plan is the following. In section 2, I present McGowan's framework with more detail. In section 3, I argue via examples that how powerful norms are depends on who enacts them. Section 4 develops this claim and considers two potential objections.

2. McGowan on speech and harm

Speech can harm. This much is common ground. Words can be used, among other things, to enact discriminatory laws and to insult. In a series of papers and a book, McGowan has explored a mechanism by which speech can harm (McGowan, 2004, 2009, 2019).³ According to her, it can do so in two different ways: it can cause harm, and it can constitute harm. Someone's words cause harm when they change the beliefs of a group and, as a consequence, the group adopts some kind of harmful behaviour—e.g. the group discriminates or marginalizes certain people. Harm constitution is McGowan's technical term for a specific way of causing harm. Speech *constitutes* harm when it enacts norms that, when followed, bring about harm (McGowan, 2019, 23–25).

Some speech acts clearly constitute harm in this sense. Imagine the CEO of a company who declares that women are no longer eligible for work. This speech act enacts a norm that makes it impermissible to hire women. If the norm is followed, it will bring about discrimination. Thus, the CEO's speech act constitutes harm. In Austin's speech act theory, this kind of speech act is called an *exercitive*. Austin defines exercitives as a kind of speech act in which speakers exercise powers, rights or influence (Austin, 1962, Lecture XII).⁴ Exercitives enact facts about what is permissible (in a given domain), and they do so via the speaker's authority. In the example, the utterance enacts facts about whom it is permissible to hire (in the company). It manages to do so because the CEO has the authority to implement hiring policies for the company. Let us call this type of speech act a *standard exercitive*.

Together with standard exercitives, one's words can also constitute harm in more ordinary ways. For instance, sexist and racist remarks made by ordinary speakers under ordinary circumstances may also discriminate and subordinate. McGowan has identified a mechanism by which ordinary speech acts can constitute, and not merely cause, harm: *conversational exercitives*. Conversational exercitives enact permissibility facts for the conversation (McGowan, 2019, 27–49). The idea is that the rules governing conversations make it the case that speech acts routinely enact facts about what is permissible in the conversation. Most contributions to a conversation modify what is appropriate to do afterwards. It is in

3 Here I will follow McGowan (2019).

4 Austin's taxonomy has been criticized, but it is still used in debates about harmful speech.

this sense that they are exercitives. Some utterances, such as sexist remarks, may make it appropriate to discriminate or subordinate women. Such utterances may constitute harm.

McGowan follows here Lewis' view on conversational dynamics (Lewis, 1979). Lewis identifies some similarities between conversations and games such as baseball. To begin with, both are activities governed by rules. Moreover, just like baseball games have a score, conversations can be thought of as activities whose development is captured by a conversational score. The conversational score is more inclusive than a game's score. It doesn't keep track of who wins, but of everything that is relevant to the development of the conversation. This includes at least what has been said, the topic under discussion, presuppositions, and whatever is needed to resolve context-sensitivity (standards of accuracy, salient objects, etc.). By definition, the score keeps track of everything that is relevant to determine which moves are appropriate, or permissible, at a given point of the conversation.

Since the score determines which moves are permissible, any speech act that updates the score enacts new permissibility facts. Following McGowan, it enacts a norm that establishes that certain moves are permissible whereas others are impermissible. Here is an example. Suppose that in a conversation about dogs Mary says "My dog loves chewing on shoes". This utterance makes Mary's dog the most salient dog, and this is captured by the score. The utterance establishes new permissibility facts. For example, it is now conversationally permissible to refer to Mary's dog with the expression "the dog". Therefore, the utterance is a conversational exercitive.⁵

There is another way in which an utterance can modify what is permissible at a given point in the conversation. Unlike the rigid rules of games, conversations are governed by rules of accommodation. Rules of accommodation update the conversational score so that the actual moves of the participants count as appropriate. Suppose that Mary meets a stranger who is walking her dog. The stranger's dog starts chewing on her shoes, and Mary says "My dog also loves chewing on shoes". This utterance introduces the presupposition that Mary has a dog. This piece of information was not part of the score, but it is necessary for Mary's utterance to be appropriate. If the utterance is accepted, the score is updated with the presupposition that Mary has a dog (via a rule of accommodation) (McGowan, 2019, 36). Because of that, it enacts facts about what is permissible in the conversation. For example, it is impermissible to question the presupposition afterwards. Again, the utterance is a conversational exercitive.

Both standard and conversational exercitives enact norms establishing what is permissible, but there are important differences between them. Unlike what happens with standard exercitives, speakers' intentions and hearers' recognition do not seem to play a relevant role in conversational exercitives. A speaker who successfully performs a standard exercitive (e.g., the CEO that enacts a new hiring policy) has the intention to enact permissibility facts. However, one can perform a conversational exercitive without conscious intention to do so. The intention to contribute to the conversation can be enough (McGowan, 2019, 56). Another important difference concerns authority. According to McGowan, conversational

5 The utterance is also an assertion. In McGowan's framework, utterances can make multiple contributions at once. A speech act can be an assertion and, on top of that, enact a change to the conversational score, thus also counting as a conversational exercitive (McGowan, 2019, 45).

exercitives do not require authority either (McGowan, 2019, 66-67). Although I will go back to this point later, let me note here that the kind of example discussed above suggests that any participant in a conversation is capable of establishing new permissibility facts. By contrast, standard exercitives require authority—perhaps even official authority. Finally, there are differences in scope. Whereas standard exercitives enact norms that can still be in place after the conversation (e.g., hiring policies), the norms enacted by conversational exercitives are restricted to the conversation.

It is worth mentioning that conversational exercitives enact norms of a specific kind. According to McGowan, the norms governing a given activity are of two kinds (McGowan, 2019, 85–86). First, some norms, called *g*-norms, are global and govern all instances of the activity. For conversations, *g*-norms include rules of grammar and Grice's Cooperative Principle and maxims. As an example, think of Grice's Maxim of Quality. According to this maxim, one ought not to say what one believes to be false or unjustified. This norm governs all conversations (or, at least, all cooperative conversations). Second, some norms, called *s*-norms, are token-activity-specific. These norms are enacted during the performance of the activity-token and govern only that activity-token. For conversations these include, for instance, norms about the topic under discussion or what is salient. Think again of a conversation about dogs. In it, it is permissible to talk about dogs and impermissible to discuss other topics. Unlike Gricean maxims, this norm (an *s*-norm) does not apply to conversations in general. Conversational exercitives enact *s*-norms, that is, norms of limited duration that apply only to the particular conversation.

The notion of a conversational exercitive is a useful tool to analyse norm-governed activities beyond conversations, such as informal social interaction, dancing with a partner or playing improvisational jazz (McGowan, 2019, 83). In these activities, some moves count as appropriate or permissible and others count as inappropriate or impermissible. Here as in games and conversations we can think of a score determining which moves are appropriate. What is interesting about this generalization is that, as McGowan notes, an utterance can be a move both in a conversation and in another non-conversational activity. Moreover, it can, and often does, enact *s*-norms, or permissibility facts, in both activities.

This framework explains how ordinary utterances like sexist or racist remarks can constitute harm. Despite not being standard exercitives, they can enact *s*-norms that prescribe harm. Moreover, they can enact harmful norms both in the conversation and in the activity of social interaction (or a more specific activity, as interaction with work colleagues or interaction during a date). Here is one of the examples that McGowan discusses. Imagine the following exchange by two co-workers at the employee lounge (McGowan, 2019, 110):

Example 1:

John: So, Steve, how did it go last night?

Steve: I banged the bitch!

Steve's utterance is a contribution to two activities and enacts *s*-norms in both of them. As a contribution to the conversation, it is a conversational exercitive that makes it permissible to use derogatory terms for women. Besides this, the utterance is a contribution to broader social practices. Here McGowan considers the practice of oppression. As a contri-

bution to the practice of oppression, the utterance also enacts s-norms. The mechanism is the same as in conversations: the g-norms and the state of the score determine which moves are permissible at a given point. Although she acknowledges that we would need to know more about the g-norms of oppression and the score in order to specify which s-norms are enacted, McGowan's hypothesis is that Steve's utterance makes it permissible to degrade women (at the environment of the utterance, for a limited time).⁶ It makes women count as second-class citizens. This makes certain behaviours that would harm women permissible, such as undervaluing women's work. Thus, Steve's utterance can constitute harm: it enacts an s-norm such that, if it is followed, women will be harmed. Alternatively, we can see this kind of remark as a contribution to the activity of social interaction at the workplace.

McGowan's framework explains how a system of oppression can be activated in a local context. The speech act does not create the structural oppressions of our societies. Nonetheless, McGowan's point is that a more or less ordinary remark uttered by an ordinary speaker might "bring the latent force of that gender system to bear in the local context", as Simpson summarizes it (Simpson, 2013, 563).

3. Conversational exercitives and standing

In what follows I argue that the social identity or role of the speaker can affect the strength of the s-norms enacted, even depriving her of the ability to perform conversational exercitives at all. In my view, not all speech acts that can be said to enact harmful norms (including, but not limited to, uses of slurs) are equally harmful. Although McGowan herself does not explore the ways in which the identity or role of the speaker can increase or decrease the normative impact of a conversational exercitive, her framework can be easily modified so as to account for such differences.

Let us start by considering what McGowan says about standing (McGowan, 2019, 66–70). Although conversational exercitives are not authoritative speech acts, McGowan herself notes two ways in which the standing of the speaker is relevant. First, in order to be a participant in a conversation one must have a certain standing. For instance, in order to be recognized as a participant in a conversation about local politics one might need to have some knowledge about the topic. A foreigner might be excluded on the basis that she lacks the required knowledge. Second, expertise, and possibly other things, can play a role in the conversation and contribute to the participant's standing. If a speaker is recognized as an expert in a certain topic, then her contributions are probably believed and not only accepted for the purposes of the conversation.

I hold that something similar happens with harmful speech. The social identity of the speaker or the role she has in a certain context can contribute to her conversational standing. This, in turn, has consequences for the strength of the s-norms she enacts—how they interact with other norms, whether they also enact norms in the activity of social interaction—as well as on her ability to reject a conversational exercitive or even to enact new s-norms.

6 Although McGowan sees the difficulty in identifying specific s-norms as an epistemic problem (McGowan, 2019, 111), one could argue that this is a case of underdetermination. In many ordinary situations, the words uttered, and perhaps even the words uttered plus the score, might underdetermine exactly which norms are enacted. The situation would be analogous to so-called semantic underdeterminacy (Picazo, forthcoming).

Moreover, this explains why some speakers' words have more harmful consequences than others', and why some speakers seem to be communicatively impaired as a result of other speakers' behaviour.

I will add a second example before analysing how the standing of the speaker affects the strength of the enacted s-norms.

Example 2:

The community of the Capital Slam Poetry contest has always been open, constructive, and free of misogynistic speech. However, one day a poet goes on stage and starts talking about women with venom, saying things like "fucking feminists". (McGowan, 2019, 117)

Classifying the speech acts in Examples 1–2 as conversational exercitives allows us to explain why they are harmful. As I said in the previous section, Steve's utterance in Example 1 might make it permissible to treat women as inferior at the work place—to slur them in conversation and perhaps also to undervalue their work. Similarly, the poet's misogynistic speech in Example 2 might make it permissible to rank women as inferior in the contest.

The way in which McGowan introduces conversational exercitives suggests that the enacted norms are equally powerful regardless of who the speaker is. However, I think that it is important to note that the standing of the speaker has consequences at the local contexts in which s-norms are enacted. Although in cases of oppression we can say that the main work is being done by the general rules of oppression (by the activity of oppression to which the harmful utterance is a contribution), the general rule needs nonetheless to be activated via an s-norm. My point is that the normative landscape that emerges at the local context where the s-norm is enacted is sensitive to the speaker's identity.

To see this, consider two versions of Example 1. Both happen during the lunch break, where workmates and managers are chatting about the weekend. In the first version, Steve and John are young interns and, although their workmates treat them with respect, they do not in general attribute much credence to their contributions or pay special attention to their anecdotes. In the second version, Steve is the manager, and the workers pay special attention to what he says, even when it is unrelated to work. I deem it plausible to think that the same words uttered by the manager would have a bigger impact on the work environment than they would if uttered by the intern. The differences can be appreciated if we focus either on the conversational or non-conversational activity. Regarding conversation, it seems that the manager's utterance would make it more permissible to use derogatory terms for women. As for the non-conversational activity, the relevant activity could be that of social interaction at the workplace. I take it that, again, the manager's speech act is more likely to make it permissible to be disrespectful with women.⁷

⁷ I am assuming here that there is no equivalent norm already active. Otherwise, it is not clear what the effect of the utterance would be. Compare with assertion. Following Stalnaker, we can understand the effect of an assertion as an update on the context set (Stalnaker, 1978/1999). Imagine now that A asserts p and, minutes later, B asserts again p. Given that p was already common ground, it is not clear that this second assertion operates an update on the context set. Similarly, if the norm that Steve's utterance enacts is already enacted, it is not clear how the score would be updated or whether Steve's utterance should be regarded as an exercitive.

Similarly, we can image two versions of Example 2. In the first, the one described above, it is a young, unknown poet who starts using misogynistic speech. In the second, it is the organizer of the contest who goes on stage and reads a misogynistic poem. Again, I think that the illocutionary effect is different whether it is the unknown poet or the organizer who reads the poem. For instance, in the first version the audience could be puzzled by the fact that the unknown poet was allowed to read such a poem, but be hesitant about whether hate speech was permissible. By contrast, in the second version it makes more sense for the audience to assume that hate speech is allowed—after all, it is the organizer himself who is using hate speech. It seems that who the speaker is—her standing—is relevant to how powerful the enacted norm is.

I capture this through the idea that norms can be stronger or weaker, that is, that they can have different degrees of strength. We can find examples of norms with different degrees of strength in the activity of social interaction. For instance, there is a norm that makes it impermissible to make bad comments about the physical aspect of our interlocutors, such as making a remark that they have a blemish on their nose or that they have gained weight. This norm is very strong when it comes to strangers and probably weak when it comes to close friends or family. This might happen also with orders or standard exercitives. When a mother orders her child to do his homework, she thereby enacts a strong norm. If it is an elder brother who utters the same words, we can say that he also enacts a norm, but that this norm is somewhat weaker. My view is that something similar can be said about conversational exercitives and s-norms. The s-norm that it is permissible to use derogatory terms for women is strong when it is enacted by the manager but weaker when enacted by the intern. This does not mean that only managers, or those with official authority, can perform conversational exercitives, but only that not all conversational exercitives are alike. In a conversation among a group of friends who see each other as conversational peers, perhaps all utterances are conversational exercitives. However, this might not always be the case.

Examples 1–2 show that s-norms can be more or less powerful depending on who enacts them. They also suggest that some speakers might even be unable to enact certain s-norms, despite of being recognized as participants in the conversation. A plausible explanation why the unknown poet's move fails to clearly have the same consequences as the organizers is that the first, but not the second, falls short of enacting a sexist s-norm. Although this conclusion can be disputed for Examples 1–2, there are more radical cases. In what follows I propose to interpret the phenomenon of silencing⁸ as showing that some speakers (such as women) are sometimes unable to enact s-norms.

McGowan argues that her framework provides a mechanism by which pornography (or, better, uses of it) can silence the speech of women. Consider a third example:

Example 3:

A man and a woman are on a first date and he starts a pornographic film in which the actress says “No” as a way to sexually excite the actor. Afterwards, the woman says “No” with the intention to refuse sex, but the man doesn't take her to seriously intend to refuse. Even though she says things like “No! I mean it!” he thinks she is

8 See Langton (1993).

doing that as a way to sexually excite him. (This example is based on examples by McGowan (2019) and Langton (1993)).

The woman in the example is communicatively disabled (Maitra, 2009). Although she intends to refuse, and utters the words that are typically used to refuse, she is not recognized as being serious in her refusal, that is, as really intending to refuse. She is silenced. Silencing is a term of art. A silenced speaker is not unable to utter words, but she is nonetheless unable to be recognized as seriously performing the speech act that she intends to perform.⁹

In the example, we can think of the pornographic film as enacting an s-norm that prescribes silencing. This norm makes it permissible to interpret a woman saying “No” as not being serious, and it makes it impermissible to interpret her words as a refusal. Or, to put it otherwise, the norm establishes that “No” doesn’t count as a refusal.¹⁰ This is McGowan’s explanation of the case. However, this cannot be the whole story. What remains to be explained is why the woman is unable to enact a new s-norm that overrides the silencing norm, that is, why is it that (the use of) the pornographic film is able to enact s-norms but the woman who says things like “I mean it” is unable to reverse them. My hypothesis is that, because of her social identity, the woman’s conversational standing is insufficient to perform certain conversational exercitives. In the situation described, the pornographic film is more authoritative than the woman. Whereas the film is recognized by the man as enacting an s-norm about how to interpret “No”, the woman is not equally recognized as rejecting that norm or as enacting a new s-norm to the effect that her use of “No” is a refusal.

In their analysis of the mechanisms by which pornography can silence women, Langton and West draw a useful parallelism with Lewis’ example of a conversation between a master and a slave (Langton & West, 1999). In a language game between a master and a slave the difference in power results in a difference in the speech acts they can correctly perform. When the master says things like “It’s now permissible for you to do x” he thereby enacts a new permissibility fact. This move is correct. But the slave cannot do the same. He lacks the authority to modify what is permissible. According to Langton and West, the speech of women might in some contexts resemble that of the slave—although they are able to utter words, they lack the power to perform certain speech acts they ought to be able to perform. For instance, they might be unable to refuse.

This difference in power is what is lacking in McGowan’s analysis. In the silencing cases the silenced speaker is not on a par with the other participants. She is not recognized as a conversational peer. We can interpret Example 3 in two different ways. First, the

9 There are other forms of silencing. The audience might also fail to recognize that the speaker has the required authority, that she is sincere and that she is expressing her true feelings. See McGowan (2019) and Caponetto (2021).

10 It could be objected that Example 3 describes a defective context, that is, a context in which the interlocutors disagree about the state of the score. The man takes the score to include the silencing norm whereas the woman takes the score to include a norm according to which utterances of “No” are serious refusals. I think we should seriously consider the possibility that some situations are such that they involve multiple scores. However, it is also possible to imagine cases involving several participants in which it is only the silenced speaker who disagrees about which norms are active at a given point in the conversation. See Mikkola (2011) for such an example. Moreover, Example 3 can be re-described so that the woman is aware that the silencing norm is still active after her utterance.

power imbalance can result in the woman being unable to reject the speech act that enacts the silencing norm. This is what happens if, right after watching the pornographic film, the woman says something like “Women never say ‘No’ meaning ‘Yes’” and her utterance is ignored. Second, we can consider that the woman is unable to enact new norms that allow her to make her communicative intentions manifest. Because of the time lapse between the enactment of the silencing norm by the film and the woman’s attempt to reverse the situation, I think that Example 3 fits better the second interpretation. When she says “I mean it”, the woman tries to enact a new positive s-norm to the effect that her saying “No” counts, in that context, as a serious and sincere refusal. This norm is in conflict with the one enacted by the film and therefore, if the woman’s norm is enacted, the other one will be overridden. Using McGowan’s framework plus conversational standings, we can see the woman’s move as an attempt to perform a conversational exercitive that is unsuccessful due to lack of conversational standing.¹¹

It is important to note that this inability to reverse the situation is a central feature of silencing. The problematic instances of silencing do not involve misinterpretations that are quickly corrected. The phenomenon is typically persistent. In this sense, imagine a version of Example 3 in which after saying “No” the woman adds “I mean it! Stop!”, and the man stops and apologizes. Although we can imagine that the man took “No” not to be a refusal (even if only for a few seconds), he soon realizes that the s-norm enacted by the pornographic film is no longer valid (if it ever was). In this case one could perhaps say that the woman is silenced for a short moment, but since the communicative disablement is not persistent I take it to be a borderline case. The more problematic instances of silencing seem to involve some kind of persistence.

The conclusion is that the identity of the speaker matters both to the strength of the norm enacted and to the ability to perform conversational exercitives. We can capture this through the notion of *conversational standing*. The conversational standings of the participants reflect the ranking among them for the purposes of the conversation.¹² It reflects the fact that some (probably many) conversations involve a hierarchy. This hierarchy can be inherited from social hierarchies, and certainly social identities and roles play an important part, but I leave it open how the conversational standing of a participant is determined and on what exactly it depends. For instance, the conversational standing could be affected by the participant’s performance in past conversations. If she has proved to be trustworthy, that might grant the speaker a high standing regardless of her social identity. Moreover, it is interesting to note that one’s conversational standing might vary across conversations. For instance, being a woman might be relevant in some conversations where the participants

11 This conception of silencing is similar to Ayala’s view on discursive injustice (Ayala, 2016, 2018). Ayala provides a structural approach in which a speaker’s speech capacity depends on her position in a social structure. McGowan’s framework can be seen as an analysis of a specific speech-related capacity—namely, the ability to enact permissibility facts. Here I show that a speaker’s social position can affect her speech capacity in two ways: it can prevent her from making certain moves (as in the example of silencing), but it can also grant the speaker the power to enact harmful norms. My view is also similar to Kukla’s (2014). According to Kukla, the members of disadvantaged groups are unable to perform certain speech acts. Here I hold that the same is true of conversational exercitives.

12 Conversational standing is similar to what Popa-Wyatt and Wyatt call ‘discourse role’ (Popa-Wyatt & Wyatt, 2018).

have sexist attitudes, but not in others where sexist beliefs and stereotypes play no significant role. Also, one can be an expert in one topic but not in others and thus have a higher standing in conversations about the first topic and a lower standing in conversations about the second. Anyway, we can assume that who a speaker is (a man or a woman, the manager or the intern, an unknown poet or the organizer of a contest, someone who has proved to be trustworthy or a liar) affects the power she has over the conversational score and, consequently, over what is permissible or appropriate in the conversation and other activities the utterance might be a contribution to. In this sense, we can say that conversational exercitives are sensitive to the conversational standing of the performer. The next section explores how conversational standing has an impact on the s-norms enacted.

4. Conversational standing and s-norms

In order to make sense of the examples in the previous section, we can think of the enacting of s-norms as dependent on conversational standing. My proposal is that we think of conversational exercitives as involving a threshold for standing and then allow for s-norms to have different degrees of strength, where the degree of strength of a particular s-norm depends again on the standing of the speaker who enacts it.¹³

Concerning the *threshold*, we have already seen that it takes a certain standing to be able to enact s-norms, i.e., to perform conversational exercitives (or some specific kind of conversational exercitive). The participant's conversational standing must meet a given threshold. The standing of the silenced speaker in Example 3 falls below the threshold. Silenced speakers are not recognized as conversational peers, and are unable to enact the s-norms that would allow them to make their communicative intentions available. Minimally, as McGowan notes, one has to be recognized as a participant in the conversation. But often something else is needed. One's utterances must also be recognized as contributions to the conversation, and not ignored, considered irrelevant or pointless. In order to perform conversational exercitives we need that the participants in the conversation are "willing and capable of hearing us" (Dotson, 2011, 238).

As far as *degrees of strength* are concerned, we have already seen that the strength of an enacted s-norm depends on the standing of the speaker. Some norms are stronger than others. To put it otherwise, some utterances make some actions more permissible than others. For instance, the organizer of the Slam Poetry contest's use of misogynistic speech makes it more permissible to use derogatory terms for women than the use of the same words by the unknown poet. In general, the higher the conversational standing of a participant, the stronger the s-norms she enacts will be.

A possible way to see how the strength of an s-norm matters is to explore how the norm interacts with other norms that are in place during the conversation or in the non-conversational activity. We can imagine a workplace in which usual courtesy norms apply in everyday conversations. According to these norms, it is impermissible to disrespect others, e.g., to insult them or use derogatory terms. Take now the first version of Example 1, in which Steve

¹³ I have argued in Picazo (2021) that communication should be understood as a gradual phenomenon. Here I extend this idea to conversational exercitives.

is an intern. Steve's subordinating s-norm is in conflict with courtesy norms. I think that it might be the case that the new s-norm is not strong enough to override those courtesy norms. All things considered, it is impermissible to use derogatory terms for women, even though there is a sense in which Steve's utterance might have made it permissible. However, things might be more complicated when it is the manager who utters the sexist remark. In that case, the s-norm might be strong enough to override the courtesy norm at stake or to take priority over norms of courtesy. All things considered, it might be permissible, after the manager's utterance, to use derogatory terms for women.

Something similar can be said about the non-conversational activity. Imagine now a workplace in which social interaction is governed by an egalitarian norm to the effect that all workers are to be treated in equal terms, all equally respected. In this workplace, the s-norm enacted by the intern might be insufficient to cancel this norm. But again, the manager's s-norm, being stronger, could result in non-egalitarian social interaction, with women counting as second-class workers.¹⁴ However, it will not be strong enough to override anti-discriminatory laws. Laws, that is, legal norms, are stronger than the kind of norms that are enacted in conversation. Not only is their duration longer and their scope wider, but they also take priority in case of conflict. A manager cannot enact an s-norm that overrides a law. What the manager can do is make certain actions socially acceptable.

Finally, once the standing is in the picture and it is admitted that conversational exercitives can enact stronger and weaker norms, one can see the relation between conversational and standard exercitives as a matter of degree. The more authority the speaker has in a given domain the stronger will be the norms she is able to enact. In the paradigmatic examples of standard exercitives, a special kind of authority will be required.

In what follows I discuss two potential objections to the view presented here. The first objection concerns the threshold. I have claimed that some speakers lack enough standing to perform conversational exercitives, that is, to enact s-norms. The example I used in support of this claim is one in which the speaker is silenced, using Langton's terminology. It could be objected that whether or not her interlocutors are aware of it, and regardless of their subsequent behaviour, the woman has enacted an s-norm. She has performed a conversational exercitive despite the fact that her interlocutor fails to recognize the new permissibility facts. This is because the enactment of norms depends only on the rules of conversation and not on the attitudes of the participants in the conversation. Given the rules of conversation, she has enacted an s-norm. This objection touches upon difficult questions regarding the notion of conversational score (and of common ground) and the determination of its state at a given point of the conversation. Two broad strategies can be followed here: one can, in the Gricean spirit, explain the score through the propositional attitudes of the participants, or one can, in a more Austinian tradition, think of communication as a conventional practice governed by rules.¹⁵ The objection assumes that the second strategy is preferable here, or at least that it should not be discarded. However, my point can be seen as concerning the way in which the theoretician can identify the rules of conversation. A reasonable way to

14 Being an s-norm, it would only apply to the token-specific activity. In this case, it would be difficult to identify when the activity begins and when it ends — does it last only for the lunch break? Is the conversation reassumed afterwards, or the day after, or during work?

15 See Stalnaker (2014).

find out whether all speech acts enact s-norms is to consider whether all speech acts are recognized as modifying what is permissible. If it turns out that some speech acts are not treated as changing the permissibility facts, then one reasonable explanation is that the rules of conversation are such that those speech acts do not enact s-norms. In the silencing cases, and in particular in Example 3, the woman's words are intended to make it impermissible to interpret an utterance of "No" as a non-serious refusal. However, they are treated as not enacting this s-norm but, rather, as re-enacting a previous s-norm to the effect that it is permissible to interpret utterances of "No" as not being genuine refusals. Thus, using this example as a way to discover the rules of conversation, there are reasons to think that the rules of conversation are such that some speakers cannot perform certain conversational exercitives (in certain contexts).

Still, it could be objected that it is not the case that the speaker lacks the standing to enact an s-norm but, rather, that the norm is enacted and then rejected. Generalizing this idea, one could hold that all speakers perform conversational exercitives but that some of them are automatically rejected. I think, however, that this would be an odd description of Example 3. In that example, the woman's words ("No! I mean it!") are not rejected, but rather reinterpreted as not being serious. It is not the same to misinterpret the force or the seriousness of a speech act as to reject it.

A more problematic case would be one in which the speaker is not silenced but her words nonetheless fail to enact new permissibility facts. I suggested that this could happen with the intern's utterance in Example 1 or with the unknown poet in Example 2.¹⁶ I think it would not be implausible to hold that some sexist remarks are not powerful enough to enact subordinating s-norms, and, in this sense, that some speakers lack the standing to perform conversational exercitives (in some contexts). The objection would apply to these cases more straightforwardly than in the cases of silencing. Here, it can make sense to interpret the example as involving an s-norm that is enacted and then rejected. However, it is not completely clear what the dynamics would be, given that, as Examples 1–2 are described, we can imagine that there is no explicit rejection and that what simply happens is that the other interlocutors do not see themselves as being allowed to use derogatory terms for women.

The second objection is that the thesis that s-norms can be more or less powerful, and that some speakers might even be unable to enact s-norms, rests on an illocutionary/perlocutionary conflation. Illocutionary acts are acts like asserting, making a request or warning. They concern the force of the utterance. Following Austin (1962), the performance of an illocutionary act is the performance of an act in saying something. Perlocutionary acts concern the effects of the illocutionary act, what we bring about by saying something. For instance, an illocutionary act can convince or frighten. One could object to the view that I have presented that what I identify as differences in degrees of permissibility or strength of norms is something that belongs to the perlocutionary act. The idea would be that there is a difference between the utterances of the intern and the manager, of the unknown poet and the organizer of the contest. Perhaps the manager and the organizer not only enact a

16 Alternatively, one can treat the utterance as enacting a weak s-norm, as I also do. I think there is no easy way to decide which option is preferable here. My impression is that there can be two kinds of cases. In some examples the s-norm is not enacted whereas in other it is enacted but is weak.

discriminatory or subordinatory norm, but also persuade their interlocutors to discriminate or subordinate women. This might very well be the case. Certain speakers, because of their position, are more persuasive than others. However, I think that McGowan's claims, and the intuitions her examples elicit, concern the permissibility facts enacted, not whether the audience is persuaded. I see no reason for treating the intuitions elicited by the two versions of Examples 1–2 sketched in section 3 differently. Thus I take it that, when the intuitions concern whether two utterances make it equally permissible to discriminate or subordinate, or whether one is more powerful than the other, then those intuitions concern the permissibility facts enacted, and not how persuasive the utterance is or other similar effects.

5. Conclusions

I have argued that not all speakers are able to enact the same permissibility facts. Whether some words enact a strong or a weak s-norm, or no s-norm at all, does not only depend on the words uttered and the rules of conversation. In general, a speaker's ability to enact s-norms, as well as the strength of the s-norms enacted, depends on her conversational standing. Differences in conversational standing can explain differences in discursive power. A high conversational standing allows the speaker to enact strong s-norm, whereas a low standing might only allow her to enact weak s-norms or even deprive her of the ability to perform the intended conversational exercitive. When it comes to hate speech, who the speaker is usually matters to how harmful her words are. Thus, we can see the difference between standard and conversational exercitives as a matter of degree.

References

- Austin, J. L. (1962), *How to Do Things with Words*, Oxford: Clarendon Press.
- Ayala, S. (2016), "Speech affordances: A structural take on how much we can do with our words", *European Journal of Philosophy*, 24(4), 879–891.
- Ayala, S. (2018), "A structural explanation of injustice in conversations: It's about norms", *Pacific Philosophical Quarterly*, 99(4), 726–748.
- Caponetto, L. (2021), "A comprehensive definition of illocutionary silencing", *Topoi*, 4(1), 191–202.
- Dotson, K. (2011), "Tracking epistemic violence, tracking practices of silencing", *Hypathia*, 26(2), 236–257.
- Kukla, R. (2014), "Performative force, convention, and discursive injustice", *Hypatia*, 29(2), 440–457.
- Langton, R. (1993), "Speech acts and unspeakable acts", *Philosophy and Public Affairs*, 22(4), 293–330.
- Langton, R. & West, C. (1999), "Scorekeeping in a pornographic language game", *Australian Journal of Philosophy*, 77(3), 303–319.
- Lewis, D. (1979), "Scorekeeping in a language game", *Journal of Philosophical Logic*, 8(1), 339–359.
- Maitra, I. (2009), "Silencing speech", *Canadian Journal of Philosophy*, 39(2), 309–338.

- McGowan, M. K. (2004), “Conversational exercitives: Something else we do with our words”, *Linguistics and Philosophy*, 27(1), 93–111.
- McGowan, M. K. (2009), “Oppressive speech”, *Australasian Journal of Philosophy*, 87(3), 389–407.
- McGowan, M. K. (2019), *Just Words: On Speech and Hidden Harm*, Oxford: Oxford University Press.
- Mikkola, M. (2011), “Illocution, silencing and the act of refusal”, *Pacific Philosophical Quarterly*, 92(3), 415–437.
- Picazo, C. (2021), “Homophonic reports and gradual communication”, *Pacific Philosophical Quarterly*, DOI: 10.1111/papq.12368.
- Picazo, C. (forthcoming), “Are utterance truth-conditions systematically determined?”, *Inquiry*.
- Popa-Wyatt, M. & Wyatt, J. (2018). “Slurs, roles and power”, *Philosophical Studies*, 175(11), 2879–2906.
- Simpson, R. M. (2013), “Un-ringing the bell: McGowan on oppressive speech and the asymmetric pliability of conversations”, *Australasian Journal of Philosophy*, 91(3), 555–575.
- Stalnaker, R. (1978/1999), “Assertion”, in *Context and Content*, Oxford: Oxford University Press.
- Stalnaker, R. (2014), *Context*, Oxford: Oxford University Press.



Exactly, what do you mean?*

Exactamente, ¿qué quieres decir?

MANUEL ALMAGRO**

NEFTALÍ VILLANUEVA***, ****

Abstract: The purpose of this paper is to explore the boundaries of a subset of the evaluative uses of language: offensive speech. More in particular, our goal is twofold. Firstly, we want to chart the relationship between context and evaluative utterances, as it stands in the current literature. Secondly, we focus on the experimental study of a particular interaction between contextual information and our evaluative claims –when the context is able to turn a seemingly descriptive utterance into an evaluative one. For this second purpose, we argue, certain recent positive proposals, in spite of their merit, come a bit short.

Keywords: Evaluative language, Offensive speech, Affective attitudes, Context-dependence, Expressive content.

Resumen: El propósito de este trabajo es explorar los límites de un subconjunto de los usos evaluativos del lenguaje: el discurso ofensivo. Nuestro objetivo es doble. Primero, introducimos la relación que hay entre el contexto y las preferencias evaluativas, tal y como puede rastrearse en la literatura reciente acerca de la cuestión. Segundo, nos centramos en el estudio experimental de una interacción particular entre la información contextual y nuestras afirmaciones evaluativas: cuándo el contexto es capaz de convertir una preferencia aparentemente descriptiva en una evaluativa. Para este segundo propósito, argumentamos, ciertas propuestas positivas recientes, a pesar de su mérito, son insuficientes.

Palabras Clave: Lenguaje evaluativo, Discurso ofensivo, Actitudes afectivas, Contexto-dependencia, Contenido expresivo

Recibido: 02/06/2021. Aceptado: 21/06/2021.

* This work was funded by the Spanish Ministry of Economy (Project FFI2016-80088-P, FPI Predoctoral Fellow BES-2017-079933), the Spanish Ministry of Science (PID2019-109764RB-I00), Junta de Andalucía (B-HUM-459-UGR18), and the University of Granada (FiloLab Excellence Unit).

** FPI Predoctoral Researcher, Department of Philosophy I, University of Granada (Spain). Email: malmagro@ugr.es. My work focuses on political polarization, mental state attribution and the evaluative use of language. Recent publications: Almagro, M., Navarro, L. & Pinedo, M. (forthcoming), «Is testimonial injustice epistemic? Let me count the ways», *Hypatia: A Journal of Feminist Philosophy*, and Almagro, N. & Villanueva, N. (2021), «Polarización y tecnologías de la información: radicales vs. extremistas», *Dilemata: International Journal of Applied Ethics*, 34, 51–69.

*** Professor and Head of the Department of Philosophy I, University of Granada (Spain). Email: nef@ugr.es. Most of my work focuses on the applications of the philosophy of language to classical questions in the history of philosophy and to political and social practical concerns. Recent publications: Almagro, N., Hannikainen, I. & Villanueva, N. (forthcoming), «Whose words hurt? Contextual determinants of offensive speech», *Personality and Social Psychology Bulletin*, and Almagro, M. & Osorio, J. (forthcoming), «Weaponized testimonial injustice», *Las Torres de Lucca*.

**** This is a wholly collaborative paper, so order of authorship is arbitrary.

1. Introduction

Despite the fact that you have the same sabbatical schedule as everyone else, he says, you are always on sabbatical. You are friends so you respond, easy.

What do you mean?

Exactly, what do you mean? (Rankine, 2014, 52)

At first sight, the previous exchange does not seem to count as one where the speaker is being offensive: the sentence “You are always on sabbatical” seems to aim at a *neutral* possible situation, which can be or not be the case. But as Rankine asks: exactly, what is the meaning communicated by the speaker through their claim in the above linguistic exchange? Does it really count as a simple description of how things are? Is it just a false description? The claim appears to mean something else. Is the speaker making an evaluation? Is it a joke? Is what has been said offensive in any way? Exactly, what does her friend mean? How can we know it? More context is needed to decide. But in what sense is the meaning we convey through our words context-sensitive?

The purpose of this paper is to explore the boundaries of a subset of the evaluative uses of language: offensive speech. More in particular, our goal is twofold. On the one hand, we want to chart the relationship between context and evaluative utterances, as it stands in the current literature. Forcefully, we will highlight certain recent findings, and say way less than it is necessary about other important phenomena. We will also focus, on the other hand, on the experimental study of a particular interaction between contextual information and our evaluative claims—when the context is able to turn a seemingly descriptive utterance into an evaluative one. For this second purpose, we will argue, certain recent positive proposals, in spite of their merit, come a bit short.

Evaluative uses of language seem to be importantly different from descriptive ones concerning the information they convey, despite the fact that the sentences commonly uttered to make claims from both language domains are structurally similar, even identical sometimes. We will highlight four different traits to showcase this difference: what is expressed, commonality of standards, retraction and practicality. First, descriptive claims seem to be embedded in practices where there is a high level of agreement in relation to how our surroundings are, while evaluative ones often exhibit lesser prior agreement on it. If I say that my room’s door is closed, I’m simply describing how things around me are in an objective way, so to speak. But if I say that my room’s door is *cool*, I’m also *expressing myself*, and my interlocutor might have a very different appreciation about the door (Cepollaro, Soria & Stojanovic, 2021) or, more importantly, about what a *cool door* amounts to. Second, descriptive and evaluative claims seem to give rise to very different situations of disagreement. If you and I disagree on whether the door of my room is closed, when the claim is just descriptive, then we seem to agree from the outset on how the dispute can be settled. In most scenarios, speaker and audience share common assumptions on how the claim could be verified. If they do not in fact share these assumptions, the situation stops making sense. If I suddenly discover that your preferred and only way to settle whether the door is closed is by asking the crystal skull that hangs from your necklace, we’re

not involved in a discussion about our surroundings anymore (perhaps in one about the shortcomings of your upbringing, or none at all). This happens far less in situations where the disagreement revolves around an evaluative claim (Field, 2009, 2018). It's possible to discover that your musical standards are radically different from mine, and still continue discussing the relative greatness of Gesualdo and Palestrina. Third, retraction—understood as a linguistic move aimed at reassessing, from the actual context, a claim that we took to be true in the past—seems to work differently in descriptive and evaluative contexts: we demand that someone retracts evaluative statements more often than we do descriptive ones, while, perhaps surprisingly, it seems to be much more difficult to retract the former than the latter (Bordonaba & Villanueva, forthcoming). Finally, and most importantly for us here, evaluative claims exhibit a closer connection to action. If I say that my room's door is cool, I'm presenting myself as someone from whom certain courses of action can be reasonably expected, while claiming that the door is closed is not particularly action-guiding when it just conveys information about our surroundings (Gibbard, 2012, 44).

Of these four features, perhaps retraction and practicality require a bit more attention, given the pivotal role that they have played in the literature. Evaluative utterances are often assumed to carry an action-guiding component under their sleeve (see, for instance, Williams, 1985, 140; Gibbard, 1990, 33; Finlay, 2004; Strandberg, 2012, 89; Charlow, 2014; Chrisman, 2017, 27)—they provide reasons for action, they are conceptually linked to ulterior courses of action. We do not check and discover that our evaluative utterances were false, as it might be the case with descriptive utterances. I'm not simply wrong when I fail to live up to the courses of action that can be reasonably expected from my evaluations, I might be deemed irrational, a hypocrite, or someone who lacks the necessary linguistic competence. This, in turn, affects my ability to effectively retract my previous evaluations. By expressing affective attitudes, by making evaluations, I'm often inviting my audience to adopt a view on how things should or should not be that can have practical implications for those living around us. Harm can be inflicted, prejudices can be reinforced, in a way that it's difficult to counter whenever I change my mind and want to take back what I said. This is not to say that nothing can be done about our previous evaluative utterances whenever we change our mind and feel the urge to say so, we just need to be aware that making things right might go beyond simply saying that we no longer think what we previously thought.

As we will see, much attention has been paid to the context-dependent dynamics of evaluative language uses, more specifically to the contextual factors that might determine that a given object is positively or negatively evaluated. That is, how the parameters that determine our appreciation of an object might change from person to person, from time to time, from situation to situation. Much less attention has been devoted, however, to the contextual cues that determine that a piece of speech is interpreted evaluatively, rather than as mere description of our surroundings—whether or not a piece of speech is able to express the kind of affective attitudes that might turn out to be offensive, to contribute to the discrimination of disenfranchised groups. Only evaluative uses of language can turn out to be offensive, and only offensive pieces of discourse can express hatred, even if not every evaluation is offensive, and not every offense expresses hatred. We will explore in this paper both the context-dependency of those parameters that might make an evaluation offensive, and those

that contribute to the evaluative—and potentially offensive—evaluation of a seemingly neutral utterance, where a descriptive interpretation is also available.

Amongst the recent studies that have tackled the question of whether or not a given piece of discourse is offensive, and why, the role of context has been brought to the fore, at least at two different levels. Firstly, status and speakers' motives have been pointed out as deciding factors as to whether a single piece of discourse is offensive or it's simply a neutral description of a given situation (Gibson et al., 2019). Secondly, the syntactic position of the words has been highlighted as one of the factors that might determine whether a word is interpreted as a slur. This line has given rise to unexpected consequences (Cepollaro, Sulpizio & Bianchi, 2019): slurs are perceived as more offensive than other kinds of insults when they are presented in isolation. However, expressions of the form "Someone is an X" turn out to be perceived to be *less offensive* when X is a slur than when it is a non-slurring insult. Not only that, the offensive character of certain expressions only diminishes in indirect speech, in expressions of the form "Bilal believes that Lois is X", but doesn't disappear. Why should *I* be deemed as offensive for reporting on the thoughts of somebody else? To complicate matters even further, it remains an open question whether quoted slurs are potentially harmful, and thus rightfully deemed offensive, or not (Cepollaro & Zeman, 2020; Moreno & Pérez-Navarro, manuscript).

The studies just mentioned show how to go beyond the truism that offensiveness is context-dependent. Claiming that whether our utterances are offensive or not depends on the context is by now less of a claim than it is a cue to provide more relevant information.¹ This being so, incidentally, can provide a reasonable basis to derive a conversational implicature out of those cases in which supposedly knowledgeable people make that very same claim, as if it was all the explanation that a certain action needs to be understood. Unlike the examples in the studies above, our goal in this paper is not to examine the behavior of terms or phrases that are taken to be offensive by default, but rather to ask about expressions which are not usually identified as offensive, those that might ignite the most stubborn conversations about the nature of language, the intentions of the speakers, and so forth. We will review some of the contextual features that might determine how and when an otherwise perfectly descriptive expression can be used to express affective attitudes, to contribute to the stereotypes that harm disenfranchised minorities. One of the relevant determinants of offensive speech, as we will see, seems to be the speaker's status. If we turn back to Rankine's example knowing that the speaker is a white person and the target person is not white, our initial intuitions regarding the meaning communicated will probably vary. We will also, in § 3, explain why we should care about these examples at all. Before any of that, more has to be said about the context-dependent nature of evaluative language (§ 2).

2. Offensiveness and context-dependence

A friend argues that Americans battle between the "historical self" and the "self self." By this she means you mostly interact as friends with mutual interest and, for the most part, compatible personalities; however, sometimes your historical selves,

¹ Cryptic? True. Potentially offensive? Also true.

her white self and your black self, or your white self and her black self, arrive with the full force of your American positioning. Then you are standing face-to-face in seconds that wipe the affable smiles right from your mouths. What did you say? Instantaneously your attachment seems fragile, tenuous, subject to any transgression of your historical self. And though your joined personal histories are supposed to save you from misunderstandings, they usually cause you to understand all too well what is meant. (Rankine, 2014, 20)

The purpose of this section is to review some contributions that emphasize, in different ways, the context-dependent nature of the evaluative. We will start by making some remarks about the context-dependence of evaluative judgments. Next, we will review some linguistic tests that, even though they are not exhaustive, point in the direction of the radical difference between the descriptive and the evaluative in terms of their context-dependence. This second type of context-dependency is slightly different: evaluative judgments are highly context-sensitive, but determining whether a judgment is descriptive or evaluative, a prior step, is also highly context-dependent. Finally, we will focus on the influence exerted by different contextual factors in the process of determining whether the meaning of a claim is descriptive or evaluative/offensive.

First, evaluative uses of language are context-dependent in at least two different senses. One of the characterizing features of evaluative judgments seems to be its gradable and multidimensional nature. As McNally and Stojanovic highlighted, when we evaluate an object, an event, a situation, a person, etc., the evaluation can come in degrees, and we might hinge on different features of the context constituting the “threshold of applicability” (McNally & Stojanovic, 2017, 21). Evaluative adjectives—adjectives whose use by default is evaluative—are gradable and multidimensional. For instance, when evaluating whether a natural enclave is beautiful or whether someone is healthy, we can say that it is *more* or *less* beautiful or healthy. Moreover, we can give consideration to the regularity of the terrain, the sort of vegetation, the color of the sky, the number of things found there, etc., or to the state of their cardiovascular system, immune system, habits, cholesterol, etc. Thus, the gradability of an expression requires that a parameter is contextually fixed, and the parameter might vary from one speaker to another, from one situation to the next. Its multidimensionality, for its part, is characterized by a contextually-sensitive list of features that determine the threshold, the bar that an object has to jump over, so to speak, in order to receive positive or negative valence. In this sense, the features that determine whether an evaluation is going to be positive or negative depend on the context.

Furthermore, the relative weight of each of the dimensions involved in setting the threshold might vary from context to context (McNally & Stojanovic, 2017; Cepollaro, Soria & Stojanovic, 2021). For instance, I might take into consideration the regularity of the terrain much more than any other parameter when deciding which place is more beautiful to take a walk, but give more weight to the color of the sky when deciding which place is more beautiful to lay down and watch the clouds passing by. And I might give much more weight to a person’s cardiovascular state when deciding who is healthy enough to play on a sports team, but give less value to that parameter in another situation. It might even be the case that

the relative weight of each dimension varies when considering two different places while deciding which one is more beautiful for a walk.

Second, a number of theories have accommodated value-related information within the realm of pragmatics, a highly context-dependent layer of meaning (see, for instance, Corredor, 2014). This realm includes approaches that place certain contextual information beyond what we literally say, in the range of what we implicate or presuppose by saying what we say. That is, our utterances can convey information that isn't part of the literal meaning of the sentences we utter, and do so by exploiting certain linguistic mechanisms and conversational principles. The main theoretical notions within this strand are *presuppositions* and different types of *implicatures*: conventional, generalized conversational and particularized conversational implicatures.

Several authors have defended a presuppositional account of different sorts of evaluative judgments and expressions (see, for example, Macià, 2002; Cepollaro, 2015, 2016, 2017a, 2017b; Cepollaro & Stojanovic, 2016). The main point of this approach is that it situates the evaluative meaning at the level of what is taken for granted. For instance, if I evaluate Loyiso Gola's special *Unlearning* by saying that it is *powerful*, I might literally mean that the show has a strong effect on people's thought, and trigger the evaluative presupposition that things that have a strong effect on people's thought are good precisely because they do so. That is, our evaluations, at least in some cases, implicitly take for granted a certain perspective, a certain set of beliefs concerning what is good and what is bad, which is situated at the level of what is presupposed.

Other authors situate the meaning conveyed by evaluative uses of language at the level of what is conventionally implicated (see, for example, Stenner, 1981; Potts, 2005, 2007; Copp, 2009; Whiting, 2007, 2013; Williamson, 2009; McCready, 2010; Gutzman, 2011). Implicatures are pieces of information conveyed by what we say, that is, information that goes beyond the literal meaning expressed through an utterance. Some implicatures are part of the conventional meaning of what is literally said. As Abbot claims, presuppositions and conventional implicatures are "close neighbors" (Abbott, 2006, 2). In contrast to presuppositions, conventionally implicated information is not taken for granted but inferentially conveyed. For instance, if I evaluate Loyiso Gola's show *Unlearning* by saying that it is powerful, I conventionally implicated the proposition that Loyiso Gola's show is good, given the conventional meaning of "powerful".

The two further types of implicatures, particularized and generalized conversational ones, are implicated pieces of information that exploit conversational rules. They allow us to restore the conversational order when some basic rule is apparently bent. The main difference between them is their degree of context-dependence. Particularized implicatures require some special contextual features, while generalized ones arise in the absence of particular circumstances that would defeat it (see Väyrynen, 2013). As an example of a conversational implicature, take the utterance of the sentence "It lasts more than an hour" as a response to the question "How was the show?" to implicate that it was not good.

As we have seen, the specific valence of our evaluations depends on the context in at least two ways: the features that determine the positive or negative valence of a judgment depend on the context, and the relative weight of those features might also vary from context to context. Moreover, a good number of theories have accommodated the

information conveyed through our evaluative judgments within the realm of pragmatics, a highly context-dependent layer of meaning. When a judgment is identified as an evaluative one, then it is highly context-dependent, in different senses. However, there is a third issue that might also need the intervention of context. By way of example: dogwhistles. Dogwhistles are a sort of political manipulation that some carry out to covertly obtain certain political benefits. Succinctly, they are pieces of speech that convey certain information which is consciously or unconsciously received just by a subset of the general audience (Saul, 2018; Torices, 2021). Only under certain very specific circumstances can a piece of public discourse be interpreted as a dogwhistle, taken to express affective attitudes only to a given proper subset of the target audience, while remaining “neutral” to the rest of it. Only with respect to this subset of the audience is it able to contribute to the promotion of negative stereotypes, to be offensive. This selective expression of hatred carried out by dogwhistling is also highly context-dependent. Note, however, that context plays a role here not only once the judgment is identified as an evaluative one, but in a previous step. The claim will express affective attitudes or it will not, will be an evaluative claim or it will not, depending on the context.

So, there are linguistic phenomena characterized by promoting certain affective attitudes in a subset of the general audience whose success is also highly context-dependent. In this sense, the role played by the context in the determination of the meaning conveyed through an utterance generalizes a bit further. Determining whether a piece of discourse is evaluative or descriptive, potentially offensive or “neutral”, also requires taking into account several features of the context. At the beginning of this paper, we have already introduced some intuitive and powerful tests to differentiate a factual claim from an evaluative one: what is expressed, commonality of standards, retraction and practicality. Let’s add two more tests to be able to spot gradable and multidimensional predicates, the kind of expressions that are typically involved in evaluative claims. A descriptive claim such as “Gola’s show was recorded at 3 PM” does not admit of degrees—e.g., “# Gola’s show was *very / slightly / more* recorded at 3 PM than the previous one”. Neither it is possible to distinguish different respects playing a role—e.g., “# Gola’s show was recorded at 3PM *in every / some respect / with respect to*”. However, an evaluative claim such as “Gola’s show was nice” does admit degrees—e.g., “Gola’s show was *very / slightly* nicer than the previous one—and multidimensionality—e.g., Gola’s show was nice *in every / some respect / with respect to* the performance, but not with respect to the plot (Cepollaro, Soria & Stojanovic 2021). Being gradable is a cue of being evaluative, though it does not warrant it. Not every gradable and multidimensional adjective, or predicative expression, is used to make evaluations, and not every evaluation is made with the aid of these expressions. Firstly, any multidimensional and gradable adjective can be used descriptively in a context in which the standards are fixed. Secondly, given the appropriate context, almost any predicative expression can be used to make an evaluation, once it is established that possessing that particular feature is associated with a valence, positive or negative.

Soria and Stojanovic have distinguished some additional tests (Soria & Stojanovic, 2019). The first one is the *juxtaposition with “although”-type connectives* test, based on the action-guidance of the evaluative. The idea is that a complex sentence composed of two sentences connected by the conjunction “although”, establishing a contrast between

a claim and a practical attitude, seems felicitous only when the claim is evaluative. For instance, the evaluative claim “Nassim thinks that eating meat is wrong, although she consumes twice more meat than average” seems felicitous, while the descriptive claim “Nassim thinks that eating meat is unpopular, although she consumes twice more meat than average” does not. Another test is the *lack of epistemic justification*. The idea is that specific evaluative judgments, such as claiming “I’ve just watched Gola’s show and it’s amazing”, won’t admit the question “How do you know?”, while descriptive claims such as “Gola’s show lasts one hour and a half” does. Finally, we also have the *lack of lying potential* test. Evaluative claims do not admit the reply “That’s a lie”, in part because they are a genuine expression of our own perspective. For instance, the claim “This show is amazing” cannot be replied by saying “That’s a lie”, while a descriptive one such as “This show is the second one of her career” can. As shown above with the first two features highlighted in this linguistic road to the evaluative, these tests also fall short from being conclusive. It’s perfectly OK to inquire into the reasons for a given evaluation, especially if the topic is important to us, or we are close to the speaker. “The White Album better than *Abbey Road*? I can’t believe you said that!”, “How can you take A over B?”, “The White Album better than *Abbey Road*? Are you sure about that?” all seem perfectly fine to us, evaluative claims where the grounds are explicitly questioned. Seemingly, instances can be found of evaluative claims followed by “That’s a lie”. Imagine you keep telling a child for years “Eat your beans, they’re delicious”. Years later, the now-adolescent discovers that you’d rather skip a meal than eating beans and protests: “You hated beans, and yet you kept telling me they were delicious, you were lying to me the whole time!!”.

The first of the tests mentioned above requires a bit more consideration. It hinges on the particular connection with our practical commitments that evaluative utterances seem to have, where the uptake seems to change the practical commitments of participants in the conversation. This rings true enough, it looks reasonable to assume that if I’m taking part in a conversation, and I let an evaluative claim be made unchallenged, the speaker has every reason to suppose that I *also* believe what they expressed. If we do not react to racist claims, we are contributing to the normalization of these words, and potentially reinforcing negative stereotypes. On the other hand, it also looks a bit excessive as a requirement for evaluative uses of language. A perfectly evaluative utterance can be understood, and be given a *pass*, even if we disagree with the claim, or with the way it is expressed. It’s too strong to demand from every participant in a conversation the attitude of a linguistic *vigilante*, ready at all times to deny the uptake to evaluative utterances, whenever there’s a disagreement. Sometimes staying quiet is required to let others find their way on their own. Speakers need to be given the chance to recognize the attitudes that they express, even if they were not aware of them before making their utterances, and this is much easier, in most contexts, if we do not adopt the attitude of linguistic *vigilantes*. Again, Claudia Rankine captures both intuitions beautifully in the following passage:

Another friend tells you you have to learn not to absorb the world. She says sometimes she can hear her own voice saying silently to whomever—you are saying this thing and I am not going to accept it. Your friend refuses to carry what doesn’t belong to her.

You take in things you don't want all the time. The second you hear or see some ordinary moment, all its intended targets, all the meanings behind the retreating seconds, as far as you are able to see, come into focus. Hold up, did you just hear, did you just say, did you just see, did you just do that? Then the voice in your head silently tells you to take your foot off your throat because just getting along shouldn't be an ambition. (Rankine, 2014, 55)

All these tests, even if they are not always completely effective, showcase the difference between the descriptive and the evaluative. But, beyond the features that the evaluative exhibits so far, such as the difference in what is expressed, commonality of standards, retraction and practicality; beyond the context-dependence, in different senses, of evaluative judgments, and beyond the linguistic particularities that are sometimes exhibited by the expressions that we use to evaluate, there is more that we need to know concerning *when* a seemingly neutral statement is indeed used to express affective attitudes. This is also highly context-dependent. The relationship between contextual information and our linguistic behavior conducive to conveying evaluations is a complex one, and the limitations of the criteria pointed above tell this story. Our purpose in this paper, as stated in the introduction, is to chart a conceptual map of the territory, exploring the benefits and shortcomings of the positive proposals contained in the recent literature, but *also* to point to a different territory, one that involves experimental philosophy and that will require more attention in the near future.

Recently, we have conducted two empirical studies on the influence of different contextual factors in assessing the same statement as simply informative or also offensive (Almagro, Hannikainen & Villanueva, 2021). In this research, we took into consideration different contextual features. In particular, we designed a first study with a set of 8 vignettes where a speaker utters a seemingly descriptive sentence, and manipulated three contextual factors: speaker status, speaker intention, and harm caused to the audience. All participants viewed eight vignettes telling a short story in which a speaker claims something, and each participant was exposed to a randomized combination of factors in each vignette. Then, we asked each participant to rate from 1 (strongly disagree) to 7 (strongly agree) the following claims after reading each scenario: "The speaker simply offers information"; "The speaker is being offensive". We added an additional question about the acceptability of the claim with two available options: "The speaker shouldn't say that kind of thing" / "I don't see any problem in the speaker saying that kind of thing". Our results showed that for participants, although the three factors' effect was significant, the speaker's status played a more important role than the intention and the outcome factors in taking the same statement as offensive and prohibited (in a second study we have disambiguated between speakers' membership and social status, and membership appears to exert a greater effect). So, according to these findings, all three factors are relevant, and speaker's status and membership are even more important contextual factors in determining whether a claim expresses affective attitudes or just conveys descriptive information. Moreover, status seemed to play a more important role than agents were willing to explicitly acknowledge, when asked directly. Thus, the context does not only affect our evaluations because our standards are diverse and might change from situation to situation; whether some utterance is deemed as evaluative—and potentially offensive—or not can only be determined by paying attention to features of the

context such as: who's speaking, what are the potential consequences of the utterance, and what intention is reasonable to attribute to the speaker.

Let's turn now to the initial example of this section. What did the speaker say by claiming that Americans battle between the "historical self" and the "self self"? Was it good or bad? Is what they said offensive? The moral of this section is that before examining the context to check what standards are in play, we need to consider the actual expressions used and their potential occurrences in other linguistic contexts, but also the status of the speaker and the audience, their intentions, and the possible consequences.

3. Why should we care?

A woman you do not know wants to join you for lunch. You are visiting her campus. In the café you both order the Caesar salad. This overlap is not the beginning of anything because she immediately points out that she, her father, her grandfather, and you, all attended the same college. She wanted her son to go there as well, but because of affirmative action or minority something—she is not sure what they are calling it these days and weren't they supposed to get rid of it?—her son wasn't accepted. You are not sure if you are meant to apologize for this failure of your alma mater's legacy program; instead you ask where he ended up. The prestigious school she mentions doesn't seem to assuage her irritation. This exchange, in effect, ends your lunch. The salads arrive. (Rankine, 2014, 19)

Although it might appear to be a simple description, saying to a non-white person that you wanted your son to go to her college but that he wasn't accepted because of affirmative action or minority something *is* offensive, even if you didn't intend it that way. Take the following case. In 2008, Phil Edwards and Ian Roberts published a short piece in the British peer-reviewed general medical journal *The Lancet*. In this article, the authors tackled the issue of global food insecurity and climate change. In particular, they made the point that reducing the demand for transportation fuel is important for reducing food prices in two senses (Edwards & Roberts, 2008). The first argument offered by the authors is as follows: motorized transport is more than 95% oil-dependent. Since oil is a key agricultural input, demand for transportation fuel increases its price, and then affects food prices. The second argument goes like this: increased car use contributes to rising food prices by promoting obesity, which increases the global demand for food. Additionally, more transportation fuel energy will be used to transport the increased mass of the obese population. Their arguments were presented under a descriptive guise:

We estimate that a population of 1 billion people with a stable mean body-mass index (BMI) of 24.5 kg/m² consumes an average 6.5 MJ of food energy per person per day to maintain basal metabolic rate, and a further 4 MJ per person per day for activities of daily living. An obese population of 1 billion people with a stable mean BMI of 29.0 kg/m² would require an average 7 MJ of food energy per person per day to maintain basal metabolic rate, and 5.4 MJ per person per day for activities of daily

living. Compared with the normal weight population, the obese population consumes 18% more food energy. (Edwards & Roberts, 2008)

From this, the authors concluded that “urban transport policies that promote walking and cycling would reduce food prices by reducing the global demand for oil, and promotion of a normal distribution of BMI would reduce the global demand for, and thus the price of, food” (Edwards & Roberts, 2008). Although this contribution might be initially seen just as a descriptive one, and it is presumably written with that intention, some media reports echoed Edwards and Roberts’ findings under headlines such as “Obese *blamed* for the world’s ills”. In the same journal, Harald Schmidt wrote a reply to the authors claiming that Edwards and Roberts’ discussion of the role of overweight people in climate change is *ethically problematic* (Schmidt, 2008). He claims that some important alternatives to the authors’ hypothesis, such as cutting unnecessary transport, using alternative fuels, lighter materials in vehicles, etc., are neither considered nor acknowledged, and that it makes it more problematic to advocate that reducing weight is a key priority in fighting against climate change and food prices. So, their contribution, he concludes, counts as an instance of victim blaming of an already marginalized and vulnerable population.

In addition to the authors’ narrow analysis that leads to focus on a single possible explanation that, as a result, blames certain people from a disenfranchised group, there are other rather problematic assumptions in play. For instance, Edwards and Roberts assume that someone’s body-mass is always proportional to the intake of a particular amount of food energy, which is a stereotype. It is also assumed that there is a *normal* body-mass index, which makes other body-mass indexes of healthy people seem abnormal. Empirical studies have shown that the use of terms such as “obese”, “fat”, etc., as “descriptive” as they might look in a scientific context, promotes negative attitudes towards people to whom such predicates are applied (see, for instance, Brochu & Esses, 2011).

Nevertheless, it can be convincingly argued that doing all this was not the authors’ intention. Presumably, the authors’ purpose was simply to genuinely make a contribution to the topic, and not to blame anyone. However, the meaning of our words is beyond our acknowledged intentions, and we seem to be especially bad at knowing our own minds. For example, we might be completely honest when we say that we consider all races to be equal, and yet still display racist attitudes in our linguistic and nonlinguistic behavior. The evaluative nature of the information we convey through our verbal and nonverbal actions gives us clues about our own personality, what to keep and what to try and keep in check.

In our study presented above, we also found some surprising results along this line. At the end of the experiment, we included a general final question in which we explicitly asked about the influence of the three factors (speaker’s status, intention, and the outcome) when considering the same statement as offensive or not. Participants had to rate from 1 (little influence) to 7 (much influence) speaker status, speaker intention, and the harm felt by the audience. Although participants had shown through their concrete case-by-case responses that the identity of the speaker, both in terms of membership and social status, is the contextual factor that most affected their responses in considering the same statement as offensive or neutral, when approached in an abstract way, participants rated the speaker’s intention as the most relevant factor, and the speaker’s identity as the least important. So,

it can be concluded that most of the participants in our study sincerely say that they think that the speaker's intention is the most relevant contextual factor in considering a claim as descriptive or offensive, but that they don't actually think so, given their responses in specific situations. This result is in line with other research suggesting that we have systematic blindspots regarding our own mental states (Nisbett & Wilson, 1977; Zajonc, 2001; Wilson, 2002; Schwitzgebel, 2008, 2011; Carruthers, 2011). Our evaluations can make us realize an aspect of our personality, of our mind, that we could have otherwise ignored.

Perhaps this realization comes a bit too late, only when we catch ourselves promoting the kind of stereotypes that reinforce the discrimination against disenfranchised groups, we do acknowledge that perhaps there was an aspect of our personality that we needed to address. It's too late to take our utterance back, and the harm is probably already done, but it might be one of the very few ways to access those dispositions of our own, and we'd better keep them in mind. Take an implicit test, trust those that are close to you, pay attention to what you say, perhaps inadvertently, especially when you are called out on that. What did you say? Exactly, what did you mean? Besides giving information about our own mind, evaluations promote negative attitudes and have other pernicious consequences, especially when they are unnoticeably made or presented under a descriptive guise. For instance, Samuel Huntington's contribution to the Report on the Governability of Democracies to the Trilateral Commission, published in 1975, was to argue that the problems of the United States in the 1970s result from "an excess of democracy". His recommendation for the United States, as Stanley puts it, was "to try to reinstall some measure of obedience to authority by making various central domains in life, ones that should be governed democratically, the domain of experts" (Stanley, 2015, 40). That is, Huntington was recommending to take certain matters of value as if they were factual ones. Appealing to experts as if they have the last word when deciding about a normative issue is covertly making an evaluation, presenting it as if it were a description. In doing so, one may appear to be making a reasonable contribution when in fact one is undermining certain democratic ideals as Stanley points out. The subtle effect exerted by different contextual factors on the meaning conveyed through a claim, together with our inability to clearly identify the influence exerted by these factors, makes us especially vulnerable to situations such as the previous ones. If Edwards and Roberts say that they didn't intend to be offensive, then someone might conclude erroneously that what they said is not offensive, and that it counts as a reasonable contribution. Similarly, one might wrongly conclude that, since Huntington appealed to experts, he was talking about a factual issue, and that he was making a reasonable contribution.

This move can be done to advance certain political agendas. As Kappel puts it, "in some cases, there might be strategic advantages to framing a disagreement as fact-dependent, even when it is not. Or conversely, disagreements may be framed as value-dependent, when they are really fact dependent, and so on" (Kappel, 2017, 317). In particular, this move generates situations of *crossed disagreements* (Osorio & Villanueva, 2019). A crossed disagreement is a situation in which each of the parties of a genuine disagreement displays clear signs of conceiving the dispute in significantly different terms, e.g., one conceives it in descriptive terms, and the other as if it were a normative dispute. One of the pernicious consequences of the systematic confusion of the descriptive and the evaluative in public debates is that they foster polarization: crossed disagreements expand the pool of arguments supporting

the views held by each party (Osorio & Villanueva, 2019; Almagro, Osorio & Villanueva, 2021), which in turn contributed to the raise of polarization. In other words, increasing the size and density of the pool of arguments that one is exposed to fosters polarization: when we are exposed to a limited group of arguments just in favor of one position, especially when they are frequently repeated, we tend to become more impervious to others' reasons (Sunstein, 2017; Barberá et al., 2015; Vicario et al., 2016). Since the rise of polarization increases distrust in public institutions and governments, it makes citizens less likely to seek different perspectives on controversial topics, and corrodes the proper functioning of democratic institutions, among other things (see Levitsky & Ziblatt, 2018; Carothers & O'Donohue, 2019), the systematic confusion of the descriptive and the normative in public debates is pretty dangerous.

In this paper we've taken a much more personal approach, focusing on cases where the speakers are expressing hatred through apparently descriptive claims. Let's briefly present other cases in this vein. In a scientific article, Simko and Ginter concluded that the uncontrolled sugar consumption leading to high rates of diabetes explains the low life expectancy of the Roma population in Europe (Simko & Ginter, 2010), ignoring many other relevant data, such as that Roma population regularly appears in the Eurobarometer as the minority to which the rest of the people in Europe feel less affinity, have less access to resources, suffer more poverty, etc. Their food choices might not be a result of their culture or their way of life, but the sheer consequence of their low income. A recent report on disparities in the risk and outcomes of COVID-19 in the United Kingdom (Public Health England, 2020), while it recognizes that non-white males have been more likely to die from a COVID-19-related condition than White males, highlights that "an analysis of over 10,000 patients with COVID-19 admitted to intensive care in UK hospitals suggests that, once age, sex, obesity and comorbidities are taken into account, there is no difference in the likelihood of being admitted to intensive care or of dying between ethnic groups", and adds that people of non-white groups "are also more likely than people of White British ethnicity to be born abroad, which means they may face additional barriers in accessing services that are created by, for example, cultural and language differences". In doing so, again, they come short of blaming people from some disenfranchised groups for their bad fortune, which is highly offensive and promotes negative attitudes towards them, despite the authors neither having this intention nor using a language that is clearly evaluative.

4. Conclusion

In this paper, we have explored the boundaries of offensive speech, a subset of evaluative uses of language. In particular, we have focused on the contextual dependence of offensive speech: the contextual cues that determine that a piece of speech is interpreted evaluatively, rather than as mere description of our surroundings, and is thus able to express the kind of affective attitudes that might contribute to the discrimination of disenfranchised groups. First, we have dwelt on how context affects evaluative meaning. As we have seen, our evaluations depend on a variety of contextual features, and the relative weight of those features also varies from context to context. Second, we have briefly reviewed some approaches to the evaluative from the realm of pragmatics, the classic

locus of context-dependence. Then, we have focused on a more general issue: how the context plays a role in determining whether a piece of speech expresses affective attitudes or remains neutral. We have revisited different tests and remarks that can help us to know whether a claim is descriptive or evaluative, and we have introduced an empirical study that we have recently conducted on the influence exerted by three contextual factors in determining whether a statement counts as a descriptive or evaluative one: speaker's status, speaker's intention and outcome. Finally, we have discussed several cases where the speaker's claim seems to be descriptive but it actually expresses affective attitudes towards certain groups of people, and showed why they are problematic and why we need to be able to recognize when a claim is an evaluative one.

The goal of this paper started from the need of making explicit certain forms of injustice that frequently go unnoticed, in order to be able to fight against them. In this sense, it belongs to *the political turn in analytic philosophy* (Pinedo & Villanueva, forthcoming). The defining feature of this political turn is taking our raw perceptions of injustice as the starting point, and bringing to the fore the capacity of our philosophical tools in detecting and intervening on oppressive and unjust practices. Our best theories will be the ones that more efficiently serve this purpose: identifying and intervening in certain social injustices.

Words work as release—well-oiled doors opening and closing between intention, gesture. A pulse in a neck, the shiftiness of the hands, an unconscious blink, the conversations you have with your eyes translate everything and nothing. What will be needed, what goes unfelt, unsaid—what has been duplicated, redacted here, redacted there, altered to hide or disguise—words encoding the bodies they cover. And despite everything the body remains.

Occasionally it is interesting to think about the outburst if you would just cry out—
To know what you'll sound like is worth noting. (Rankine, 2014, 69)

References

- Abbott, B. (2006), "Where have some of the presuppositions gone", in B. Birner & G. Ward, (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Amsterdam: Studies in Language Companion Series, pp. 1–20.
- Almagro, M., Hannikainen, I. & Villanueva, N. (2021), "Whose words hurt? Contextual determinants of offensive speech", *Personality and Social Psychology Bulletin*.
- Almagro, M., Osorio, J. & Villanueva, N. (2021). "Weaponized testimonial injustice", *Las Torres de Lucca*.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. & Bonneau, R. (2015), "Tweeting from left to right: Is online political communication more than an echo chamber?", *Psychological Science*, 26(10), 1531–1542.
- Bordonaba, D. & Villanueva, N. (forthcoming). "Retractación y contextualismo: nuevas condiciones de adecuación", in D. Pérez Chico (ed.), *Contextualismo semántico*, Zaragoza: Prensas de la Universidad de Zaragoza.

- Brochu, P. M., & Esses, V. M. (2011). “What’s in a name? The effects of the labels “fat” versus “overweight” on weight bias”, *Journal of Applied Social Psychology*, 41(8), 1981–2008.
- Carothers, T., & O’Donohue, A. (2019), *Democracies Divided: The Global Challenge of Political Polarization*, Washington, DC: Brookings Institution Press.
- Carruthers, P. (2011), *The Opacity of Mind: An Integrative Theory of Self-Knowledge*, Oxford: Oxford University Press.
- Cepollaro, B. (2015), “In defense of a presuppositional account of slurs”, *Language Sciences*, 52, 36–45.
- Cepollaro, B. (2016), “Building evaluation into language”, *Phenomenology and Mind*, 11, 158–168.
- Cepollaro, B. (2017a), “The shortcut of discrimination”, *Rivista di Estetica*, 64, 53–65.
- Cepollaro, B. (2017b), *The Semantics and Pragmatics of Slurs and Thick Terms*, PhD thesis, Paris: PSL Research University.
- Cepollaro, B. & Stojanovic, I. (2016), “Hybrid evaluatives: In defense of a presuppositional account”, *Grazer Philosophische Studien*, 93(3), 458–488.
- Cepollaro, B., Sulpizio, S. & Bianchi, C. (2019), “How bad is it to report a slur? An empirical investigation”, *Journal of Pragmatics*, 146, 32–42.
- Cepollaro, B. & Zeman, D. (2020), “The challenge from non-derogatory uses of slurs”, *Grazer Philosophische Studien*, 97(1), 1–10.
- Cepollaro, B., Soria, A. & Stojanovic, I. (2021), “The semantics and pragmatics of value judgments”, in P. Stalmaszczyk (ed.), *The Cambridge Handbook of the Philosophy of Language*, Cambridge: Cambridge University Press.
- Charlow, N. (2014), “The problem with the Frege-Geach problem”, *Philosophical Studies*, 167(3), 635–665.
- Chrisman, M. (2016), *What Is This Thing Called Metaethics*, London: Routledge.
- Corredor, C. (2014), “Pejoratives and social interaction”, in P. Stalmaszczyk (ed.), *Issues in Philosophy of Language and Linguistics*, Łódź: Wydawnictwa Uniwersytetu Łódzkiego, 39–55.
- Copp, D. (2009), “Realist-expressivism and conventional implicature”, *Oxford Studies in Metaethics*, 4, 167–202.
- Edwards, P. & Roberts, I. (2008), “Finding long-term solutions to the world food crisis”, *The Lancet*, 371(9622), 1389.
- Field, H. (2009), “Epistemology without metaphysics”, *Philosophical Studies*, 143(2), 249–290.
- Field, H. (2018), “Epistemology from an evaluativist perspective”, *Philosophers’ Imprint*, 18(2), 1–23.
- Finlay, S. (2004), “The conversational practicality of value judgment”, *Journal of Ethics*, 8(3), 205–223.
- Gibbard, A. (1990), *Wise Choices, Apt Feelings*, Cambridge: Cambridge University Press.
- Gibbard, A. (2012), *Meaning and Normativity*, Oxford: Oxford University Press.
- Gibson, J. L., Epstein, L. & Magarian, G. (2019), “Taming uncivil discourse”, *Political Psychology*, 41(2), 383–401.

- Gutzman, D. (2011), “Expressive modifiers & mixed expressives”, *Empirical Issues in Syntax and Semantics*, 8, 123–141.
- Kappel, K. (2017), “Fact-dependent policy disagreements and political legitimacy”, *Ethical Theory and Moral Practice*, 20(2), 313–331
- Levitsky, S., & Ziblatt, D. (2018), *How Democracies Die*, New York: Crown.
- Macià, J. (2002), “Presuposición y significado expresivo”, *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia*, 3(45), 499–513.
- McCready, E. (2010), “Varieties of conventional implicature”, *Semantics and Pragmatics*, 3(8), 1–57.
- McNally, L. & Stojanovic, I. (2017), “Aesthetic adjectives”, in J. Young (ed.), *The Semantics of Aesthetic Judgment*, Oxford: Oxford University Press.
- Moreno, A. & Pérez-Navarro, E. (manuscript), “Beyond the conversation: The pervasive danger of slurs”.
- Nisbett, R., & Wilson, T. (1977), “Telling more than we can know: Verbal reports on mental processes”, *Psychological Review*, 84(3), 231–259.
- Osorio, J. & Villanueva, N. (2019), “Expressivism and crossed disagreements”, in M. J. Frápolli (ed.), *Expressivisms, Knowledge and Truth*, Cambridge: Cambridge University Press, pp. 111–132.
- Pinedo, M. & Villanueva, N. (forthcoming), “Epistemic de-platforming”, in D. Bordonaba, V. Fernández & J. R. Torices (eds.), *The Political Turn in Analytic Philosophy: Reflections on Social Injustice and Oppressions*, Berlin: De Gruyter.
- Potts, C. (2005), *The Logic of Conventional Implicatures*, Oxford: Oxford University Press.
- Public Health England (2020), “Disparities in the risk and outcomes of COVID-19”, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf
- Rankine, C. (2014), *Citizen: An American Lyric*. Minneapolis, MN: Graywolf Press.
- Saul, J. (2018), “Dogwhistles, political manipulation and philosophy of language”, in D. Fogal, D. W. Harris & M. Moss (eds.), *New Works on Speech Acts*, Oxford: Oxford University Press, pp. 360–383.
- Schmidt, H. (2008), “Transport policy, food policy, obese people, and victim blaming”, *The Lancet*, 372(9622), 627.
- Simko, V. & Ginter E. (2010), “Short life expectancy and metabolic syndrome in Romanies (gypsies) in Slovakia”, *National Center for Biotechnology Information*, 18(1), 16–18.
- Stanley, J. (2015), *How Propaganda Works*, Princeton, NJ: Princeton University Press.
- Stenner, A. J. (1981), “A note on logical truth and non-sexist semantics”, in M. Vetterling-Braggin, (ed.), *Sexist Language: A Modern Philosophical Analysis*, New York: Littlefield, Adams & Co., pp. 299–306.
- Strandberg, C. (2012), “A dual aspect account of moral language”, *Philosophy and Phenomenological Research*, 84(1), 87–122.
- Sunstein, C. (2017), *#Republic: Divided Democracy in the Age of Social Media*, Princeton, NJ: Princeton University Press.
- Schwitzgebel, E. (2008), “The unreliability of naïve introspection”, *The Philosophical Review*, 117(2), 245–273.
- Schwitzgebel, E. (2011), *Perplexities of Consciousness*, Cambridge, MA: MIT Press.

- Torices, J. R. (2021). “Understanding dogwhistles politics”, *Theoria. An International Journal for Theory, History and Foundations of Science*. <https://doi.org/10.1387/theoria.22510>
- Väyrynen, P. (2013), *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*, Oxford: Oxford University Press.
- Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E. & Quattrocioni, W. (2016), “The spreading of misinformation online”, *Proceedings of the National Academy of Sciences*, 113(3), pp. 554–559.
- Whiting, D. (2007), “Inferentialism, representationalism and derogatory words”, *International Journal of Philosophical Studies*, 15(2), 191–205.
- Whiting, D. (2013), “It’s not what you said, it’s the way you said it: Slurs and conventional implicatures”, *Analytic Philosophy*, 54(3), 364–377.
- Williams, B. (1985), *Ethics and the Limits of Philosophy*, London: Fontana.
- Williamson, T. (2009), “Reference, inference, and the semantics of pejoratives”, in J. Almog & P. Leonardi (eds.), *The Philosophy of David Kaplan*, Oxford: Oxford University Press, pp. 137–159.
- Wilson, T. (2002), *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Cambridge, MA: Harvard University Press.
- Zajonc, R. (2001), “Mere exposure: A gateway to the subliminal”, *Current Directions in Psychological Science*, 10(6), 224–228.



I am large, I contain multitudes: Epistemic pragmatism, testimonial injustice and positive intersectionalism*

Soy grande, contengo multitudes.
Pragmatismo epistémico, injusticia testimonial e
interseccionalismo positivo

MARÍA J. FRÁPOLLI**
LLANOS NAVARRO***, ****

Abstract: We explore the compatibilities and incompatibilities between two highly successful approaches to knowledge: Brandom's epistemic pragmatism, [EP], and the view that derives from Fricker's seminal work on the ethics of knowing, [EK]. [EP] and [EK] are complementary approaches that put forward aspects of the application of the concept that deserve to be preserved. Nevertheless, their mere cumulative superposition produces dysfunctions that call for certain readjustments. We propose a positive kind of intersectionalism, [PI], that accounts for the fact that individuals simultaneously belong to diverse

Resumen: En este trabajo exploramos las compatibilidades e incompatibilidades entre dos enfoques del conocimiento distintos y de gran éxito. El primero es el análisis del conocimiento que ofrece el pragmatismo epistémico de Brandom [PE]. El segundo es el punto de vista que se deriva de la obra seminal de Fricker sobre la ética del conocimiento [EC]. [PE] y [EC] plantean concepciones complementarias de la aplicación del concepto de conocimiento que merecen preservarse. Sin embargo, su mera superposición acumulativa produce disfunciones que exigen ciertos reajustes. Proponemos un tipo de interseccionalismo

Recibido: 31/05/2021. Aceptado: 22/06/2021.

* The following people have commented previous versions of this paper, helping to improve the final result: Alba Moreno, Eduardo Pérez, Claudia Picazo, and Andrés Soria. Two anonymous referees have also contributed with their suggestions. To all of them, we are deeply indebted.

** Professor of Logic and Philosophy of Science (Department of Philosophy I, University of Granada). Contact: frapolli@ugr.es. Research interests: philosophy of logic, philosophy of language, pragmatism. Recent publications: Frápolli, M. J. and Forero-Mora, J. A. (2021), "Show me: Tractarian non-Representationalism", *Teorema*, 40(2), 63–81; and M. J. Frápolli (2019), "Propositions first: Biting Geach's bullet", *Royal Institute of Philosophy Supplements*, 86, 1–9.

*** Predoctoral fellow (Department of Philosophy I, University of Granada). Contact: llanosnavarro@ugr.es. Research interests: philosophy of language, metaethics, social epistemology. Recent publications: Navarro, L. (forthcoming), "Where are ethical properties? Predication, location and category mistake", in N. S. Miras and M. Bella (eds.), *Women in Pragmatism: Past, Present, and Future*, Berlin: Springer; and Almagro, M., Navarro, L. & Pinedo, M. (forthcoming), "Is testimonial injustice epistemic? Let me count the ways", *Hypatia: A Journal of Feminist Philosophy*.

**** The names of the two authors are in alphabetical order. Both authors have contributed equally to the paper.

groups with variable epistemic conditions, some advantageous, some disadvantageous. [EP], [EK] and [PI] make a rich and coherent picture of subjects as full epistemic agents.

Key words: Brandom, Fricker, intersectionality, knowledge, licensing, McGowan

lismo positivo, [IP], que da cuenta del hecho de que los individuos pertenecen simultáneamente a diversos grupos con condiciones epistémicas variables, algunas ventajosas y otras desventajosas. [EP], [EK] y [PI] conforman un cuadro rico y coherente sobre los sujetos como agentes epistémicos plenos.

Palabras clave: Brandom, conocimiento, Fricker, interseccionalidad, licencia, McGowan

1. The argument

Let us begin with a broad characterisation of [EP]. Being a knower is a social status whose possession requires the acknowledgement of other people. Discursive creatures, i.e. *sapientis* in Brandom's terminology, are participants in socio-linguistic games in which reasons are provided, evaluated, and consumed. Paying tribute to Sellars, Brandom identifies the game of 'making assertions and giving and asking for reasons' (Brandom, 2000, 26) as occupying the ground level. Paying tribute to Hegel, one of Sellar's inspiring heroes (Brandom, 2019, 114), he assumes that normative statuses, socially instituted, take the form of '*mutual or reciprocal* recognition' (Brandom, 2019, 12). This intuition shared by Hegel, Sellars, and Brandom is common to any pragmatist approach to language. Conceptual contents emerge—and communication occurs—in highly regulated linguistic actions where participants play specific roles for which they have the appropriate kind of licensing. We will take this point up in the next section.

This is the first claim of our overall argument. For our purposes, it can be reduced to the Hegelian-like slogan, [HS],

[HS] Knowledge requires acknowledgement.

The second claim is a development of the first: the role of knowledge, as a normative notion, is the expression of complex attitudes of the agent who attributes it and the subject to which it is attributed. Let us call this claim 'epistemic expressivism', [EE]. Brandom's epistemic expressivism makes self-attributions of knowledge logically dependent on third-person knowledge attributions. Agents learn the game of attributing knowledge to themselves, i.e. they recognise themselves as knowers, by learning the game of attributing knowledge to others. This is a possible wording of [EE]:

[EE] Mastering the practice of self-attributing the property of *being a knowledge possessor* presupposes mastering the practice of attributing it to others.

[HS] and [EE] make *conceptual* remarks. They expose necessary connections between the concepts involved and the practices of applying them. Epistemic pragmatism, [EP], is the view that results of the combination of [HS] and [EE].

[EP] directly implies that people completely deprived of the status of knower and, as a consequence, deprived of the status of knowledge attributor, cannot develop the required skills to recognise themselves as knowledge possessors.

Let us now consider the basic tenets of [EK] as we will understand it in this paper. In Fricker's analysis, epistemic injustice occurs when people are harmed in their condition of knowers because they belong to disadvantaged social groups. Other researchers (Medina, 2013; Cassam, 2019) follow Fricker's lead in stressing that members of oppressed groups are prone to develop epistemic vices caused by their disadvantaged social positions. Fortunately, oppressive situations not only produce pernicious effects in people who suffer them. Some epistemic virtues are also produced by systematic epistemic harm. And the other way around. Situations of dominance contribute to the development of epistemic vices and, sometimes, of epistemic virtues too. Thus, the ethics of knowing and virtue epistemology make room for hope, articulating the possibility of overcoming their underprivileged starting point.

There is a sharp contrast between [EP] and [EK] in terms of their relative rigidity. [EP] points to certain necessary connections between concepts and their application, whereas [EK] explains the essential plasticity of individuals, which allows them to modify, adjust, and counterbalance their epistemic positions. We contend that both sets of intuitions are grounded and that there should be room for both in a mature approach to knowledge.

The relevance of [EP] conceptual analyses for the [EK] narrative is straightforward. Among the consequences of [EP] for [EK] we count the following three. First, [EP] implies that testimonial injustice, as defined by (Fricker, 2007), essentially produces an *epistemic* kind of injustice, harming the victims in their status as epistemic subjects (Almagro et al., forthcoming).

Second, related to hermeneutic injustice, disadvantaged subjects do not possess the status to enact the norms that would facilitate the circulation of those terms and narratives that they use to explain their conditions. This way, members of dominant groups will be unable to take over the conceptual resources to understand oppressive situations.

The third one is that [EP], because of its conceptual and hence necessary nature, precludes the possibility of victims of different kinds of epistemic injustice to escape their fate by their own means.

These two last consequences of [EP] are too strong, contested by the analyses of Fricker, Medina, Cassam, and others, and refuted by empirical data. Our aim in this paper is to show how [EP] and [EK] should be enriched to accommodate the empirical data and build up a consistent proposal that gathers together conceptual analyses and political approaches. Some hypotheses whose addition we suggest have been amply discussed, as the contextual nature of knowledge, and we will not insist upon them. Others seem to have attracted less attention, at least in the version required in this paper. In any case, our proposal is largely programmatic and in need of completion. Nevertheless, we believe that making the structure of the general argument explicit is a substantial step forward towards a general approach to knowledge that benefits from [EP] and [EK].

2. Epistemic pragmatism

Attributions of knowledge are the expression of hybrid attitudes. The notions of knowledge and belief are conceptually distinct because knowledge, unlike belief, incorpo-

rates the attributor's perspective. Even if self-ascriptions of belief can sometimes be challenged from outside, considering e.g. the believer's behaviour, in ascriptions of knowledge the external perspective is an essential ingredient (Brandom, 2009, 157; Frápolli, 2019).

In this picture, knowledge attributions always give complex information that concerns the attributee's as much as the attributor's mental states. When we attribute knowledge to a subject, S , concerning a particular issue, p , as represented in scheme (1),

(1) S knows that p ,

we perform three intertwined actions: (i) we attribute to S a belief that p , (ii) we attribute to S the entitlement to her belief that p , and (iii) we express our justified endorsement of p . These conditions are the counterparts of the traditional analysis of knowledge as belief, (i), that is justified, (ii), and true, (iii).

Attributions of factive notions always involve the attributor's attitudes together with the attributee's states, they always expose the attributor's states and reveal the attributor's perspective. Because factives support the inference from their attribution to the independent assertion of their propositional contents, the attributor indirectly asserts the content she attributes to a third party. This is basically Brandom's account of the meaning of the concept of *knowledge* (Brandom, 2009, 157–8). Some examples of inferences involving factives are the following:

- (1) She knows that p , therefore p ,
- (2) Her claim that p is true, therefore p ,
- (3) She remembers that p , therefore p .

(1-3) are instances of (4),

(4) $S \phi$ s that p , $\vdash p$.

Moore-like paradoxes pop up whenever the assumption of the premise in (1-3) combines with the rejection of the conclusion.

The pragmatic layout just depicted explains Gettier-like phenomena. In Gettier's example, Smith is unable to distinguish between his justified true belief that the candidate who will get the job has ten coins in his pocket and the counterfactual situation in which knowledge of this fact could be attributed to him (Gettier, 1963).

Normative pragmatism shows that, exclusively from the first-person perspective, the speaker cannot discriminate between her credentials to assert that p , to assert that p is true, to assert that she believes that p , and to assert that she knows that p . To attribute knowledge, some kind of social triangulation is required, since speakers on their own lack perspective—which involves access to the right kind of information—to ground a principled distinction between their knowledge and their mere justified belief (Frápolli, 2019).

Thus explained, knowing is something that others attribute to you, and that you can claim for yourself as a result of being attributed. This claim has widest scope and amounts to saying that the *individual practice* of knowledge self-attributions is parasitic on the *social practice* of attributing knowledge to others, i.e. that the former cannot be made sense of

without the latter. Speakers learn the self-attribution game as a result of mastering the game of external knowledge attributions. This situation highlights the importance of the role of the others for the image that we have of ourselves as epistemic subjects.

Brandom follows Hegel's core idea, [HS], i.e. that recognition (*Anerkennung*) is the notion that grounds social practices in which other people are treated as bearing normative statuses. In his book on Hegel's *Phenomenology*, we read:

To be a self in the full normative sense is to have not only actual normative attitudes, but also actual normative statuses. It is not only to *take* one-self or others to have authority or be responsible, but actually to *have* authority or *be* responsible. To achieve such a status, a normative subject must participate in a general recognitive dyad: must actually be recognized by someone that subject actually recognizes. For only suitably socially complemented attitudes institute actual statuses. It follows that normative statuses, normative subjects of such statuses, and recognitive communities of normative subjects are all synthesized simultaneously by recognitive processes that have an appropriate structure: the structure of reciprocal recognition. (Brandom, 2019, 293–4)

Brandom's normative pragmatism is an elaboration of this intuition, which will be developed from a different perspective in the next section.

We endorse the essentially social nature of knowledge. It might be argued that sometimes subjects are in the position to attribute to themselves the knowledge that others deny them. We agree. Consider the following example: Victoria, a successful psychologist, insists, against the opinion of her colleagues, that the patient is in high risk of engaging in self-harming behaviour. She knows, she says, because she has followed the patient's health history in the past two years, and because she has been profusely exposed to similar cases. Does she know? The answer does not depend on the outcome, i. e. on whether the patient actually inflicts some harm to herself. It might happen that the patient engages in some kind of self-damaging behaviour without Victoria having prior knowledge of it, and the other way round, that the patient manages to continue with her life without any patent indication that she may cause herself any harm even though she was, at that point, a danger for herself. But still, this is a case in which we have reasons to attribute knowledge to Victoria, based on her expertise, socially acknowledged, although the attribution is subjected to the analysis of the arguments on one side and the other. Truth and justification do not suffice for the attribution of knowledge, which also needs the acknowledgement of the attributor who acts as referee of the different scores that this complex status involves (Brandom, 2000, 119). Subjective certainty is neither necessary (Wittgenstein, 1969, §§340-1; Williams, 2001, 18) nor sufficient (Ramsey, 1929, 256). And, as Gettier's cases show, objective truth also falls short of grounding knowledge.

The role of the third-person perspective is common ground in speech act theory. Explaining some recent developments that support our argument is the aim of the next section.

3. Authority and licensing

The structure of situations that involve attribution of knowledge—or its refusal—parallels the structure of other types of speech acts to which several agents contribute with

utterances that count as assertions. We follow McGowan (2019) in calling this general kind of speech exchange a ‘conversation’. McGowan argued that there are significant analogies between ordinary conversations and the performance of standard exercitives, in Austin’s sense. The CEO of a company uttering in an official meeting ‘From now on, holidays are suspended’ is an example of standard exercitive. A mother informing their children that dogs are not allowed inside the house is another, less formal, example. In exercitives, speakers enact norms that determine what is permissible in certain situations.

For the exercitive to be successful, the speaker who utters the permissibility norm has to possess a certain status. The secretary uttering exactly the same words as the CEO in the first example does not enact the norm; the child that pompously tells her siblings that dogs are not allowed inside the house does not produce a permissibility fact either. Only agents with the appropriate authority can enact norms in suitable contexts. As McGowan explains, the speaker who utters an exercitive must possess, besides authority, the intention to enact the norm, and his intention has to be recognised by the participants in the conversation. The success of the exercitive thus depends on the recognition of the speaker’s intentions, in the usual Gricean sense.

Contributions to ordinary conversations also enact permissibility norms—McGowan calls them ‘conversational exercitives’—that produce a change in the conversational score-keeping, in Lewis’s terminology (McGowan, 2019, 28ff), or in the shared common ground, in Stalnaker’s terminology (McGowan, 2019, 39ff.). This is the main thesis of McGowan (2019). Hate speech and offensive or discriminatory linguistic actions cause harm by enacting permissibility facts that can potentially add oppression to members of already oppressed groups. The effect of changing the score—or the background—is a general feature of contributions to conversational exercitives, not restricted to situations of hatred speech.

Consider the following example: Victoria and Joan are discussing the best way to get to the concert hall to attend the performance of their favourite rapper, for which they bought tickets many months in advance. In the conversation, Victoria says: ‘My boyfriend, Mark, has a car.’ After this contribution, Mark becomes a salient individual in the conversation. The pronoun in Joan’s reaction to Victoria’s utterance, ‘Tell him to get us there’, unequivocally refers to Mark as a result of the permissibility fact enacted by Victoria’s sentence. After Victoria’s intervention, speaking of ‘the car’ is also permitted, etc.

Neither intentions to enact a norm nor the recognition of intentions is a success condition for conversational exercitives. By contrast with what happens in the case of standard exercitives, speakers do not need to have the intention to enact any norm when they contribute to ordinary conversations.

A further difference between standard and conversational exercitives concerns the role of authority. As we have mentioned, standard exercitives rely on the authority of the speaker to produce their effects, but McGowan is reluctant to extend the role of authority to conversational exercitives. She acknowledges that, in a sense or another, participants exert some authority over the conversation (McGowan, 2019, 66), although of a weaker kind that she calls ‘standing’ (McGowan, 2019, 67).

There is a complex network of social norms that regulate who can participate in a conversation, i.e. who has the appropriate standing. Not everybody is allowed to join an ongoing speech act. This is something that depends on a wide variety of subtle social facts: your position in the community, your relationship to the participants, specific aspects of the

conversation at issue, etc. Adapting McGowan's example, imagine that I approach a group of women at the Beverly Yacht Club—a club of which I am not a member. I friendly greet the ladies in the group and begin to talk with the patent intention to engage in a conversational exchange. Spite my efforts, they ignore me, not looking at me or even making any sign that they have heard me. They are making it clear that I am not welcome. As McGowan says, 'I lack the requisite (in this case, social) standing' (loc. cit.).

Possessing the appropriate standing is something that others confer upon you. One might think that my lack of standing rests on an objective condition; after all, I am not a member of this elitist resort. Had I owned the exclusive membership, their reaction to my efforts would have been different. A superficial reflection will show that this is not the case. The same group of ladies could have denied acknowledgement to members of the club, e.g. to a member who supported the Green Party in the last local elections, and could have welcome non-members, e.g. the elegant wife and daughters of the new governor.

Thus, the condition of being a suitable participant in a social exchange, linguistic or non-linguistic, does not depend on the candidate's efforts. It is a normative status graciously conceded and maintained by people who already possess the standing. It is again an issue of *Anerkennung*, in Hegel's and Brandom's sense. Cases of bullying among teenagers and mobbing in working places, presently amplified by social networks, profusely illustrate how groups can ostracise people, denying them the access to participate of their activities.

Licensing is the kind of act by means of which the status to take part in a conversation is conferred upon a candidate. In standard cases, licensing is brought about by an *explicit* permission issued by others. Sometimes, nevertheless, the licensing is produced *implicitly*. A member of the group ostensibly speaking to the candidate enacts a permissibility fact to the effect of licensing the standing, thus permitting the candidate and the rest of the group to act as if she were a full member. In extraordinary situations, agents can even license themselves to act as authorities in specific contexts. In an emergency situation inside an aircraft, for instance, a passenger might assume the responsibility for controlling the passage, commanding other passengers to perform particular tasks, such as taking care of agitated passengers, calming children down, helping less able people to reach the emergency exits, etc. To do this, the accidental hero needs to be neither a physician, nor a policewoman, nor an emergency worker. Virtually anyone with the appropriate character and the will to help can take command of the situation. A further kind of self-licensing is produced by pretending: an agent can act in a situation as if she had the appropriate standing until it is recognised by the others participants (McGowan, 2019, 67). But even in cases of self-licensing, the others have to accept your authority. In the aircraft examples, the courageous passenger who steps up will only have authority if the passage accepts their commands.

Attributing knowledge to someone is considering her a full epistemic actor. The person who attributes and the person who is attributed must be licensed to participate in the epistemic conversation. People belonging to discriminated communities often lack the resources to assume and to pretend an authority role, and the possibility of being accepted by others as the bearers of authority. The effect is the *silencing* of discriminated subjects, who are most of the times ignored as interlocutors by members of dominant groups.

Licensing and silencing (Langton, 1993; Hornsby and Langton, 1998; see also Dotson, 2011) show that testimonial injustice applies to other types of speech act alongside with

giving testimony. The concept of discursive injustice coined by Kukla (2014, 441) points out that those who belong to disadvantaged groups not only suffer in their capacity as knowers but also in their capacity as discursive peers. In a similar vein, Ayala and Vasilyeva (2015) argue that the range of speech acts that a speaker can perform within a conversation, what they call her range of speech affordances, is ‘adjusted in virtue of membership in intersecting social categories, more or less salient’ (Ayala and Vasilyeva, 2015, 132). Thus, Fricker is right when she connects epistemic injustice with other kinds of injustice to which members of disadvantaged groups are subjected. Testimonial injustice is the topic of the next section.

4. Testimonial Injustice

[EP] establishes a strict connection between the attitudes of other people towards an individual and her possibilities to become a full-fledged participant in social activities of a certain kind. The result is the impossibility for the subject to overcome a disadvantaged situation by her own means since self-attribution of knowledge and self-licensing are parasitic activities dependent on the intervention of others.

Alternative approaches to the analysis of knowledge, which pay attention to individuals, their social environments, and the character traits that they develop or inhibit because of their social position and experiences, make more flexible prognoses about people’s options to mend unfavourable conditions. These analyses deserve to be taken into account, if only because of their extraordinary success. In this section and the next one, we will touch upon the main lines of the phenomenon of testimonial injustice.

Fricker defines testimonial injustice as the credibility deficit received by a speaker due to prejudices against her social identity. As she says: ‘(T)he central case of testimonial injustice is *identity-prejudicial credibility deficit*’ (Fricker, 2007, 28). To qualify as a kind of injustice, the epistemic type must be connected to other kinds of injustice—social, political, economic—inflicted to the victim as a member of her identity group. For this reason, the phenomenon of testimonial injustice requires persistent and systematic mistreat. The feature of *persistence* calls for the transversality of epistemic offence across different context and situations; the feature of *systematicity* calls for its structural character. An isolated episode of credibility deficit does not count as testimonial injustice.

The scope of the harm caused by epistemic injustice, Fricker acknowledges (Fricker, 2007, 58), must be determined by empirical research. There is abundant literature that supports the predictions of [EK] at school (Ehrlinger and Dunning, 2003; Huguet and Regner, 2007; Reuben et al., 2012; Estes and Felker, 2012) and outside, in feminist and critical race studies.

The mere mismatch between the credibility received and the credibility a witness deserves, even if systematic and persistent, is not enough for testimonial injustice to arise. As Fricker defines it, no epistemic injustice derives from a credibility excess. Other authors have challenged this claim. Medina, for instance, widens the scope of epistemic injustice and includes the harm socially caused by the perpetuation of oppressive situations (Medina, 2013, 60–1). Excess of credibility can undermine the possibilities of the subject to become a sound epistemic agent by promoting the development of epistemic vices, as we will discuss in the next section. Luzzi (2016), in turn, argues that testimonial injustice may emerge even in cases of accurate credibility attributions.

The epistemic harm inflicted to a subject can be direct or indirect. When discussing hermeneutical injustice—a kind of epistemic injustice caused by the shortage of interpretive, i.e. conceptual, resources to make sense of certain situations that the agent undergoes—Fricker acknowledges a ‘secondary’ kind of epistemic injustice that concerns the practical consequences of the ‘primary’, direct, harm. She describes the case of Carmita Wood, a worker victim of harassment that had to leave her work and could not claim any kind of benefits for the lack of an acceptable way of justifying the reasons that led her to quit her job (Fricker, 2007, 149–50).

Hermeneutical and testimonial kinds of epistemic injustice erode the confidence of subjects in their epistemic abilities. The lack of confidence caused in a subject by her systematic and persistent valuation as an epistemically defective subject is a particularly significant consequence of the epistemic injustice of the primary kind (Fricker, 2007, 174). When the subject is not equipped to understand and explain the injustices inflicted to her, she is deprived of an essential instrument to conclude that she has reasons to feel what she feels and the right to seek reparation. When the subject is systematically negated the status of knowledge giver, she is seriously undermined in her possibilities to develop an appropriate degree of confidence in her epistemic competence (Fricker, 2007, 58, Roessler, 2015).

Cases of epistemic injustice produce the development of epistemic vices and block the development of epistemic virtues, both in victims and oppressors. [EK] thus offers an internalist approach that prescribes corresponding epistemic—and ethical—virtues to balance the effect of those epistemic vices responsible for and caused by epistemic injustices. Even if Fricker stresses that ‘the human condition is, necessarily, a socially situated condition’, and that her philosophical conception aims to analyse ‘what constitutes good epistemic conduct in the socially situated context’ (Fricker, 2007, 176), her approach relies on the development of the appropriate individual attitudes to overcome unjust individual situations.

5. Epistemic Friction

Being a victim of epistemic injustice is compatible with being epistemically benefited in some other features.

In the nuanced approach that Fricker (2007), Medina (2013), and Cassam (2019) offer in their discussion of the causes and consequences of the epistemic kinds of injustice, they acknowledge complex epistemic conditions to the members of different groups. Victims and oppressors develop epistemic vices and virtues, due to their position, that explain why victims may become epistemically apt to a certain extent and oppressors epistemically diminished.

People belonging to dominant groups are prone to be confident and assertive, but also arrogant, insouciant, and narrow-minded. People in disadvantaged groups, by contrast, sometimes develop beneficial character traits, such as humility, open-mindedness and intellectual curiosity (Medina, 2013; Fricker, 2007, 20; Cassam, 2019, 150), together with epistemically disadvantageous features derived from their social position, such as lack of confidence. Medina illustrates the point by mentioning the typical frame of mind of enslaved people, according to Tocqueville’s analyses, and the patent effects in women’s self-confidence as denounced by feminist literature. In both cases, situations of systematic epistemic humiliation may end up in a lack of epistemic cognitive control on the part of the abused subject, who may end up doubting that they possess any cognitive power or even a mind.

Medina calls this extreme case of epistemic self-negation ‘ego-skepticism’: ‘a skepticism about the self, about its capacities and even about its very existence’ (Medina, 2013, 42).

Nevertheless, to draw the right conclusions from the theoretical and experimental work on undervalued populations, it is advisable to avoid both ends of an epistemic attitudinal continuum: neither should we ‘romanticized the predicament of the oppressed’ nor assume that they possess ‘impeccable epistemic characters’ (Medina, 2013, 40). And the same goes for the oppressing subject. The epistemic virtues and vices frequently detected in members of socially significant groups are neither *exclusive* of these groups, nor possessed by *all* individuals in the group, nor *automatically* developed by the mere belonging to the group (Medina, 2013, 43). Other individual traits and experiences may reshape the epistemic condition of particular agents.

Medina, following Wittgenstein, includes in his analysis the notion of epistemic friction. The epistemic character that a subject eventually builds up is substantially affected by the *resistance* that the world and other agents present to their beliefs. Note that our use of the term ‘resistance’ differs from Medina’s use. In his case, ‘resistance’ sometimes presents a political overtone that refers to the rejection by members of oppressed groups to accept the narrative with which dominant groups characterises them. In our case, it is a purely epistemic term that stresses the limits externally imposed for the development of a specific image of the world, understood in a broad sense. Not even mathematicians, who allegedly enjoy a maximal degree of freedom, develop their theories in complete absence of theoretical constraints (Frápolli, 2015, 337). Some friction is needed in order to maintain the subject’s web of beliefs grounded into the world. The willing exposition to the criticism of our own opinions reinforces epistemic responsibility (Medina, 2013, 54). People not accustomed to taking alternative epistemic perspectives into account, e.g. because of their dominant position guaranteeing their opinions to prevail, are less likely to develop the kind of humility and open-mindedness that would otherwise immunise them against what we might call ‘social blindness’. Dominant classes can afford some degree of social blindness since their survival, actual or metaphoric, does not depend on any kind of social negotiation.

Social blindness and a deficit of epistemic responsibility cause epistemic injustices according to Medina although not according to Fricker. The discrepancy is more apparent than real and rests on Fricker’s insistence in restricting epistemic injustices to those practices that cause harm to the speaker, whereas Medina widens the scope to cover also the harm socially caused in the form of perpetuation of injustices (Medina, 2013, 60–1). For our purposes, the debate is inconsequential.

In Medina’s global account, disadvantaged groups present a symmetric situation related to epistemic responsibility: because their members’ survival often depends on an understanding as accurate as possible of the oppressor’s viewpoint, it is in their own best interest to develop social lucidity (Medina, 2013, chapter 5).

The virtue of epistemic responsibility only needs a minimal amount of social and external friction, but it needs some amount. This minimal amount allows members of dominant groups, affected by some kind of social blindness, to develop some degree of responsibility and thus to be able to possess some knowledge about the life of others. On the other hand, the epistemic lucidity that members of oppressed groups develop also informs them of the lack of epistemic acknowledgement that they can expect. This meta-knowledge of the dynamics

of social relations may counterbalance the effects of the different kinds of epistemic injustice that afflict disadvantaged subjects. Thus, their situation is not as desperate as it might be had them been ignorant of the dissimilarities defining opposing social groups. Besides, self-awareness of one's capacity as a full epistemic subject can come from within your own marginalised group of belonging. An example are safe spaces, conceived as controlled communities, deliberately protected from oppressive epistemic practices, and aimed at reinforcing the epistemic agency of marginalised subjects. In safe spaces, women are treated as epistemic subjects by other women, racialised people by other racialised people. And, as we will see in the case of women reporting other people's health conditions, even by members of dominant groups, although restricted to specific topics that derive from the social stereotype of women as carers. The effect of the variety and heterogeneity of epistemic situations within marginalised (and dominant) groups also covers hermeneutical injustice. Mason (2011), for instance, argues that the marginalised group of belonging is sometimes able to provide the hermeneutical resources that dominant communities lack. That they are able to export them to non-marginalised communities is a different story, as we have suggested in §1.

6. Positive Intersectionalism

It is now time to take stock. We have surveyed and endorsed the main lines of two contrasting approaches to knowledge. [EP] sheds an externalist stance on the complex information, partly descriptive and partly normative, that knowledge attributions convey. [EK], by contrast, insists on the effects that sustained epistemic mistreatment produces on individuals.

Unlike [EP], [EK] gathers together heterogeneous positions. Cassam, Fricker, Herbert, Medina, to name some of the authors mentioned in previous sections, do not completely share their overall backgrounds. What they have in common is their interest in epistemic kinds of injustice and their impact on subjects in terms of vices and virtues. Moreover, they promote individualist strategies to repair damages, strategies that, nevertheless, are compatible with denouncing structural injustices and supporting global policies to overcome social injustice.

Part of the theoretical and practical interests of [EK] consists in identifying 'how vices of the mind are obstacles to knowledge or how, as José Medina, puts it, they "get in the way of knowledge (2013, 30)"' (Cassam, 2019, 3). Moreover, [EK] presents a therapeutical derivation that [EP] completely lacks. Cassam (2019, chapter 8), for instance, discusses how self-improvement is possible. There, Cassam argues for what he calls a 'qualified optimism': 'My aim isn't that lasting self-improvement is always possible but that it sometimes is. Some epistemic vices are resistant to self-improvement strategies but such strategies are sometimes available and effective' (Cassam, 2019, 171). A similar insistence on the importance of promoting epistemic virtues is detectable in Fricker (2007, 71) and Medina (2013, chapter 6).

Considered this way, the two positions operate at different levels and show different theoretical and practical concerns. In a relevant sense, their theoretical consequences are orthogonal and should be compatible. Nevertheless, we have been suggesting that this is not necessarily so. There are specific points at which the consequences of [EP] affect the picture developed by [EK]. On the one hand, [EP] offers a principled criterion to highlight the essentially epistemic character of testimonial injustice (Almagro et al., forthcoming),

against those who have doubted that the victims of epistemic injustice actually are wronged in their epistemic rights. On the other, the conceptual priority of third-person attributions over first-person attributions seems to preclude that the capacity to overcome unfavourable epistemic statuses lies within the power of individual agents.

To accommodate the analyses derived from [EK] in a general overview respectful to the externalist perspective developed by [EP] we suggest the addition of a pluralist and dynamic approach to the composition and evolution of social groups. In doing so, we draw on the foundational idea of intersectional feminism (Crenshaw, 1989) to understand the complex nature of social identities. Discrimination does not occur along a single categorical axis, but a variety of mutually interacting or intersecting axes often combines to shape the individual's social identity. Besides, Crenshaw argues, subjects defined by several axes of oppression suffer from a form of discrimination that goes beyond the sum of the basic oppressions.

We extend Crenshaw's insight to cover not only those axes on which discrimination occur, a kind of 'negative intersectionalism', but also the positive epistemic situations that even members of multiply marginalised communities encounter. We call our proposal 'positive intersectionalism', whose leading insight is the acknowledgement that our position as epistemic agents is most of the times a mixture of advantaged and disadvantaged conditions. The complexity of social identities potentially has, we contend, a saving role. Otherwise, as it follows from [EP], we could not grow up as epistemic agents. Besides, as it follows from [EK], all of us present epistemic vices and virtues, and also character traits and epistemic habits that lead to knowledge and get in the way of it.

Positive intersectionalism puts forward two general points. First, social groups present fuzzy boundaries. This first aspect may be difficult to spot due to the underdescribed and schematic nature of the examples sometimes used. The excessive focus on identity aspects, disregarding political ones, illustrates the point. A migrant is only a migrant in her host country, and there are poor and rich migrants, although rich migrants are sometimes referred to by a different term: 'expat'. Moreover, social identities are frequently named by vague predicates: 'being working-class', 'being educated', 'being rich', etc.

The second aspect that we want to highlight is that the relation between individuals and social groups is not one-to-one. Each of us belongs to an array of different communities, with varied epistemic statuses. This aspect has also been suggested by (Medina, 2012, 218–20). One can be a white middle-class educated woman or a wealthy homosexual black man. Positive intersectionalism intends to focus on positive conditions too. Within each group, individuals can occupy different power positions: not all members of disadvantaged groups are at the same risk of suffering from epistemic injustice. Wealthy people are always in better position than poor people, men better than women, white better than black, etc. Some black women also are physicians and mothers. With their patients and children, they are the authority figure and receive the right amount of credence to compensate the deficit they suffer when other roles are more salient.

To these two senses of [PI], one should add the context-dependence of knowledge attributions. Attributions of knowledge depend on contextual factors such as the relative epistemic position of attributor and addressee, the participants in the situation in which knowledge is attributed, but also the topic of attribution.

Manne offers some examples of the perspectival status of individuals as subjects of knowledge. Based on empirical research, Manne explains that a single subject can be respected as an expert about some contents and dismissed as unqualified about some others. This is the case with women who report on health issues of people at their care—elderly people or children—, about which they are considered ‘supremely competent’ (Manne, 2020, 89), while they are scorned when it comes to report on their own health. The studies that Manne discusses ‘suggest that when women try to testify to their pain, they are routinely dismissed by the medical establishment on both of these bases—impugned as incompetent and hysterical, on the one hand, or as dishonest malingerers, on the other’ (Manne, 2020, 86). She also presents results that confirm the diverse composition of groups, which are often treated as homogeneous: ‘According to recent estimates, black women in the United States are some three to four times more likely to die as the result of pregnancy or childbirth than their white counterparts’ (Manne, 2020, 77). Black women are seen as less credible than white women, independently of their social position or economic status, as Serena Williams’s case shows (Rankine, 2014, chapter 2; Manne, 2020, 78). Being a woman is not an epistemically homogeneous category.

[PI], in the sense that we have suggested, explains why members of disadvantaged groups are not condemned as epistemic subjects. Oppressed groups can present a sufficient degree of variability as to permit their members not to receive unjust epistemic treatment throughout all their epistemic interventions. Epistemic statuses should not be thought of as absolute states, granted once and for all and identical across all contexts, but as unstable, context-dependent properties, subject to permanent negotiation. Because epistemic injustice requires systematicity and persistence, as we seen in section 4, but not universality, [EK] can accommodate [PI].

That sound epistemic subjects can belong to marginalised groups does not neutralise the lessons learnt from [EP]. The necessary priority of third-person attributions still implies that subjects systematically deprived of any epistemic acknowledgement, i.e. victims of extreme epistemic violence, cannot become full epistemic agents. Fortunately, as positive intersectionalism suggests, absolute epistemic neglect is not frequent. The standard situation is the one in which most of us usually are, one in which we receive a variable degree of epistemic acknowledgement and epistemic contempt. Without this variability we would be condemned, as [EP] claims. Fortunately, as [EK] shows, some degree of variability that includes beneficial as much as harmful effects is (almost) always present in every human group.

References

- Almagro, M., Navarro, L. & Pinedo, M. (forthcoming), “Is testimonial injustice epistemic? Let me count the ways”, *Hypatia*.
- Ayala, S. & Vasilyeva, N. (2015), “Explaining injustice in speech: Individualistic vs. structural explanation”, in R. Dale, C. Jennings, P. Maglio, T. Matlock, D. C. Noelle, A. Warlaumont & J. Yoshimi (eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, Austin, TX: Cognitive Science Society, pp. 130–135.
- Brandom, R. (2000), *Articulating Reasons: An Introduction to Inferentialism*, Cambridge, MA: Harvard University Press.

- Brandom, R. (2009), *Reason in Philosophy: Animating Ideas*, Cambridge, MA: Harvard University Press.
- Brandom, R. (2019), *A Spirit of Trust: A Reading on Hegel's Phenomenology*, Cambridge, MA: Harvard University Press.
- Crenshaw, K. (1989), "Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics", *University of Chicago Legal Forum*, 140(1), 139–167.
- Cassam, Q. (2019), *Vices of the Mind: From the Intellectual to the Political*, Oxford: Oxford University Press.
- Dotson, K. (2011), "Tracking epistemic violence, tracking practices of silencing", *Hypatia*, 26(2), 236–257.
- Ehrlinger, J. & Dunning, D. (2003), "How chronic self-views influence (and potentially mislead) estimates of performance", *Journal of Personality and Social Psychology*, 84(1), 5–17.
- Estes, Z. & Felker, S. (2012), "Confidence mediates the sex difference in mental rotation performance", *Archives of Sexual Behavior*, 41(3), 557–570.
- Frápolli, M. J. (2015), "Non-representational mathematical realism", *Theoria: An International Journal for Theory, History and Foundations of Science*, 30(3), 331–348.
- Frápolli, M. J. (2019), "The pragmatic Gettier: Brandom on knowledge and truth", *Disputatio: Philosophical Research Bulletin*, 8(9), 563–591.
- Fricke, M. (2007), *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press.
- Gettier, E. (1963), "Is justified true belief knowledge?", *Analysis*, 23, 121–123.
- Hornsby, J. & Langton, R. (1998), "Free speech and illocution", *Legal Theory*, 4(1), 21–37.
- Huguet, P. & Regner, I. (2007), "Stereotype threat among school-girls in quasi-ordinary classroom circumstances", *Journal of Educational Psychology*, 99(3), 545–560.
- Kukla, R. (2014), "Performative force, convention, and discursive injustice", *Hypatia*, 29(2), 440–457.
- Langton, R. (1993), "Speech acts and unspeakable acts", *Philosophy and Public Affairs*, 22(4), 293–330.
- Luzzi, F. (2016), "Testimonial injustice without credibility deficit (or excess)", *Thought: A Journal of Philosophy*, 5(3), 203–211.
- McGowan, M. K. (2019), *Just Words: On Speech and Hidden Harm*, Oxford: Oxford University Press.
- Manne, K. (2020), *Entitled: How Male Privilege Hurts Women*, New York: Crown.
- Mason, R. (2011), "Two kinds of knowing", *Hypatia*, 26(2), 294–307.
- Medina, J. (2012), "Hermeneutical injustice and polyphonic contextualism: Social silences and shared hermeneutical responsibilities", *Social Epistemology: A Journal of Knowledge, Culture and Policy*, 26(2), 201–220.
- Medina, J. (2013), *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*, Oxford: Oxford University Press.
- Ramsey, F. P. (1929), "Last papers: Probability and partial belief", in R. B. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays*, London: Kegan Paul, Trench, Trubner and Company, pp. 256–257.

Rankine, C. (2014), *Citizen: An American Lyric*. Minneapolis, MN: Greywolf Press.

Reuben, E., Rey-Biel, P., Sapienza, P. & Zingales, L. (2012), “The emergence of male leadership in competitive environments”, *Journal of Economic Behavior and Organization*, 83(1), 111–117.

Roessler, B. (2015), “Autonomy, self-knowledge, and oppression”, in M. A. L. Oshana (ed.), *Personal Autonomy and Social Oppression: Philosophical Perspectives*, New York: Routledge, pp. 68–84.

Williams, M. (2001), *Problems of Knowledge: A Critical Introduction to Epistemology*, Oxford: Oxford University Press.

Wittgenstein, L. (1969), *On Certainty*, Oxford: Blackwell.



Las cualidades secundarias y el valor epistémico de la experiencia estética

Secondary Qualities and the Epistemic Value of Aesthetic Experience

NIEVES ACEDO*

Resumen: Entendido lo estético como vinculado al carácter apreciativo de la experiencia sensible, el presente artículo explora su rescate del ámbito de la mera opinión vinculándolo con la intencionalidad de la sensación. Tras definir el problema en la introducción, el texto se centra en las llamadas cualidades secundarias, objeto inmediato de la sensación. Se parte de la hipótesis de que el débil contenido epistémico que se ha atribuido a las cualidades secundarias a lo largo de la historia es responsable de la difícil valoración del juicio estético. Se procede por ello a una reformulación del tipo de noticia que recibimos de dichas cualidades, partiendo de la definición de cualidad secundaria de John Locke, y revisándola a partir de algunos textos de John McDowell y Crispin Wright, en diálogo con la teoría aristotélica de los sensibles propios. El resultado de esta revisión afectará al papel que atribuyamos, dentro del espectro de las disciplinas, a la estética y, secundariamente, al arte.

Palabras clave: teoría de la percepción, cualidades secundarias, estética empirista, epistemología de la sensación, John McDowell, Crispin Wright.

Abstract: Considering Aesthetics as linked to the appreciative nature of sensorial experience, this article uses the concept of intentionality of sensation to rescue Aesthetics from being confined into the scope of mere opinions. After introducing and defining the problem, the text focuses on the so-called secondary qualities, immediate object of the sensation. The hypothesis is that the weak epistemic content attributed to secondary qualities throughout history is responsible for the difficult assessment of aesthetic judgment. A reformulation of the kind of news we receive from these qualities is proposed reviewing John McDowell and Crispin Wright review of Lockean's secondary qualities, in dialogue with the Aristotelian theory of the proper sensible. The result of this review should influence the role we attribute, within the spectrum of disciplines, to Aesthetics and, secondarily, to Art.

Keywords: theory of perception, secondary qualities, empirical aesthetics, epistemology of sensation.

Recibido: 16/01/2021 Aceptado: 26/05/2021.

* Profesora colaboradora en el Departamento de Historia, Arte y Geografía de la Universidad de Navarra. nacedo@unav.es. Líneas de investigación: teoría del arte, estudios curatoriales, arte y sociedad. Publicaciones recientes: Acedo, Nieves (2018): "Los Tapies del Museo Universidad de Navarra: el estilo como frontera entre lo internacional y lo identitario", *Príncipe de Viana*, Año LXXIX, n.º 272, pp. 1307-1322; Acedo, Nieves (2019): "Arte contemporáneo y verdad. Estrategias para el desengaño", *La posverdad o el dominio de lo trivial*, Flamarique y Carbonell (eds.), Encuentro, Madrid, pp.180-194.

1. El problema de la circularidad arte-estética

La estética ocupa un lugar inestable e incómodo en el complejo mapa de las disciplinas debido a la cambiante definición de su objeto. Como teoría de lo bello, según la definición clásica de la belleza como trascendental del ente, formó parte de la ciencia del ser, en cuanto que la belleza es convertible con este último. Lo específico de la estética se buscaba en la diferencia de la propiedad trascendental “bello” con respecto a las otras propiedades (bueno, verdadero...) que son igualmente convertibles con el ser, pero diversas entre sí. Se encontró así que una de las notas diferenciales de lo bello es su carácter relacional. En línea con este aspecto relacional, ya en una época marcada por el giro copernicano hacia el sujeto, la idea de lo bello se buscó investigando el tipo de emoción que despierta. Así, coincidiendo con la definición de la estética como disciplina independiente en el siglo XVIII entró en juego otra idea de lo bello que, siguiendo el cauce del pensamiento moderno, tiempo después acabó por excluir la belleza natural, abriendo paso a la idea de que lo bello tiene su lugar propio en el ámbito del arte. Por su parte, desde las vanguardias, la praxis del arte quiso desvincularse de la conexión con lo bello, de la que tal vez había abusado el academicismo. De este modo, lo artístico desligado de lo bello llegó a adquirir protagonismo en los estudios de estética, hasta el punto de que hoy ya no es fácil diferenciar la estética de la teoría del arte.

En contra de esta identificación, la teoría del arte se separa de la estética precisamente porque acude a ella para evitar la disolución del hecho artístico en las cambiantes configuraciones culturales a la que conduce la mera descripción historicista. Así, el arte busca en la estética una ontología que responda a la pregunta por el modo de ser específico de lo artístico, un modo de ser o “artisticidad” en el que a su vez radica aquello “estético”, objeto de la estética.

Como se ve, la deriva moderna de la estética hacia la teoría del arte cae en un planteamiento circular: el arte queda indefinido porque no se puede clarificar a partir de una estética que busca en lo artístico su propia definición. En otros términos, el núcleo estético de lo artístico exige atribuir a la estética un objeto diverso del arte; “lo estético” no puede ser primariamente “lo artístico”, sino alguna otra cosa. Para superar la aporía de esta circularidad será útil volver al punto de partida de la moderna disciplina estética, que ya no está centrada en la belleza como trascendental, pero tampoco se limita a lo productivo (*poiesis*), sino que está interesada sobre todo en la experiencia o la apreciación (*aísthesis*). Y pensar la experiencia estética nos coloca en el ámbito de la percepción sensible.

Fenomenológicamente, la sensación aparece como base de la experiencia estética en cuanto que reporta una noticia que no conduce a la conceptualización y que, sin ser irracional —sino en todo caso “pre-comprensiva”— percibimos como próxima a lo afectivo. No significa esto una reducción de la experiencia estética a lo inmediatamente sensible, porque a las cualidades, a lo particular de la experiencia del mundo, a las formalidades del aparecer, se añade la mediación de las configuraciones culturales que permiten reconocer en las apariencias otros plexos de relaciones distintos al inmediato manifestarse del objeto que también configuran la complacencia o el rechazo. Se podría definir así un logos de lo estético en el que cada experiencia particular, además de permitir el conocimiento según una razón científica, se formaliza para seguir mediando de modo acumulativo el placer y el displacer de nuestra relación con el mundo. El receptáculo que somos no es por tanto neutro, sino ya configurado y siempre configurándose de un modo que es histórico-contingente.

Dicho de otro modo, la sensación no se agota en la operación de tipo inmanente que es el conocimiento, sino que es materia para procesos transitivos en los que se formaliza y se almacena como material que media las subsiguientes sensaciones.

Ahora bien, aunque transitiva y no inmanente, la percepción sensible no es pasiva, sino que es también operación, inmutación activa. Este carácter activo lo encontramos ya en Aristóteles, para quien la sensación es operación o acto segundo del alma (*DA*, 418^a)¹ y en Tomás de Aquino, que abunda en ello al exigir, para que se dé la sensación, un tipo de alteración activa que llama “espiritual” (*ST*, I, 78, 3). La cuestión es importante porque, como se sabe, no hay objeto intencional sin acto. Por lo tanto, del reconocimiento del carácter activo de la sensación depende la intencionalidad de su objeto. El empirismo representacionista de John Locke desplaza la parte activa de la sensación a la reflexión que acompaña y refuerza la impresión (*E*, II, 9, 15) de tal modo que el pensamiento posterior con frecuencia da por hecho su pasividad. Cabe sin embargo, como veremos, postular la intencionalidad del objeto de la percepción, lo que afecta no sólo al alcance del conocimiento sensible, sino también a la adecuada comprensión del aspecto estético.

Como dijimos, la experiencia estética viene también definida por el componente apreciativo que se advierte en la sensación. Así, por ejemplo, el aprecio por la específica configuración sensible del momento presente, que percibo con un agrado que ignora el simple bienestar: esa vista y su luz cambiante, el peculiar timbre, tono y volumen de esas voces, inseparables en mi experiencia de este frío y este olor particular, aquí, en esta sala. De modo que entendemos lo estético como el aspecto que dota a la experiencia sensible de un elemento apreciativo y judicativo, que tiene su base, como el conocimiento, en la inmutación activa que es la sensación. La cuestión es que mi juicio estético aprobatorio del momento presente, por seguir con el ejemplo anterior, aparece vinculado por una parte con lo particular sensible —resistente a la generalización— y por otro con la mencionada formalización de mi sensibilidad a partir de configuraciones culturales —históricamente relativas—, y ambos factores parecen relegar tal juicio al ámbito meramente subjetivo de las opiniones, donde no es posible encontrar ningún contenido que fundamente la carga de verdad que se pone en juego al juzgar.

Ahora bien, puesto que tanto en el juicio teórico como en el práctico comparece un elemento estético, y siendo así que la vida política, la convivencia social, el reconocimiento de las identidades, etc., están atravesados por cuestiones estéticas, importa llegar a definir el papel de la experiencia estética en el proceso de juzgar. Teniendo en cuenta para ello lo señalado sobre la intencionalidad de la sensación, nos disponemos a utilizar este elemento como palanca de rescate del juicio estético del ámbito de la opinión.

Para ello en lo que sigue nos centraremos en las llamadas cualidades secundarias, que la escuela empirista definió como el objeto inmediato de la sensación, procediendo en primer lugar a describir la taxonomía de las cualidades sensibles de John Locke. En segundo lugar, nos referiremos a la reciente revisión de esta taxonomía por parte de filósofos analíticos, en particular John McDowell y Crispin Wright.

1 En los autores de los que existe una edición canónica de la obra completa se ha optado por citar el texto en castellano (según las ediciones que aparecen en la bibliografía), acompañado de la numeración canónica correspondiente.

Pero antes de proceder a este análisis conviene señalar que la defensa de la objetividad o intersubjetividad de la experiencia estética no es una preocupación nueva o aislada. De hecho, lo general en la reflexión filosófica ha sido no resignarse a la idea de que la determinación estética sea sólo un aspecto concomitante del juicio y, por tanto, potencialmente perturbador, obstáculo a su rectitud. Así, el empeño en rescatar el juicio estético del relativismo subjetivo fue una de las principales tareas a las que se dedicó la estética moderna. Especialmente destacados fueron los esfuerzos en el entorno empirista que intentaron demostrar la posibilidad de un “placer desinteresado” para justificar la imparcialidad del juicio de gusto (desinteresado en cuanto que la voluntad no busca poseerlo).

La teoría kantiana se desmarcó en parte del empirismo, pero asumió algunas de sus soluciones, como la del mencionado “placer desinteresado” que se despierta ante ciertos objetos o representaciones y que permite distinguir lo bello de lo bueno. Así en la *Crítica del juicio* (1790) se dice: “Gusto es la facultad de juzgar un objeto o una representación mediante una satisfacción o un descontento, sin interés alguno. El objeto de semejante satisfacción llámase bello” (KU, §5).

Kant se distancia del empirismo porque sustituye el sujeto empírico por un sujeto trascendental, con lo que integra el juicio de gusto en su esquema trascendental y supera el límite psicológico propio de la teoría del gusto practicada por los empiristas. En la filosofía crítica, el esfuerzo por dotar de principios estables e intersubjetivos al sentimiento de placer forma parte de la búsqueda fundamental de un principio de unidad para lo diverso que dé razón de la teleología del mundo y sirva de puente entre la razón teórica y la razón práctica. Pero la teoría kantiana, por más que alcance la generalización del juicio de gusto a partir del placer desinteresado, se queda en un nivel hipotético que carece de valor epistémico. En la *Crítica del juicio* los juicios de gusto —reflexionantes— y los que establecen conexiones causales en la naturaleza —determinantes— se relacionan formalmente —comparten la estructura teleológica—, pero en la práctica del ejercicio de juzgar permanecen inconexos, por lo que no acaban de superar el nivel de concomitancia arriba aludido.

En el afán de preservar el valor objetivo o transubjetivo de la experiencia estética se ha revelado también útil la descripción fenomenológica, orientada a la percepción de lo que de hecho hay, de lo real en cuanto dado en la experiencia. Sin embargo, su reconocimiento implícito de la intencionalidad de la percepción no basta para superar el relativismo del juicio estético, pues es posible describir lo cultural-artístico desde una visión naturalista que no alcanza a explicar la captación de valor que comparece en la experiencia estética. La antropología cultural, por ejemplo, es prolija en un proceder descriptivo desjerarquizado en el que cualquier resto material es considerado por igual —todo es cultura, dirían Tylor y tantos otros a partir de él—, dando por sentado el relativismo y el evolucionismo².

2 El relativismo y el evolucionismo suponen una renuncia a la intuición multiseccular de que la cultura es algo distinto de la naturaleza. Según un modo de entender las cosas diverso de estas propuestas, la cultura se relaciona con la naturaleza en el sentido de que la cumple. Lejos de ser el mero conjunto acumulado de toda la producción humana, es susceptible de ordenación según el papel que cumpla en la configuración de lo humano, por más que dicha configuración no sea única, ni nunca definitiva.

2. Las cualidades secundarias en la tradición empirista

Abordamos ahora la clasificación de las cualidades en primarias y secundarias que, ya presente en Descartes y Galileo, fue desarrollada y popularizada por el empirismo, especialmente a partir de la clasificación de John Locke recogida del libro II del *Ensayo sobre el entendimiento humano* (1610). La epistemología de Locke responde a un modelo representacionista en el que la sensación media entre las cosas y las ideas simples. Estas últimas son la base de todo el edificio del conocimiento de un modo radical, ya que Locke niega cualquier tipo de idea innata, incluidos los primeros principios, que él considera ideas (*E*, I, 3,1)³. Los sentidos, afectados materialmente por los objetos externos (*E*, II, 8, 12), transmiten a la mente ideas que llamamos cualidades sensibles. Estas son de dos tipos: cualidades primarias o “ideas que adquirimos a través de más de un solo sentido [que] son las del espacio o extensión, de la forma, del reposo y del movimiento” (*E*, II, 5, 1); y cualidades secundarias o “poderes de producir en nosotros diversas sensaciones (...) como son colores, sonido, gustos, etc.” (*E*, II, 8, 14). Para diferenciarlas se vale del concepto de “semejanza”. Las cualidades primarias son semejantes a los cuerpos que las producen, las secundarias son *desem*ejantes a los mismos. Se podría decir que la cualidad primaria es la cualidad en el cuerpo, que el sujeto recibe en la idea simple como cualidad secundaria: “lo que en idea es dulce, azul o caliente, no es, en los cuerpos así llamados, sino cierto volumen, forma y movimiento de las partes insensibles de los cuerpos mismos” (*E*, II, 8,15). No obstante, es importante destacar que en ninguno de los dos tipos hay identidad de la cualidad, ni con el cuerpo (primarias), ni con la idea (secundarias), sino sólo semejanza. El problema que plantea el representacionismo empirista, y que pronto se pondrá de manifiesto, es que establece una brecha entre las cosas y la mente que la idea de sensación como mediación no acaba de salvar. De hecho, ¿con qué compararemos las cualidades primarias o secundarias para poder predicar su semejanza o *desem*ejanza respecto a su referente real, si ellas son el único modo de acceder a las cosas mismas?⁴ Ya en 1710 Berkeley señala en su *Tratado sobre los Principios del Conocimiento Humano* (I, 8) que las ideas sólo se pueden parecer a otras ideas.

John Locke, empeñado en dejar atrás el dogmatismo, explica el entendimiento humano estrenando una nueva terminología que refunda conceptos básicos de la tradición filosófica. Por ejemplo, habla de “cualidades” y no de “sensibles”, sin diferenciar la cualidad de la cantidad e ignorando las categorías aristotélicas. Tampoco el uso que hace del término “idea” tiene nada que ver con el de la tradición platónica. Se trata de nociones nuevas cuyo significado es necesario discernir en unos textos que las usan de un modo no del todo homogéneo. De ahí cierto carácter equívoco en algunas de sus afirmaciones. En cualquier caso, su terminología es la que ha prevalecido en el tema que nos ocupa y la que emplean autores como McDowell y Wright, a quienes nos referiremos después.

3 También para Aristóteles lo sensible es la única fuente de conocimiento: “puesto que, a lo que parece, no existe cosa alguna separada y fuera de las magnitudes sensibles, los objetos inteligibles (...) se encuentran en las formas sensibles. De ahí que, careciendo de sensación, no sería posible ni aprender ni comprender (*DA*, 432a5).

4 En la formulación de Aristóteles, por el contrario, la diferencia entre potencia y acto permite no sólo afirmar la semejanza de la potencia sensible con su objeto, sino incluso su identidad (Cfr. *DA*, 431b20- 25 y 432a5).

Aun suponiendo un modelo epistemológico distinto, la noción de cualidades primarias y secundarias coincide en sus rasgos generales con la división aristotélica entre sensibles propios y comunes y con las categorías de cualidad y cantidad. La sensación para Aristóteles consiste en una alteración cuyo objeto inmediato es el llamado sensible propio o el sensible común, siendo el primero el que es perceptible por un solo sentido, tal como “la visión del color, la audición del sonido y la gustación del sabor” (DA, 418^a20), y el segundo el que es común a todos los sentidos “el movimiento, la inmovilidad, el número, la figura y el tamaño. (...) El movimiento, en efecto, es perceptible tanto al tacto como a la vista” (Ibid.). El sensible propio equivale, por tanto, a la cualidad secundaria y el sensible común a la cualidad primaria.

A su vez, cabe cierta correspondencia de la moderna “cualidad secundaria” con la categoría aristotélica “cualidad”, así como de la “cualidad primaria” con la categoría “cantidad”. Así lo explicitan autores como Descartes o Galileo, que utilizan esta terminología antes que Locke y Berkeley. Señalamos esta equivalencia porque permite advertir el problema que sobreviene cuando se traslada a las categorías aristotélicas el modelo de relación entre cualidades primarias y secundarias que establece el empirismo, y que parece convertir la cantidad en el contenido inteligible de la cualidad. Esta visión prescinde de que la magnitud o cantidad, por más que se manifieste mediada por la cualidad, es también un sensible (sensible común en la clasificación aristotélica), diverso de los aspectos estrictamente inteligibles⁵. Es importante esta aclaración por las implicaciones ontológicas que tiene: o bien se entiende la abstracción racional como de lo sensible en toda su dimensión (propio y común, cualidades primarias y secundarias) o se considera que sólo se abstrae la cualidad para reducirla a extensión. En el segundo caso se acaba por hacer coincidir esta última con la sustancia.

Tal resultado, de importantes consecuencias para la consideración de lo estético, arranca de una confusión entre el orden metodológico de las ciencias experimentales y el orden ontológico. La cantidad, que es objeto de la física matemática, es ciertamente accesible a través de la cualidad. O lo que es lo mismo, la cantidad se percibe cualitativamente. Esto implica la precedencia gnoseológica de la cualidad, entendida como cierta disposición o potencia en el objeto que por la inmutación se actualiza en la sensación, lo que no supone una correlativa precedencia ontológica por parte de la cualidad primaria. Y sin embargo, el énfasis en la objetividad de las cualidades primarias (asimilables en general a los aspectos mensurables de la sensación) como fundamento de la verdad del mundo material, unido al reduccionismo positivista, ha llevado a olvidar el valor intencional de las cualidades secundarias y, con ellas, el de la experiencia estética. El carácter relacional de la cualidad hace que las ciencias positivas prescindan de ella en su estudio de las cosas, alcanzando lo cuantitativo siempre que puede por medios experimentales más fiables. Pero afirmar el carácter disposicional de la cualidad secundaria —causa de su discontinuidad— no implica negarle realidad en la cosa o su reducción a entidad mental. La abstracción que la ciencia hace de la cualidad es

5 Que la magnitud se percibe también sensiblemente lo explica Tomás de Aquino al afirmar que “el sentido se altera de manera distinta por una superficie grande o por una pequeña, pues incluso de la blancura se dice grande o pequeña” (ST, I, 78,3).

un recurso metodológico perfectamente legítimo que no se puede elevar justificadamente a una negación con alcance ontológico.

Como se ve, la diferenciación entre cualidades primarias o secundarias en el marco de la teoría empirista del conocimiento y vinculada al método de la ciencia moderna, deriva en la cultura positivista en una comprensión de lo estético-artístico como opuesto a lo científico y a cierto criterio de verdad.

Por ello, en nuestra búsqueda del valor real de la experiencia estética debemos revisar precisamente el contenido de la sensación, cuyo objeto como se ha dicho es la cualidad. La revisión de la clasificación de las cualidades primarias y secundarias a la que ha procedido la filosofía analítica insiste precisamente en este aspecto.

3. La revisión de las cualidades secundarias por John McDowell

Recurrimos en concreto a la conocida controversia entre John McDowell y Crispin Wright a propósito de la objetividad moral, que se apoya precisamente en el modelo que le presentan las cualidades secundarias.

Ambos autores, que forman parte de la tradición analítica, responsable de la pervivencia en el siglo XX de cierta herencia terminológica del empirismo, retornan a la división de lo sensible en cualidades primarias y secundarias.

McDowell se interesa por la percepción del valor, que considera inherente a la experiencia de la cualidad secundaria, como justificación del valor moral. Wright por su parte lo critica, no tanto por la consideración que hace de las cualidades secundarias como por su aplicación en el caso de la ética. Pero lo que nos interesa aquí no son las cuestiones relativas al valor moral, sino el análisis que hacen de lo que supuestamente es su modelo: la relevancia de las cualidades secundarias en la apreciación estética, inseparable de su valor intencional.

John McDowell participó en 1985 en un volumen escrito como homenaje al desaparecido J. L. Mackie, destacado defensor del escepticismo moral, con un texto titulado “Values and Secondary Qualities”. En ese artículo, McDowell trata de mover el peso de la epistemología del valor de un modelo basado en las cualidades primarias, a otro sustentado por las cualidades secundarias. Lo hace mediante un diálogo crítico con el pensamiento de John Mackie, para quien “the primary-quality model turns the epistemology of value into mere mystification” (1985, 111). El resultado es una profunda revisión de la taxonomía de las cualidades recibida de Locke por la corrección de las ideas de subjetividad y semejanza.

McDowell parte de afirmar que si John Mackie se centra en las cualidades primarias es porque da por bueno el modelo de Locke, para quien las cualidades secundarias inducen a error a la conciencia pre-filosófica, al tender nosotros a atribuirles un papel que sólo es aplicable a las primarias. Pero esta concepción de la experiencia, entiende McDowell, es errónea. Por el contrario, señala que las cualidades secundarias tienen un valor de infalibilidad próximo al del sensible propio aristotélico, para quien sensible propio es “aquel objeto que no puede ser percibido por ninguna otra sensación y en torno al cual no es posible sufrir error” (DA, 418a20). De modo similar, afirma McDowell:

A secondary quality is a property the ascription of which to an object is not adequately understood except as true, if it is true, in virtue of the object’s disposition to

present a certain sort of perceptual appearance: specifically, an appearance characterizable by using a word for the property itself to say how the object perceptually appears. (1985, 111)

Y continua poco después:

Secondary-quality experience presents itself as perceptual awareness of properties genuinely possessed by the objects that confront one. And there is no general obstacle to taking that appearance as face value. An object's being such as to look red is independent of its actually looking red to anyone on any particular occasion; so, notwithstanding the conceptual connection between being red and being experienced as red, an experience of something as red can count as a case of being presented with a property that is there anyway —there independently of the experience itself. (1985, 112)

Este modo de entender la cualidad secundaria, continúa, es coherente con la definición de Locke de estas cualidades como las que tienen “powers to produce various sensations in us”. Pero McDowell diferencia la apariencia perceptiva y la causa física de la cualidad. La cualidad secundaria (la disposición en el objeto) corresponde a la apariencia perceptiva y no debe ser confundida con su causa física. Así, “[n]o doubt it is true that a given thing is red in virtue of some microscopic textural property of its surface; but a predication understood only in such terms —not in terms of how the object would look— would not be an ascription of the secondary quality of redness” (1985, 112). Si la causa física no es la cualidad secundaria, debe haber en esta última algo no reductible a dicha causa física.

Para Mackie, que básicamente sigue en esto a Locke, lo que justifica el calificativo de rojo es una cualidad primaria objetiva. Por lo que las cualidades secundarias siempre son reductibles a cualidades primarias. La consecuencia es que las cualidades secundarias no dicen nada. Pero, arguye McDowell, si esto fuera así, en primer lugar los colores aparecerían como neutrales, cuando en realidad la apreciación de la cualidad secundaria se experimenta siempre de modo valorativo y, en segundo lugar, los colores deberían entonces permitirnos reconstruir las cualidades primarias del objeto del que son apariencia, lo que es imposible, como demuestra el hecho de que en el laboratorio científico se considere a las cualidades secundarias como poco fiables para este tipo de reconstrucción. McDowell afirma que existe una conexión conceptual entre el hecho de que un objeto sea rojo y que aparezca como rojo a un observador y que no hay motivo suficiente para acusar a la experiencia de engañosa. Lo que justifica que yo llame rojo a un objeto, repetimos, no es nada distinto de su capacidad de aparecer como rojo.

En esta argumentación, como se ve, McDowell está reivindicando la intencionalidad y el valor apreciativo que acompaña siempre la percepción de la cualidad secundaria. De modo que hay una especificidad en la cualidad secundaria que está vinculada, no con la propiedad física del objeto, sino con el valor, que corresponde a otro tipo de conocimiento o experiencia. La percepción del valor que acompaña la experiencia de las cualidades secundarias es, según él, su modo específico de ser, aquello no reductible a las cualidades primarias.

Para justificar la objetividad del juicio de valor estético (que pretende, como veremos, aplicar al juicio ético), McDowell tiene que superar la idea de que las cualidades secunda-

rias inducen a error. Idea que se ha recibido del empirismo como un contenido inseparable de su propia definición porque está implícito en la idea de *desemejanza*. Ya hemos visto cómo insiste en afirmar algo así como la infalibilidad de los sensibles propios. Se detiene también a explicar en qué sentido la percepción de las cualidades secundarias es subjetiva. Comenta McDowell que entender lo subjetivo como irreal o ilusorio es un error típico de Locke, asumido por autores como Mackie.

Now this contrast between objective and subjective is not a contrast between veridical and illusory experience. But it is easily confused with a different contrast, in which to call a putative object of awareness ‘objective’ is to say that it is there to be experienced, as opposed to being a mere figment of the subjective state that purports to be an experience of it⁶. (1985, 113-114)

¿Por qué lo subjetivo puede ser real?, porque la relación entre cómo son las cosas y cómo aparecen, a diferencia de lo que afirma el representacionismo empirista, no es una relación de semejanza, sino —cuando es verídica— de identidad (1985, 115). Encontramos de nuevo aquí la afirmación de algo ya presente en Aristóteles, para quien el *cognoscente* y lo conocido son uno en acto. Como se ha dicho arriba, la cualidad secundaria se da en el *cognoscente* en virtud de la operación. En McDowell, la experiencia subjetiva no es opaca, lo aparente o fenoménico no es su límite, sino que está abierta a la realidad.

If we are to achieve a satisfactory understanding of experience’s openness to objective reality, we must put a more radical construction on experience’s essential subjectivity. And this removes an insidious obstacle (...) that stands in the way of understanding how secondary- quality experience can be awareness, with nothing misleading about its phenomenal character, of properties genuinely possessed by elements in a not exclusively phenomenal reality. (1985, 117)

Una vez señaladas en las cualidades secundarias las notas de infalibilidad, intencionalidad, valor e irreductibilidad, McDowell da el salto de la estética a su principal punto de interés, que es la ética. La argumentación que desarrolla en este sentido nos interesa aquí por la crítica que contiene al proyectivismo como único método de verificación, que afecta también a las afirmaciones de Crispin Wright a las que nos referiremos después. En el método proyectivista, indica McDowell, “having one’s ethical or aesthetic responses rationally suited to their objects would be a matter of having the relevant mechanism functioning acceptably” (1985, 122). Es decir, la verificabilidad se basa en una descripción adecuada del proceso que exige al sujeto situarse en un imposible punto de partida externo al juicio. Pero la percepción de la cualidad secundaria y la valoración que cuaja en el juicio estético no responde a ningún proceso. McDowell critica que, en cuestión de ética o estética, este tipo de objetivación termina cayendo en definiciones externalistas de lo bello y de lo bueno.

6 En un trabajo anterior McDowell había abordado con más detenimiento la crítica a la noción de objetividad que emplea Mackie en su argumentación, poniéndola en relación con las tesis de Bernard Williams (Cfr. McDowell, 1983).

How, then, are we to understand this pictured availability of the processing mechanism as an object for contemplation, separated off from the world of value? Is there any alternative to thinking of it as capable of being captured, at least in theory, by a set of principles for superimposing values on to a value-free reality? The upshot is that the search for an evaluative outlook that one can endorse as rational becomes, virtually irresistibly, a search for such a set of principles: a search for a *theory* of beauty or goodness. One comes to count ‘intuitions’ respectable only in so far as they can be validated by an approximation to that ideal. (1985, 122)

Este es el resultado del intento de objetivar la subjetividad. A este método proyectivista McDowell opone un modo internalista que sería el adecuado para los juicios de tipo ético o estético. Según él los valores, del mismo modo que las cualidades secundarias, no son inteligibles fuera de la experiencia que se tiene de ellos, por lo que la atribución de valor ético a una acción surge, como en el caso de la percepción de la cualidad, de cierta modificación del sentido o la sensibilidad (1985, 117). La imposibilidad de error en el caso de la cualidad secundaria es transferida luego al juicio ético para alcanzar una forma de verificación no cognitivista con la que responder al escepticismo empirista de Mackie.

La alternativa al proyectivismo que McDowell propone le sirve tanto para evitar el *anything goes* como la generalización de los valores. Su posición cuestiona la posibilidad de establecer una teoría normativa del buen gusto ético y estético, pero está igualmente lejos de defender la arbitrariedad inapelable del gusto subjetivo.

4. El modelo de verificación proyectivista de Crispin Wright

La idea de una verificación del juicio no cognitivista de McDowell fue contestada en 1992 por Crispin Wright, quien en *Truth and Objectivity* intentó llevar las tesis de McDowell a un modelo proyectivista, tratando de estandarizar, por un lado, al sujeto que emite el juicio y, por otro, la experiencia de la que emerge. Wright reconoce con McDowell que es necesario superar a Locke y dejar de identificar la subjetividad de las cualidades secundarias con su carácter ilusorio o falto de realidad⁷. Acepta también la imposibilidad de una verificación “a posteriori”, pero intenta llegar a una determinación de las condiciones “a priori”. Conviene detenerse brevemente en este intento para mostrar cómo termina de nuevo en otra forma de normalización o estandarización del juicio estético que deshace en cierto modo el carácter siempre particular y “creativo” que McDowell había acertado a ver.

El capítulo de *Truth and Objectivity* titulado “The Euthyphro Contrast” se ocupa expresamente de la teoría de las cualidades secundarias, dejando de lado su valor ontológico para centrarse en los aspectos lógicos o predicamentales de la cualidad. Wright valora los intentos de autores como McDowell por superar el escepticismo antirrealista, pero insiste en la necesidad de formalizar un método de verificación para las cualidades secundarias aceptable en el marco de la filosofía analítica.

7 En este sentido admite Wright: “an object’s secondary qualities constitute material for cognition” (2003, 156-157).

En el diálogo platónico *Eutifrón*, Sócrates presenta a Eutifrón el dilema entre dos alternativas que Wright convierte en modelos de verificación. A saber:

- Algo es piadoso porque Dios lo aprueba, o bien,
- lo que es piadoso, Dios lo aprueba.

La primera, correspondiente a la opción del sofista Eutifrón, seguiría un modelo que Wright llama “proyectivista”, aplicable a las cualidades secundarias; la segunda, defendida por Sócrates en el diálogo platónico, equivaldría a un modelo de verificación “detectivista”, adecuado para las cualidades primarias. Es decir el peso probatorio del juicio del sujeto sobre lo sensible es diverso en uno y otro caso. El tipo de proposición que expresa el juicio sobre una cualidad secundaria (o cualidad, a secas) según Wright es un bicondicional del tipo “ x is Q iff for any S : if S were perceptually normal and were to encounter x in perceptually normal conditions, S would experience x as Q ”.

Si Q es una cualidad secundaria, la mejor opinión (es decir, el juicio realizado en condiciones de percepción estándar por un sujeto con capacidad perceptiva estándar) es la base conceptual para atribuir la verdad dentro de un discurso. En este caso la atribución de verdad es necesaria. En el caso de que Q sea primaria (modelo detectivista), la mejor opinión no determina necesariamente la verdad, por lo que la atribución de verdad a lo propuesto por S es contingente.

Por consiguiente, en este planteamiento es fundamental establecer las condiciones que permiten entender si el juicio particular sobre una cualidad secundaria representa la mejor opinión, y para ello se procede a la determinación suficiente de las condiciones que ha llamado “estándar”, que según él deben ser describibles. El resultado es, finalmente, que no se encuentra un criterio mejor para establecer esas condiciones que la mera estadística, de modo que “it is the judgements of those who are *actually* statistically standard observers, in what are *actually* statistically standard conditions of observation, which count” (1992, 114). La falta de una descripción más segura de la “mejor opinión”, obliga a conceder que la diferencia entre cualidades primarias y secundarias no puede radicar en la contingencia o la necesidad del juicio del sujeto, tal como había establecido al comienzo de su argumentación, puesto que no es posible atribuir necesidad a un modelo estadístico. La diferencia entonces consiste en que las condiciones estándar para el mejor juicio sobre una cualidad, en el caso de las secundarias, pueden ser establecidas *a priori*.

Así, estableciendo que las cualidades secundarias, dependientes de respuesta (response-dependent), son verificables según su concordancia con verdades construidas a priori (1992, 136), y al hacer de dichas construcciones a priori un modelo externo de verificación, Wright cae de nuevo en el tipo de objetivación heterónoma que, según McDowell, no es aplicable a las valoraciones estéticas o morales.

Podemos entonces concluir que la formalización de un modelo analítico de verificación para la predicación de las cualidades secundarias no ayuda a ampliar el estrecho campo de certezas que permite alcanzar el positivismo. Y es que, como se ha señalado, el tipo de verdad atribuible al juicio estético no puede depender de proposiciones unívocas y bivalentes, entre otras cosas porque mantiene siempre su dependencia del carácter particular y posicionado de la experiencia. Se trata entonces de un tipo de verdad cuya predicación no

es verificable de modo teórico, similar en esto a la verdad práctica que se dice de la acción moral. Otra cosa es que sea pertinente el paralelismo que establece McDowell entre el tipo de reacción sensible que comparece en uno y otro tipo de juicio (el moral y el estético). Sin entrar a fondo en esta disquisición, baste señalar que para Wright no existe, en el juez que decide sobre lo bueno, un sistema receptor seguro e infalible (que sí reconoce por tanto en el caso de las cualidades secundarias), sino una competencia adquirida⁸. En línea con la crítica de Wright cabe añadir que la “naturalidad” con la que la sensibilidad reacciona al valor ético no exige ser explicada mediante la existencia de un problemático “sentido moral”, puesto que la noción de “segunda naturaleza” propia del concepto aristotélico de virtud ya da cuenta de ello de modo suficientemente satisfactorio.

5. Cualidad secundaria y verdad

Estamos ya en condiciones de dar un paso más en nuestra búsqueda del valor epistémico de la cualidad secundaria. Si, como hemos visto, la cualidad secundaria cuenta con un sistema receptor seguro e infalible, intencional, que no es sólo material para el conocimiento racional abstracto, podemos ya preguntarnos qué otro tipo de noticia nos trae. Dicho de otro modo, ¿de qué se predica esa verdad encontrada en la experiencia estética?, puesto que no se dice de la acción, contenido del tipo de verdad (práctica) a la que, al descartar su verificabilidad lógica, hemos dicho que se parece la del juicio estético.

En primer lugar, partimos de reconocer que existe en la cualidad secundaria un algo previo distinguible de los conceptos. En Wright se trata de un material que se registra en la conciencia, pero que no tiene valor epistémico alguno mientras no comparece el concepto. La especificidad que Wright encuentra en la sensación de las cualidades secundarias se limita al registro “crudo” (*raw*) que no lleva noticia alguna si no es “cocinado” con el concepto⁹. Por su parte, McDowell ha continuado en toda ocasión defendiendo la epistemología de las cualidades secundarias. Así, por ejemplo, en su *Aquinas Lecture* de 2011, publicada con el título *Perception as a Capacity for Knowledge*.

En segundo lugar, consideramos la defensa que McDowell hace de la valoración como un aspecto intrínseco y definitorio de la percepción cualitativa. Para explicarla nos serviremos

8 En 2003 Wright publicó *Saving the Differences*, en cuyo capítulo “Moral Values, Projection and Secondary Qualities” (pp. 155-182) criticaba la comparación entre cualidades secundarias y valores que McDowell había puesto de moda en los ochenta, porque implica la existencia de un tipo especial de experiencia que sería la experiencia moral, con una especificidad análoga a la que existe para experimentar las cualidades secundarias. La idea de la existencia de un sentido moral ha sido común desde Hume, que decía: “When you pronounce any action or character to be vicious, you mean nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it. Vice and virtue, therefore, may be compared to sounds, colours, heat and cold, which according to modern philosophy, are not qualities in objects but perceptions in the mind” (2003, 156).

9 Lo explica mediante el ejemplo de la reacción de un bebé a un ruido intenso. Y concluye: “If our experience of secondary qualities provides a model of anything, then it is of a notion of experience which is, up to a point at least, raw. (...) It is not merely what we are able to say or judge about it but the experience itself what actually registers in consciousness, which depends, in some measure, on the concepts brings to bear. The point remains: if our experience of secondary qualities is to be the model, then not everything in the manifest character of an experience can be seen as owed to concepts brought to bear” (2003, 167).

de la teoría aristotélica de los sentidos, la idea de proporción y su vinculación con la *órexis* o la estimativa. A continuación explicamos estas nociones, que clarifican y soportan aspectos principales de lo que McDowell defiende.

Como es sabido, Aristóteles define el sentido como una cierta proporción, pero entender correctamente esta definición exige enmarcarla en algunas distinciones previas. En primer lugar, hay que diferenciar lo percibido sensiblemente —propriadamente la cualidad— de la cosa: “el sentido sufre también el influjo de cualquier realidad individual que tenga color, sabor o sonido, pero no en tanto que se trata de una realidad individual, sino en tanto que es de tal cualidad y en cuanto a su forma”. En segundo lugar, dentro del sentido diferencia la potencia del órgano: “El órgano primario es, por su parte, aquel en el que reside semejante potencia”. Aclara no obstante que la diferencia entre ambos (potencia sensible y órgano) es inteligible, no real:

Desde luego que la potencia no se distingue realmente del órgano, pero su esencia es distinta: en caso contrario, el ser dotado de sensibilidad sería, en cuanto tal, una magnitud; y, sin embargo, ni la esencia de la facultad sensitiva ni el sentido son magnitud, sino más bien su proporción idónea y su potencia (DA, 424a25).

Este último párrafo indica, en primer lugar, que lo que define como proporción es la esencia de la facultad o potencia sensitiva, y en segundo lugar que la proporción no es magnitud (lo que llevaría de nuevo a la reducción de la cualidad cuestionada más arriba), sino el sentido de su adecuación. Ahora bien, es sabido que la causa de este sentido de adecuación —o proporción— se suele llamar armonía. Definir el sentido como proporción, y ésta como adecuación introduce la apreciación en la esencia del mismo. Porque, ¿qué es el sentido de la adecuación sino ese afecto o apreciación valorativa que se manifiesta en el placer¹⁰ (o en su contrario)? Podemos decir ya entonces que la percepción sensible manifiesta —actualiza— nuestra adecuación con el mundo¹¹, lo que es coherente con el carácter relacional de la cualidad y con su intrínseco sentido valorativo. La sensación de la cualidad por tanto, aun dependiente de la relación, tiene carácter epistémico y da noticia precisamente de esa relación, pues informa tanto sobre el sujeto como sobre el mundo.

A diferencia de lo que algunos entendieron, la idea aristotélica de armonía, que se manifiesta en la sensación/proporción como placer o displacer, no permite para lo estético el tipo de objetivación normativa que McDowell reprobaría por heterónoma. La proporción es susceptible de modificación pues, como se ha dicho, la alteración de la sensación es origen de procesos transitivos —y no sólo de operaciones inmanentes—: la capacidad de

10 De hecho, en Aristóteles es el placer lo que vincula la sensación con el deseo (*órexis*), lo que no implica contradicción con la idea moderna de placer desinteresado. Tomás de Aquino nombra de modo diverso la *órexis* en los seres irracionales y en los racionales: la *estimativa* como relativa al instinto animal, la *cogitativa* relacionada con la capacidad humana de establecer comparación entre lo almacenado en la memoria. Dicha distinción, igualmente aplicable a la consideración de un placer desinteresado de la existencia de su objeto, sirve también a la defensa del libre arbitrio (ST, I, 78,4).

11 En esta afirmación está implícito el reconocimiento de la apertura universal de la sensibilidad, que es a su vez necesario en la defensa de la sensibilidad como fuente única de conocimiento. En Tomás de Aquino, la clasificación de los sentidos, en razón de la semejanza que permite la relación cognoscitiva, responde a una clasificación previa y real de las cualidades existentes (Cfr. ST, I, 87,1).

placer o disgusto se va ampliando gracias a las mediaciones culturales que modifican constantemente nuestro modo de relación con el mundo. La adecuación se capta como valor por la intermediación de la sensación o por la expansión de las posibilidades de adecuación que permite la mediación cultural. Por eso, a pesar de que el límite orgánico —el exceso o defecto de lo que podemos ver o no ver, oír o no oír— es objetivable, la proporción del sentido que define la armonía no lo es, o al menos no es universalizable.

De modo que aquel “afecto pre-comprensivo” de que se habló al principio, que McDowell relaciona con la captación de valor y que sería lo específico de la experiencia estética, se explica y clarifica a partir de la definición del sentido como proporción. Esta apreciación coincide con lo particular de la sensación, con un material que no se podría llamar *crudo*, si por *crudo* entendemos *neutro*. Si cabe pensar la experiencia estética desde la felicidad o el placer, como hace la teoría estética moderna, es por la unidad del síntoma —el placer— con lo que significa, cierto sentido de adecuación o conveniencia con el mundo que es tan *real* como *relacional*.

6. Conclusión

En las páginas precedentes nos hemos interesado por el estudio de las cualidades secundarias como un medio para desatar la circularidad que nace de la mutua dependencia entre estética y teoría del arte.

Nos hemos servido para ello del enfoque de John McDowell que, por un lado, reivindica al igual que Wright la intencionalidad que Locke y Mackie niegan a las cualidades secundarias, y que, por otro, señala la apreciación como elemento intrínseco y no añadido a la percepción de la cualidad.

La comparación con la teoría aristotélica de los sensibles propios y del sentido entendido como proporción refuerza la intuición contenida en las tesis de McDowell, en cuyo enfoque la percepción sensible queda definida por sus notas de intencionalidad y valoración.

La disolución de la problemática interdependencia de arte y estética a partir de este enfoque se sustenta en dos premisas, asumidas a lo largo de la exposición precedente. En primer lugar, que la estética es el estudio de la *aísthesis*, aspecto de la experiencia cuyo principio es la percepción de la cualidad. En segundo lugar, que lo estético es el principal elemento definitorio de lo que llamamos “artístico”. Hecho esto, a partir del análisis realizado de las cualidades secundarias es posible aventurar la siguiente hipótesis, que deberá ser abordada en estudios subsiguientes: que la “noticia” sensible sobre la adecuación o inadecuación de sujeto y mundo sea la base, esto es, el mínimo necesario, sobre el que establecer el valor artístico.

El arte y la estética han sido relegados en su valor epistémico por la fría razón productiva heredera de la modernidad. Este haber sido relegados encuentra en nuestra “vida alienada” su expresión positiva en el anhelo de mediaciones que nos permitan recuperar el sentimiento de la adecuación. En la necesidad de poseer el presente, sin dilapidarlo en anticipaciones y aplazamientos, concurre la experiencia estética, que no viene para cubrir con un velo de *cualidad* y placer la dura realidad *cuantitativa*, sino a desvelar la verdad de nuestro lugar en el mundo, la posibilidad de habitarlo *adecuadamente*.

Referencias bibliográficas

- Aristóteles, (2000) *Acerca del alma* (citado como *DA*), Traducción de Tomás Calvo Martínez, Gredos.
- Kant, I. (2007), *Crítica del juicio* (citado como *KU*), traducción de Manuel García Morente, Tecnos.
- Locke, J. (1956), *Ensayo sobre el entendimiento humano* (citado como *E*), traducción directa de Edmundo O’Gorman, Fondo de Cultura Económica.
- McDowell, J. (1983), “Aesthetic Value, Objectivity, and the Fabric of the World”, en Eva Schaper (ed.), *Pleasure, Preference and Value. Studies in Philosophical Aesthetics*, Cambridge U. Press, pp. 1-16.
- McDowell, J. (1985), “Values and Secondary Qualities”, en Ted Honderich (ed.), *Morality and Objectivity. A Tribute to J.L. Mackie*, Routledge & Kegan Paul, pp. 110-129.
- McDowell, J. (2003), *Mente y mundo*, Ediciones Sígueme.
- McDowell, J. (2011), *Perception as a Capacity for Knowledge*, Marquette University Press.
- Tomás de Aquino (1956), *Suma Teológica* (citado como *ST*), BAC.
- Wright, C. (1992), *Truth and Objectivity*, Harvard University Press.
- Wright, C. (2003), *Saving the differences. Essays on themes from “Truth and Objectivity”*, Harvard University Press.



Una concepción pragmatista sobre las reglas

A pragmatist conception of rules

PEDRO ANTONIO GARCÍA JORGE*

Una regla, en la medida en que nos interesa, no actúa a distancia.

El lenguaje no ha surgido de un razonamiento.

Wittgenstein

Resumen: La defensa de la normatividad del significado ha derivado en el debate sobre la prioridad metafísica de las reglas o del significado (cfr. Glüer and Wikforss, 2018). Sin embargo, la defensa de la prioridad de las reglas no es más que una variante de intelectualismo y, por ende, está sujeta a las mismas críticas que éste, mientras que la defensa de la prioridad del significado deja sin respuesta a la pregunta metasemántica ¿cómo es qué el lenguaje es significativo? Una concepción pragmatista sobre las reglas permite superar el debate evitando el intelectualismo y proporcionando una respuesta a la pregunta metasemántica.

Palabras clave: reglas, intelectualismo, tesis de la normatividad, significado, pragmatismo.

Abstract: Defending the normativity of meaning has led to the debate about the metaphysical priority of either rules or meaning (cfr. Glüer & Wikforss, 2018). On the one hand, defending priority of rules is just a variant of intellectualism and, therefore, it is subject to the same criticisms. On the other hand, defending priority of meaning leaves the meta-semantic question: how is it that language is significant? unanswered. A pragmatist conception of rules makes it possible to overcome the debate by avoiding intellectualism and providing an answer to the meta-semantic question.

Keywords: rules, intellectualism, normativity thesis, meaning, pragmatism.

Recibido: 10/11/2019 Aceptado: 06/03/2021.

* Profesor del Programa Académico de Filosofía, Universidad Autónoma de Nayarit (UAN), estudiante del doctorado interuniversitario en Lógica y Filosofía de la Ciencia, Universidad de Salamanca (USAL). Correo electrónico: pedroangar@usal.es. Líneas actuales de investigación: Filosofía del lenguaje (expresivismo, inferencialismo, pragmatismo). Publicaciones recientes: Sambrotta, M., & García Jorge, P. A. (2018). Expressivism without Mentalism in Meta-Ontology. *International Journal of Philosophical Studies*, 26(5), 781–800. <https://doi.org/10.1080/09672559.2018.1542278>. Este trabajo fue posible gracias a la financiación para el proyecto: Expresivismos contemporáneos y la indispensabilidad del vocabulario normativo: alcance y límites de la hipótesis expresivista, Ministerio de Economía y Competitividad, Gobierno de España (FFI2016-80088-P). Agradezco a María José Frápolli, Nefalí Villanueva, Manuel de Pinedo, Mirco Sambrotta, a los revisores de *Daimon* y a los integrantes de los grupos de investigación: *Filosofía y Análisis* (UGR) y *Actio linguae* (UAN) sus comentarios y observaciones.

Introducción

La publicación de *Wittgenstein on rules and private language* (Kripke, 1982) posicionó en la discusión filosófica sobre el lenguaje en el siglo XX la tesis de la normatividad del significado. De acuerdo con esta tesis, para explicar la significatividad del lenguaje natural es preciso incluir en dicha explicación el papel que las normas o las reglas¹ juegan al momento de usar *correctamente* una expresión lingüística.

Una forma común de enunciar dicha tesis es:

(TN): sin normas/reglas semánticas no habría significado

Mientras que las reglas en cuestión se suelen enunciar mediante prescripciones del tipo:

(N) Para cualquier hablante *H*, y cualquier tiempo *t*, si ‘verde’ significa *verde* para *H* en *t*, entonces *H* **debe** aplicar ‘verde’ a un objeto *x* sii *x* es verde en *t*.² (Wikforss, 2017, p. 58, énfasis agregado. Cfr. Hattiangadi, 2006, p. 225)

(TN) se puede entender como indicando que las normas son condición previa y necesaria para la determinación del significado (*meaning determining (MD) normativity*). No obstante, es posible pensar que la relación es a la inversa, es decir, que el significado es previo y que, por ende, del hecho de que hay expresiones lingüísticas con significado, ciertas reglas entran en vigor para los hablantes al momento de que usan el lenguaje (*meaning engendered (ME) normativity*). Poder hacer ambas lecturas de (TN), como señalan Glüer & Wikforss (2018)³, es lo que ha dado lugar al debate mencionado. Sin embargo, una vez planteado en estos términos surgen los siguientes problemas:

a) El Problema de la Naturaleza de las Reglas

Si la versión correcta de (TN) es *MD normativity*, se proporciona respuesta a la pregunta metasemántica ¿cómo es que las expresiones lingüísticas adquieren significado? Sin embargo, las reglas se presentan como entidades previas al significado, con cierta autonomía y cuya naturaleza resulta cuestionable.

a) El Problema Metasemántico

Si la versión correcta de (TN) es *ME normativity*, la pregunta que queda sin responder es ¿cómo es que las expresiones lingüísticas adquieren el significado que tienen? Lo único que importa es que del hecho de que hay significado, ciertas reglas entran en vigor para los hablantes.

1 Usaré ambos términos sin distinción.

2 Traducción sugerida por uno de los revisores. El resto de las traducciones para los textos que solo existen en inglés son más.

3 Glüer & Wikforss elaboran su explicación para el caso del significado lingüístico y del contenido mental. Me enfocaré solamente en el primer caso.

La caracterización anterior permite apreciar que no solo el debate acerca de la relación entre normatividad y significado, sino también los problemas derivados comparten un presupuesto común, a saber: que el significado y la normatividad, si bien están relacionados, se pueden entender de forma independiente. Tal presupuesto, sin embargo, resulta incompatible con (TN) en la forma en que fue desarrollada por Wittgenstein⁴ a lo largo de su producción post-tractariana. Desde una perspectiva wittgensteniana lo que (TN) destaca es que ni el significado existe con independencia de la normatividad, ni la normatividad con independencia del significado (*SC* §62), sino que ambos son dos caras de la misma moneda y son, por tanto, inseparables. Brandom (1994), defiende (TN) en este tenor retomando la concepción que Wittgenstein mantuvo sobre las reglas y caracterizándola como una concepción pragmatista (p. 21) para defender: a) que usar un lenguaje tiene que ver más con la adquisición de habilidades sociales o con *saber cómo* usar correctamente dicho lenguaje que con el conocimiento previo de reglas semánticas (*saber proposicional*); b) que las reglas vigentes al momento de usar el lenguaje se encuentran implícitas en nuestras prácticas lingüísticas que no son sino prácticas sociales instituidas a partir nuestro modo de usar palabras y conceptos y de las actitudes que adoptamos respecto a la corrección o incorrección de eso que hacemos.

Lo que defenderé es que aplicar la Concepción Pragmatista Sobre las Reglas (CPSR en adelante) al debate sobre la prioridad de las normas o el significado permite su disolución y, por ende, no da lugar al problema de la naturaleza de las reglas, el cual surge debido a que *MD normativity* no es más que un tipo de intelectualismo, ni deja sin respuesta a la pregunta metasemántica. Mi objetivo, por tanto, es mostrar la utilidad y las ventajas de CPSR frente a cualquier concepción metafísica de las reglas y el significado y, en particular, frente al intelectualismo implicado por *MD normativity*. Esto es lo que haré. En la primera sección argumentaré que *MD normativity* es una variante de intelectualismo aplicado al significado. Para ello, retomaré la caracterización del intelectualismo elaborada por Ryle (1969), así como las críticas que planteó a dicha doctrina. En la segunda sección, presentaré la defensa contemporánea del intelectualismo (*intelectualismo razonable*) por parte de Stanley y Williamson (Stanley and Williamson, 2001) y Stanley (2011a, 2011b, 2011c) y defenderé que las modificaciones que introducen para hacer frente a las críticas de Ryle no logran su cometido. En la tercera sección, consideraré algunos de los argumentos ofrecidos en favor del intelectualismo razonable y presentaré algunas objeciones que se han elaborado o se pueden hacer a estos argumentos. Tales objeciones hacen inviable al intelectualismo, luego, si hay razones para abandonar el intelectualismo, las mismas razones valen para abandonar *MD normativity*, lo cual implica, a su vez, rechazar cualquier concepción de las reglas como algo que precede al significado. En la cuarta sección, presentaré más a detallé CPSR indicando cómo es que, de acuerdo con Brandom (1994), las reglas explícitas (en forma de prescripciones) dependen de una forma más fundamental de reglas implícitas en las prácticas sociales o en lo que hacemos y cómo esto conlleva una respuesta a la pregunta metasemántica.

4 Atribuir a Wittgenstein la defensa de cualquier tesis en filosofía puede inquietar a algunos wittgenstenianos ortodoxos. Sin embargo, permítaseme la licencia para destacar mi compromiso con la lectura que Kripke, así como Brandom, hacen de Wittgenstein, sin que esto implique que Brandom suscribe plenamente la lectura de Kripke.

1. Intelectualismo

En esta sección ilustraré en qué sentido la defensa de *MD normativity* implica una concepción intelectualista respecto a las reglas y su relación con el significado lingüístico. Lo primero que haré será caracterizar el intelectualismo, así como los problemas que implica, posteriormente defenderé que, si hay razones para abandonar el intelectualismo, las mismas razones valen para abandonar *MD normativity*.

La caracterización clásica del intelectualismo fue presentada por Ryle en *El concepto de lo mental* (1969/2009b, pp. 14–20). De acuerdo con Ryle, el intelectualismo sostiene que cada vez que un agente realiza una acción inteligente o que manifiesta inteligencia (así como responsabilidad de tal acción), en la mente del agente sucede una acción adicional de carácter teórico-intelectual, a saber, la acción de consultar, implícita o explícitamente, ciertas proposiciones (principios, máximas o reglas) que *guían* su acción. El intelectualismo implica, por tanto, que las reglas preceden a la acción. Las reglas, mediante su comprensión y aplicación, dictan qué acción realizar y cómo realizarla y, por ende, cumplen también con la función de justificar (las reglas como razones) o explicar (las reglas como causas) dicha acción. Aplicado al caso particular de significado lingüístico, el intelectualismo explicaría la competencia lingüística de los hablantes del siguiente modo: cada que un hablante utiliza correctamente el lenguaje lo hace porque tiene conocimiento de las reglas que determinan el significado particular de una expresión lingüística al momento de ser proferida. Es decir, porque *sabe que*, p. ej., debe aplicar ‘verde’ a un objeto *x* sii *x* es verde. Ahora bien, esto último es precisamente lo que *MD normativity* sostiene. Luego *MD normativity* no es más que un tipo de intelectualismo respecto a las reglas y su relación con el significado lingüístico o, más simple, *MD normativity* implica intelectualismo.

Lo anterior permite ilustrar otra característica clave del intelectualismo: la defensa de que *saber cómo* hacer algo, poner en práctica cierta habilidad, consiste en saber un determinado conjunto de reglas (proposiciones) que sirven de guía para la acción (Ryle, 2009b, p. 18).

Podemos ahora resumir el intelectualismo en dos tesis:

*I*₁: La realización de una acción inteligente y de la cual se es responsable presupone que previamente se ha llevado la acción teórica intelectual de consultar las proposiciones (reglas) que *guían* dicha acción.

*I*₂: El *saber cómo* presupone y puede ser reducido al *saber proposicional*.

El problema con *I*₁ es que implica una regresión al infinito (cfr. Ryle, 1969/2009b, p. 19, Wittgenstein IF §201), mientras que el problema con *I*₂ es que, al identificar el saber hacer con el saber proposicional, se asume que pertenecen a la misma categoría o tipo de conocimiento. Es decir, se comete un error categorial.

La identificación de ambos tipos de saber se debe, señala Ryle, a que el intelectualismo no es más que una variante del dogma del fantasma en la máquina o de la idea de corte cartesiano según la cual, si bien lo mental se distingue de lo físico, hay una interacción causal entre ambos. Frente a tal concepción, Ryle contraargumenta que hay acciones inteligentes que no son reguladas por el conocimiento previo, implícito o explícito, de los principios o normas que las rigen (2009a, p. 228-230) y que, por tanto, *saber cómo* hacer algo no es lo

mismo que *saber que algo es el caso* o saber determinadas proposiciones. Que ambos tipos de saber pertenecen a distintas categorías, señala Ryle, se aclara si consideramos cada uno presenta características propias. En primer lugar, se puede destacar que para aprender/saber cómo hacer algo se requiere llevar a cabo la acción de manera constante y perfeccionar su ejecución, mientras que, para aprender/saber que algo es el caso, basta con, p. ej., memorizar lo que alguien más nos dice o lo que consignan los manuales, libros de texto, lo que dice algún experto, etc. Por ejemplo, decimos que alguien sabe cómo hacer algo, si ejecuta dicha acción de forma exitosa, si es hábil en eso que hace, mientras que decimos de alguien que sabe que *p* simplemente porque puede declararlo con pretensión de veracidad. En segundo lugar, mientras el *saber cómo* admite grados, el *saber que* no presenta tal característica. Debido a esto, es normal decir de alguien que sabe hacer algo en cierto grado, p. ej.: jugar ajedrez, mientras que no parece tener mucho sentido decir de alguien que sabe que *p*, aunque solo lo sabe en cierto grado, tal y como se aprecia en las siguientes oraciones

- (1) Naomi es una jugadora principiante de ajedrez
 (1a) Naomi sabe cómo jugar ajedrez, aunque solo lo sabe en cierto grado (es principiante)⁵
 (1)* Naomi sabe que la reina se puede mover en cualquier dirección, aunque solo lo sabe en cierto grado⁶.

En tercer lugar, alguien puede ser un buen argumentador, p. ej., en tanto que no argumenta falazmente, sin ser capaz de dar cuenta de las reglas de inferencia, del mismo modo que un hablante competente no es aquel que puede dar cuenta de las reglas semánticas o gramaticales, sino aquél que usa el lenguaje adecuadamente aun sin consultar tales reglas o saber que reglas está aplicando.

Respecto de la regresión al infinito implicada por II, el argumento es el siguiente:

Argumento de la regresión:

Si para llevar a cabo una acción se requiere consultar la regla o proposición que la autoriza, entonces tal acción de consultar la regla requiere, a su vez, consultar la regla que autoriza a llevar a cabo la acción de consultar la regla anteriormente consultada y así sucesivamente.

Luego, si la defensa del intelectualismo tiene esta consecuencia indeseable, habría que buscar alternativas que permitan explicar de un mejor modo nuestro actuar.

Una forma de satisfacer este *desideratum* es defender, con Ryle, que saber cómo hacer algo no depende de saber determinadas proposiciones y que esto es así porque ambos tipos de

5 Aunque (1) es más idiomática, (1a) sería equivalente y permite apreciar el contraste entre una atribución de saber cómo y una atribución de saber proposicional.

6 A lo anterior se podría objetar que las atribuciones de saber proposicional son verdaderas o falsas de acuerdo con ciertos estándares y que, por tanto, su verdad no es absoluta tal y como parece presuponer la distinción entre el *saber cómo* y el *saber que* basada en que el primero admite grados. Concedo que las atribuciones de saber proposicional no son absolutas *per se*, sino que lo son respecto a ciertos estándares, no obstante, lo normal es que tales atribuciones se hagan de manera absoluta aun con referencia a los estándares involucrados.

saber son distintos e irreducibles. Sin embargo, otra opción es llevar el anti-intelectualismo de Ryle un paso más allá y defender lo opuesto a lo que defiende el intelectualismo: que el *saber que* es un tipo de *saber cómo*. Esto último es precisamente lo que CPSR defiende (Brandom, 2011, p. 9)⁷. Sin embargo, antes de ver con mayor detalle en qué consiste esta concepción, es preciso decir algo respecto al intelectualismo en su versión contemporánea y sus problemas.

2. El Intelectualismo razonable de Stanley

El intelectualismo ha sido defendido recientemente por Stanley y Williamson (2001) y Stanley (2011a, 2011b, 2011c), aunque con la introducción de algunas modificaciones que buscan hacer frente a las críticas de Ryle.

La tesis I_I es reafirmada por Stanley cuando afirma:

Por ejemplo, cuando aprendiste cómo nadar, lo que pasó es que aprendiste algunos hechos [proposiciones]⁸ sobre la natación. El conocimiento de estos hechos es lo que te proporcionó el conocimiento sobre cómo nadar. Algo similar ocurrió con todas las demás actividades que sabes hacer en la actualidad (2011a, p. vii).

Sin embargo, Stanley (2011a, p. 14) considera que es inadecuado atribuir literalmente I_I al intelectualista dado que 1) no es más que un hombre de paja construido por Ryle y 2) el intelectualismo es compatible con defender que la *guía* que el conocimiento proposicional otorga a las acciones inteligentes es “automática y sin reflexión” (p. 24). Por tal motivo, señala, lo único con lo que un intelectualista *razonable* estaría comprometido es con defender que las acciones inteligentes lo son en tanto que están *guiadas* por conocimiento proposicional, mas no con defender que “toda acción inteligente sea precedida por acciones de decirse a uno mismo determinadas proposiciones” lo cual, para Stanley, sería además de falso, absurdo (2011c, p. 14; cfr. Stanley & Williamson, 2001, p. 415). Así, según el *intelectualismo razonable* propuesto por Stanley, I_I se transforma en:

I_{r_I} : La realización de una acción inteligente y de la cual se es responsable presupone ciertas proposiciones (reglas) que sirven de *guía* a dicha acción.

El paso de I_I a I_{r_I} tiene como finalidad señalar que el argumento de la regresión al infinito no es sólido dado que asume una tesis falsa y, por ende, la amenaza de la regresión al infinito desaparece. Ahora bien, señala Stanley, dado que el regreso al infinito no es inherente al intelectualismo, entonces sigue siendo una opción plausible. No obstante, resulta interesante

7 La mayor parte del tiempo, en efecto, Ryle parece comprometido con defender que ambos tipos de saber son distintos e independientes. Sin embargo, como señala Fantl (2017), por momentos parece comprometerse con la tesis más fuerte de que el saber proposicional es un tipo de saber cómo: “el saber cómo es un concepto que es lógicamente previo al concepto del saber proposicional” (Ryle, 2009a, p. 225). Agradezco a uno de los revisores el hacerme notar está diferencia fundamental entre el pragmatismo/anti-intelectualismo moderado de Ryle y el pragmatismo/anti-intelectualismo fundamental de Brandom.

8 Una proposición verdadera es un hecho para Stanley (2011a, p. vii)

la forma en que Stanley aboga a favor de su propuesta ya que lo que hace es pedir igualdad de condiciones mediante el siguiente argumento *a pari*

Si a Ryle le fue permitido asumir, sin argumento de por medio, que el saber cómo puede manifestarse sin ningún acto previo —sea éste el de considerar una máxima o regla, o la aparición de una representación— entonces, lo mismo debe serle permitido al intelectualista respecto al saber proposicional. (2011c, p. 17)

Decir que Ryle no ofrece argumento alguno es pasar por alto los análisis de casos⁹ que realiza acerca de qué es lo que queremos decir mediante una atribución de saber cómo. Por otra parte, incluso si se concede que, en efecto, el saber proposicional de un agente se manifiesta de forma directa, sin actos de contemplación, esto por sí mismo no prueba que el saber proposicional sirve de *guía* al saber cómo y que, por ende, el intelectualismo es correcto. De hecho, siguiendo el argumento *a pari* al que apela Stanley, se podría concluir que la tesis correcta es la contraria, es decir, que el saber proposicional es un tipo de saber cómo, que es lo que CPSR defiende, en tanto que ambos se manifiestan mediante determinado comportamiento del agente.

Veamos ahora en qué consiste la modificación de I_2 . Para ello mencionaré, primero, una diferencia importante entre el intelectualismo y CPSR relacionada con el uso de las atribuciones de conocimiento. Para CPSR mediante una atribución de conocimiento lo que hacemos es proponer a aquel a quien atribuimos conocimiento como alguien fiable o alguien de quien cabe esperar cierto tipo de comportamiento, motivo por el cual una atribución de conocimiento tiene un uso evaluativo, antes que descriptivo. El intelectualismo, por el contrario, parte del supuesto de que toda atribución de conocimiento es usada para describir una relación entre un agente y una proposición. Este supuesto, atribuido erróneamente a Ryle¹⁰ y sus seguidores, es respaldado por Stanley cuando afirma que

De acuerdo con los seguidores de Ryle, el saber cómo es una relación que se da entre una persona y una acción-tipo [una habilidad] lo cual lo hace distinto del *saber [proposicional]*, que es una relación que se da entre una persona y una proposición. (2011b, p. 226)

Lo que Stanley señala es que para Ryle el saber cómo y el saber proposicional serían dos tipos de estados mentales distintos, siendo el primero de ellos un estado mental no proposicional, dado que aquello con lo que se relaciona es una habilidad, y el segundo un estado mental proposicional, dado que aquello con lo que se relaciona es una proposición. Sin embargo, lo que Stanley y Williamson defienden es que saber cómo hacer x y saber determinadas proposiciones son el mismo tipo de estado mental cuyo contenido es una proposición p sin que esto implique, tal y como establece I_2 , una reducción del primero al segundo (Stanley & Williamson, 2001, p. 434). Por tal motivo, al eliminar dicho aspecto reduccionista, el intelectualismo razonable transforma I_2 en la tesis más débil:

⁹ Como los presentados en la sección anterior.

¹⁰ Tal error está relacionado con el hecho de que para Stanley (2011c, p. 2), Ryle es un conductista lógico.

*Ir*₂: El *saber hacer* presupone y puede ser definido mediante el *saber proposicional*.

Defender esta tesis más débil es lo que lleva a Stanley y Williamson a afirmar que el saber cómo es un tipo de saber proposicional (2001, p. 430) y que, por ende, saber cómo hacer algo es estar en posesión de cierto contenido proposicional (Stanley, 2011b, p. 212). La única diferencia es que, en el saber cómo el contenido proposicional p ¹¹ incluye una forma (*way*) w en la cual uno podría hacer algo. Así, decir:

a sabe montar en bicicleta

equivale, por definición, a decir

a sabe que hay una forma w en la cual a podría montar en bicicleta (Stanley and Williamson, 2001, p. 427).

Lo mismo valdría en el caso de la competencia lingüística ya que:

dado un término t con cierto significado s , que un agente a sepa cómo usar t con el significado s equivale [...] a que a sepa que hay una forma w en la cual a podría usar t con el significado s (Stanley and Williamson, 2001, p. 444).

Luego, la dicotomía ryleana entre *saber cómo* y *saber que* es incorrecta (Stanley, 2011b, p. 207- 209). La única *diferencia* entre ambos tipos de saber es que en el primero tal proposición p (que incluye una forma w de hacer x) es considerada mediante un *modo de presentación práctico*¹² (Stanley & Williamson, 2001, p. 430), mientras que en el segundo es considerada mediante un modo de presentación descriptivo. Ahora bien, señala Stanley “considerar algo bajo un modo práctico de presentación es tener cierto complejo de disposiciones hacia esa cosa” (2011b, p. 211). Es decir, el conocimiento o la consideración de una proposición mediante su modo de presentación práctico se *manifiesta* en el actuar de los agentes en tanto que saben cómo hacer algo o tienen ciertas habilidades, p. ej: conducir una bicicleta (Stanley, 2011a, p. viii) o usar el lenguaje con determinados significados.

En el plano lingüístico, por su parte, lo que se dice mediante una atribución de *saber cómo* es que alguien considera o piensa una proposición p mediante tal modo de presentación. Se atribuye un determinado contenido proposicional mediante un modo de presentación práctico y por ello toda atribución de saber cómo es en realidad una atribución de saber proposicional (se atribuye al agente que sabe que hay una forma w de hacer x) y no una atribución de una habilidad. Por otra parte, al decir que A sabe hacer x con un grado mayor de complejidad que B , lo que hacemos es comparar los modos en que tales agentes saben cómo hacer x (los modos de presentación prácticos) y declarando uno mejor

11 Donde p es una proposición russelliana, es decir una entidad estructurada o compuesta por objetos y propiedades.

12 Stanley (2011a) prefiere *practical way of thinking* en lugar de *practical mode of presentation*, sin embargo, con ambas expresiones se alude al modo en que una proposición es considerada, a saber, como una proposición/regla que guía la acción.

que el otro. Por tanto, saber cómo hacer x no consiste en tener una habilidad práctica, sino en una relación entre un agente y una proposición, en concreto la relación de considerar (*entertain*) ciertas proposiciones sobre dicha actividad (Stanley, 2011b, p. 212) mediante un modo de presentación práctico.

3. Problemas del intelectualismo razonable

En esta sección consideraré algunos de los argumentos presentados a favor del intelectualismo razonable, así como algunas objeciones que se han presentado o se pueden hacer a dichos argumentos.

En apoyo de la equivalencia por definición entre las atribuciones de saber cómo y de saber proposicional, Stanley destaca que su propuesta es compatible con los análisis estándar que la lingüística contemporánea¹³ ofrece en términos puramente proposicionales para las oraciones que incluyen una pregunta indirecta¹⁴ (2001, p. 417-420, 2011b, p. 221-226), como es el caso de las atribuciones de saber cómo. De acuerdo con este análisis, una oración del tipo “ a sabe cómo montar en bicicleta” no tiene como constituyente a la expresión “sabe cómo” y, por tanto, su complemento no es un verbo en infinitivo mediante el cual se atribuye una habilidad. Antes bien, el constituyente de tal oración sería “ a sabe” mientras que su complemento sería aquello que a sabe, es decir, una proposición p (que en este caso sería “hay una forma w en la cual a podría montar en bicicleta”). Por tanto, postular, como hace Ryle, una distinción entre ambos tipos de saber sería ir en contra de las teorías sintácticas y semánticas contemporáneas¹⁵ sobre las preguntas indirectas. Debido a ello, señala Stanley, para mostrar que el intelectualismo es incorrecto se tendría que mostrar que existe una explicación en la que las atribuciones de saber cómo no sean consideradas como preguntas indirectas (2011b, p. 221).

Bastaría pues con mostrar que hay, al menos, una explicación distinta desde las teorías semánticas y sintácticas contemporáneas a la explicación que desde el intelectualismo razonable se da del saber cómo en términos puramente proposicionales. La evidencia de que estas explicaciones existen y que, por ende, la explicación adoptada por Stanley ni es tan estándar como supone, ni brinda evidencia concluyente a su favor, ha sido presentada por Abbott (2013). En este trabajo, Abbot menciona, entre otros, el análisis que, p. ej., desde la lógica intensional de Montague, se puede hacer de las oraciones que atribuyen saber cómo. Desde tal perspectiva es posible explicar por qué estas oraciones, a diferencia de las atribuciones de saber proposicional, admiten grados de comparación, como cuando decimos de alguien que sabe demasiado acerca de cómo hacer algo, p. ej.: jugar tenis, con la intención de destacar que es sumamente hábil.

13 Lo cual no deja de ser sorprendente ya que si algo hemos aprendido de los desarrollos en filosofía del lenguaje es que la forma gramatical puede ser engañosa. Dejarnos guiar exclusivamente por la forma gramatical de una oración es lo que ha llevado a asumir que, dado que hay términos que aparecen como sustantivos en una oración, tales términos tienen que referir a una sustancia o entidad. Un ejemplo claro de ello es la noción de verdad que al aparecer como sustantivo en oraciones del tipo “La verdad nos hará libres” nos lleva a pensar que hay una entidad, *la verdad*, que tiene la propiedad de conducirnos hacia la libertad.

14 Debido a que incluyen lo que en inglés se denomina una *wh-word*.

15 Por ejemplo, la teoría de Karttunen (1977) y la de Groenendijk and Stokhof (1982, 1984).

(2) Federer sabe demasiado acerca de cómo jugar tenis

En casos como (2) la expresión *cómo jugar tenis*, al ir después de la preposición *acerca de*, tendría, en lugar de una función oracional, una función nominal dado que indica aquello en lo cual el agente es sumamente hábil y, además, sirve de complemento para la preposición *acerca de* (Abbott, 2013, pp. 4-5). Por tanto, dado que hay explicaciones en términos no proposicionales de las atribuciones de saber cómo, el intelectualismo, incluso en su versión razonable, sigue siendo, si no falso, al menos cuestionable.

En cuanto a los modos de presentación se puede señalar lo siguiente. Decir que las proposiciones involucradas (aquellas que incluyen una forma w en la cual uno podría hacer x) en una atribución de saber cómo requieren ser pensadas o considerados por un agente mediante un modo de presentación práctico remite inevitablemente a la tesis I_1 pese a la modificación que de esta tesis hace Stanley. Por otra parte, la introducción de los distintos modos de presentación (práctico y descriptivo), además de ser poco clara, parece introducir una *diferencia* entre el saber proposicional y el saber cómo (Noë, 2005, p. 287), lo cual choca con la tesis intelectualista de que no hay diferencia entre ambos.

Otro problema adicional, reconocido por Stanley, es que “es difícil definir un modo práctico de presentación” (2011b, p. 212; cfr. Stanley & Williamson, 2001, p. 429), aunque, en su opinión, son necesarios para explicar por qué las atribuciones de saber cómo son de fondo atribuciones de saber proposicional. Sin embargo, parece que su función es la misma que la que el intelectualismo en la caracterización de Ryle otorgaba a las reglas. Por tal motivo, una explicación que prescindiera de ellos y que a la vez tenga al menos el mismo o, incluso, mayor poder explicativo resulta más plausible. Como veremos más adelante, la adopción de CPSR permite satisfacer tal *desideratum* al explicar el saber cómo mediante la práctica constante o el entrenamiento de los agentes que interactúan en sociedad.

Otro de los argumentos presentados por Stanley *versus* la distinción entre el *saber proposicional* y el *saber cómo* es el que explota la posibilidad de que las atribuciones de *saber cómo* tengan un uso distinto al de atribuir habilidades. El argumento procede como sigue:

Si de acuerdo con Ryle, lo que permite distinguir el saber cómo del saber proposicional es que el primero consiste en tener determinadas habilidades y no en saber que p , entonces, al atribuir saber cómo mediante una oración, estaríamos usando dichas oraciones exclusivamente para atribuir habilidades. Sin embargo, “es a todas luces falso que toda atribución de saber cómo se use para atribuir habilidades” (Stanley and Williamson, 2001). Luego, no hay distinción entre ambos tipos de saber.

Lo primero que se puede señalar ante tal argumento y la conclusión extraída es que lo que de hecho es falso, o al menos cuestionable, es que Ryle defendiera que toda atribución de saber cómo sea usada para atribuir habilidades. Defender que hay diferencias entre el saber cómo y el saber proposicional es compatible con defender que las atribuciones de saber cómo no tengan un uso uniforme y que, por tanto, haya casos en que se usen para atribuir saber proposicional. Partiendo de lo anterior, lo que habría que decir es que mediante una atribución de saber cómo, lo que hacemos (generalmente, pero no siempre) es decir que alguien tiene determinada habilidad que pondrá en práctica en las circunstancias adecuadas. Por tanto, es un error pensar, como hacen Stanley y Williamson, que por el hecho de que en una oración aparezca la expresión *sabe(r) cómo* dicha oración sea usada,

en toda situación, del mismo modo. Otro uso posible de estas oraciones es el de hacer una atribución de saber proposicional en el sentido de que mediante una atribución de saber cómo lo que se dice, de forma abreviada, es que alguien *sabe todo lo que hay que saber acerca de cómo hacer algo*. Lo anterior es distinto a decir, como hace el intelectualista, que toda atribución de saber cómo equivale a atribuir saber proposicional.

La consideración de la posibilidad de usos no estándar de las atribuciones de saber cómo tiene la ventaja de que permite explicar por qué hay casos en los que alguien puede saber todas las reglas o principios implicados en una actividad y, pese a ello, carecer de la habilidad para realizarla. Por ejemplo, en ciertos contextos podemos decir de Juan, un físico especializado en hidrodinámica, que sabe cómo nadar, dado que sabe todos los principios hidrodinámicos que hay que saber para nadar eficientemente, aunque Juan no sea capaz de nadar. Por supuesto, el intelectualista podría decir en su defensa que lo que *él* defiende es que para que Juan sepa cómo nadar, no basta con que sepa todos los principios hidrodinámicos implicados en la natación, sino que además debe considerar dichos principios mediante un modo de presentación práctico. Sin embargo, ello implicaría que Juan, que es incapaz de nadar, solo considera o piensa tales proposiciones mediante un modo de presentación descriptivo antes que práctico. Ante ello, resulta más intuitiva la explicación de que, aunque Juan está en posesión de cierto saber proposicional, Juan no es hábil nadando, mientras que la habilidad de Luis, un nadador experimentado, pero que no sabe gran cosa de hidrodinámica, es resultado de la práctica constante antes que de su conocimiento de ciertas proposiciones sobre la natación¹⁶. Luego, *pace* Stanley y Williamson, resulta poco intuitivo que el saber proposicional sea necesario para el saber cómo y que este sea una especie del primero.

4. Concepción pragmatista sobre las reglas

Tras haber caracterizado el intelectualismo y señalado los problemas que conlleva, incluso en la versión *razonable* de Stanley, retomaré ahora el esbozo de CPSR presentado al final de la introducción. El objetivo será indicar con más detalle en qué consiste la concepción de Wittgenstein sobre las reglas y de acuerdo con la cual Brandom propone que las normas explícitas (en forma de prescripciones) dependen de una forma más fundamental de normas que están implícitas en lo que hacemos, antes que en lo que declaramos o decimos.

Aunque Wittgenstein no hace uso de la distinción ryleana entre *saber cómo* y *saber que*, la idea de que una acción inteligente es el resultado del saber cómo antes que del saber proposicional es un punto de partida común de ambos. A partir de esta idea Wittgenstein desarrolla su análisis pragmatista (anti-intelectualista) sobre el seguimiento de reglas en gran

16 Este ejemplo es una adaptación del presentado por Fridland (2015, p. 9-10). Fridland considera el caso de Bela Karoli, entrenador de Nadia Comăneci y destaca que, si bien Karoli poseía todo el conocimiento proposicional sobre la gimnasia, no por ello sabía cómo hacer gimnasia. Sin embargo, el caso de Karoli es compatible con que el saber que poseía fuera más un tipo de saber cómo derivado de la observación acerca de cómo hacer ciertos movimientos en la gimnasia, p. ej., un salto en el potro. Es decir, Karoli no necesariamente sabía los principios (físicos, biomecánicos, etc.) implicados en la ejecución de la gimnasia, aunque sí que sabía cómo enseñar a hacer gimnasia.

parte de su producción post-tractariana, siendo §232 de *Investigaciones Filosóficas (IF)* una muestra clara de su anti-intelectualismo respecto al seguimiento de reglas¹⁷:

Supón que una regla me inspira cómo debo seguirla; es decir, cuando recorro la línea con los ojos, una voz interior me dice: «¡Tira así!» — ¿Cuál es la diferencia entre este proceso de seguir una especie de inspiración y el de seguir una regla? Pues seguro que no son lo mismo. En el caso de la inspiración aguardo instrucciones. No podré enseñarle a otro mi ‘técnica’ de seguir la línea. A no ser que le enseñe una especie de saber escuchar, de receptividad. Pero entonces no puedo, naturalmente, exigir que él siga la línea como yo.

Aunque §232 no hace mención directa del intelectualismo se puede afirmar que lo que Wittgenstein destaca es que, si tal postura fuera correcta, entonces cuando decimos de alguien que está siguiendo una regla, estaríamos diciendo que hay una regla que inspira (guía) su actuar. Esto implicaría que no habría manera de determinar que alguien sigue correctamente una regla, sino que cada uno tendría su propia interpretación de la regla. Pero, dado que hay casos en los cuales determinamos la correcta aplicación o seguimiento de una regla, hay que descartar el intelectualismo y presentar una explicación alternativa respecto a cómo seguimos reglas.

Wittgenstein pone de manifiesto su concepción pragmatista sobre las reglas analizando diversas acciones en las que es factible decir que hay reglas vigentes al realizarlas, p ej.: jugar ajedrez, continuar una serie numérica, leer (mecánicamente), usar una palabra o una expresión lingüística compleja —oraciones o frases—, con la intención de significar algo, etc. Mediante estos análisis destaca que lo común a cada una de esas actividades es que: a) pueden ser consideradas como juegos cuya realización exhibe un patrón de comportamiento¹⁸; b) requieren de adiestramiento, formal o informal, sobre cómo han de realizarse; c) su realización suele estar sujeta a revisión por parte de aquellos que llevan a cabo el adiestramiento.

Por ejemplo, señala Wittgenstein, cuando aprendemos a continuar una serie numérica, se nos enseña a continuar la serie de una forma y no de otra, se nos dice que *continuar de esta forma*: p. ej., obtener 102 después de 100 y 104 después de 102 es la forma *correcta* de continuar la serie de 2 en 2. Sin embargo “el de esta forma (en “continuar de esta forma”) se designa mediante una cifra, un valor. Pues en este nivel, la expresión de la regla se explica por el valor, no el valor por la regla” (Z §301). Es decir, obtener 104 después de 102, al continuar con la serie, explica la regla de continuar de 2 en 2 y no a la inversa y esto es así debido a que el entrenamiento en el uso de la regla es la forma en la cual se aprende una

17 El tema es abordado, principalmente, en *Observaciones sobre los fundamentos de la matemática* y en *Investigaciones filosóficas*. Sin embargo, aparece también en obras previas — *Cuadernos Azul y Marrón* — y posteriores — *Zettel (Z)* §§293-321 y *Sobre la certeza (SC)* §§ 28-29, 45-47, 61-62, 95 —. No obstante, es usual que al abordarlo se haga referencia a *IF* §§138-242 de, aunque es introducido ya en §§80-86 y aparece a lo largo de esta.

18 Lo cual no implica que siempre que una acción sea acorde con una regla, sea una instancia de aplicación o seguimiento de reglas. Una acción puede estar de acuerdo con una regla por mera casualidad. Por ello, la regularidad o patrón de conducta es un criterio necesario, aunque no único, para excluir tales casualidades.

regla (*IF* §201, 208, *SC* §§28-29, 45-48, *Z* §318; Kripke, 1982, p. 87). Luego, seguir una regla, no consiste en decir que regla se está aplicando (*SC* §§431-434), sino en saber cómo hacer algo (Ryle, 2009a, p. 227; Ryle, 2009b, p. 283).

Así, mientras que para la concepción intelectualista sobre las reglas lo que hacemos es aprender/comprender/interpretar y *representarnos*¹⁹ la regla para luego aplicarla (siendo esto un logro individual), para CPSR lo que aprendemos no son las reglas, sino su aplicación; aprendemos a usar la regla, no lo que dice la regla acerca de cómo ha de ser aplicada en la práctica, y esto constituye, no un logro individual, sino un logro del individuo que interactúa en sociedad. Las reglas son algo que se explica, ejemplifica y modifica mediante la regularidad exhibida en nuestras prácticas sociales (*IF* §83, *Z* §§ 301-305). Por tal motivo, seguir una regla es tanto una costumbre —algo que está instituido—, como una práctica social —algo para lo cual somos adiestrados mediante ejemplos y ejercicios— y que implica un tipo de certeza pragmática en tanto que al seguir una regla llevamos a cabo una acción que no requiere de justificación adicional a la acción (*SC* §§ 110, 307, 511; Kripke, 1982, p. 87).

No obstante, que seguir una regla sea una manera de actuar que no requiere de justificación adicional a la acción no implica que en ocasiones no recurramos a explicaciones del tipo ‘pensé en cómo hacerlo...’, ‘recordé que era mejor...’, ‘sentí que era lo que tenía que hacer’, ‘en el manual de procedimientos se indica que...’, etc., y que tales justificaciones no desempeñen papel alguno en determinadas circunstancias. Lo único que CPSR destaca es que:

- (a) las explicaciones de corte internista (basadas en nuestras experiencias o sensaciones individuales) no fundamentan, ni justifican que una acción está de acuerdo con una regla
- (b) en algún momento estas justificaciones se nos acabarán
- (c) llegado ese momento, solo nos quedará decir, así es como he aprendido a actuar (*IF* §217).

Partiendo de lo anterior, se puede afirmar que el criterio de comprensión de una regla es su aplicación *correcta* en un número considerable de casos y en determinadas circunstancias. La *evaluación* de que nuestro actuar es acorde con la regla en cuestión viene dada por los demás y es por ello por lo que se trata de un criterio de índole externa, no interna (*IF* §§ 143, 146). Son los otros los que juzgan si nuestra conducta es acorde con una regla determinada o si estamos aplicándola *correctamente*. Sin embargo, la acción de juzgar la conducta de alguien más como correcta/incorrecta, no es llevada a cabo mediante la acción adicional de consultar el catálogo de reglas para ver si mi conducta es acorde a las mismas, ni tampoco porque haya una regla que sirva de guía a dicha acción. Juzgar una acción o conducta como correcta/incorrecta es hacer algo, no saber que algo es el caso.

Antes de finalizar, y para ahondar en el carácter normativo de las reglas, conviene aclarar que, aunque CPSR toma como punto de partida las regularidades presentes en nuestras prácticas sociales, ello no quiere decir que las reglas se reduzcan y equiparen

19 El intelectualismo podría englobarse en lo que Matthews (2007, pp. 19–26) denomina *the received view* en tanto que las reglas tienen que ser representadas en el sistema cognitivo del individuo.

con dichas regularidades. Si fuera así, de ello se seguiría que cualquier comportamiento regular de cualquier objeto en el mundo natural, p. ej., la caída de los cuerpos atraídos por la gravedad hacía el centro de la tierra, tendría que ser considerado como una práctica normativa o, por el contrario, que las reglas solo serían relevantes para describir tales regularidades y, por ende, no tendrían carácter normativo (Brandom, 1994, pp. 26–30). Para evitar tal lectura hay que destacar que lo que CPSR defiende es que las reglas y su carácter normativo son el resultado de nuestra concepción o reconocimiento implícito de las regularidades de tales prácticas o comportamientos sociales. Son pues, la explicitación en términos lingüísticos y proposicionales de las regularidades exhibidas en lo que hacemos, es decir, en nuestras prácticas sociales (Brandom, 2000). Derivamos, instituímos, las reglas a partir de nuestras regularidades de comportamiento, pero en estas regularidades no hay regla alguna (SC §130-131). El carácter prescriptivo de las reglas se deriva de la actitud normativa que adoptan los integrantes de una comunidad²⁰ (agentes externos) ante las prácticas de esa comunidad. Adoptar dicha actitud es, a su vez, hacer algo, es juzgar que ciertas actividades son las que *deben* de ser continuadas (Brandom, 1994, pp. 29–35, 60). Es adoptar una actitud de compromiso ante las prácticas sociales establecidas²¹, siendo una de tales prácticas la de usar el lenguaje con ciertos significados. Las reglas semánticas, no son, por tanto, una simple descripción sobre cómo se usan las expresiones lingüísticas, sino que mediante estas los hablantes prescriben cómo *debemos* usar el lenguaje²². Sin embargo, reconocer cierto tipo de fuerza normativa a las reglas semánticas no implica que éstas determinen cómo actuaremos de hecho, ya que sí así fuera, se eliminaría la posibilidad de usar incorrectamente una expresión lingüística y, por ende, el carácter normativo del significado. (Brandom, 1994, p. 14).

5. Conclusión

La aplicación de CPSR al caso particular del significado lingüístico remite inevitablemente a la tesis de que usar un lenguaje con determinados significados es, como cualquier otra habilidad, una habilidad adquirida y desarrollada mediante la práctica. Adoptar este punto de partida, a diferencia del punto de partida del intelectualismo implicado por *MD normativity*, respeta la intuición, que como hablantes tenemos, de que aprender a hablar un lenguaje es aprender una serie de habilidades sociales, no un sistema de reglas/proposiciones previas sobre el lenguaje. Lo anterior es especialmente notorio en el caso de los niños cuando aprenden a hablar y en cómo están sujetos, por parte del resto de hablantes, a la evaluación o corrección acerca del uso que hacen del lenguaje. La reflexión acerca de cómo y para qué usamos el lenguaje es una actividad posterior. Mediante tal reflexión se formulan y hacen explícitas las reglas sobre el uso del lenguaje. Esta explicitación de las reglas requiere su enunciación en forma de prescripciones, mientras que la práctica de

20 No se trata por tanto de un enfoque comunitarista de acuerdo con el cual la comunidad determina la corrección/incorrección de su actuar. La comunidad como tal carece de actitudes normativas. Son los individuos pertenecientes a una comunidad los que adoptan actitudes normativas (Brandom, 1994, pp. 38–39)

21 Sin que esto implique que las prácticas sociales sean inmutables o incuestionables.

22 No es que exista La Forma Correcta de usar el lenguaje, sino que reconocemos y aceptamos determinados usos del lenguaje como correctos de acuerdo con ciertas *prácticas locales* (Pennycook, 2010, pp. 1-10)

usar un lenguaje conlleva un reconocimiento implícito de que la forma en la que de hecho es usado el lenguaje es la forma en que *debería* ser usado por el resto de los hablantes. Luego, las reglas semánticas no son previas al significado, ni lo constituyen, *pace MD normativity*, y, por ende, el problema de la naturaleza de las reglas desaparece. Aunado a lo anterior, otra consecuencia positiva de adoptar CPSR es que, *pace ME normativity*, es posible derivar una respuesta a la pregunta metasemántica ¿cómo es que las expresiones lingüísticas adquieren o tienen el significado que tienen? De acuerdo con CPSR el significado lingüístico es instituido a partir de nuestras prácticas sociales en general y de nuestras prácticas lingüísticas en particular.

Referencias

- Abbott, B. (2013) 'Linguistic solutions to philosophical problems: the case of knowing how', *Philosophical Perspectives*. doi: 10.1111/phpe.12019.
- Brandom, R. (1994) *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- Brandom, R. (2000) *Articulating Reasons: an introduction to inferentialism, Articulating Reasons*. Harvard University Press. doi: 10.2307/j.ctvjghvz0.
- Brandom, R. (2011) *Perspectives on Pragmatism: Classical, Recent, and Contemporary*. 1st edn. Cambridge, Massachusetts: Harvard University Press.
- Fantl, J. (2017) 'Knowing How', *The Stanford Encyclopedia of Philosophy*. Fall 2017. Edited by E. N. Zalta. Metaphysics Research Lab, Stanford University.
- Fridland, E. (2015) 'Knowing-how: Problems and Considerations', *European Journal of Philosophy*, 23(3), pp. 703–727. doi: 10.1111/ejop.12000.
- Glüer, K. and Wikforss, Å. (2018) 'The Normativity of Meaning and Content', *The Stanford Encyclopedia of Philosophy*. Spring 2018. Edited by E. N. Zalta. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/spr2018/entries/meaning-normativity/>.
- Hattiangadi, A. (2006) 'Is meaning normative?', *Mind and Language*, 21(2), pp. 220–240. doi: 10.1111/j.0268-1064.2006.00312.x.
- Kripke, S. A. (1982) *Wittgenstein on rules and private language: an elementary exposition*. Cambridge, Massachusetts: Harvard University Press.
- Matthews, R. J. (2007) *The Measure of Mind: Propositional Attitudes and their Attribution*. Oxford University Press.
- Noë, A. (2005) 'Against intellectualism', *Analysis*, 65(288), pp. 278–290. doi: 10.1111/j.1467-8284.2005.00567.x.
- Pennycook, A. (2010) *Language as a Local Practice*. New York, NY: Routledge.
- Ryle, G. (2009a) 'Knowing how and knowing that', in *Collected Essays 1929 - 1968: Collected Papers Volume 2*. London: Routledge, pp. 222–235.
- Ryle, G. (2009b) *The Concept of Mind: 60th Anniversary Edition*. London: Routledge.
- Stanley, J. (2011a) *Know How*. 1st edn. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199695362.001.0001.
- Stanley, J. (2011b) 'Knowing (How)', *Noûs*, 45(2), pp. 207–238.

- Stanley, J. (2011c) 'Ryle on Knowing How', in *Know How*. New York, NY: Oxford University Press, pp. 1–35. doi: 10.1093/acprof:oso/9780199695362.003.0001.
- Stanley, J. and Williamson, T. (2001) 'Knowing How', *The Journal of Philosophy*, 98(8), pp. 411–444.
- Wikforss, Å. (2017) 'Davidson and Wittgenstein : A Homeric Struggle?', in Verheggen, C. (ed.) *Wittgenstein and Davidson on Language, Thought, and Action*. Cambridge University Press, pp. 46–68.
- Wittgenstein, L. (2003) *Investigaciones Filosóficas*. México: UNAM-IIF.
- Wittgenstein, L. (1991) *Sobre La Certeza*. Barcelona: Gedisa.
- Wittgenstein, L. (1981) *Zettel*. Edited by G. E. M. Anscombe and G. H. von Wright, 2nd ed., Basil Blackwell.



Enactivismo y valoración. Cómo superar la querrela entre teorías somáticas y cognitivas de las emociones*

Enactivism and appraisal. Overcoming the dispute between somatic and cognitive theories of emotions

*ANDREA F. MELAMED***

Resumen: En este trabajo me propongo mostrar que es posible abordar el fenómeno emocional atendiendo especialmente a los focos de conflicto entre cognitivistas y no cognitivistas pero poniendo en duda el marco a partir del cuál se han erigido. El recorrido que propongo apunta a establecer una nueva manera de abordar el antagonismo entre enfoques somáticos y cognitivos de las emociones, que tanta influencia ha tenido sobre la investigación de las emociones, a la luz de uno de los mayores problemas que enfrenta el enfoque somático: el problema de la variabilidad. Buscaré mostrar cómo ciertos conceptos adquieren nuevo significado en el marco postcognitivista y permiten dar nuevas (y mejores) respuestas a viejos problemas en lo que respecta a los fenómenos emocionales.

Palabras clave: cognición, emoción, enactivismo, valoración.

Abstract: In this work my aim to show that it is possible to address emotions by paying special attention to the sources of conflict between cognitivists and non-cognitivists but by questioning the framework within which they have been erected. The path I propose aims to establish a new way of approaching the antagonism between somatic and cognitive approaches to emotions, which has had so much influence on the investigation of emotions, in light of one of the greatest problems that the somatic approach faces: the problem of variability. I will show how certain concepts acquire new meaning in the post-cognitivist framework and allow new (and better) answers to old problems regarding emotional phenomena.

Keywords: appraisal, cognition, emotion, enactivism.

Recibido: 31/03/2020. Aceptado: 16/01/2021.

* Este trabajo fue realizado en el marco del proyecto UBACyT 20020170100215BA. Agradezco la rigurosa revisión recibida por los/las evaluadores/as anónimos/as, cuyos comentarios han contribuido a mejorar este trabajo.

** Universidad de Buenos Aires; Becaria Posdoctoral del Instituto de Investigaciones Filosóficas (SADAF/CONICET). Contacto: afmelamed@filo.uba.ar.

Introducción

En las últimas décadas mucho se ha escrito acerca de las emociones —qué son, cómo funcionan, cómo entenderlas, cómo describirlas, etc.— subsanando en parte el desinterés que padeciera el tema durante tanto tiempo. Motivado por una larga tradición que nos condujo a concebir a las emociones como enemigas de la razón, parte del renovado interés por ellas gira precisamente en torno a la revisión de la dicotomía entre razón y emoción.¹ Con tal objetivo, se han propuesto algunas teorías, tanto en el marco de la filosofía como de la psicología, que buscando acomodar a las emociones dentro del ámbito de lo racional sugirieron la identificación de las emociones a partir de un componente “cognitivo”. Este componente sería un elemento constitutivo de los estados emocionales, en tanto que además de ser un elemento necesario en todo estado emocional, sería aquel responsable de especificar el tipo de emoción que se experimenta. Las teorías que defienden la existencia de un componente de carácter cognitivo en las emociones se han agrupado bajo lo que se llama un “enfoque cognitivo” (Arnold, 1960; Gordon, 1987; Lazarus, 1991; Lyons, 1980; Solomon, 1976), enfrentándose al así llamado “enfoque perceptual” (Charland, 1997; Goldie, 2000; James, 1884; Prinz, 2004; Zajonc, 1980) y que he preferido llamar “enfoque somático” en virtud de lo que considero su tesis más fundamental —como quedará claro en el siguiente apartado—. Así, la perspectiva cognitiva, por un lado, y la perspectiva somática, por otro, se han presentado como dos modos incompatibles e irreconciliables de concebir a las emociones, disputa que en gran medida ha obstruido la posibilidad de ofrecer una exploración adecuada del fenómeno emocional. Asumiendo el marco conceptual de la ciencia cognitiva —y sus nociones clave de representación mental (interna de objetos externos) conjuntamente con algunas dicotomías (percepción/cognición; percepción/acción; emoción/razón etc.)— la investigación emocional se estructuró en torno a la disputa acerca de su presunto carácter cognitivo o no cognitivo. Durante buena parte de la investigación científica reciente sobre las emociones, combatir los presupuestos de la teoría cognitiva de las emociones significó hacerlo dentro de los límites de la ciencia cognitiva clásica u ortodoxa, de modo que negar el conjunto de implicaciones cognitivistas para las emociones, implicó tomar posiciones del tipo somático/perceptual, que valiéndose de las mismas nociones cognitivas, instrumentaba ciertas deflaciones para acomodar las intuiciones o ideales biologicistas (Ekman, 1984; Zajonc, 1980) proponiéndose para este fin valoraciones afectivas (Robinson, 2005), valoraciones corporizadas (Prinz, 2004), etc.

En este trabajo me propongo mostrar que es posible abordar el fenómeno emocional atendiendo especialmente a los focos de conflicto entre cognitivistas y no cognitivistas pero de un modo ligeramente distinto, sin buscar acomodar o subsumir el fenómeno emocional bajo los estándares previos, sino poniéndolos en duda. Con esto me refiero al cuestionamiento que la ortodoxia cognitivista ha recibido en los últimos años —por diversas razones, independientes del estudio de las emociones— y que poco a poco se ha erigido como un nuevo campo de confluencia intelectual —no por ello homogéneo— que busca poner en tela de juicio a los presupuestos cognitivistas y que podríamos resumir con bajo el término

1 Antonio Damasio estuvo a la vanguardia del cuestionamiento de la dicotomía cristalizada con su libro *El error de Descartes* (1994). Véase Pérez & Melamed (2020) trabajo que continúa la línea de investigación inaugurada por Damasio.

paraguas de “postcognitivism”, incluyendo bajo esta denominación teorías tan diversas como las presentadas en Calvo & Gomila, 2008; Gallagher, 2005; Noë, 2004; Rowlands, 2010; Varela, Thompson, & Rosch, 1991. Específicamente, buscaré mostrar cómo ciertos conceptos adquieren nuevo significado en el marco postcognitivista y permiten dar nuevas (y mejores) respuestas a viejos problemas en lo que respecta a los fenómenos emocionales.

El recorrido que propongo no busca brindar nuevas teorías acerca de la naturaleza de las emociones, ni de sugerir nuevas taxonomías para organizarlas, sino apunta a establecer una nueva manera de abordar uno de los conflictos que mayor influencia ha tenido sobre la investigación de las emociones: el antagonismo entre enfoques somáticos y cognitivos de las emociones. En ese sentido, este trabajo, de carácter elucidatorio, se propone en primer lugar precisar el desacuerdo entre las posiciones somáticas y cognitivas sobre las emociones. Para ello comenzaré ofreciendo una reconstrucción sucinta de las dos posiciones generales, centrándome en lo que considero es la mayor dificultad que enfrenta la perspectiva somática (el problema de la variabilidad) y que el cognitivismo consigue superar (aunque lo hace con el altísimo costo de la hiperintelectualización de las emociones). En segundo lugar, a partir de la presentación de este problema, que supondrá una relectura del debate entre perspectivas somáticas y cognitivas, defenderé que el giro postcognitivista es capaz de hacer lugar a aquellas intuiciones cognitivas de modo más minimalista: permite dar cuenta de la variabilidad sin arrastrar los problemas del cognitivismo, es decir, sin asumir sus compromisos. De este modo, la indagación en horizontes no-clásicos, o post-cognitivos y principalmente el enfoque enactivo (Colombetti, 2010, 2014; Colombetti & Thompson, 2007; Varela et al., 1991), me permitirán mostrar que el abandono del marco teórico cognitivista y consiguientemente la superación de ciertas limitaciones conceptuales que conllevaba la adopción de ese marco teórico, harían posible la caracterización del fenómeno emocional de manera integral, haciendo lugar al espíritu de las motivaciones que tuvieron los enfoques somáticos y cognitivos, y que sin embargo, no podían resolverse en su propia arena. Específicamente, mostraré cómo la propuesta de Colombetti y Thompson (2007), al redefinir la noción de valoración permite resolver el problema de la variabilidad de modo más adecuado.

1. De qué hablamos cuando hablamos de las emociones: la disputa entre teorías somáticas y cognitivas

Uno de los problemas que enfrentamos al intentar ofrecer una caracterización de las emociones deriva de cómo acomodar *todos* los aspectos que comúnmente se asocian a ella:

1. Cierta evaluación de la situación (del entorno y de mí mismo en contexto).
2. Cambios corporales.
 - a. Fisiología: sistema endócrino, sistema autonómico, etc.
 - b. Movimientos músculo esqueléticos y faciales.
 - c. Conducta abierta observable.
3. Fenomenología: experiencia, sensación, sentires.
4. Tendencias de acción.

¿Son todos estos elementos necesarios o constitutivos de toda emoción? Si la respuesta es negativa, ¿cuáles son los componentes esenciales que dan lugar y caracterizan a una

emoción? Ya sean estos concebidos como componentes esenciales o como meros acompañamientos ¿cómo se coordinan o acomodan todos ellos para dar lugar a *una* emoción? Las diversas respuestas que las preguntas anteriores recibieran a lo largo de la historia de la filosofía y luego de la psicología, constituyeron distintos enfoques, vinculados a diversos modos de *pensar* acerca de las emociones. Por un lado, el enfoque que hace hincapié en el aspecto somático, que concibe a las emociones como un proceso esencialmente fisiológico o corporal, enfoque que agrupa diversas teorías bajo la denominación de ‘teorías somáticas’. Siguiendo a William James (1884, 1890), los cambios fisiológicos (también a nivel facial y corporal (2), conllevan siempre un sentir [*feeling*] (3), y eso *es* la emoción. Es decir, luego de la percepción de un evento relevante, los cambios corporales se siguen de manera directa (i.e. sin intervención de la cognición) y *sentir* aquellos cambios mientras ocurren *es* la emoción: en virtud de lo cual James afirma que “nos sentimos asustados porque temblamos; no es que temblemos porque tenemos miedo” (James, 1884, p. 190).²

En cambio, siguiendo a Schachter y Singer (1962), habría que decir que el miedo tiene una base fisiológica (2), pero ésta es sólo un *estado de excitación general*, que alcanza la especificación necesaria para resultar en una emoción de miedo recién cuando identifico que las alteraciones en mi cuerpo obedecen a la cercanía del peligro que altera mis expectativas de bienestar (1). Es decir, encontramos una segunda perspectiva que subraya el carácter cognitivo o mental de las emociones, esto es, sostiene que las emociones se encuentran atravesadas o constituidas por elementos cognitivos.³ Hay, sin embargo, un conjunto de teorías muy diversas asociadas por esta tesis cognitiva. Algunos pensadores se vieron seducidos por la idea de definir a las emociones a partir del elemento racional por excelencia: el pensamiento o la actividad cognitiva, buscando corregir en cierto modo la suerte que corriera el tratamiento de las emociones a lo largo de la historia de la filosofía, esto es, como aquello que se oponía a la razón o la distorsionaba. De esta manera, las teorías cognitivas parecían revertir la presunta irracionalidad de las emociones. Pero de la mano de esta propuesta nacía un nuevo conjunto de dificultades. La principal consecuencia de este movimiento fue que estudiar las emociones deviniera casi exclusivamente en estudiar la cognición que las dispara o constituye (ya sea que se tenga una posición cognitiva causal o cognitiva constitutiva). Y ésta fue la tarea que llevaron adelante los defensores de la consabida “corriente cognitiva”: el estudio y tipificación de las creencias/juicios que constituyen una emoción, enfrentándose a las actividades que venía llevándose adelante desde la psicología experimental de base biologicista, que buscaba identificar los mecanismos somáticos que subyacen a los procesos emocionales. De modo que dentro de la ciencia cognitiva clásica se fueron afianzando estos dos modos muy diversos de concebir la tarea de investigación acerca de las emociones, propiciada por la heterogeneidad característica del fenómeno y sus múltiples aristas: el enfoque cognitivo, por un lado; el enfoque, somático

2 El miedo se convirtió en una de las emociones paradigmáticas dentro de las emociones básicas, tanto por su ubicuidad en la literatura sobre emociones, como por haber sido estudiada en mayor profundidad por la neurobiología. Por ejemplo, Joseph LeDoux defiende la existencia de un mecanismo dual de procesamiento emocional a partir del desarrollo de investigaciones empíricas específicas del miedo (LeDoux, 1996).

3 La caracterización adecuada de ‘lo cognitivo’ es tan espinosa como la de ‘lo emocional’. Basta con notar que aquí ‘cognitivo’ se opone a los procesos corporales meramente fisiológicos, tales como los que ocurren en el sistema endócrino, el aparato digestivo o el sistema respiratorio.

por el otro. Se configuró así una dinámica peculiar en la investigación sobre emociones y con ella se profundizó el pretendido antagonismo entre estas perspectivas.⁴

Con el fin de precisar cuál es el espíritu del cambio teórico en el pasaje de la perspectiva somática hacia la cognitiva, en el próximo apartado presentaré sucintamente a la perspectiva somática, concentrándome en el que considero es el problema más importante: el problema de la *variabilidad* de las respuestas emocionales. En la siguiente sección (2.) defenderé que la transición hacia un marco postcognitivista permitiría superar la disputa entre las perspectivas somáticas y cognitivas de las emociones, y resolver el problema de la variabilidad sin pagar los costos que las teorías cognitivas asumieron. Así, este trabajo puede verse como proveyendo razones adicionales a favor de la adopción de un marco enactivo, en virtud de su capacidad para lidiar más adecuadamente con ciertos problemas intrínsecos al marco cognitivo clásico.

1.1. Las limitaciones del enfoque somático de las emociones

La perspectiva o enfoque somático de las emociones podría sintetizarse a partir de la siguiente tesis fundamental: es falso que las emociones *causen* cambios corporales, más bien, son los cambios corporales quienes desempeñan un rol constitutivo en la generación de emociones. Es decir, luego de la percepción de un evento relevante, los cambios corporales se siguen de manera directa (i.e. sin intervención de la cognición) y *sentir* aquellos cambios mientras ocurren *es* la emoción (James, 1884). Uno de los ejemplos paradigmáticos de emociones es el miedo, aquello que siente el excursionista que se encuentra sorpresivamente con una serpiente en el bosque.⁵ Según el sentido común, el sujeto sentiría miedo frente al reptil, que causaría o se *manifestaría* en temblores y conductas del tipo congelarse, huir o luchar. La perspectiva jamesiana revoca esa cadena causal: el miedo *es* el temblor y/o la conducta de huida/congelamiento/lucha. Es decir, no existe un estado mental previo que luego se manifiesta o expresa corporalmente. La emoción *es* ese sentir corporal. En virtud de que siento que tiemblo, tengo miedo.

Antes de pasar a las objeciones que recibió este abordaje, nótese que el análisis de James se realiza en el seno de la distinción que él mismo traza entre emociones básicas (*coarser*) y emociones sofisticadas (*subtler*) (James, 1890, p. 743).⁶ En este sentido, mantiene su análisis circunscripto al ámbito de lo que canónicamente se conoce como el de las “emociones básicas”. Cabe aclarar que, aunque no sea un objetivo de este trabajo, no descarto que puedan brindarse nuevas taxonomías que reemplacen a la dicotomía clásica entre emociones básicas y emociones sofisticadas/derivadas/complejas/culturales/intelectuales/no-estándar. Por el momento baste notar cuáles son límites del análisis que inaugura William James

4 En los últimos años se han ofrecido algunos intentos de reconciliación entre estos enfoques (Charland, 1997; Prinz, 2004; Robinson, 2005) lo que refuerza la idea de que el antagonismo entre estos enfoques no es tan profundo como parecía.

5 Quizás el ejemplo suene extraño, puesto que en nuestro contexto ecológico (el de los grandes centros urbanos) rara vez nos topamos con una serpiente en las calles; sin embargo, me parece interesante mantener este ejemplo por razones que se verán a más adelante.

6 Recuérdese que James en “*What is an emotion?*” ya había excluido explícitamente a las últimas: “antes que nada debería decir que las únicas emociones que me propongo expresamente considerar aquí son aquellas que tienen una expresión corporal distintiva” (James, 1884, p. 189)

y se extiende hasta el presente, según el cual lo que habitualmente conocemos como las ‘emociones básicas’, y que Griffiths llama ‘programas afectivos’ (siguiendo a Ekman), está vinculado con nuestra condición biológica, que conecta nuestra vida emocional con la de otras especies animales, con las que estamos ligadas por obra de la evolución. Es decir, se relaciona con la tesis de la universalidad de las expresiones emocionales (circunscriptas a estas básicas), idea propuesta por Darwin (1872) y profundizada por Izard (1971), Tomkins (1962) y Ekman (1984). Estas emociones denominadas ‘básicas’ son seis: el miedo, la alegría, el asco, la sorpresa, la ira y la tristeza.⁷ El enfoque ha recibido múltiples objeciones, sin embargo, por cuestiones de espacio, mencionaré las 3 más importantes tanto en el contexto del debate en general, como en particular, a los efectos de este trabajo.

1.1.1. Objeción a partir la pretendida la univocidad fisiológica

Esta es una de las críticas más usuales contra la teoría somática. Aborda un corolario de la tesis somática, que establece que puesto que las emociones son sentires de cambios corporales, es preciso que exista un subconjunto de cambios corporales distintivo para cada emoción. Esta fisiología distintiva tiene el doble propósito de demarcar, por un lado, estados emocionales de estados no emocionales, y, por otro lado, distintos tipos de estados emocionales entre sí. No obstante, Cannon (1927) encuentra que los mismos cambios viscerales ocurren en emociones muy diferentes, así como también en estados no emocionales, concluyendo la inadecuación de la tesis somática, a partir de la refutación de su corolario. Debe señalarse que, con posterioridad a la crítica de Cannon, se han desarrollado numerosas investigaciones, cuyos resultados al menos ponen en duda esta imposibilidad de diferenciación. En ese sentido, distintos trabajos de Paul Ekman exhiben evidencia a favor de que existen patrones de activación distintiva del sistema nervioso autónomo, para ira, miedo y asco (Ekman, 1977; Ekman, Levenson, & Scherer, 1983), aunque por supuesto esta cuestión dista de haber quedado zanjada ni conceptual ni empíricamente.

1.1.2. Sobre la imposibilidad de la inducción artificial de una emoción

Es un hecho que la adrenalina actúa en el cuerpo imitando la acción de los impulsos nerviosos simpáticos, de manera tal que, si se inyecta directamente en el torrente sanguíneo, induce la dilatación de los bronquiolos, la constricción de los vasos sanguíneos, la liberación del azúcar del hígado, la suspensión de las funciones gastrointestinales, entre otros. Si las emociones son (consecuencia de) cambios viscerales, y han de identificarse a partir ellos, la inducción artificial de esos cambios viscerales debería necesariamente dar lugar a una emoción. Pero experimentos como el de Maraño (1924) — y su continuación, por Schachter y Singer (1962), parecen indicar lo contrario.⁸

7 Téngase en cuenta que aquí el sentido de ‘básico’ no implica una naturaleza simple, sino que apunta a enfatizar el rol que tuvo la evolución en el moldeado tanto de las características peculiares como de los rasgos comunes que exhiben estas emociones, así como también de su función (Ekman, 1994, p. 15).

8 Una objeción adicional que el enfoque recibió apunta a que una caracterización de las emociones a partir del sentir de los cambios corporales no puede acomodar la intuición de que las emociones son acerca de algo; es

1.1.3. El problema de la variabilidad emocional

Tomando en cuenta estas objeciones de modo global, es fácil notar que esta perspectiva luce demasiado rígida. Me refiero a que, puesto que para diferenciar emociones es preciso que éstas sean suficientemente distintas (1.1.1.) y tan estables que siempre que estos cambios acontezcan, las emociones también estarán (1.1.2.) las emociones parecen ser reacciones inflexibles. El problema es que, así concebidas, no parece haber sitio para la variabilidad interpersonal, no podemos explicar cómo es posible que un mismo evento sea responsable de reacciones emocionales diferentes en distintas personas. Podemos ilustrar este punto nuevamente con el ejemplo de la serpiente. Si me encuentro en el bosque con una serpiente, la percepción del animal desencadenará una serie de alertas que (con suerte) me permitirán huir, y el *sentir* aquéllas *será* el miedo. Este mecanismo adaptativo funciona en exceso, es decir, a veces me encuentro huyendo de ramas, pero aun con estos “falsos positivos” esta conducta de huida reporta grandes ventajas para la supervivencia⁹. Ahora bien, resulta que mi amigo Antón, amante de las mascotas exóticas, adoptó una boa constrictor. Y contra las predicciones que la teoría somática haría, cuando se encuentra frente a ella, no se desencadenan en él aquellos mecanismos que señalan peligro: Antón no tiene miedo, por el contrario, él ama a su mascota. ¿Cómo es posible entonces que frente al mismo estímulo Antón sienta alegría y ejecute tareas de cuidado, mientras que yo siento miedo agobiante?

1.2. La perspectiva cognitiva de las emociones

La “teoría cognitiva de las emociones” en rigor no es *una* teoría, sino una multiplicidad de teorías, filosóficas y psicológicas, con un espíritu en común, que a su vez recoge distintas posiciones emparentadas.¹⁰ En líneas generales, la perspectiva cognitiva de las emociones

decir, no puede dar cuenta de la intencionalidad de las emociones. Esta objeción hace mella particularmente en las versiones tempranas de las teorías somáticas (la de William James en todas sus versiones —débil, fuerte— y la de Robert Zajonc (1980, 1984). Sin embargo, existen versiones somáticas más recientes, de inspiración jamesiana, que han hecho lugar a la objeción y se han ocupado de dar cuenta de la intencionalidad de las emociones, tal es el caso de Goldie (2000). Naturalmente, no es la única —y probablemente tampoco la mejor— manera de sortear esta dificultad. Pero basta aquí mencionarla para mostrar cuál es el alcance de la objeción al enfoque somático, pretendidamente ubicua. Los detractores externos a la perspectiva somática, esto es, los teóricos cognitivos, han hecho especial énfasis en esta objeción, juntamente con la objeción por la imposibilidad de inducir artificialmente emociones, para hacer énfasis en la necesidad de un aspecto o dimensión (según sea el caso) adicional a los cambios viscerales.

- 9 Debido a la rapidez de procesamiento de la vía subcortical, las representaciones que arroja resultan difusas, inacabadas, generando a veces reacciones emocionales inadecuadas para determinados estímulos. El costo de un falso positivo (i.e. por una rama que genera una señal de alerta) es ínfimo en comparación con el beneficio que implica para su huida y supervivencia. Este error luego será subsanado cuando, por su parte, la vía cortical, al precio de tomarse un tiempo más prolongado, le proporcione al sujeto las representaciones más definidas o precisas, para que el sujeto sea capaz de descartar la información irrelevante y estar listo para atender a lo que sí concierne a su bienestar.
- 10 Esta perspectiva reúne lo que en la literatura encontramos bajo los nombres de “teorías cognitivas” (Charland, 1997; Lyons, 1980; Prinz, 2004), “teorías valorativas” [*appraisal theories*] (Lazarus, 1991; Moors et al, 2013) y “la escuela de las actitudes proposicionales” (Griffiths, 1997).

es aquella que le otorga a la cognición un lugar central¹¹ en el proceso emocional. En ese sentido, para las teorías cognitivas “los mecanismos neurales y fisiológicos que subyacen a la experiencia emocional sólo proveen un estado de excitación inespecífico, donde la dirección de la excitación es determinada por la aprehensión cognitiva de la situación externa que da lugar a la excitación” (Gazzaniga & LeDoux, 1978, p. 152). Esta presentación preliminar ya deja entrever uno de los puntos sustanciales sobre los cuales se ha articulado el debate sobre la naturaleza del fenómeno emocional. A continuación, me concentraré en una de las propuestas cognitivas que más impacto ha tenido en la literatura: la teoría valorativa de las emociones.

Valoración y emoción

Como mencioné en el apartado anterior, las teorías somáticas de las emociones encontraron una limitación fundamental: ¿cómo explicar la variación emocional entre distintos sujetos? Se puede ver cómo las diversas propuestas que caen bajo el ámbito de las teorías cognitivas comparten este punto de partida (aunque luego se distancien entre sí al brindar precisiones): las emociones involucran un elemento valorativo (del entorno o la situación específica del sujeto en el entorno) que permite dar cuenta de tales diferencias: no todos los sujetos necesariamente valoran los estímulos de modo idéntico.¹²

Magda Arnold en su texto *Emotion and Personality* (1960) introduce el término “valoración” para referirse a un componente de las emociones, que también incluyen otros elementos (cambios en el sistema nervioso autónomo, tendencias motoras, y sentir). Una de las cuestiones más importantes sobre las que se debe llamar la atención es que Arnold realiza una distinción dentro del ámbito de las valoraciones. Ella distingue entre:

- (1) **Juicios sensoriales** [*Sense judgments*]: caracterizados por ser directos, inmediatos, no reflexivos, no intelectuales, automáticos, instintivos e intuitivos.
- (2) **Juicios reflexivos** [*Reflective judgments*] o evaluaciones secundarias: deliberadas, valoraciones conscientes: a su vez son raros y dependientes de las valoraciones intuitivas.

Que las valoraciones sean parte intrínseca de las emociones tiene una consecuencia doble. En primer lugar, resulta que no es posible que ocurra una emoción sin una correspondiente evaluación (de la relación del sujeto-entorno, concerniente al bienestar del sujeto). Pero, en segundo lugar, la dependencia valorativa es más profunda, pues el resultado de la valoración es el que hace que un estado no solo sea emocional, sino que sea de un tipo particular de emoción. Lazarus (1991) da cuenta de la diferenciación de tipos emocionales a partir de la diferenciación de dimensiones valorativas. Una primera valoración nace de cómo es vista la relación sujeto-ambiente, i.e. como beneficiosa o dañina. De ésta surge una especialización

11 Para algunas versiones del cognitismo, la “cognición” es constitutiva de la emoción (teorías cognitivas impuras), para otros es la emoción (teorías cognitivas puras).

12 El concepto de ‘valoración’ [appraisal] es una noción tan amplia como ambivalente. La principal fuente de confusión proviene del hecho de que ‘valoración’ se utilice tanto para referirse al estado final (lo valorado) como al acto o proceso a través del cual tiene se obtiene la valoración-estado (Lazarus, 2001, p. 42). Asimismo, señala Nico Frijda (1993), el concepto de “valoración” se ha utilizado en la literatura de una manera dual: por un lado, para referirse al contenido de la experiencia emocional, y por otro lado, también para hablar de los antecedentes cognitivos de la emoción.

preliminar de las emociones, que las divide en emociones positivas y negativas, en virtud de si valoran la relación con el ambiente como una relación beneficiosa o dañina para el bienestar del sujeto, respectivamente. El enriquecimiento teórico que realiza Lazarus implica que luego de aquella dimensión inicial (positivo/negativo), se lleve a cabo una especialización ulterior, a partir de las distintas configuraciones que tales situaciones beneficiosas o dañinas pueden representar, dando lugar a distintas categorías concretas de emoción (cólera, ansiedad, culpa, vergüenza, tristeza, disgusto, felicidad, esperanza, orgullo, amor, alivio son algunas de las mencionadas por el autor). De modo que estos patrones de valoración de la relación sujeto-ambiente se fusionan en lo que Lazarus (1991) denomina “temas relacionales centrales” [*core relational themes*] donde cada género de emoción se caracteriza por tener un *tema relacional central* distintivo. El punto entonces es que ‘valorar’ implica considerar la significación del ambiente para el sujeto, esto es, el impacto que la situación que lo rodea, el entorno, tendría sobre él, no sólo sobre su cuerpo sino, en especial, sobre sus intereses [*concerns*], la satisfacción u obstrucción de sus metas (Frijda, 1986), la adecuación a sus creencias, etc.

Si bien no existe acuerdo con respecto a cuáles son los procesos específicos a través de los cuales tienen lugar estas valoraciones, en la literatura existe cierto consenso en cuanto a la existencia de una *multiplicidad* de mecanismos, de diverso nivel¹³. Lazarus distingue *dos modos de valorar*: uno automático,¹⁴ no-reflexivo e inconsciente o preconsciente; el otro, consciente y deliberado (Lazarus, 1991, p. 128). Y destaca que “no es posible decir con seguridad qué proporción de valoraciones y emociones están basadas en cada modo de actividad cognitiva, y quizás la mayoría de las valoraciones de los adultos involucra una mezcla de ambas” (Lazarus, 1991, p. 155). Esta distinción entre mecanismos valorativos de diverso nivel, también es asumida por otros importantes teóricos de las emociones, como Paul Ekman, quien admite que “habitualmente la valoración no solamente es rápida sino que sucede sin percatación, de modo que debo postular que el mecanismo de valoración es capaz de operar automáticamente” (Ekman, 1977, p. 58). En un trabajo más reciente Moors *et al* (2013) reportan cierto consenso entre quienes adoptan una perspectiva valorativa tanto respecto de la multiplicidad de mecanismos que subyacen a las valoraciones, como de la amplitud en la gama (el formato) de representaciones con las que opera: conceptuales y/o proposicionales, perceptuales y/o corporizadas; localizadas o distribuidas. Las valoraciones frecuentemente tienen lugar de modo automático (inconsciente, rápido, no sujeto al control voluntario), pero a veces también puede proceder de modo no-automático (Moors, 2010).

13 Moors (2013) advierte que su perspectiva encuentra fundamento en el marco de los niveles de análisis de Marr (1982), que opera con nociones como las de mecanismo y representación (i.e. en el nivel algorítmico). En este sentido, se imprimen algunos constreñimientos: los mecanismos capaces de sustentar a las valoraciones deben estar mediados o resultar en una representación, donde ‘representación’ se entiende de modo estrictamente funcional, como aquello que es invocado para explicar relaciones variables de *input-output*. En un sentido análogo, Reisenzein presenta su teoría valorativa, remarcando que por su carácter puramente estructural, carece de presupuestos respecto de la forma de los procesos de las valoraciones en cuestión: “cualquier modo de determinar el valor de una función valorativa *V* para un objeto *O* cuenta como un modo de valorar *O* en la dimensión valorativa correspondiente” (Reisenzein, 2001, p. 189), acentuando asimismo que esta perspectiva estructural/funcional es completamente compatible con diversas teorías de procesamiento (algorítmico).

14 Advierte Lazarus que tal carácter automático no debe ser equiparado con una condición primitiva, puesto que el procesamiento automático admite “significados [*significances*] complejos, abstractos y simbólicos, que a través de la experiencia pueden ser condensados en significado instantáneo” (Lazarus, 1991, p. 155).

Las teorías valorativas mantienen su vigencia en la actualidad, con algunas novedades: las teorías valorativas contemporáneas (dentro de la ciencia cognitiva clásica) conciben a las valoraciones como *procesos*, distinción que parece haber sido omitida en las discusiones previas. Moors et al (2013) reconstruyen las teorías valorativas contemporáneas como teorías componenciales, puesto que conciben al episodio emocional como un complejo, formado por distintos subsistemas:

1. Un componente valorativo con evaluaciones del ambiente y de la interacción organismo-ambiente;
2. Un componente motivacional con tendencias de acción u otras formas de disposición para la acción;
3. Un componente somático con respuestas fisiológicas periféricas;
4. Un componente motor con conducta expresiva e instrumental;
5. Un componente afectivo, con experiencias subjetivas o sentires.

La primera cuestión sobresaliente es que, así entendidas las teorías valorativas, resulta patente que no son pura o estrictamente cognitivas, sino más bien, resultan ser teorías híbridas o mixtas: incluyen aspectos cognitivos —en la medida que se siga concibiendo a la valoración como algo diferente a la percepción— así como también elementos somáticos. Entonces, ¿qué es lo que hace que una teoría sea propiamente una ‘teoría valorativa’? Como ya dije, ‘valorar’ implica tomar en consideración la significación del ambiente para el sujeto, (sus intereses [*concerns*], la adecuación a metas, creencias, etc). Y para Arnold, estimar cómo el mundo nos afecta personalmente parece requerir “un paso más allá de la percepción, que no puede ser la función de ninguna modalidad sensorial sola, ni de todas ellas juntas” (Arnold, 1960, p. 188). En ese sentido, se dice que las valoraciones son intrínsecamente transaccionales (Moors *et al.*, 2013). Pero lo que las define como teorías valorativas, en particular, radica en la función que se le asigna a la valoración: defender una posición valorativo-cognitiva implica reconocer la existencia de componentes valorativos, pero más específicamente, consiste en atribuirle una función específica a la valoración:

las emociones son producidas [*elicited*] y diferenciadas sobre la base de la evaluación o valoración [*appraisal*] subjetiva del individuo de la significación que un evento, situación u objeto, tiene para su persona, de acuerdo con un número de dimensiones o criterios (Scherer, 1999, p. 637).

Es decir, no alcanza con conceder que hay valoraciones, es necesario atribuirle un rol especial: las teorías valorativas asumen que la valoración dispara y especifica los episodios emocionales, a través de la operación de numerosos cambios sincrónicos en diversos componentes. Aunque existan otros componentes relevantes, es la valoración la que prescribe, y la que coordina a los demás. En tal sentido, las teorías valorativas especifican los criterios o variables relevantes para diferenciar emociones, algunas de los cuales ya hemos mencionado: relevancia de metas, congruencia de metas, certeza, agencialidad (o responsabilidad sobre el hecho), posibilidad de control, etc. De esta manera, un cambio en alguna de estas variables sería suficiente para explicar el comportamiento diverso de distintos sujetos frente al mismo objeto del mundo.

Objeciones a la perspectiva cognitiva

Las teorías cognitivas también han recibido numerosas objeciones, algunas de ellas son transversales a todo el *enfoque*, otras son atinentes a algunas posiciones específicas. Debe notarse que son precisamente las críticas que el enfoque cognitivo recibió, las que han motivado ulteriores elaboraciones teóricas, o el perfeccionamiento de algunas ideas, como sucede con tantos otros casos de progreso científico-conceptual. Mencionaré brevemente algunas de estas objeciones.

1.2.1. Sofisticación del componente cognitivo

Los distintos autores que han contribuido con la corriente cognitiva han sostenido diversos compromisos. Algunas de estas versiones asumen un alto grado de complejidad del componente cognitivo, por ejemplo, equiparándolo a creencias. Este punto es blanco de críticas debido a su alta exigencia para el sujeto, y en este sentido es una objeción que hace mella en especial en versiones de la teoría que no he revisado aquí, como la de Robert Solomon (1976): si tener miedo es juzgar que algo es peligroso, tener una emoción o estar en un estado emocional estaría requiriendo un nivel de consciencia por parte del sujeto emocionado, un tanto elevado. Adicionalmente, en caso de mantener una versión sofisticada del componente cognitivo, muchos sujetos que supondríamos alcanzados por la teoría resultarían excluidos de su ámbito de aplicación. Sostener que para tener miedo necesito tener la creencia (en sentido ocurrente) de que hay peligro (actitud proposicional con contenido semánticamente evaluable), excluye directamente la posibilidad de que animales no lingüísticos, así como también niños pre-lingüísticos, tengan emociones. Como ha quedado claro en el apartado anterior, algunas versiones del enfoque cognitivo han podido eludir esta crítica a partir de la deflación del componente cognitivo, admitiendo así valoraciones automáticas.

1.2.2. La impenetrabilidad cognitiva de las emociones

En términos de Goldie (2000) un sistema es cognitivamente penetrable si puede ser afectado por las creencias del sujeto. Ya sea que se tome al elemento cognitivo como condición necesaria o suficiente, lo que está claro es que, en caso de modificarse el elemento cognitivo (sea ésta una creencia, un juicio, o una valoración automática) consecuentemente sufrirá alteraciones la emoción también. Análogamente, se espera que, en caso de operar una modificación en la emoción, exista un cambio sobre el cual repose el cambio emocional (por ejemplo, una revaloración de la situación, un cambio en las creencias del sujeto, etc.). Sin embargo, es un hecho indiscutido que muchas veces, muy a pesar de nuestras creencias, algunas emociones son *persistentes*. O, dicho de otro modo, es muy frecuente tener emociones que son contrarias a nuestras creencias. Por ejemplo, teniendo plena conciencia del carácter inofensivo de las abejas, tan pronto una aparece, el terror me sobrecoge. Ni siquiera las múltiples experiencias que he tenido con abejas que no me hacen daño han podido quebrar el vínculo entre ‘abeja’ y ‘terror’ para mí. En el mismo sentido Jenefer Robinson se refiere a la ‘inercia emocional’ (Robinson, 1995, p. 68) para señalar cómo los

juicios desapasionados que pueda tener acerca de las arañas no consiguen evitar su respuesta temerosa (su miedo) a las mismas.

2. Superando el antagonismo clásico

En la sección anterior defendí que el antagonismo entre las posiciones somáticas y cognitivas puede reducirse al problema de la variabilidad: una objeción que el enfoque somático no parece superar y que el cognitivo resuelve. Sin embargo, aquella solución conlleva un altísimo costo. Ya sea como consecuencia de la existencia de componentes cognitivamente tan sofisticados que como resultado excluyen a animales no lingüísticos y niños pre-lingüísticos de la experimentación de emociones. O bien, alternativamente, puestos a eludir las consecuencias nefastas señaladas por aquellas críticas, proponen mecanismos valorativos automáticos e inconscientes que implican que el límite entre la cognición y la percepción se desdibuje. En otras palabras, el enfoque cognitivo de las emociones resuelve el problema de la variabilidad, dejando fuera del ámbito emocional a animales pre-lingüísticos, o aceptando valoraciones inconscientes, no alterables por otros procesos cognitivos, que parecen no encajar adecuadamente con la noción clásica de cognición.

Ahora bien, esta deflación de la valoración nos anima, o más bien, nos demanda realizar una reconceptualización de la percepción, la cognición, la relación entre ellas (y consecuentemente de las emociones) y el rol de cada uno de ellos. Esta es, ni más ni menos, una motivación suficientemente sólida como para poner en tela de juicio la confianza depositada en el paradigma cognitivo clásico, y buscar modos alternativos de concebir aquello que veníamos llamando ‘cognición’, ‘percepción’ y ‘emoción’.

La propuesta entonces apunta a mostrar que la adopción de un marco teórico alternativo nos brindará un modo novedoso de abordar la problemática, a la vez que nos proveerá soluciones a algunos problemas existentes en las descripciones del paradigma cognitivo. Esta reconceptualización de la mente nos permitirá proveer una mejor pintura del funcionamiento de las emociones, sustituyendo la perspectiva inadecuada (por las razones recién expuestas) de la ciencia cognitiva clásica u ortodoxa. El modo en que la ‘nueva ciencia de la mente’ se opone a la ortodoxia cognitivista, no es unívoco, y aunque la reconstrucción de los diversos modos constituye un espacio conceptual plural, sumamente rico e interesante, excede los límites de este trabajo.¹⁵ En lo que sigue, consideraré un modo de ofrecer una alternativa, cuyo espíritu es el de superar las limitaciones impuestas por la metáfora del ordenador, y sus implicancias conceptuales ligadas a la tesis funcionalista de la realizabilidad múltiple: el cuerpo no es simplemente *una* implementación del ‘programa’ emocional (entre otras posibles).

Susan Hurley ofrece un interesante diagnóstico de la ciencia cognitiva clásica y sus problemáticas. Así, caracteriza a la ciencia cognitiva a partir de lo que denomina el ‘modelo sandwich’ de la mente (Hurley, 1998), puesto que asume la doble disociación de la percepción y la acción, dónde la primera es completamente pasiva y la segunda puramente activa. La visión input/output a la que se opondrá, queda exquisitamente resumida en las siguientes líneas:

¹⁵ Para un análisis detallado de las tesis en juego y la discusión de su interacción o incompatibilidad pueden consultarse los trabajos de Rowlands (2010) y Burdman (2015).

Una visión de la mente sostenida con mucha frecuencia tiene dos componentes principales (...) La cognición es virtualmente central, aun si la mera implementación de los procesos cognitivos fuese distribuida. La mente se descompone verticalmente en módulos: la cognición hace de interfaz entre la percepción y la acción. Percepción y acción no sólo están separadas la una de la otra, sino también de los procesos superiores de la cognición. La mente es un tipo de sándwich, y la cognición es el relleno (Hurley, 1998, p. 401).

Esta imagen capta el espíritu de la ortodoxia cognitiva, así como también explica por qué la incorporación del estudio de las emociones dentro del paradigma de la ciencia cognitiva se tradujo en una fuerte *cognitivización* del fenómeno. En esta concepción clásica la cognición es un sistema central que sirve de mediador y transforma la información *recibida* por los sistemas periféricos de entrada (input), y arroja resultados a través de sus sistemas de salida (output), es decir, la acción. Nótese que el rol mediador de la cognición gravita, a su vez, sobre una concepción —también cristalizada— de la percepción, como sistema receptivo, pasivo, que incorpora de modo transparente (objetivo) cierta información que el mundo le provee, valiéndose de representaciones. De modo que además de la consideración del *modelo sándwich de la cognición*, nótese que se está asumiendo una tesis del “*ideal transparente de la percepción*”, cuyo cuestionamiento también será fundamental a fin de proveer una presentación adecuada de qué son y cómo tienen lugar las emociones.

Esta visión segmentada de la mente es la que combatirán los enfoques heterodoxos de la mente, tal como expresan Varela, Thompson y Rosch, en el emblemático libro *The Embodied Mind*, buscando “sortear esta geografía lógica de “interno/externo” estudiando la cognición ni como recuperación ni como proyección, sino *acción corporizada*” (Varela, Thompson, & Rosch, 1991, p. 202). La redefinición de la cognición, de su rol y sus relaciones con otros ‘subsistemas’ de la mente, invita asimismo a una reconceptualización de su relación con el mundo, del modo en el que se constituye el significado. En este contexto, el significado ya no resulta provisto por el mundo de manera acabada o completa. Varela, Thompson, & Rosch, (1991) sostienen que el organismo y el medio ambiente están mutuamente plegados de múltiples modos, y que “lo que constituye el mundo de un organismo es enactuado por la historia de acoplamiento estructural de ese organismo” (Varela, Thompson, & Rosch, 1991, p. 235). Compatible con su enfoque de la acción guiada perceptivamente o acción corporizada, recuperan así el espíritu del enfoque de Gibson (1986) según la cual el entorno exhibe ciertas *affordances* —i.e. oportunidades o facilitaciones para ciertas interacciones, vinculadas a las posibilidades sensoriomotoras del organismo—. De manera que el concepto clave para dar cuenta de esta constitución de significado perceptivo es la noción de ‘oportunidad para la acción’ [*affordance*], que implica la postulación de propiedades no-físicas, es decir, propiedades que no están en el mundo físico, sino que emergen de la interacción del organismo (que posee una historia específica de interacciones anteriores y determinadas capacidades sensoriomotoras) con el ambiente.¹⁶

16 Existen algunas diferencias entre los seguidores de Gibson respecto de los compromisos ontológicos que asumen las ‘oportunidades para la acción’. Independientemente de cuál fuese el modo adecuado de interpretar a Gibson, mi interés aquí responde exclusivamente a la recuperación que hicieran Varela, Thompson y Rosch, aun cuando pudiera objetarse por no ser fiel al espíritu de Gibson.

La percepción directa de oportunidades para la acción implica la percepción directa de significados, y es precisamente aquello lo que posibilita que un mismo objeto físico pueda ser percibido de modo diverso por diferentes animales, o por dos seres humanos con distintos intereses. Todas estas ventajas y lesiones, seguridades y peligros, estas oportunidades para la acción (negativas y positivas) “son propiedades de las cosas tomadas con referencia a un observador, pero no propiedades de las experiencias del observador. No son valores subjetivos; no son sentires de placer o dolor añadidos a las percepciones neutrales” (Gibson, 1986, p. 129). Aquí Gibson nos está brindando la llave para entrar en el universo de los valores y al mismo tiempo, la clave para sortear el problema de su estatus ontológico: las oportunidades para la acción no pertenecen ni al mundo físico ni al fenoménico, no son exclusivamente materiales, ni tampoco mentales. Y esa es la clave de la perspectiva corporizada moderna de la mente, que asume una relación transaccional y recursiva entre el cuerpo y la mente, y que al hacerlo no se apoya sobre el dualismo cartesiano que trata a la mente y al cuerpo como elementos separados e independientes de las emociones (Barrett & Lindquist, 2008, p. 246).

Dentro del paradigma postcognitivista se ha utilizado al enactivismo para repensar las emociones. En términos generales, el enactivismo (Stewart, Gapenne, & Di Paolo, 2010; Thompson, 2007; Varela, Thompson, & Rosch, 1991; Di Paolo, 2005) se caracteriza por defender una concepción de la mente en la que:

1. Los seres vivos son agentes autónomos que activamente generan y mantienen sus identidades, y así enactúan o crean sus propios dominios cognitivos.

Esta noción de ‘agente autónomo’ se propone fundamentalmente poner en duda la imagen habitual de un sistema cognitivo (el sujeto) que tiene una relación epistémica con un objeto: el mundo, que exhibe información que se encuentra ya disponible, de modo completo, para quien quiera percibirla. Esa información preexistente sería meramente capturada por el individuo (es información que ingresa a través de los órganos *receptivos*), es decir, es aprehendida por todos los individuos del mismo modo. En su lugar, el enactivismo defiende una posición más bien constructivista, donde el individuo es *agente*, esto es, tiene un rol activo en la conformación del objeto que ‘percibe’. La información, el mundo, surge de las continuas interacciones del sujeto con el ambiente: es *transaccional*. De ello se sigue que:

2. El mundo de los seres cognitivos no es un terreno externo, que ha sido especificado previa e independientemente al individuo y que éste se *representa* internamente. Por el contrario, se trata de un dominio relacional, llevado a cabo por la agencia del ser autónomo, y su peculiar modo de vincularse con el ambiente.
3. La cognición es una forma de *acción* corporizada: El acoplamiento sensorio motor entre el organismo y el medio ambiente modula, pero no determina, la formación de patrones endógenos y dinámicos de actividad neural. Esta actividad, a su vez, informa el acoplamiento sensorio motor, de modo que todo el organismo corporizado puede ser visto como un sistema autónomo auto-organizado que crea significado.

En la revolución conceptual que sugiere el enactivismo que ellos abrazan se reúnen tanto la tesis de la mente corporizada (la idea de que la características del cuerpo son constitutivas

de los procesos mentales, rechazando la visión funcionalista de la mente como una entidad autónoma e independiente de las determinaciones corporales del organismo), como la de la mente extendida (i.e. la idea de que los recursos ambientales desempeñan un rol necesario (constitutivo) en la cognición Clark, 1997; Clark & Chalmers, 1998). En resumidas cuentas, el punto es que el significado y la experiencia son creados por o llevados a cabo [*enacted through*] a través de la continua interacción recíproca del cerebro, el cuerpo y el mundo (Colombetti & Thompson, 2007, p. 56).

Inevitablemente, este cambio conceptual tendrá consecuencias directas sobre la concepción de las emociones, que se manifiestan en particular, a partir de la redefinición de la noción de valoración. Desde esta perspectiva, la valoración no es ya un proceso cognitivo de evaluación subjetiva ‘en la cabeza’, y la excitación y la conducta no son concomitantes corporales objetivos. Más bien, los eventos corporales son constitutivos de la valoración, tanto estructural como fenomenológicamente (Colombetti & Thompson, 2007, p. 58). La novedad que opera Colombetti radica en su concepción de la naturaleza de las valoraciones, en tanto que “evaluar [*evaluating*] el mundo y responder emocionalmente a él no son procesos distintos” (Colombetti, 2014, pp. 111). Desde su perspectiva enactiva, ella concibe el proceso de valorar [*appraising*] “como una actividad del organismo, no separada sino superpuesta con lo que usualmente son vistos como componentes corporales, no cognitivos de las emociones” (Colombetti, 2014, p. 112).

Lo interesante de esta propuesta, y la razón por la que la encuentro más adecuada que las versiones híbridas o mixtas dentro de la ortodoxia cognitiva, es precisamente que la valoración no se postula para dar solución a un problema. En efecto, en las teorías cognitivas, desde Arnold hasta Lazarus, así como también en la teoría mixta de Robinson (2005), la valoración se introduce en la economía conceptual para atender a una necesidad específica, responder el *enigma* al que hice referencia más arriba: cómo dar cuenta de la variabilidad de reacciones emocionales, frente a estímulos idénticos; cómo dar cuenta de las reacciones emocionales opuestas que tenemos Antón y yo frente a su mascota, la boa. Manteniendo incuestionable la neutralidad de la percepción, la variabilidad debe explicarse por la participación de *otro* mecanismo. Recordemos las palabras de Arnold:

El estimar cómo nos afecta personalmente parece requerir un paso más allá de la percepción que no puede ser la función de ninguna modalidad sensorial sola ni de todas ellas juntas (Arnold, 1960, p. 188).

En el contexto del postcognitivismismo es el organismo, en/a través de su corporalidad, el que aprecia, valora o juzga, cuando responde afectivamente. No hay dos instancias ni cosas distintas, sólo está mi cuerpo reaccionando ante determinadas señales y valiéndose de determinados recursos (por ejemplo, de los marcadores somáticos de Damasio, 1994). Más aún, descomponer esa valoración, evaluación o juicio en 4, 6 u 8 dimensiones, da como resultado una versión superintelectualizante, hiper cognitiva de la emoción y de lo que sucede de hecho cada vez que una emoción tiene lugar en el organismo. Lo que busco remarcar es que el rol protagónico aquí lo tiene el organismo, como un todo, ya que es el organismo como un todo quien *valora*: quien *ve* al entorno y sus potencialidades, en relación con sus intereses, metas, necesidades, etc. dando lugar a ‘peligrosidades’, ‘alegrías’ y demás.

Colombetti (2010) brinda algunas precisiones adicionales respecto del modo en que el organismo asigna valores, valiéndose de la noción enactiva de *construcción de sentido* [*sense-making*]¹⁷ de Di Paolo (2005).¹⁸ Esto implica adoptar una posición distintiva respecto al surgimiento del significado: el mundo no está ahí afuera, completo, esperando que el organismo lo conozca —se lo represente internamente—. No hay un mundo pre-dado, el entorno es siempre aprehendido *como* significativo, y ese significado es relativo a la perspectiva o punto de vista de cada sistema vivo. En ese sentido, recuperan la noción de *Umwelt*¹⁹, o mundo egocéntrico [*self-centered world*] que rescata la posibilidad de que dos organismos vivos que comparten el mismo ambiente físico-natural tengan, no obstante, distintos mundos significativos. De esta manera, la respuesta al problema de la variabilidad está implícita en el marco teórico alternativo adoptado, no depende de un agregado *ad hoc* para resolver un problema teórico.

Conclusión: hacia un abordaje enactivo de las emociones

Siendo el objetivo principal de este trabajo contribuir a la comprensión de las emociones, me propuse revisar la querrela entre teorías somáticas y cognitivas de las emociones. Presenté el problema de la variabilidad de las emociones y defendí que se trata una dificultad que el enfoque somático no supera en sus propios términos, al mismo tiempo que la corriente cognitiva parece nacer *para* resolverlos, no sin dar lugar a un nuevo conjunto de problemas. Todos estos problemas nos exhortan a repensar las distinciones que vienen estructurando la investigación, desde que Platón ubicara las emociones en las antípodas de la razón. En ese sentido, esboqué cómo se podría realizar la transición hacia un marco postcognitivista y cómo ésto permitiría superar la disputa entre las perspectivas somáticas y cognitivas de las emociones. He dicho que el modo en que la ‘nueva ciencia de la mente’ se opone a la ortodoxia cognitivista, no es unívoco, cuestión que se ve agravada por tratarse de un espacio en pleno desarrollo. No obstante, en este trabajo me encargué de examinar *uno* de los modos en que se da esta rebelión contra la ortodoxia cognitiva clásica, a partir de la reivindicación de nociones como la de corporización, que siendo tan antigua como William James, ahora revisitada desde el postcognitvismo, configura el espacio enactivista.

La postulación original de Arnold y los desarrollos subsiguientes de la noción de valoración cognitiva, implicaban que la reacción emocional estaba constituida por ciertas apreciaciones que el sujeto realizaba del entorno, y su situación peculiar en él (sus intere-

17 Son aquellas ‘actividades de construcción de sentido’ las que permiten distinguir entre encuentros netamente físicos y encuentros en tanto sistemas propiamente cognitivos: sólo en los últimos el organismo establece una interacción con el entorno y regula su acoplamiento a él: “Los intercambios con el mundo son inherentemente significativos para el que *conoce* y esto es una propiedad definicional para el sistema cognitivo: la creación y apreciación de sentido o ‘construcción de sentido’ en breve” (De Jaegher & Di Paolo, 2007, p. 488).

18 Aunque Di Paolo no se ocupe explícitamente de las emociones, Colombetti defiende que las *actividades de construcción de sentido* pueden ser naturalmente entendidas como el ‘reconocimiento del carácter emocional constitutivo de la enacción’ (Colombetti, 2010, p. 147). En este sentido, la elaboración adecuada de estas nociones no sólo arroja luz sobre la naturaleza de las emociones, sino que adicionalmente, contribuiría a una mejor comprensión del carácter *emotivo* de la cognición y la relación organismo/ambiente.

19 El término *Umwelt* fue introducido por von Uexküll (1921) para señalar al entorno en tanto experimentado por el organismo vivo.

ses, metas, etc.). Esta formulación inicial implicaba mantener separados los dominios de la percepción y la cognición, puesto que esta valoración cognitiva (constitutiva de la emoción) sobrevendría ulteriormente a la percepción del objeto —percepción que así mantenía su ideal objetivo—. Es decir, la valoración cognitiva nacía para atender el problema de la variabilidad de las respuestas emocionales, frente a los hechos objetivos del mundo (así como también, las teorías cognitivas en su vertiente etiquetadora, brindaba una solución al problema de la unicidad fisiológica en tanto que productora de emociones). La propuesta de Colombetti apunta a proveer una visión enactiva de los aspectos valorativos de las emociones. Esto es, a mostrar que el elemento cognitivo pretendidamente ‘irreducible’ de la visión cognitiva ortodoxa —la valoración—, queda naturalizada en el marco de la ciencia enactiva, a la luz de la noción de la ‘construcción de sentido’.

Finalmente, quisiera destacar dos cuestiones que considero sumamente interesantes. La primera es que la ‘construcción de sentido’ rige para los organismos vivos en sentido amplio, desde los organismos unicelulares, hasta los seres humanos racionales. Concebir la vida y la cognición así, resuelve (de un modo excesivo quizás) el problema en torno a la atribución de emociones a animales no humanos, y a niños prelingüísticos. La segunda cuestión que vale la pena subrayar es que aquello que surgía como una consecuencia indeseable o problemática en la discusión original entre enfoques cognitivos y somáticos (dentro del marco cognitivista clásico) a saber, el desvanecimiento de los límites precisos entre los dominios cognitivos, emotivos y perceptivos, en este contexto se convierte en una fortaleza, en tanto se conforma como la tesis central desde el punto de partida de la concepción enactiva, la reivindicación explícita de la continuidad entre cognición y emoción, en las dos direcciones: las emociones son cognitivas, pero además, la cognición es emocional.

Referencias

- Arnold, M. (1960). *Emotion and Personality*. New York: Columbia University Press.
- Barrett, L. F., & Lindquist, K. A. (2008). The Embodiment of Emotion. En *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches* (pp. 237–262).
- Burdman, F. (2015). El Post-cognitvismo en cuestión: Extensión, corporización y enactivismo. *Principia*, 19(3), 475–495.
- Calvo, P., & Gomila, T. (2008). Directions for and Embodied Cognitive Science: Toward an Integrated Approach. En P. Calvo & T. Gomila (Eds.), *Handbook of Cognitive Science. An Empodied Approach* (pp. 1–25). Elsevier.
- Cannon, W. B. (1927). The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, 39(1), 106–124.
- Charland, L. C. (1997). Reconciling Cognitive and Perceptual Theories of Emotion: A Representational Proposal. *Philosophy of Science*, 64(4), 555.
- Clark, A. (1997). *Being There: Putting Brain Body and World Together Again*.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58, 10–23.
- Colombetti, G. (2010). Enaction, sense-making and emotion. En J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction: Toward a new paradigm for cognitive science* (pp. 145–164). MIT Press.

- Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. The MIT Press.
- Colombetti, G., & Thompson, E. (2007). The Feeling Body: Towards an Enactive Approach to Emotion. En W. F. Overton, U. Müller, & J. Newman (Eds.), *Developmental perspectives on embodiment and consciousness* (pp. 45–68). New York: Lawrence Erlbaum.
- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset/Putnam.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.
- De Jaeger, H., & Di Paolo, E. A. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429–452.
- Ekman, P. (1977). Biological and Cultural Contributions to Body and Facial Movement. En J. Blacking (Ed.), *The Anthropology of the Body* (pp. 39–84). London: Academic Press.
- Ekman, P. (1984). Expression And The Nature Of Emotion. En K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 319–343). Hillsdale, NJ: Lawrence Erlbaum.
- Ekman, P. (1994). All Emotions Are Basic. En P. Ekman & R. J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions* (pp. 15–19). New York: Oxford University Press.
- Ekman, P., Levenson, R. W., & Scherer, K. R. (1983). Autonomic Nervous System Activity Distinguishes Among Emotions. *Science*, 221(4616), 1208–1210.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition & Emotion*, 7(3–4), 357–387.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Gazzaniga & LeDoux (1978) *The Integrated Mind*. New York: Plenum Press.
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. New York/London: Talyon & Francis Group.
- Goldie, P. (2000). *The Emotions: A Philosophical Exploration* (Vol. 53). Oxford University Press.
- Gordon, R. M. (1987). *The Structure of Emotions: Investigations in Cognitive Philosophy*. Cambridge, MA: Cambridge University Press.
- Griffiths, P. E. (1997). *What emotions really are: The problem of psychological categories*. University of Chicago Press Chicago.
- Hurley, S. L. (1998). *Consciousness in Action*. Harvard University Press.
- Izard, C. E. (1971). *The face of emotion*. Appleton-Century-Crofts.
- James, W. (1884). What is an emotion? *Mind*, 9, 188–205.
- James, W. (1890). *Principles of Psychology*. Chicago: Enciclopedia Britannica, Inc.
- Lazarus, R. S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.
- Lazarus, R. S. (2001). Relational Meaning and Discrete Emotions. En K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, Methods, Research*. New York: Oxford University Press.
- LeDoux, J. E. (1996). *The Emotional Brain. The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.

- Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.
- Marañón, G. (1924). Contribution à l'étude de l'action émotive de l'adrenaline. *Revue Française d'Endocrinologie*, 21, 301–325.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.
- Moors, A. (2010). Automatic constructive appraisal as a candidate cause of emotion. *Emotion Review*, 2(2), 139–156.
- Moors, A. (2013). On the causal role of appraisal in emotion. *Emotion Review*, 5(2), 132–140.
- Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, 5(2), 119–124.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: The MIT Press.
- Pérez, D. I., & Melamed, A. F. (2020). La razón en las pasiones y las pasiones en la razón. Un entretreído desatendido. En J. Vidal & C. Muñoz (Eds.), *Perspectivas de la subjetividad* (pp. 213–230). Editorial de la Universidad de Concepción.
- Prinz, J. J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.
- Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions. En K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 187–201). New York: Oxford University Press.
- Robinson, J. (1995). Startle. *Journal of Philosophy*, 92(2), 53–74.
- Robinson, J. (2005). *Deeper than Reason*. New York: Oxford University Press.
- Rowlands, M. (2010). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge MA: MIT Press.
- Schachter, S., & Singer, J. (1962). Cognitive, Social, and Physiological Determinants of Emotional State. *Psychological review*, 69(5).
- Scherer, K. R. (1999). Appraisal theory. En T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). New York: John Wiley and sons.
- Solomon, R. C. (1976). *The Passions: The Myth and Nature of Human Emotion*. New York: Anchor Press, Doubleday.
- Stewart, J., Gapenne, O., & Di Paolo, E. A. (Eds.). (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. Book. MIT Press.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Tomkins, S. S. (1962). *Affect, Imagery, Consciousness: Vol. I. America*. New York: Springer.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–175.



La autoridad de la primera persona como criterio de determinación y reconocimiento de la identidad de género

First-Person Authority as a Criterion for Gender Identity Determination and Recognition

JORGE SUÁREZ MUÑOZ*

Resumen: El interés principal de este artículo versa acerca de la posibilidad de justificar la autodeterminación como criterio fundamental para la determinación y el reconocimiento social de la identidad de género de un individuo. La reciente revisión de los criterios médicos abre el debate de su reemplazo por criterios de autopercepción y vivencia en primera persona. Por tanto, tomando como base la concepción de la autoridad como servicio de Joseph Raz y el fenómeno de la autoridad de la primera persona, se argumenta a favor de que los testimonios en primera persona son candidatos justificados para sustituir a los actuales criterios.

Palabras clave: autoridad, autodeterminación, sexo, género, reconocimiento, identidad.

Abstract: The focus of this paper is the possibility of justifying self-determination as the fundamental criterion for social determination and recognition of the individuals' sexual identity. The recent review of medical criteria has started a debate about their replacement by criteria based on self-perception and first-person experience. Therefore, taking as a basis Joseph Raz's service conception of authority and the first-person authority phenomenon, it will be argued that first-person statements are justifiable substitutes for the current criteria.

Keywords: authority, self-determination, sex, gender, recognition, identity.

En el presente artículo se aborda la cuestión del reconocimiento de la identidad de género y los criterios que se emplean social, médica e institucionalmente para su determinación, además de la autodeterminación como principal alternativa que en los últimos años está siendo reivindicada como la sucesora de estos criterios instaurados hoy en una gran mayoría de sociedades. El objetivo principal consiste en analizar críticamente esta propuesta a fin de darle una justificación teórica y comprobar si constituye un mejor sistema de determinación de la identidad de género de los individuos.

Recibido: 08/12/2020. Aceptado: 27/02/2021.

* Estudiante de investigación doctoral del Departamento de Filosofía, área de Filosofía Moral y Política, de la Universidad de Málaga (suarezmjorge@uma.es). Líneas de investigación principales: filosofía moral y política, con especial atención a los fenómenos de autoridad epistémica y a los estudios sobre migraciones en el contexto europeo.

Para ello, los dos primeros apartados hacen crítica al sistema de determinación actual. En el primero de ellos se expone la situación actual al respecto de la determinación del sexo a nivel legal e institucional, además de mostrar cuáles son los movimientos sociales que en las últimas décadas están promoviendo cambios. A continuación, en el segundo apartado, se aborda la cuestión de la determinación del sexo a nivel social, es decir, planteando cómo en la cotidianidad de las relaciones sociales se suele reconocer la identidad de género de una persona. Esto va acompañado de una crítica de los criterios basados en la genitalidad de las personas para realizar esta determinación, así como una defensa de criterios basados en el testimonio en primera persona acerca de la identidad de uno mismo.

Tras esto, en el resto del artículo se ofrece una justificación para la alternativa de la autodeterminación. Así, en el tercer apartado se toma como punto de partida la concepción de la autoridad como servicio de Joseph Raz para dicha justificación, presentando sus tesis fundamentales y su relevancia en la explicación de la autoridad epistémica. En el último apartado, esta concepción de autoridad se conecta con el fenómeno de la autoridad en primera persona. Con todo ello se concluye que la autodeterminación puede ser justificada como un criterio de autoridad válido para determinar la identidad de género de una persona y consiguientemente reconocerle su identidad así expresada, constituyendo un sistema justificado y que resuelve problemas propios de los criterios tradicionales.

1. Autodeterminación y reconocimiento legal en España

La autodeterminación de género es una de las vindicaciones actuales más repetidas como medio para la completa promoción de los derechos de las personas transexuales¹ e intersexuales² y, más en concreto, del derecho al libre desarrollo de la personalidad, tal y como viene recogido en el artículo 10.1 de la Constitución Española (Alventosa del Río 2016, 169-170).

El concepto de autodeterminación referido al género de la persona ha jugado un papel clave en las reclamaciones de diversas asociaciones y colectivos promotores de dichos derechos. Esto ha quedado reflejado en un creciente número de marcos legales actuales en que se hace mención explícita a la “soberanía [de] la voluntad humana sobre cualquier otra consideración física”, citando la Ley Integral de Transexualidad de Andalucía³. O también el derecho al reconocimiento de la personalidad jurídica de la “identidad de género que cada persona defina para sí”, según lo expuesto en los Principios de Yogyakarta (ICJ 2007), uno

1 La terminología es variable y en los últimos años se están consolidando en diversos ámbitos los términos “transgénero”, “trans” o “trans*”. En el presente artículo, dado que se citan documentos legales, se empleará “transexual” o “transexualidad” para facilitar la homogeneidad terminológica. Entiéndase dentro de estos el resto de posibles acepciones que cubren los términos alternativos señalados.

2 En este artículo se pone el foco en las vivencias de las personas transexuales a fin de acotar el marco de estudio. No obstante, las personas intersexuales viven sus propias experiencias de violencia derivadas de la imposición del paradigma de la genitalidad sobre sus cuerpos desde el mismo momento de su nacimiento a través de cirugías no consentidas u otras prácticas que actualmente reciben grandes críticas por parte de colectivos sociales e instituciones internacionales. Para un análisis enfocado en la intersexualidad, véase García López (2015).

3 Ley 2/2014, de 8 de julio, integral para la no discriminación por motivos de identidad de género y reconocimiento de los derechos de las personas transexuales de Andalucía.

de los documentos de referencia en derecho internacional en lo relativo a la interpretación de los Derechos Humanos en el colectivo LGBTI.

La autodeterminación del género surge en las últimas décadas como la propuesta político-social alternativa a la perspectiva médica y psiquiátrica que, desde la segunda mitad del siglo XX, se había instaurado como paradigma de acercamiento al fenómeno de la transexualidad. En 1977, la Organización Mundial de la Salud clasificó la transexualidad como síndrome médico. No fue sino hasta junio de 2018 que el mismo organismo la eliminó de la lista de patologías psiquiátricas, pasando a ser incluida en la lista de condiciones relativas a la sexualidad en la undécima publicación de la Clasificación Internacional de Enfermedades (CIE-11), bajo el título de “incongruencia de género”⁴. Asimismo, ha pasado también por diversas consideraciones que ha tenido en las varias ediciones del *Diagnostic and Statistical Manual of Mental Disorders* (DSM) de la Asociación Estadounidense de Psiquiatría, acabando por desaparecer del listado de trastornos en su quinta edición, aprobado en 2012⁵.

Este proceso suele ser denominado como despatologización de la transexualidad. Como puede verse, hasta la década pasada no se ha culminado en la desaparición completa de los principales manuales médicos. Sin embargo, como en muchas otras ocasiones, las instituciones públicas no reflejan los cambios sociales en un marco legal hasta que dicho cambio se ha consolidado social y políticamente. Por ello, mientras que en muchos países siguen vigentes leyes aún fundamentadas en la visión de la transexualidad como un fenómeno perteneciente al ámbito médico, en otros, como la pionera Argentina, o en Malta y Noruega en el contexto europeo, ya se han realizado cambios para adaptar su legislación a la visión no medicalizada (ILGA World 2020; ILGA-Europe 2020). Además, en los últimos años, instituciones internacionales como la Comisión Europea promueven la adaptación de leyes basadas en la autodeterminación entre los estados miembros que aún no hayan realizado dicho cambio (Comisión Europea 2020).

En España, ciertamente se reconoce el fenómeno de la transexualidad, se permite el correspondiente cambio registral de nombre y de género y se protegen algunos derechos de las personas transexuales⁶. Ahora bien, todo el proceso de transición legal pasa por el sistema de salud, que debe certificar ciertos requisitos psiquiátricos y hormonales ante los tribunales y registros para validarlo. En España, la ley 3/2007⁷ es la que regula que la petición de cambio registral debe incluir dos documentos médicos⁸: un diagnóstico psicológico de disforia de género y un certificado de tratamiento médico durante al menos 2 años, sin requerir cirugía

4 El País, 19 de junio de 2018, «La OMS saca la transexualidad de la lista de enfermedades mentales». Accesible en: https://elpais.com/internacional/2018/06/18/actualidad/1529346704_000097.html (consultado el 29 de diciembre de 2020).

5 El País, 5 de diciembre de 2012, «Los transexuales ya no son enfermos mentales». Accesible en: https://elpais.com/sociedad/2012/12/04/actualidad/1354628518_847308.html (consultado el 29 de diciembre de 2020).

6 Por ejemplo, se condenan ataques transfóbicos o se da derecho a los individuos a ser tratados acorde a su identidad de género y nombre sentidos en algunas instituciones públicas, aunque no en todas. No obstante, continúan sucediendo numerosos casos de, por ejemplo, exclusión institucional, imposición de intervenciones médicas o abusos policiales a personas transexuales e intersexuales (cf. FRA 2020).

7 Ley 3/2007, de 15 de marzo, reguladora de la rectificación registral de la mención relativa al sexo de las personas.

8 Cabe señalar que el artículo primero establece criterios formales de legitimación: nacionalidad española, mayoría de edad y capacidad suficiente. Estos criterios suponen la exclusión de base de personas no nacionales, menores o con diagnósticos psiquiátricos adicionales al de disforia de género.

genital, solo hormonación. Conviene aclarar que este último requisito está exento solo en caso de riesgo para la salud de la persona paciente. Todo ello según el artículo 4 de dicha ley.

Así, existen dos reivindicaciones sociales convergentes. La primera, como señala la Comisión Internacional de Juristas (ICJ), que la despatologización de la transexualidad culmine en la prohibición de obligar a alguien a “someterse a procedimientos médicos, incluyendo la cirugía genital, la esterilización o la terapia hormonal, como requisito para el reconocimiento legal de su identidad de género” (ICJ 2007, 12). La segunda, que la autodeterminación reemplace a todo criterio médico como requisito para su reconocimiento social y legal.

De esta forma es como se han enunciado recientes propuestas como la Proposición de Ley contra la Discriminación por Orientación Sexual, Identidad o Expresión de Género y Características Sexuales, y de Igualdad Social de Lesbianas, Gais, Bisexuales, Transexuales, Transgénero e Intersexuales, presentada el 12 de mayo de 2017 en el Congreso de los Diputados. Dicha iniciativa legal propone la modificación del artículo 4 de la Ley 3/2007, eliminando los dos requisitos arriba mencionados y sugiriendo que:

La efectividad del derecho al reconocimiento de la identidad de género y, en su caso, la rectificación registral de la mención del sexo no vendrá supeditada a la existencia de un diagnóstico de disforia de género ni a la acreditación de haberse sometido a ninguna terapia o tratamiento médico o psicológico ni a ninguna intervención quirúrgica de reasignación, total o parcial⁹.

A pesar de que esta propuesta no fuese aprobada en su momento, resulta interesante remarcar lo que casi año y medio después se estableció desde la Dirección General de los Registros y el Notariado. En la Instrucción sobre cambio de nombre en el Registro Civil de personas transexuales¹⁰, tomándose como punto de partida que la aceptación de una ley a nivel estatal está retrasándose en su discusión y aprobación, se afirma que “mientras eso llega, hay situaciones actuales que demandan una solución urgente”.

La Instrucción establece la posibilidad de presentar una solicitud de cambio de nombre —no de sexo— que será tramitada para adecuarlo a la identidad de género sentida por la persona, aunque sea discrepante con el sexo registrado. La justificación dada se basa en la vigente Ley de Registro Civil de 1957, que obliga a que el nombre registrado no induzca a confusión del sexo de la persona. La novedad, no obstante, radica en entender que el sexo de la persona denomina a la identidad sentida y no al sexo registrado, como podemos leer en la instrucción referida:

Debe analizarse cuál es el verdadero sexo correspondiente a las personas con disonancia de género, si el que viene dado por sus órganos genitales, que determinó que al nacer se le inscribiera como perteneciente al mismo, o el verdadera y profundamente sentido por dichas personas, y parece que la respuesta debería ser que este

9 Proposición de Ley 122/000097, publicada en el Boletín Oficial de las Cortes Generales (BOCG) el 12 de mayo de 2017, núm. 122-1.

10 Instrucción de 23 de octubre de 2018, de la Dirección General de los Registros y del Notariado, sobre cambio de nombre en el Registro Civil de personas transexuales.

último, dada la apuntada prevalencia de los factores psicosociales. De este modo, no se produciría una absoluta contradicción con la exigencia del art. 54, de impedir el error en cuanto al sexo¹¹.

Se ve, por consiguiente, el intento de adaptación del marco legal al proceso de despatologización señalado, otorgando prioridad al testimonio directo del sujeto, sin necesidad de que medie el personal sanitario o se establezcan otros criterios más allá de la autopercepción del sujeto. Por supuesto, la tramitación de la llamada “Ley Trans”, que se está llevando a cabo en la actualidad, está completamente alineada con la idea de dotar de fuerza legal a la autodeterminación de género como criterio.

Esto puede ser conceptualizado como un cambio en la forma de determinación de la identidad de género que de forma general se emplea social e institucionalmente. Para esclarecer este aspecto, en el siguiente apartado se aborda el modo en que de forma cotidiana las personas reconocen la identidad de otras, además de las instituciones. En primer lugar, desde un punto de vista descriptivo, para luego poder afrontar una crítica del sistema así descrito.

2. La determinación del sexo: la genitalidad como criterio

Como se señalaba al final del apartado anterior, la observación de los genitales externos en el momento del nacimiento constituye el modo automático de registro del sexo de un neonato. A lo largo de la vida del individuo, su identidad de género tiene una presencia muy importante, tanto en la percepción de sí mismo como en la percepción que otras personas tienen de él.

Esta asignación de acuerdo con la genitalidad, que se realiza de forma tan natural en la mayoría de las sociedades, ha sido señalada por Talia Mae Bettcher como el paradigma de determinación social del sexo. De hecho, esta va más allá del momento del nacimiento, alcanzando virtualmente todos los momentos de la vida de un individuo. A juicio de la autora, la presentación del género de una persona va ligada irremediabilmente a una presentación de su genitalidad, que juega un papel fundamental a la hora de determinar el sexo biológico, incluso cuando no es compartido por otras características como pudiera ser el cariotipo (Bettcher 2009, 106).

Siguiendo este modelo, frente a la declaración de una persona del tipo “soy un hombre” o “soy una mujer”, se estaría interpretando que posee, respectivamente, pene o vagina, ya que eso fue lo que en el momento de su nacimiento marcó su registro de sexo. Ahora bien, no parece que, cuando preguntamos por el género de alguien, estemos interesándonos por su estatus genital. En este sentido, está completamente fuera de lugar que, tras la declaración de la identidad de género de una persona, se le exija cualquier tipo de verificación de sus genitales. Más aún, cualquier intento de realizar tal acción supone una violación clara del derecho a la intimidad de la persona.

Desde una perspectiva descriptiva, el criterio de determinación del sexo de un bebé estaría basado en su genitalidad visible, y este criterio se perpetuaría de forma incuestionada aun cuando pierda su componente de constatación empírica. Como se aclara más abajo,

11 *Ib.*

esta perspectiva estaría estableciendo una relación causal donde solo hay una correlación de hechos. Aún más, desde una perspectiva crítica, cabe plantearse qué nivel de validez puede llegar a tener este criterio de determinación, considerando que es o bien incompatible con el derecho a la intimidad en caso de que debiera ser constatado, o bien no constatable empíricamente, en caso de que no se exija la revelación genital.

Esto es, precisamente, lo que Bettcher señala como una de las formas de discriminación más extendidas al concebir a las personas transexuales como “farsantes” (*pretenders*) o “impostores” (*deceivers*). Por ejemplo, cuando a una mujer transexual se le cuestiona y esta declara públicamente su anatomía genital, a menudo puede recibir declaraciones como: “en realidad es un hombre” (Bettcher 2014: 391). Esto esconde un sistema de verdad, en el que el criterio de verificación de la identidad de género de una persona es su estatus genital. De este modo, las personas cissexuales¹² serían quienes cuentan con la correspondencia *correcta* entre identidad de género y genitales, mientras que las personas transexuales se ubicarían en una correspondencia *errónea o desviada*. De ahí que se empleen formas adverbiales como “en realidad”, “realmente” o “verdaderamente” en declaraciones similares a la escrita anteriormente.

Otro ejemplo muy ilustrativo es expuesto por Julia Serano en referencia a la caracterización de personajes transexuales en medios como el cine. La autora señala que en conocidas películas y series existe un arquetipo de personaje, que es secretamente transexual, y que acaba teniendo un momento de “revelación de la verdad” en la trama. Tras este momento de revelación, su identidad sentida queda anulada por el descubrimiento de su anatomía física, que acaba imponiéndose como su supuesta identidad real (Serano 2007, 37).

Por último, a esto se puede añadir las cifras de homicidios violentos de personas transexuales. Según apuntan Kristen Schilt y Laurel Westbrook, solo el 33 por ciento de estos homicidios tenían una causa no relacionada con la condición de transexualidad de la víctima. A su vez, hasta el 56 por ciento del total resultó de interacciones sexuales íntimas en que el agresor se sintió de algún modo “engañado para realizar prácticas homosexuales” y un 6 por ciento por otros tipos de homofobia y transfobia (Schilt y Westbrook 2009, 452). El elevado número de agresores que alegó sentirse conducido a prácticas *homosexuales*¹³ refuerza la idea de que la identidad de género de la víctima es revelada acorde a sus genitales, imponiéndose a su propia identidad sentida.

Por tanto, la vindicación directa que surge tras haber señalado esta conexión es la necesidad de poner en crítica el papel de la genitalidad de una persona en el reconocimiento social del género. Legitimar un mecanismo social en que los genitales puedan ser un criterio para establecer la identidad de género, más allá de la etapa como bebé, es difícilmente conciliable con el respeto al derecho de la intimidad de las personas. No obstante, la suficiencia de la autodeterminación como criterio ha estado enfrentada con críticas por parte de un número importante de filósofos.

12 El término “cissexual” o “cisexual” fue introducido como antónimo de “transexual”. Se refiere a personas cuyo sexo sentido se corresponde con el sexo registrado al nacer. Al igual que ocurre con “transexual”, podemos encontrar las variantes “cis” o “cisgénero”.

13 En varios países, y especialmente en Estados Unidos, este tipo de defensa ha sido empleada exitosamente para rebajar la dureza de las sentencias. Esta estrategia de defensa se ha nombrado *trans panic*, por analogía y similitud con la defensa *gay panic*. Un estudio más detallado sobre el tema puede encontrarse en Wodda y Panfil (2015).

Entre otros, Pablo de Lora ha expuesto recientemente bastantes reservas a la hora de aceptar el cambio de paradigma hacia la autodeterminación. Para él, el tratamiento social y la creencia personal pueden diferir en lo relativo a la identidad de género de otra persona:

Tengo para mí que de un modo análogo [a como reconocemos como hijos naturales a los adoptivos] deberíamos interactuar con quienes se reclaman hombres o mujeres por mera voluntad, aunque en nuestro fuero interno creamos que no es tal identificación sino lo que la naturaleza haya brindado lo que finalmente tenga que contar (De Lora 2019, 193).

De este modo, aunque se acepte la identidad autopercebida en el trato cortés cotidiano con esa persona, podría objetarse, como hace Pablo de Lora, que esto debiese alcanzar la esfera de nuestras categorías y el modo en que entendemos el mundo. Existiría, por tanto, una distinción entre el plano ético del tratamiento y el plano epistémico del reconocimiento.

Evidentemente, existen razones de peso para tratar a una persona acorde con su identidad declarada. Actuar de forma contraria en un contexto cotidiano a menudo será reprochado como innecesaria e injustamente violento e irrespetuoso. Pero actuar de un modo determinado por razones de convivencia y respeto no necesariamente va ligado a una creencia al respecto de algo. La libertad religiosa podría ofrecer un símil ilustrativo. Alguien podría pensar que ciertos dogmas de algunas religiones son absurdos e incluso guarde para sí la consideración de que son creencias irracionales. Ahora bien, ante una persona que se declare profundamente creyente en un contexto cotidiano, acusarla de irracional o absurda sería algo reprochable de acuerdo con valores de convivencia y respeto que deberían primar.

De este modo, la determinación del género de los individuos humanos en el plano epistémico-ontológico es independiente y queda desconectado de las exigencias morales que se hagan en el trato cotidiano a personas transexuales. En la esfera de las creencias y el conocimiento, para Pablo de Lora, la genitalidad heredada de la persona tendría primacía sobre la autopercepción del individuo. La posición de este filósofo, que expone a lo largo de su obra, estaría indicando que la determinación del género sucede de acuerdo con características biológicas y es concluyente, es decir, que no puede ser modificada a lo largo de la vida del individuo. Para él, el dimorfismo sexual “es un hecho biológico, no una construcción cultural ni una realidad institucional, pues no depende de nuestros estados mentales, sino de cómo es la naturaleza” (De Lora 2019, 178). Esta tesis sería la opuesta a la de Judith Butler, para quien un cuerpo u otro no es determinante en la construcción de la identidad de mujer u hombre (Butler 1986, 35; 1990, 6).

Mientras que, a juicio de Pablo de Lora, el dimorfismo sexual constituye diferencias esenciales entre hombres y mujeres, para Judith Butler son características meramente accidentales y no conectadas de forma necesaria con la identidad de género de cada cual. Siguiendo la idea de Butler, entre la genitalidad de las personas y su identidad de género solo se puede establecer una correlación no causal, meramente probabilística. Así, el hecho de que al neonato se le registre un sexo u otro según la observación de sus genitales solo se basa en probabilidades de acierto: la mayoría de los hombres tienen pene y la mayoría de las mujeres tienen vagina. No obstante, de ningún modo, siguiendo a Butler, podríamos decir

que existe una relación causal entre genitalidad e identidad de género, sino que el cuerpo de cada persona sirve como ocasión de construcción de una identidad de género que en todo momento puede ser de distintas formas (Butler 1986, 46; 1990, 111-112).

Ahora bien, se puede ver que la tesis de Pablo de Lora contraviene a lo expuesto anteriormente acerca de la forma normal en que se suele descubrir la identidad de género de una persona. Esto es, preguntándole. Pongamos el caso de que alguien dice identificarse como mujer. Sin embargo, juzgando su apariencia física y basándose en creencias personales, uno considera que esa persona “es un hombre por naturaleza”, siguiendo la tesis de Pablo de Lora. Por supuesto, en todo momento sería posible hacer cavilaciones internas e, incluso, el más imprudente podrá lanzarse a preguntarle directamente, aun siendo esta una pregunta de lo más íntima¹⁴. Ahora bien, si la persona preguntada se acoge a su derecho de no hacer público su estatus genital u otras condiciones anatómicas, genéticas o biográficas, como generalmente se consideraría que debería suceder, ¿no resulta que uno debería acogerse al testimonio de la propia persona y desistir de su indagación al respecto? De este modo, dado que el estatus anatómico y genético de la persona es de ámbito privado, la identidad sentida y declarada parece una apuesta mucho más accesible y segura para cerrar la deliberación sobre la identidad de género de esta.

De hecho, se entiende que el estatus genital de una persona solo se debería revelar en el ámbito íntimo de la persona o por motivos médicos relevantes. Por consiguiente, dado que la identidad de género de una persona afecta fundamentalmente a su vida en el trato social e institucional que reciba día a día, resulta coherente reducir la relevancia de la condición corporal de la persona a su ámbito privado y priorizar la identidad que expresa públicamente en el resto de su vida.

Por ello, la autoridad de la primera persona¹⁵ se puede plantear como una característica privilegiada para establecerse como una vía alternativa para el reconocimiento de la identidad de género de una persona. La autoridad de la primera persona permite analizar el acceso asimétrico que los agentes tienen a los contenidos mentales propios respecto de los ajenos. En línea con las ideas expuestas en párrafos anteriores, la identidad de género de una persona sería accesible, si no de forma exclusiva por ella misma, al menos sí de forma mucho más inmediata y eficaz.

Sin embargo, antes de seguir profundizando en este ámbito, resultará útil establecer una base sobre la que poder explicar y justificar el fenómeno de la autoridad en un sentido más amplio. De este modo, tras un análisis más detallado acerca de qué supone tener o ser una autoridad sobre algo, se podrá entender mejor de qué forma puede establecerse la autodeterminación del género como un criterio más adecuado que el estatus genital.

14 Son muchas las personas que señalan esto como una de las discriminaciones más normalizadas y extendidas que sufren las personas transexuales, puesto que es una pregunta que típicamente reciben estas por el hecho de ser transexuales, y que les pone en una situación social delicada y de presión. Ozturk lo conceptualiza como una violación de la intimidad prohibitiva (Ozturk 2017, 147).

15 Este término proviene del inglés “*first-person authority*”, y debe entenderse como un privilegio de acceso epistémico en la relación entre el sujeto y sus propios estados mentales. No debe confundirse con que la autoridad de la primera persona solo se aplica a la propia persona y no al resto de agentes (Cf. Davidson 1984, 101-102).

3. Razones y autoridad

Para introducir la argumentación sobre autoridad y creencias, merece la pena tomar como punto de partida el célebre capítulo 10 de la *Investigación sobre el entendimiento humano*. David Hume afirmó en él que el modo normal en que un sujeto conforma sus creencias consiste, resumidamente, en un balance de todas las razones pertinentes al respecto de algo, optando por la interpretación que parece más sólida:

Un hombre sabio [...] sopesa las experiencias contrarias, considera qué posibilidad es la apoyada por el mayor número de experiencias, se inclina por esta posibilidad con dudas y vacilaciones, y cuando, finalmente, ha fijado su juicio, la evidencia no excede a lo que, hablando con propiedad, llamamos *probabilidad*. Toda probabilidad, por tanto, supone una oposición de experiencias y observaciones, encontrándose que una posibilidad aventaja a las otras (Hume 1988, 135).

En la terminología epistemológica actual, es posible entrever en Hume lo que hoy se denominan “razones para creer” (*reasons for belief*). Algunas de las experiencias de las que el filósofo habla no necesariamente son observaciones directas o sensaciones percibidas en primera persona, sino también testimonios o información aportada por otros sujetos creíbles. Las razones para creer serían, por tanto, hechos¹⁶ que hablan en favor de sostener una cierta creencia como verdadera.

De forma normal, es posible apuntar que los agentes puedan tener creencias muy dispares al respecto de un mismo fenómeno o cosa. También pueden discrepar en que, ante una circunstancia en la que se deba tomar una decisión, la acción adecuada puede ser una u otra. Por supuesto, los agentes sostienen una posición u otra porque es la que consideran correcta. Algunos autores, como Linda Zagzebski, explican esto como el “deseo natural por la verdad” (Zagzebski 2012, 33), mientras que otros aluden a argumentos pragmatistas para justificar esta posición basándose en la consecución de fines loables (Williams 2002). Pero aun ofreciendo explicaciones diferentes, lo más importante es destacar que los agentes ofrecerán razones para creer cuando deban rendir cuentas de por qué creen lo que creen. De este modo, tanto nuestras decisiones como nuestras creencias descansan sobre aquellos hechos que podemos aportar como explicación y justificación. El tipo de razones de interés para el presente artículo son las razones de autoridad. Estas se pueden identificar como aquellas que se defienden de forma espontánea mediante la afirmación “es una razón correcta porque lo ha dicho A”, siendo A un agente con autoridad en el asunto implicado.

En el campo de la autoridad, la concepción desarrollada por Joseph Raz es de interés para establecer la base teórica de la autoridad epistémica. La concepción de la autoridad como servicio de Raz constituye una explicación del fenómeno de la autoridad de notable influencia en la actualidad. Raz parte de un proyecto de análisis de otras explicaciones de la autoridad, como la crítica frontal al anarquismo de Wolff o la revisión a la concepción de

16 Podría ser discutido si las razones son solo hechos o estas incluyen ciertos estados psicológicos e intencionales. A efectos de limitar la extensión, en este artículo tomo como base la idea de que las razones son hechos objetivos del mundo y, por tanto, pueden ser reconocidas y compartidas por todos los agentes (Cf. Raz 1975, 17).

autoridad de Hart. A partir de ahí, este filósofo propone que la forma normal de justificar la autoridad se basa en que supone una forma de acertar mejor respecto a qué debería ser llevado a cabo. Según afirma, la autoridad satisface dos tesis: dependencia y justificación normal. De este modo, en *The Morality of Freedom*, señala que cualquier mandato de autoridad se refiere a razones que se aplican al caso con anterioridad al propio mandato (dependencia) y es más eficaz para alcanzar una deliberación correcta que la que podría alcanzar de forma independiente el sujeto que lo recibe (justificación normal) (Raz 1986, 42-57).

Para que se aprecie adecuadamente la relevancia de esta concepción para el propósito de este artículo, resulta necesario concretar mejor lo que resulta de interés de estas dos tesis. En primer lugar, cuando una autoridad articula un mandato, debe ser esperable que las razones a las que alude ya eran aplicables a la acción que prescribe con anterioridad. Dicho de otra forma, la autoridad debe emplear y señalar razones que ya eran pertinentes, pero que quizás no estaban siendo consideradas de la forma adecuada a la hora de tomar una decisión. En esto consiste la tesis de la dependencia. De este modo, una autoridad no puede legítimamente realizar mandatos caprichosos o injustificados, sino que tiene que atender a las circunstancias y variables que afectan a la deliberación (Raz 1986, 47).

Por otra parte, los mandatos de autoridad no son meras declaraciones lanzadas al aire, sino que aspiran a ser obedecidos por sus receptores. Esto implica que debe existir una justificación para tal obediencia. A juicio de Raz, ello se condensaría en la tesis de justificación normal, según la cual la obediencia a un mandato de autoridad se justifica por ser una mejor estrategia para actuar del mejor modo (Raz 1986, 53).

Una figura de autoridad se ubicaría, por tanto, en una posición más adecuada que quien obedece para deliberar y sopesar las razones que se aplican al caso. De este modo, la concepción de la autoridad como servicio señala la importancia de entender que la naturaleza de este fenómeno es servir para acertar más eficazmente en las acciones que se realizan. En otras palabras, una autoridad es merecedora de ser obedecida por su efectividad a la hora de prescribir lo que es más acertado. Por lo tanto, la posibilidad de la desobediencia puede verse como una estrategia ineficaz.

Aunque la concepción de Raz esté descrita en casi toda su obra como una explicación de la autoridad en sentido práctico, aplicada al ámbito de las acciones, resulta útil su traslación al ámbito de las creencias. Por ejemplo, parece lógico pensar que una persona experta en nutrición será tomada como una autoridad sobre, entre otras cosas, el impacto esperable de la ingesta de un determinado alimento sobre el cuerpo de quien lo consuma. Si esta persona afirma que el consumo frecuente de dicho alimento puede provocar trastornos en el metabolismo, sería prudente creerle, incluso aunque para uno sean desconocidas todas las investigaciones realizadas o no tenga conocimiento científico alguno. Esto parece satisfacer las dos tesis de Raz señaladas anteriormente. En primer lugar, las razones empleadas por la persona experta serán hechos empíricos que podrían ser verificados por cualquiera, y no consideraciones u opiniones personales —tesis de dependencia—. En segundo lugar, su afirmación actúa con autoridad ya que es una forma más eficaz de obtener conocimiento verdadero que intentar deducirlo por uno mismo —justificación normal— (Zagzebski 2012, 109-110).

Una vez se ha desarrollado la concepción de la autoridad como servicio y la posibilidad de extenderla al ámbito de las creencias, esto debe ser ahora contrastado con la determinación de la identidad de género. Si la autoridad de la primera persona demuestra ser una forma eficaz

de determinarla, entonces la concepción de la autoridad como servicio constituiría una posibilidad de justificación apropiada para defender la autodeterminación como criterio fundamental.

4. La autoridad de la primera persona en la identidad de género

La autoridad de la primera persona ha sido conceptualizada de diversas formas, de entre las que se pueden destacar tres: epistémica, ética y negociativa. La concepción epistémica se basa en el tipo de relación que se establece entre un sujeto y sus propios estados mentales, frente a la relación de terceras personas con estos. En la valoración de la afirmación de una persona del tipo “la casa está en llamas”, Donald Davidson señala que existen al menos tres aspectos a evaluar: si la casa está en llamas, si el hablante tiene la creencia de que la casa está en llamas y el modo en que el fuego ha causado su creencia (Davidson 1984). La autoridad de la primera persona se hace manifiesta en el segundo de los aspectos, que se justifica en breve más abajo.

En segundo lugar, la concepción ética pone en valor las profundas dudas que se puede tener al considerar la frecuencia con la que se producen autoengaños, sesgos o negaciones inconscientes. Ante esto, la posición ética de la autoridad de la primera persona resta relevancia al factor epistémico del acceso privilegiado, para asentar la justificación sobre la responsabilidad individual de quien expresa estados psicológicos propios. De acuerdo con Krista Lawlor, la autoridad de uno sobre sus propias emociones y estados no descansa sobre si tiene razón o no, sino sobre su responsabilidad sobre ellos (Lawlor 2010, 550). De este modo, cuando alguien comunica que tiene o no un deseo de hacer algo, la respuesta del resto de personas debería ser acorde con ello. Por consiguiente, una persona no puede asumir el papel de asignar a otra sus propios estados mentales. Por más que pueda preguntarle acerca de ellos, la última decisión de establecer cuáles son y rendir cuentas de ellos recae siempre sobre la primera persona (Bettcher 2009, 101-103).

Sin embargo, la concepción ética implica una obligación por parte de terceras personas de no rechazar las afirmaciones desde la autoridad de la primera persona que, para algunos filósofos, resulta injustificadamente tajante (Ozturk 2017, 140). Por ello, la teoría negociativa se plantea como una versión más flexible que la anterior, además de más ajustada al ámbito de las identidades personales. Según Burkay Ozturk, quien formula y da nombre a esta concepción de la autoridad de la primera persona, existen conceptos identitarios que se definen de forma comunitaria, no individual, tales como la nacionalidad, la religión o la propia identidad de género. Así, para Ozturk, la adscripción de una identidad comienza desde la primera persona, pero no concluye aquí. La identificación de alguien podría aún ser puesta en duda por terceras personas, siempre que esta duda no viole tres límites éticos: no dañar, no invadir la privacidad y respetar la dignidad (Ozturk 2017, 145-150).

No obstante, Davidson señaló un aspecto de la concepción epistémica de la autoridad de la primera persona que bien puede conectar con las otras dos concepciones. También permite conectar con la concepción de la autoridad como servicio, como se expone más adelante. Este aspecto se basa en que la interpretación del mensaje por parte del receptor está necesariamente mediada por esquemas mentales propios y asunciones. Por el contrario, para el emisor, que quiere ser entendido, no procede la duda acerca de si realmente quería decir lo que dice (Davidson 1984, 110). Por supuesto, el emisor siempre puede equivocarse

en el significado las palabras escogidas, pero esto no elimina la asimetría de la adscripción de estados mentales entre el emisor y el receptor.

De este modo, las tres perspectivas de la autoridad de la primera persona, aun difiriendo en diversos puntos, se mostrarían de acuerdo en que existen razones para justificar la asimetría entre la relación de los estados psicológicos con uno mismo y con otros. A partir de esto, difieren al respecto de si esta asimetría debe entenderse en términos éticos, sociolingüísticos o epistémicos (Gertler 2020). Para poder conectar adecuadamente con la concepción de la autoridad como servicio de Raz, la perspectiva de Davidson definida en el párrafo anterior puede resultar suficiente y apropiada para mostrar cómo la autoridad de la primera persona puede justificarse.

Joseph Raz ya señala que, a pesar de que su teoría esté desarrollada teniendo en el foco a la autoridad en sentido práctico, el fenómeno de la autoridad en el ámbito epistémico debe tener la misma estructura básica (Raz 1986, 53). En definitiva, lo definitorio de la autoridad como fenómeno es que funciona en el ámbito de las razones, por tanto, en el campo deóntico. Aunque la autoridad práctica y la epistémica difieran en sus contenidos, su justificación es común y radica en el qué *debe* uno hacer o creer (Raz 1986, 30). Por consiguiente, si se acepta que las autoridades epistémicas tienen plena cabida en la concepción de la autoridad como servicio, la justificación de la autoridad de la primera persona debería satisfacer las dos tesis de justificación que fueron expuestas más arriba.

En cuanto a la tesis de la dependencia, la autoridad de la primera persona debería funcionar siempre señalando razones para creer algo que ya existía. Cuando la comunidad científica ha estado defendiendo el cambio climático, siempre ha hecho uso de diversas observaciones, mediciones y hechos que son relevantes y ya existían. Aunque uno no sepa los métodos por los que se puede calcular el incremento de la temperatura media global, no significa que tal incremento no sea real, y la comunidad científica se limita a señalarlo como una razón a favor del cambio climático.

Del mismo modo, ante la afirmación de una persona del tipo “soy un hombre”, las razones que se pueden ofrecer para sostener esa afirmación deberían satisfacer esta tesis. Estas razones pueden variar según la concepción de la conformación del sexo que se suscriba. Como se ha señalado, para algunos autores es una cuestión anatómica. Para otros, es una identidad construida a partir de vivencias internas o comportamientos sociales (Cf. Mikkola 2019). Ahora bien, como señala Ozturk, la identidad de género a menudo es presentada como multifactorial y las razones pertinentes para defenderla son variables. Sin embargo, la pertinencia de estas para poder ser sostenidas como buenas razones radica en gran medida en que pueda ofrecerse una justificación razonable (Ozturk 2017, 150-151).

Tómese un caso hipotético a menudo planteado por los críticos de la autodeterminación: un hombre que, con la finalidad de beneficiarse de medidas de acción positiva para la igualdad entre mujeres y hombres, decide presentar una petición de cambio de sexo para ser considerado como mujer. ¿Qué razones son esperables que aporte al afirmar que “es una mujer”? Más allá de las razones anatómicas que algunos podrían señalar, la vivencia experiencial que este sujeto pueda tener como “ser mujer” diferirá radicalmente de la que una mujer transexual normalmente tiene.

De este modo, “sentirse mujer” sería una razón que se adapta a la tesis de dependencia solo en el caso de la mujer transexual, pues es un hecho que esta mujer tenía bien formado y

con anterioridad a su declaración de identidad de género (Cf. Bettcher 2014). Esto mismo, por supuesto, no ocurre en el caso del hombre que pretende defraudar, pues su supuesta vivencia de “ser mujer” surge *a posteriori* para intentar defender su fingida declaración de género.

Por otro lado, la tesis de justificación normal tiene su cumplimiento en la asimetría ya señalada por Davidson y expuesta antes en este mismo apartado. Tanto si se cuentan las razones anatómicas como las experienciales como razones válidas, el cumplimiento de esta tesis debería comenzar por un mejor acceso a estas. En lo relativo a las experiencias y vivencias de las personas, ya ha quedado señalado que desde la primera persona existe un privilegio de interpretación. Esto justifica colocar al testimonio de la propia persona como un modo mejor de conocer su propia identidad que intentar interpretar por uno mismo desde un punto de vista externo (Gertler 2020).

En lo relativo al cuerpo de las personas, la justificación no radica en la capacidad de observación empírica, sino en las limitaciones éticas de invasión de privacidad. Dicho de otro modo, solo la propia persona debe poder decidir a quién comunica o muestra sus genitales u otras consideraciones corporales o biográficas. Por tanto, la determinación de género mediante la genitalidad debería rechazarse y abogar por la autoridad de la primera persona que, en este sentido, estaría justificada por las razones morales señaladas (Bettcher 2009, 107). Así, la tesis de justificación normal estaría satisfecha en la medida en que el testimonio de una persona sobre su propia identidad de género es una mejor forma de crear un juicio al respecto que emplear percepciones y razonamientos en tercera persona.

5. Conclusiones

El objetivo fundamental en este artículo se centraba en la validez y justificación de la autodeterminación de la identidad de género, que acaban siendo demostradas mediante las tesis de justificación de la autoridad y el fenómeno de la autoridad de la primera persona.

Para alcanzar esta demostración, los dos primeros apartados sirven para analizar y exponer la forma tradicional en que se determina la identidad de género, tanto a nivel institucional como social. A rasgos generales, la determinación del sexo de los individuos se realiza de forma prácticamente hegemónica basándose en criterios de genitalidad u otras características físicas. No solo en el momento del nacimiento, sino también en el resto de las etapas de la vida. En sentido opuesto, en las últimas décadas se han comenzado a mover vindicaciones relacionadas con el reemplazo por criterios de autodeterminación, que prioricen la vivencia en primera persona del individuo sobre otras características físicas.

Los criterios tradicionales conllevan diversos problemas de conciliación con la despatologización de la transexualidad, con el derecho a la intimidad y con la relevancia de la autoridad de la propia expresión de la identidad. Por el contrario, la autoridad de la primera persona, fenómeno que ha sido extensamente estudiado y desarrollado en epistemología, supone una herramienta que evita estos inconvenientes y que resulta privilegiada para otorgar una justificación apropiada a la visión de la autodeterminación de la identidad de género. Desde esta perspectiva, la identidad de género de una persona puede reconocerse socialmente basándose en el testimonio que el individuo ofrece acerca de su autopercepción. Este testimonio tendría primacía epistémica sobre las interpretaciones que externamente se

puedan hacer sobre su identidad, de acuerdo con la asimetría del acceso en primera persona a los contenidos mentales propios.

Así, se abre camino a la conexión de la autodeterminación con una teoría de justificación de la autoridad como la de Joseph Raz. Desde la concepción de la autoridad como servicio, se puede ver cómo la autoridad de la primera persona satisface las tesis necesarias para dar una justificación apropiada a la autodeterminación de la identidad de género. Por un lado, se satisface la tesis de dependencia al mostrar cómo las razones ofrecidas por los individuos para rendir cuentas de su identidad son vivencias y autopercepciones que ya existían con anterioridad al momento del testimonio público. Por otro lado, la tesis de justificación normal queda satisfecha también dado que dicho testimonio supone una forma más adecuada de reconocerle a alguien su identidad de género correcta.

De esta forma, la justificación y la validez de la autodeterminación que se buscaban quedan cubiertas y conectadas mediante dicha concepción de la autoridad. La autodeterminación, en virtud de estar basada en la autoridad de la primera persona, se constituye como un modo mejor, más accesible y más confiable de determinar y reconocer la identidad de género de una persona, evitándose los problemas asociados a otros criterios.

Referencias bibliográficas

- Agencia de los Derechos Fundamentales de la Unión Europea (FRA) (2020), *A long way to go for LGBTI equality*, Luxemburgo: Publications Office of the European Union. Accesible en: <https://fra.europa.eu/en/publication/2020/eu-lgbti-survey-results> (consultado el 29 de diciembre de 2020).
- Alventosa del Río, J. (2016), «Menores transexuales. Su protección jurídica en la Constitución y legislación española», *Revista Española de Derecho Constitucional*, 107, pp. 153-186.
- Bettcher, T. M. (2009), «Trans Identities and First-Person Authority», en: Shrage, L. (ed.): *You've Changed: Sex Reassignment and Personal Identity*, Nueva York: Oxford University Press, pp. 98-120.
- Bettcher, T. M. (2014), «Trapped in the Wrong Theory: Re-Thinking Trans Oppression and Resistance», *Signs: Journal of Women in Culture and Society*, 39(2), pp. 383-406.
- Butler, J. (1986), «Sex and Gender in Simone de Beauvoir's Second Sex», *Yale French Studies*, 72, pp. 35-49.
- Butler, J. (1990), *Gender trouble: feminism and the subversion of identity*, Londres: Routledge.
- Comisión Europea (2020), *Union of Equality: LGBTIQ Equality Strategy 2020-2025*. Accesible en: https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/lesbian-gay-bi-trans-and-intersex-equality/lgbtiq-equality-strategy-2020-2025_en (consultado el 29 de diciembre de 2020).
- Comisión Internacional de Juristas (ICJ) (2007), *Principios de Yogyakarta: Principios sobre la aplicación de la legislación internacional de derechos humanos en relación con la orientación sexual y la identidad de género*. Accesible en: <https://www.refworld.org/es/docid/48244e9f2.html> (consultado el 29 de diciembre de 2020).
- Davidson, D. (1984), «First Person Authority», *Dialectica*, 38(2-3), pp. 101-111.

- De Lora, P. (2019), *Lo sexual es político (y jurídico)*, Madrid: Alianza Editorial.
- García López, D. (2015), «La intersexualidad en el discurso médico-jurídico», *Eunomía. Revista en Cultura de la Legalidad*, 8, pp. 54-70.
- Gertler, B. (2020), «Self-Knowledge», en: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy*. Accesible en: <https://plato.stanford.edu/archives/spr2020/entries/self-knowledge/> (consultado el 29 de diciembre de 2020).
- Hume, D. (1988), *Investigación sobre el entendimiento humano*, Jaime de Salas Ortueta (tr.), Madrid: Alianza Editorial.
- ILGA World (2020), *Trans Legal Mapping Report 2019: Recognition before the Law*. Accesible en: <https://ilga.org/trans-legal-mapping-report> (consultado el 29 de diciembre de 2020).
- ILGA-Europe (2020), *Annual Review of the Human Rights Situation of Lesbian, Gay, Bisexual, Trans and Intersex People in Europe and Central Asia 2020*. Accesible en: <https://www.ilga-europe.org/annualreview/2020> (consultado el 29 de diciembre de 2020).
- Lawlor, K. (2010), «Elusive Reasons: A Problem for First-Person Authority», *Philosophical Psychology*, 16(4), pp. 549-564.
- Mikkola, M. (2019), «Feminist Perspectives on Sex and Gender», en: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy*. Accesible en: <https://plato.stanford.edu/archives/fall2019/entries/feminism-gender/> (consultado el 29 de diciembre de 2020).
- Ozturk, B. (2017), «The Negotiative Theory of Gender Identity and the Limits of First-Person Authority», en: Halwani, R., Soble, A., Hoffman, S. y Held, J. (eds.): *The Philosophy of Sex*, New York: Rowman & Littlefield, pp. 139-159.
- Raz, J. (1975), *Practical Reason and Norms*, Oxford: Oxford University Press.
- Raz, J. (1986), *The Morality of Freedom*, New York: Oxford University Press.
- Schilt, K. y Westbrook, L. (2009), «Doing Gender, Doing Heteronormativity: 'Gender Normals,' Transgender People, and the Social Maintenance of Heterosexuality», *Peer Review Articles*, 7, pp. 440-464.
- Serano, J. (2007), *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity*, Emeryville: Seal.
- Williams, B. (2002), *Truth & Truthfulness: An Essay in Genealogy*, Princeton: Princeton University Press.
- Wodda, A. y Panfil, V. R. (2015), «Don't Talk to Me about Deception: The Necessary Erosion of the Trans* Panic Defense», *Albany Law Review*, 78(3), pp. 927-971.
- Zagzebski, L. (2012), *Epistemic Authority: A Theory of Trust, Authority and Autonomy in Belief*, New York: Oxford University Press.



Investigación filosófica y emociones: la ontología histórica de Ian Hacking como alternativa

Emotions Research and Philosophy: Ian Hacking's historical ontology as an alternative*

JUAN MANUEL ZARAGOZA BERNAL**

Resumen: Los estudios filosóficos sobre emociones se pueden dividir en dos grandes grupos: 1) aquellos que consideran que las emociones son naturales, resultado de nuestra evolución; 2) aquellos que, interesados en el pensamiento de autores concretos *respecto* a las emociones, no sostienen ninguna posición sustantiva acerca de ellas.

En este artículo exploramos una tercera opción: entender que las emociones son históricas, contextuales y socialmente definidas. Abrimos así la posibilidad a una nueva forma de investigación filosófica, que partiría de la *ontología histórica* del filósofo Ian Hacking como herramienta desde la que abordar el estudio filosófico de las emociones *en* la historia.

Palabras clave: emociones; clases naturales; Ian Hacking; ontología histórica; experiencia.

Abstract: Emotions research in Philosophy could be classified into two main groups: 1) Those that considers emotions as a natural feature, the result of our evolutionary history; 2) Those that are interested in the study of what a particular philosopher *said about* the emotions, and, therefore, do not sustain any substantive claim about the nature of the emotions.

I explore a third option: to understand emotions as historical, contextual, and socially constituted. This new approach opens up the door to a new form of philosophical inquiry on emotions in their historical contexts, based on Ian Hacking's historical ontology.

Keywords: emotions; natural class; Ian Hacking; historical ontology; experience.

Recibido: 07/01/2020. Aceptado: 28/07/2020.

* Este trabajo es resultado del proyecto de investigación “Historia y filosofía de la experiencia. Elementos objetivos y subjetivos del bienestar (wellbeing) en la historia, la cultura y la práctica clínica”, financiado por el MINECO (FFI2016-78285-R).

** Doctor por la Universidad Autónoma de Madrid y profesor asociado en la Facultad de Filosofía de la Universidad de Murcia. Correo electrónico: jm.zaragozabernal@um.es. Líneas de investigación: Historia y filosofía de la experiencia y de las emociones; *Ontological turn*: nuevo materialismo y ontología histórica. Publicaciones recientes: Zaragoza, JM (2021), “A Change of Pace. The History of (Emotional) Experiences”. *History of Psychology*; Zaragoza, JM (2022), *Componer un mundo un común. Una introducción al pensamiento de Bruno Latour*. Madrid: Lengua de Trapo.

1. Introducción

En los últimos años hemos asistido a un renovado interés por las emociones que se ha manifestado, de forma transversal, tanto en las disciplinas científicas como en las humanísticas. Así, hemos contemplado la aparición de una serie de estudios sobre emociones que desbordaban a la disciplina que, tradicionalmente, les había prestado mayor atención: la psicología (Gendron y Barrett 2009; Reizenzein 2015). Al decidido avance de las *neurociencias afectivas* (Panksepp 2004) se unieron la sociología (Hopkins 2009), la antropología (Leavitt 1996), la economía (Loewenstein 2000) y la historia (Reddy 2001), en lo que se ha dado en llamar un *giro afectivo* (Clough y Halley 2007). La filosofía no ha sido ajena a esta tendencia, y así hemos podido ver cómo el interés por las emociones se extendía desde su primer núcleo de desarrollo alrededor de la filosofía de la mente (Gordon y Gordon 1990), para encontrarla en la epistemología (Brun, Doguoglu, y Doguoglu 2016), la política (Kassab 2016), la ética (Bagnoli 2011) o la estética (Alcaraz 2011).

Es en los primeros años de la década de los 2000 cuando se publican algunas de las principales obras que tratan de ofrecer una teoría de las emociones que facilite su estudio por parte de las humanidades (Nussbaum 2001; Reddy 2001). El principal problema al que debían enfrentarse, como ya habían señalado anteriormente los antropólogos, consistía en la necesidad de superar la idea de que las emociones no eran más que una parte de nuestra realidad biológica y que, por tanto, no eran un objeto apropiado para las humanidades y las ciencias sociales (Lutz y White 1986). Encontraron para ello un aliado en la psicología cognitiva, que sostenía que las emociones tenían un papel en nuestros procesos cognitivos (Frijda 1986). Sin embargo, no es este el lugar para seguir esta historia, que ya hemos narrado en otros textos (Zaragoza 2013, 2021). Baste por tanto con señalar que la aproximación a las emociones de la psicología cognitiva, pasada por el tamiz de Nussbaum, es una de las principales influencias en la actual filosofía de las emociones, sobre todo en la ética y en la filosofía política (Nussbaum 2013), pero no exclusivamente. La epistemología, por poner un ejemplo, viene desde hace años trabajando en la relación entre emoción y conocimiento (Gordon 1969).

Esta no es, sin embargo, la única forma de entender las emociones que ha tenido un impacto en la filosofía. Encuadrados en una posición opuesta, encontramos a autores como Jesse Prinz. El libro que le dio a conocer, *Gut Reactions*, tenía como subtítulo *A Perceptual Theory of Emotion*, y proponía explícitamente una teoría alternativa a la cognitiva que basaba en William James (1884). La propuesta de Prinz difiere en aspectos fundamentales respecto a la que plantea Nussbaum, principalmente porque niega la necesaria vinculación entre juicio y emoción que esta sostiene, pero también por otros factores secundarios acerca de su funcionamiento. Prinz entiende que su propuesta es una alternativa que pasa por “poner al día” la teoría jamesiana de las emociones, de forma que los cambios corporales que son parte de la emoción cumplirían el papel que, en la teoría cognitiva, cumplían los juicios, distinguiendo así entre *embodied appraisals* y *cognitive appraisals*.

Sin embargo, más allá de lo que separa a cada una de estas aproximaciones al estudio filosófico de las emociones, hay algo que las une: su posición acerca de la “naturaleza” de las emociones. O, dicho de otra forma, ambas sostienen que las emociones son *clases naturales*. En el caso de Prinz es evidente, ya que dedica todo un capítulo de su libro a sos-

tener este aspecto (Prinz 2004, 79-102), y es conocida su posición al respecto¹. En el caso de Nussbaum es un poco más complejo, pero como bien señala Paul Griffiths, en tanto que Nussbaum considera que las emociones serían miembros “de una única clase” que comparten ser “*intelligent responses to the perception of value*” (Nussbaum 2001, 1), es decir, en tanto que comparten un *contenido* que sería igual o similar, se estaría adscribiendo a una de las muchas definiciones posibles de clase natural aplicada a las emociones. El objetivo de este artículo es, precisamente, poner en cuestión la tesis común de que las emociones son clases naturales. Una tesis que podríamos resumir, en palabras de Feldman Barret, como sigue:

It is assumed that each kind of emotion can be identified by a more or less unique signature response (within the body) that is triggered or evoked by a distinct causal mechanism (within the brain). As a result, it should be posible to recognize distinct emotions in other people, identify them in oneself, and measure them in the face, physiology, and behaviour. (Barrett 2006b, 30)²

En lo que sigue, haremos una breve introducción a la visión clásica de las emociones entendidas como clases naturales, que ejemplificaremos a través de la obra de Jesse Prinz. Escogemos a Prinz porque, como veremos, sostiene que las emociones son clases naturales en un sentido fuerte, esencialista. Pasaremos, posteriormente, a cuestionar esta idea para proponer un cambio en nuestra comprensión de las emociones. Nuestra propuesta, en este aspecto, es que debemos entender las emociones no como clases naturales, sino como clases *humanas*, un concepto desarrollado por Ian Hacking. A partir de esta crítica y de esta definición, esbozaremos un programa de investigación, basado en la ontología histórica de Hacking, que abra nuevas vías para el estudio filosófico de las emociones.

2. Las emociones como clases naturales

Es ampliamente conocido que la doctrina de las clases naturales evoluciona a partir de los trabajos de Whewell y Mill sobre lógica e inducción (Whewell 1847; Mill 1872), y en concreto de aquellas partes en que ambos discuten acerca de la función y finalidad de la *clasificación* (Hacking 1991a). Sin embargo, la importancia de las clases naturales en el debate filosófico de finales del siglo XX y principios del XXI se debe a las aportaciones de Kripke y Putnam, realizadas en la década de los años 70, a la semántica de las clases naturales³. La que se ha conocido como la teoría Kripke/Putnam de las clases naturales viene a decir que es la esencia de estas lo que hace que un objeto pertenezca a la clase, y no nuestra descripción del mismo. El ejemplo más claro de esta idea nos lo proporciona la

1 Es justo señalar que la posición de Prinz ha variado con el tiempo, y que actualmente su propuesta es mucho más matizada de lo que era en 2004 (Eickers, Loaiza & Prinz 2017).

2 Feldman Barrett hace una excelente análisis de cómo la teoría de los *appraisals*, que influye tanto a Nussbaum como a Reddy, mantiene los compromisos mínimos derivados de esta visión naturalista de las emociones: la modularidad, la conexión causal, el automatismo de la respuesta, etc. Ver Barrett 2006b, 30-31.

3 Aunque, en rigor, fueran Russell y Quine los que introdujesen la noción de clase natural en la filosofía moderna, en el caso de Quine como respuesta al problema de la inducción propuesto por Goodman a través del concepto de *grue*. Se encontrará una excelente historia del problema de las clases naturales en (Lemeire 2015).

tabla periódica. Lo que la tabla periódica nos presenta no es otra cosa, de acuerdo a este nuevo esencialismo, que los candidatos mejor posicionados para ser considerados clases naturales, en tanto que es la estructura y organización de sus átomos lo que determina sus propiedades observacionales. Como en el famoso experimento mental de las tierras gemelas, que una cosa sea o no sea agua no lo decide nuestro estado mental, sino que su composición sea H₂O o no lo sea (Putnam 1975)⁴. La hipótesis Kripke/Putnam, sin embargo, presenta muchas más dificultades cuando se intenta extender su uso a las ciencias “especiales” (Dupré 1981). Así, en filosofía de la biología el consenso extendido es que los taxones no son el resultado de ninguna *esencia* previa, sino de un grupo de características X compartido por una proporción significativa de los miembros: “[...] *species are distinguished by clusters of phenotypic traits that tend to co-vary [...] not by essential characters*” (Okasha 2002, 196). Sería precisamente el estudio del funcionamiento de las clases en estas “*special sciences*” lo que sostendría propuestas alejadas del esencialismo, como la de Richard Boyd acerca de las clases naturales como un grupo de propiedades homeostáticas (Boyd 1989).

Distinguiamos así entre una versión esencialista de las clases naturales y otra que, sin abandonar esta tradición, sí rebaja considerablemente sus exigencias, digámoslo así, ontológicas. Una de las formas más evidentes (aunque no la única) de la visión esencialista de las clases naturales es aquella que implementa, con todos los matices que queramos poner, la teoría de las emociones básicas.

2.1. Emociones básicas

La teoría de las emociones básicas es una teoría psicológica enunciada por Paul Ekman en su libro de 1972 *Emotion in the human face* (Ekman, Friesen, y Ellsworth 1972). En este libro, Ekman y su equipo analizaban los resultados de más de 30 años de investigación previa en la expresión de las emociones, en diversos contextos culturales, y llegaban a la conclusión de que todas las investigaciones analizadas proporcionaban evidencias suficientes para afirmar la existencia de seis emociones básicas: felicidad, sorpresa, miedo, tristeza, ira y disgusto combinado con desprecio (Ekman 1992, 550). Esta lista ha ido variando con el tiempo, conforme se ha incorporado evidencia. Así, en un artículo publicado en 2011, Ekman aumentaba su número hasta siete y ofrecía un listado de diez potenciales candidatas a incorporarse en un futuro, incluyendo emociones como los placeres sensoriales, la diversión o el entusiasmo (*excitement*), pero también otras mucho más complejas como la denominada *naches*, una palabra *yiddish* que hace referencia a “*the feeling a parent/caregiver, or teacher, feels when witnessing the achievements of their offspring*” (Ekman y Cordaro 2011, 365). Lo que no ha variado, o al menos no ha variado *excesivamente*, son las bases de esta teoría. En este mismo texto de 2011, Ekman señala las trece características que compartirían todas estas emociones “básicas”. Nos centraremos únicamente en la primera de ellas: presentar claras evidencias de ser universales.

La investigación de Ekman parte de esta hipótesis doble: todos los seres humanos compartimos un grupo básico de emociones y somos capaces de identificar esas emociones bási-

4 Debemos tener en cuenta, sin embargo, la evolución posterior de Putnam respecto a las clases naturales, que le alejaría de posiciones esencialistas (Hacking 2007).

cas porque las expresamos mediante expresiones faciales idénticas o muy similares (Ekman y Friesen 1969, 1971). Ekman desarrollaba así las teorías de su mentor, Silvan Tomkins, uno de los pioneros de la *Affect Theory*, que plantea la existencia de una base biológica transhistórica en ciertos *afectos* (una expresión que hace referencia explícita a la parte de nuestro comportamiento emocional que sería innata y automática, y por tanto precognitiva) que estarían conectados a ciertas expresiones faciales (Boddice 2018, 111). La teoría de las emociones básicas pasaría a formar parte, en la década de 1990, del núcleo de la psicología evolutiva (Buss 1995; Starkey 2008): las emociones básicas serían universales porque son el resultado de nuestra historia evolutiva como especie y tendrían una firme base fisiológica y neurológica (Vytal y Hamann 2010). Es en este contexto, tanto filosófico como científico, en el que debemos situar la propuesta de Jesse Prinz de una *embodied theory of emotions*.

2.2. Reacciones viscerales

Prinz desarrolla su análisis de la emoción como clase natural en el capítulo 4 de su libro *Gut Reactions*, que lleva por título *Basic Emotions and Natural Kinds* (Prinz 2004, 79-102). El capítulo se divide en dos partes. La primera, titulada *The Disunity Thesis*, presenta los argumentos que sostienen que las emociones no son una clase natural y que, por tanto, no es ni posible ni deseable la existencia de una “ciencia de las emociones”. Si bien Prinz cita en el inicio del capítulo a diversos autores, lo cierto es que pronto pasa a centrarse en el trabajo de Paul Griffiths, *What Emotions Really Are* (Griffiths 1998). En esta obra, el filósofo neozelandés proponía una revisión de los estudios sobre emociones, tanto filosóficos como psicológicos, de forma que pasasen a adoptar una perspectiva evolutiva. En lo que respecta a la emoción como clase natural, Griffiths se adhiere a la definición de clase natural de Boyd, entendida como un conjunto de propiedades homeostáticas, de forma que la posesión, por parte de un individuo, de un número suficiente de estas propiedades lo convertiría en miembro de la clase. Qué cualidades sean esas y cuántas se consideres “suficientes” se plantea como un *a priori* (Boyd 1989, 15). Lo que resulta evidente, en todo caso, es que las clases, así definidas, no precisan de ninguna esencia que garantice su unidad.

La ventaja que esta forma de definir las clases naturales tiene para Boyd es su *vaguedad*, en el sentido de que se asocia con una indeterminación intensional (Boyd 1989, 16). Al contrario que otras definiciones, esta no nos habla de condiciones necesarias y suficientes, y esto es lo que hace que sea tan adecuada para esas “ciencias especiales” como la biología.

Griffiths elabora a partir de esta definición de Boyd su propia teoría de las emociones. Esta filiación debería hacer más sencillo defender que la emoción es una clase natural, ya que no sería necesario predicar una esencia común a todas ellas, sino que bastaría con proponer un conjunto compartido de propiedades causalmente homeostáticas, sin embargo, Griffiths concluye lo contrario: las emociones no serían, ni siquiera con este criterio voluntariamente vago, clases naturales.

Griffiths distingue entre dos tipos de “emociones”: un conjunto compuesto por lo que llama “programas afectivos”, cuya principal característica es ser modulares (son rápidos, asociados a redes neuronales específicas, operan obligatoriamente y están informacionalmente encapsulados), y que se corresponderían grosso modo con las emociones básicas de Ekman. El segundo grupo estaría compuesto por aquellas emociones que él llama *higher*

cognitive emotions: la envidia, la culpa, el orgullo... aquellas emociones que, al contrario que el primer grupo, sí tendrían un contenido cognitivo. Estas emociones serían difíciles de relacionar con los programas afectivos, ya que, nos dice Griffiths, no compartirían suficientes características como para decir que forman *una misma clase natural*: los mecanismos por los que entran en funcionamiento serían totalmente diferentes, así como su historia evolutiva. No habría, por tanto, una categoría natural de emoción, sino dos categorías bien diferenciadas.

La segunda parte del capítulo la dedica Prinz a refutar esta posición. Para ello, parte de aquello que comparte con Griffiths: las emociones básicas (los programas afectivos) sí cumplirían los requisitos para ser una clase natural. El siguiente paso, por tanto, sería relacionar estas emociones básicas con el otro grupo. Prinz plantea dos caminos para ello. El primero sería pensar que las “altas emociones” son el resultado de combinar dos o más emociones básicas (Prinz 2004, 92). Esta teoría, sin embargo, no sería capaz de dar cuenta de todas nuestras emociones “complejas”, por lo que debemos dar un paso más y aceptar que algunas de estas mezclas también implican elaboraciones cognitivas de emociones básicas (Prinz 2004, 93). Se aporta así un argumento a favor de la unidad de ambas categorías, ya que incluso las emociones más complejas esconden en su núcleo una categoría básica.

Prinz responde a continuación a las críticas, algunas de calado, planteadas por diversos autores a esta idea de “emociones complejas”, reduciéndolas a un único argumento: si los programas afectivos (o *embodied appraisals* para Prinz) son estados de un sistema modular y las emociones compuestas (o altas emociones cognitivas) son estados de ese sistema modular con el añadido de elaboraciones cognitivas, entonces sus propiedades psicológicas serán muy diferentes entre sí (Prinz 2004, 97). Para resolver este problema, que no tiene una solución sencilla, Prinz da un salto adelante: nos propone pensar las “altas emociones cognitivas” no como un compuesto de los programas afectivos y elaboraciones cognitivas, sino que estas últimas se verían excluidas del contenido de la emoción. O, dicho de otra forma, eso que llamamos celos no es más que la percepción de una serie de *evaluaciones corporales* (*embodied appraisals*) que han sido *recalibradas* mediante juicios que las relacionan con el contexto. Pero en ningún caso estos juicios forman parte de la emoción de los celos, de la misma forma que no decimos que la huella del oso, por rescatar el clásico ejemplo jamesiano, *forme parte* de nuestra emoción del miedo.

A continuación, Prinz postula la existencia de una serie de *archivos de calibración* (“calibration files”). Estos archivos tienen la función de servir de enlace entre los juicios de una clase determinada y las evaluaciones corporales. Estos archivos serían estructuras de datos localizados en una “long-term memory”. Cada uno de estos archivos contendría un conjunto de representaciones que causarían la misma pauta de respuestas corporales, o una muy similar. La percepción de estas respuestas corporales, causadas por el conjunto de representaciones contenidas en nuestros *archivos de calibración* es lo que llamamos una emoción. Estos *archivos* estarían presentes tanto en las que conocemos como emociones básicas como en las que hemos llamado *complejas*, y el mecanismo de funcionamiento de ambos tipos sería el mismo que acabamos de describir. Lo que varía, de acuerdo con Prinz, es el origen del conjunto de representaciones contenido en esos *archivos*: mientras que las emociones básicas serían calibradas mediante archivos que han sido creados merced a la selección natural, los *archivos de calibración* de las emociones complejas o no básicas se

crean bien a través de la mezcla de dos emociones básicas (y de sus archivos), bien mediante el “desvío” de otras emociones que ya estaban presentes en nuestro repertorio emocional.

Este es el argumento final de Prinz a favor la emoción como una clase natural. Cree haber resuelto, de forma elegante, todos los problemas planteados por Griffiths, a través de esta puesta al día de la propuesta de William James, que es capaz de tener en cuenta cómo los contenidos mentales modifican nuestra experiencia emocional sin por ello modificar el contenido mismo de la emoción, que sigue siendo corpórea, modular y encapsulada. Prinz concluye este capítulo de forma triunfal:

The case for the unity is actually so overwhelming. When one reflects on the many things that emotions share in common, it is rather surprising that many philosophers have been tempted to divide up the category [...] Emotions are a natural kind in a strong sense. They share a common essence. It is rare for nature (and folk psychology) to offer such a neat category. (Prinz 2004, 102).

2.3. *¿Y si las emociones no son clases naturales... de nuevo?*

Para otros, sin embargo, la evidencia no será tan concluyente, y el problema de las emociones como clases naturales sigue abierto. En este apartado nos vamos a limitar a señalar dos aportaciones al debate: en primer lugar, la lectura que hace Griffiths del libro de Prinz, que se mantendría dentro de los límites en los que se venía desarrollando el debate hasta esos momentos; y, en segundo lugar, el cambio radical de paradigma propuesto de la psicóloga Lisa Feldman Barrett, que desborda la concepción de las emociones *estándar* para proponer una nueva visión de su naturaleza y funcionamiento.

Griffiths se encuentra entre los primeros en criticar la obra de Prinz. En una amplia reseña, publicada en 2008, señala las importantes aportaciones de su obra, pero se muestra crítico en ciertos aspectos clave. Uno de ellos tiene que ver, precisamente, con la apuesta de Prinz por calificar a las emociones como una clase natural. Griffiths se centra en la afirmación final de Prinz (“*emotions are natural kinds in a strong sense*”) para señalar que, de hecho, tener una esencia no implica, necesariamente, ser una clase natural. Y que incluso puede ser antiproducente, pues una clase natural, si aceptamos la propuesta de Boyd, como Prinz y él mismo hacen, en realidad no sería otra cosa que una categoría sobre las que podemos hacer descubrimientos científicos. Esto implica, dice Griffiths, que una clase natural “fuerte” sería aquella que reuniese las siguientes condiciones: “*a large domain of properties over which it can be used for induction, the reliability of those inductions, and the theoretical fruitfulness of the research program(s) in which it is embedded*” (Griffiths 2008, 560). Nada de esto nos lo asegura el mero hecho de suponer la existencia de “una esencia” que unificase la categoría. Así, insiste, hay clases, como la de “objetos supralunares”, que no son clases naturales, pero tienen unas condiciones de pertenencia (necesarias y suficientes) muy bien definidas; mientras que otras, como las especies biológicas, tan difíciles de definir, estarían mucho mejor posicionadas en principio para ser calificadas como naturales.

Lo interesante de esta crítica, en nuestra opinión, no es tanto los argumentos concretos que ofrece como la afirmación de que, en última instancia, Prinz yerra el tiro. El problema real, nos dice Griffiths, no radica en que seamos capaces o no de ofrecer una categoría uni-

ficada de emoción, sino si esa categoría es o no productiva (“*the theoretical fruitfulness*”) para los programas de investigación que la emplean.

La segunda crítica a la concepción de las emociones como clases naturales la realiza la psicóloga Lisa Feldman Barrett, en su artículo de 2006: *Are Emotions Natural Kinds?* Barrett realiza una revisión exhaustiva de los estudios sobre emociones realizados durante la segunda mitad del siglo XX, analizando los diversos argumentos que sostendrían que las emociones son clases naturales: los estudios realizados sobre la existencia de mecanismos causales, sobre la relación entre el gesto y la emoción, acerca de las diferencias entre percibir una emoción y expresar una emoción, etc. En todos los casos la conclusión es prácticamente la misma: no existen evidencias suficientes para sostener la hipótesis planteada. No existen evidencias suficientes, por ejemplo, para sostener que las expresiones faciales proporcionan información sobre estados emocionales discretos del emisor, algo que ha sido aceptado tradicionalmente como una evidencia de que existen clases naturales de emociones y de su universalidad (Barrett 2006b, 40). La conclusión de Barrett es, como mínimo, tan contundente como la de Prinz:

Many of the most influential models in our science assume that emotions are biological categories imposed by nature, so that emotions categories are recognized, rather than constructed, by the human mind. And the most optimistic read of the existing evidence is that it is inconclusive, by whatever criteria are used. (Barrett 2006b, 47)

La conclusión, para Barrett, es clara: necesitamos un nuevo paradigma. Un nuevo paradigma que plantee nuevas preguntas, que adopte una aproximación inductiva, y que dé cuenta de la evidencia acumulada hasta el momento. Un paradigma, nos dice Barrett, que abandone la fe en la capacidad explicativa de las emociones, que deben dejar de ser parte del *explanans* para ocupar un espacio en el *explanandum* (Barrett 2006b, 49): “*In this view, emotions do not refer to the things being classified, but rather are classification schemes that people impose on their world during perception*” (Barrett 2006b, 46).

2.4. Un nuevo paradigma

Feldman Barret ha desarrollado su teoría a lo largo de los últimos 15 años en diversos artículos (Barrett 2006a; Barrett y Simmons 2015; Barrett 2018b; Gendron, Crivelli, y Barrett 2018) y libros (Barrett y Russell 2014; Barrett 2018a). Este nuevo paradigma, que Barrett llama teoría de la emoción construida, parte de la crítica a la visión tradicional de las emociones, que hemos visto en el apartado anterior, y que la autora relaciona con la tradicional división de nuestras categorías mentales en diversas “facultades”, que englobarían a las facultades cognitivas, volitivas, y afectivas. Esta visión, dice Barrett, que ha dado forma a la investigación sobre emociones de los últimos cuarenta años, es insostenible con los avances en nuestra comprensión del funcionamiento del cerebro (Barrett 2017). Barrett resume su nuevo paradigma como sigue:

Cuando los científicos dejan de lado la visión clásica y se limitan a los datos, emerge una explicación de las emociones radicalmente diferente. En pocas palabras, vemos

que las emociones no son monolíticas, sino que están hechas de componentes más básicos; que en lugar de ser universales varían de una cultura a otra; que no son provocadas sino que las creamos nosotros; que surgen de una combinación entre las propiedades físicas del cuerpo, un cerebro flexible cuyas conexiones reflejan el entorno en el que se desarrolla, y la cultura y la educación que ofrecen ese entorno. Las emociones son reales, pero no en el mismo sentido objetivo que las moléculas o las neuronas. Son reales en el sentido en que lo es el dinero, es decir, no son una ilusión, pero sí un producto del consenso humano (Barrett 2018a, 18)

Dentro de este nuevo paradigma, las emociones no son clases naturales, pero tampoco necesitan serlo para cumplir su función. Esto no quiere decir, por supuesto, que las emociones sean ilusiones, como no lo es el dinero. Quiere decir, entre otras muchas cosas, que las emociones son un producto de nuestro devenir en la historia. Que lo que llamamos “miedo” no es lo mismo que lo que llamaban “miedo” los campesinos filipinos del siglo XII. No sólo porque nuestra cultura sea distinta, sino porque nuestro cuerpo, como bien señalaba Foucault, *tampoco es el mismo* (Foucault 1977). Esto abre una nueva vía de investigación filosófica que la visión clásica de las emociones excluía taxativamente: la de atender a su papel como un elemento en el devenir de la historia.

3. Las emociones como clases humanas

Las emociones no serían, de acuerdo a este nuevo modelo propuesto por Barrett, clases naturales. No podemos encontrar en la naturaleza un mecanismo automático, encapsulado, automático y derivado de nuestra historia evolutiva que podamos llamar “felicidad”, “tristeza”, “ira”, “miedo” o “asco”. En su lugar, tenemos una visión de las relaciones entre nuestro cerebro y nuestra acción mucho más compleja, plagada de factores con base biológica, pero también por los nichos culturales en que nos desarrollamos como individuos, todo ello produciendo cambios reales en nuestro cuerpo, a través de la idea de la plasticidad del cerebro (Atzil et al. 2018). Ahora bien, si las distintas emociones no son clases naturales, entonces, ¿qué tipo de clasificación son? Y aún más, señalar que las emociones dejan de ser parte del *explanans* para pasar a serlo del *explanandum* ¿quiere decir que el estudio filosófico de las emociones es inviable? Todo lo contrario. Esta nueva aproximación provee a la filosofía con nuevas posibilidades, permitiendo la puesta en marcha de nuevos programas de investigación en los que, además, la disciplina contará con un mayor protagonismo.

3.1. Clases naturales y clases relevantes

Ian Hacking emprendió en la década de 1990 un camino similar al de Barrett. Al igual que ella, desarrolla un análisis del paradigma imperante acerca de las clases naturales para terminar presentando una propuesta alternativa. Una propuesta que articula frente a aproximaciones como la de Boyd, que buscan extender el alcance del concepto. La intervención de Hacking en este debate se cifra en dos aspectos. En primer lugar, identifica una determinada tradición filosófica dentro de la cual el debate sobre clases naturales tiene sentido: aquella que reflexiona sobre las formas de clasificar el mundo empleadas por la ciencia moderna (Hacking

1991a). Su segunda aportación tiene que ver con lo que en este primer momento llama “clases sociales” y que posteriormente llamará “humanas” y también “interactivas” (Khalidi 2010).

Hacking empieza su análisis de las clases naturales señalando dos aspectos: en primer lugar, que el debate sobre las clases naturales suele convertirse en una disputa interminable alrededor de oscuros argumentos, lo que lo convierte en una discusión estéril; y en segundo, que el debate tan sólo le interesa en tanto que le sirve para definir mejor su idea de clases humanas. Esto no implica negar la existencia de clases naturales, que pueden ser de utilidad “*so long as it is kept modest*” (Hacking 1991a, 110). Hacking identifica, a continuación, una serie de principios que compondrían lo que él llama la “tradicción” de las clases naturales (Hacking 1991a, 110-11), en la que las clases naturales se caracterizarían por:

- i) *Estar definidas mediante un conjunto de propiedades necesarias y suficientes;*
- ii) *La posesión de estas propiedades indicaría la posesión de un número mayor de propiedades que resultarían de interés metodológico;*
- iii) *Estas propiedades definitivas serían naturales, no sociales*⁵.

El concepto de clase natural desarrollado por esta tradición se caracterizaría por su *modestia*, dejando espacio no sólo a diversas formas de definir las clases naturales, sino que no intentaría ir allí más allá del terreno para el que han sido pensadas. Es decir: si hay ciertas clasificaciones (como las de la biología) que no se corresponden con nuestra definición de clase natural, ¿qué obtenemos tratándolas como tales? ¿Qué obtenemos modificando la definición de clase natural de forma que se adapte (o al menos lo intente) a ellas? Hacking lo tiene claro: “*obscure philosophy*” (Hacking 1991a, 109).

Lo mismo ocurre, dice Hacking, con aquellos intentos de tratar categorías que incumplen la condición (iii) como clases naturales, y aquí incluiríamos todas aquellas categorías de las ciencias humanas y sociales, entre las que Hacking incluye a parte de la medicina (Hacking 1999). Tratar este tipo de conceptos como “clases naturales” sólo nos conduce a callejones sin salida que no estamos obligados a visitar. Lo que nos interesa, dice Hacking, es enfrentarnos a conceptos *relevantes*, a aquellos que operan en el mundo, sin necesidad de vernos obligados a decir si son, o dejan de ser, iguales a las clases naturales. Al igual que pasa con Barrett, una vez que abandonamos el lenguaje de las clases naturales podemos empezar a plantear nuevas preguntas, a abordar nuevos problemas.

3.2. Clases humanas y clases interactivas

Hacking, por tanto, no está interesado en el debate de las clases naturales, sino en cómo funcionan las categorías que empleamos en las ciencias sociales y en las humanidades. Estas categorías no son naturales, como no lo es, por volver a Goodman, la forma en que categorizamos las distintas obras de música clásica (sonata, oratorio, concierto, sinfonía, etc.). Pero sí son *relevantes*, en tanto que tienen un efecto claro y preciso en nuestra percepción de dichas obras. Ser categorizado como refugiado, por proponer un ejemplo menos banal, tiene un impacto mayúsculo en la vida de una persona que huye de la guerra y todos coin-

5 Tomo este resumen de la posición de Hacking de Boyd 1991, 127. Véase la respuesta de Hacking a las críticas de Boyd en Hacking 1991c.

cidiremos en que no existe nada de natural en este proceso. Estas categorizaciones siguen su propia lógica y tienen un funcionamiento totalmente distinto a las *naturales*. Hacking las llama categorías humanas o *interactivas* y dedicará gran parte de su obra más reciente a analizar su funcionamiento y el impacto que tienen en nuestra experiencia.

De forma muy resumida, diremos que las categorías interactivas se caracterizan por las siguientes características:

- i) El fundamento de estas categorías no reside en su correspondencia con una “realidad” subyacente previa, sino en su historia de uso: “*The historical incidence of utterance is what distinguishes healthy predicates from unhealthy predicates*” (Hacking 1993, 294). Son, por tanto, históricamente situadas y socialmente producidas.
- ii) No todas las categorías que aplicamos a los humanos son interactivas, tan sólo aquellas sobre las que tenemos un conocimiento sistemático: “*classifications that could be used to formulate general truths about people; generalizations sufficiently strong that they seem like laws about people, their actions, or their sentiments*” (Hacking 1995, 352)⁶.
- iii) El *efecto bucle* (*loop effect*). Este efecto bucle funciona de la siguiente forma: la categoría que se impone sobre una persona, introduce cambios en cómo esa persona se percibe a sí misma y, a partir de ahí, también modifica su comportamiento. Esa modificación en la autopercepción y en el comportamiento del individuo puede hacer variar, consecuentemente, el contenido de la categoría, que se vería así modificada, reiniciándose el proceso.

En su análisis de la aparición y desarrollo del abuso infantil, Hacking lo resume como sigue:

[...] as we evolve an idea about a kind of person or of human behavior, people change, behaviors change. Children experience their hurt differently. They are more self-conscious about when and how emotional and sexual abuse is painful [...] New kinds of people come into being that don't fit the wisdom just acquired, less because the recent knowledge was wrong than because of a feedback effect. (Hacking 1991b, 254)

Al estudio de la formación de estos conceptos, de su impacto en la experiencia de los clasificados y del efecto bucle que se produce lo llama Hacking, siguiendo a Foucault, *ontología histórica* (Hacking 2002; Foucault 1984). La ontología histórica implica un uso *filosófico* de la historia (Hacking 1990) y permitiría el estudio de la aparición, a lo largo del devenir humano, de diversas *clases de personas* (Hacking 2006) con distintas posibilidades de experiencia. En definitiva, de la aparición de nuevas formas de *ser en el mundo*.

6 Más adelante especifica aun más las condiciones que una categoría debe cumplir para que considerarse como una clase humana: “*When I speak of human kinds, I mean (i) kinds that are relevant to some of us; (ii) kinds that primarily sort people, their actions, and behaviour, and (iii) kinds that are studied in the human and social sciences, i.e. kinds about which we hope to have knowledge. I add (iv) that kinds of people are paramount*” (Hacking 1995, 354).

Nuestra propuesta es que debemos entender las emociones como clases *humanas* o interactivas. Esto no sólo se ajusta a la propuesta de Barrett del nuevo paradigma de la *emoción construida*, sino que permite desarrollar un programa de investigación filosófico novedoso, capaz de ofrecer aportaciones relevantes más allá de la propia disciplina⁷.

4. La ontología histórica de las emociones

Pensar las emociones como “clases humanas” no es más que asumir el nuevo paradigma propuesto por Barrett, en el que las emociones ya no nos hablan de la naturaleza, sino que nos hablan, sobre todo, de nosotros mismos⁸. Un buen ejemplo de esto nos lo proporciona la felicidad. Una emoción con una larga historia, que podríamos remontar hasta los griegos y su búsqueda de la *εὐδαιμονία*. Sin embargo, si pensamos dentro del antiguo paradigma de las emociones como clases naturales esta historia no tendría nada que ver con la “felicidad”, sino con aquellos accesorios que la *colorean*, que la *calibran*, pero que no son ella. De la misma forma que la huella del oso no forma parte del “miedo”. No en vano “happiness” es una de las emociones básicas de Ekman. Pero si aceptamos el nuevo paradigma, entonces nuestra capacidad de análisis se multiplica. No sólo porque, de repente, qué tuviera que decir Aristóteles sobre la *εὐδαιμονία* era relevante para cómo los griegos la experimentaban, sino porque el análisis de la felicidad se convierte en una tarea que sólo puede abordarse *transdisciplinariamente*, en un espacio en que lo que digamos los humanistas tiene, nuevamente, un impacto no en aquello que “rodea” a la categoría natural, sino en el mismo núcleo de la categoría que ahora es (vuelve a ser) humana. Un excelente ejemplo de este tipo de trabajo es el último libro de Eva Illouz y Edgar Cabanas en el que analizan lo que llaman la *economía de la felicidad* (Cabanas y Illouz 2019), situando esta emoción, que nos presentaban como un programa aislado que respondía ante determinados eventos del mundo de forma automática, nuevamente en la *matriz* en que se vuelve *relevante* para nuestras vidas (Hacking 2001, 32-34).

A partir de esta constatación es posible esbozar el programa de investigación de la ontología histórica de las emociones, que contendría las siguientes líneas de investigación, estrechamente relacionadas entre sí:

1. **Ontología histórica de emociones concretas.** Al igual que el ejemplo que hemos señalado de la felicidad, sería posible estudiar la aparición de emociones particulares (o la modificación de otras existentes) que darían lugar a nuevas formas de ser en el mundo. Un buen ejemplo de este posible análisis filosófico (otro de los términos empleados por Hacking) sería aplicarlo a la idea de *empatía*. Un concepto que aparece a finales del siglo XIX en la obra de un oscuro historiador del arte alemán (Robert

7 El concepto de clase humana de Ian Hacking ha recibido múltiples críticas en los últimos años, al tiempo que ha sido empleado para la puesta en marcha de programa de investigación ambiciosos en diversos campos del conocimiento: la psicología, la historia y, por supuesto, la filosofía. La literatura es numerosa y aquí recogemos únicamente una selección: Sparti 2001; Cooper 2004; Ereshefsky 2004; Khalidi 2010, 2015; Lemeire 2015; Hauswald 2016; Millard 2017.

8 Sin embargo, el proyecto resulta independiente del de Barrett. Incluso si finalmente se demostrase que Barrett está equivocada, la ontología histórica de las emociones ofrecería resultados interesantes, si bien su alcance sería más limitado.

Vischer) como *Einfühlung*, que se traduce posteriormente al inglés como *empathy* y se emplea en el análisis de las obras de arte, hasta que a principios del siglo XX Theodor Lipps la emplea en psicología, de donde la toma Freud (Pigman 1995). La historia de cómo esta categoría estética pasa a convertirse en un criterio diagnóstico de la psicopatía, además del fundamento de nuestras actitudes morales hacia el otro, es una excelente candidata para esta ontología histórica de las emociones.

2. **Ontología histórica de situaciones emocionales complejas.** Nuestra segunda línea de investigación tiene que ver con el papel desempeñado por las emociones en situaciones complejas. Un buen ejemplo de este posible uso sería el trabajo realizado por Fay Bound Alberti alrededor del concepto de *soledad* (Alberti 2019). La soledad, al igual que la felicidad, es un concepto con una historia, tanto intelectual como cultural, compleja. Si para los Padres del Desierto, en el siglo IV, la soledad era una experiencia positiva, en nuestra sociedad contemporánea no sólo se considera como algo negativo, sino que empezamos a hablar de una “epidemia” de soledad (Bound Alberti 2018).
3. **Ontología histórica de la experiencia.** La tercera línea de investigación estaría relacionada con un estudio de nuestras posibilidades de experiencia. William James definió la *experiencia* como una elaboración posterior de aquellos elementos que conforman el *flujo de la existencia*: “*the immediate flux of life which furnishes the material to our later reflection with its conceptual categories*” (James 1912, 93). Esta línea de investigación analizaría las relaciones concretas, históricamente situadas y socialmente constituidas, entre las categorías y la experiencia. Un posible ejemplo de este trabajo sería el estudio que sobre las relaciones entre experiencia subjetiva y la consideración subjetiva del bienestar hemos realizado con anterioridad (Moscoso & Zaragoza 2014).

Estas tres líneas de trabajo nos permiten vislumbrar una serie de posibilidades para el estudio filosófico de las emociones que no tenían cabida en el viejo paradigma, que nos limitaba a discutir acerca de aquello que *rodeaba* la emoción, sin llegar nunca a tocar el *núcleo* de la misma. El nuevo paradigma de Barrett, en conjunción con la propuesta de Hacking, nos brinda acceso a ese núcleo. La ontología histórica de las emociones, entendida como un *uso filosófico del pasado*, no se limita por tanto narrar una historia más o menos anecdótica, sino que ofrece conclusiones relevantes para entender cómo se constituyen nuestras emociones y cómo estas alumbran, en algunos casos, nuevas *clases de personas*, nuevas formas de ser y de experimentar que pasan a formar parte de nuestro mundo. Y esto, se mire como se mire, es una tarea profundamente filosófica.

Bibliografía

- Alberti, Fay Bound (2019), *A Biography of Loneliness: The History of an Emotion*. Oxford, New York: Oxford University Press.
- Alberti, Fay Bound (2018), «This “Modern Epidemic”: Loneliness as an Emotion Cluster and a Neglected Subject in the History of Emotions». *Emotion Review* 10 (3): 1-13. <https://doi.org/10.1177/1754073918768876>.

- Alcaraz, Leon Maria J. (2011), «Contrary Feelings and the Cognitive Significance of Art». *Eстетика: The Central European Journal of Aesthetics* 48 (1): 64-81.
- Atzil, Shir, Wei Gao, Isaac Fradkin, y Lisa Feldman Barrett (2018), «Growing a Social Brain». *Nature Human Behaviour* 2 (9): 624-36. <https://doi.org/10.1038/s41562-018-0384-6>.
- Bagnoli, Carla (2011), *Morality and the Emotions*. Oxford: OUP Oxford.
- Barrett, Lisa Feldman (2006a), «Solving the Emotion Paradox: Categorization and the Experience of Emotion». *Personality and Social Psychology Review* 10 (1): 20,46. https://doi.org/10.1207/s15327957pspr1001_2.
- Barrett, Lisa Feldman (2006b), «Are Emotions Natural Kinds?» *Perspectives on Psychological Science* 1 (1): 28-58. <https://doi.org/10.1111/j.1745-6916.2006.00003.x>.
- Barrett, Lisa Feldman (2017), «The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization». *Social Cognitive and Affective Neuroscience* 12 (1): 1-23. <https://doi.org/10.1093/scan/nsw154>.
- Barrett, Lisa Feldman (2018a), *La vida secreta del cerebro: Cómo se construyen las emociones*. Traducido por Genís Sánchez Barberán. Barcelona: Ediciones Paidós.
- Barrett, Lisa Feldman (2018b), «Seeing Fear: It's All in the Eyes?» *Trends in Neurosciences* 41 (9): 559-63. <https://doi.org/10.1016/j.tins.2018.06.009>.
- Barrett, Lisa Feldman, Kristen A. Lindquist et al. (2007), «Of Mice and Men: Natural Kinds of Emotions in the Mammalian Brain? A Response to Panksepp and Izard». *Perspectives on Psychological Science* (2), 3, 297-312.
- Barrett, Lisa Feldman, y James A. Russell, eds. (2014), *The psychological construction of emotion*. New York: The Guilford Press.
- Barrett, Lisa Feldman, y W. Kyle Simmons (2015), «Interoceptive Predictions in the Brain». *Nature Reviews Neuroscience* 16 (7): 419-29. <https://doi.org/10.1038/nrn3950>.
- Boddice, Rob (2018), *The History of Emotions*. Manchester: Manchester University Press.
- Boyd, Richard (1989), «What Realism Implies and What It Does Not». *Dialectica* 43 (1-2): 5-29. <https://doi.org/10.1111/j.1746-8361.1989.tb00928.x>.
- Boyd, Richard (1991), «Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 61 (1/2): 127-48.
- Brun, Georg, Ulvi Doguoglu, y Ulvi Doguoglu (2016), *Epistemology and Emotions*. Routledge. <https://doi.org/10.4324/9781315580128>.
- Buss, David M. (1995), «Evolutionary Psychology: A New Paradigm for Psychological Science». *Psychological Inquiry* 6 (1): 1-30. https://doi.org/10.1207/s15327965pli0601_1.
- Cabanas, Edgar, y Eva Illouz (2019), *Happycracia: Cómo la ciencia y la industria de la felicidad controlan nuestras vidas*. Madrid: Planeta.
- Clough, Patricia Ticineto, y Jean O'Malley Halley, eds. (2007), *The Affective Turn: Theorizing the Social*. Durham, N.C. : Chesham: Duke University Press.
- Cooper, Rachel (2004), «Why Hacking is Wrong about Human Kinds». *The British Journal for the Philosophy of Science* 55 (1): 73-85. <https://doi.org/10.1093/bjps/55.1.73>.
- Dupré, John (1981), «Natural Kinds and Biological Taxa». *The Philosophical Review* 90 (1): 66-90. <https://doi.org/10.2307/2184373>.

- Eickers, Gina, Juan R. Loaiza, J. & Jesse Prinz (2017), «Embodiment, Context-Sensitivity, and Discrete Emotions: A Response to Moors». *Psychological Inquiry*, 28(1), 31–38.
- Ekman, Paul (1992), «Are there basic emotions?» *Psychological Review* 99 (3): 550-53.
- Ekman, Paul, y Daniel Cordaro (2011), «What Is Meant by Calling Emotions Basic». *Emotion Review* 3 (4): 364-70. <https://doi.org/10.1177/1754073911410740>.
- Ekman, Paul, y Wallace V. Friesen (1969), «The repertoire of nonverbal behavior: Categories, origins, usage, and coding». *semiotica* 1 (1): 49–98.
- Ekman, Paul, y Wallace V. Friesen (1971), «Constants across cultures in the face and emotion.» *Journal of personality and social psychology* 17 (2): 124.
- Ekman, Paul, Wallace V. Friesen, y Phoebe Ellsworth (1972), *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon Press.
- Ereshefsky, Marc (2004), «Bridging the Gap between Human Kinds and Biological Kinds». *Philosophy of Science* 71 (5): 912-21.
- Foucault, Michel (1977), «Nietzsche, Genealogy, History». En *Language, Counter-Memory, Practice: Selected Essays and Interviews*, editado por D.F. Bouchard, 139-64. Ithaca: Cornell University Press.
- Foucault, Michel (1984), «What is Enlightenment?» En *The Foucault Reader*, editado por Paul Rabinow. New York: Pantheon.
- Frijda, Nico H. (1986), *The emotions*. Studies in emotion and social interaction. Cambridge ; New York : Paris: Cambridge University Press ; Editions de la Maison des sciences de l'homme.
- Gendron, Maria, y Lisa Feldman Barrett (2009), «Reconstructing the Past: A Century of Ideas About Emotion in Psychology». *Emotion review* 1 (4): 316-39. <https://doi.org/10.1177/1754073909338877>.
- Gendron, Maria, Carlos Crivelli, y Lisa Feldman Barrett (2018), «Universality Reconsidered: Diversity in Making Meaning of Facial Expressions». *Current Directions in Psychological Science* 27 (4): 211-19. <https://doi.org/10.1177/0963721417746794>.
- Gordon, Robert M. (1969), «Emotions and Knowledge». *The Journal of Philosophy* 66 (13): 408-13. <https://doi.org/10.2307/2024422>.
- Gordon, Robert M., y Robert Morris Gordon (1990), *The Structure of Emotions: Investigations in Cognitive Philosophy*. Cambridge University Press.
- Griffiths, Paul E. (1998), *What Emotions Really Are: The Problem of Psychological Categories*. Edición: 1. Chicago, Ill: University of Chicago Press.
- Griffiths, Paul E. (2002), «Is Emotion a Natural Kind?» En *Thinking About Feeling: Contemporary Philosophers on Emotions*, editado por Robert C. Solomon. Oxford University Press.
- Griffiths, Paul E. (2008), «Jesse Prinz Gut Reactions: A Perceptual Theory of Emotion». *The British Journal for the Philosophy of Science* 59 (3): 559-67. <https://doi.org/10.1093/bjps/axn020>.
- Hacking, Ian (1990), «Two Kinds of “New Historicism” for Philosophers». *New Literary History* 21: 343-64.
- Hacking, Ian (1991a), «A Tradition of Natural Kinds». *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 61 (1/2): 109-26.

- Hacking, Ian (1991b), «The Making and Molding of Child Abuse». *Critical Inquiry* 17: 253-88.
- Hacking, Ian (1991c), «On Boyd». *Philosophical Studies* 61 (1): 149-54. <https://doi.org/10.1007/BF00385838>.
- Hacking, Ian (1993), «On Kripke's and Goodman's Uses of 'Grue'». *Philosophy* 68 (265): 269-95. <https://doi.org/10.1017/S003181910004122X>.
- Hacking, Ian (1995), «The Looping Effect of Human Kinds». En *Causal Cognition: A Multidisciplinary Debate*, editado por Dan Sperber, David Premack, y Ann James Premack, 351-83. New York: Oxford University Press.
- Hacking, Ian (1999), *Mad Travellers: Reflections on the Reality of Transient Mental Illness*. London: Free Association.
- Hacking, Ian (2001), ¿La construcción social de qué? Barcelona: Paidós.
- Hacking, Ian (2002), *Historical Ontology*. Cambridge, Mass. & London: Harvard University Press.
- Hacking, Ian (2006), «Kinds of People: Moving Targets». *Proceedings of the British Academy* 151: 285-318.
- Hacking, Ian (2007), «Putnam's Theory of Natural Kinds and Their Names Is Not the Same as Kripke's». *Principia: An International Journal of Epistemology* 11 (1): 1-24.
- Hauswald, Rico (2016), «The Ontology of Interactive Kinds». *Journal of Social Ontology* 2 (2): 203-221. <https://doi.org/10.1515/jso-2015-0049>.
- Hopkins, Debra (2009), *Theorizing Emotions: Sociological Explorations and Applications*. Campus Verlag.
- James, William (1884), «What is an Emotion?» *Mind* 9 (34): 188-205.
- James, William (1912), *Essays in Radical Empiricism*. New York: Longmans, Greens & Co.
- Kassab, Hanna Samir (2016), *The Power of Emotion in Politics, Philosophy, and Ideology*. New York: Palgrave Macmillan US.
- Khalidi, Muhammad Ali (2010), «Interactive Kinds». *The British Journal for the Philosophy of Science* 61 (2): 335-60.
- Khalidi, Muhammad Ali (2015), *Natural Categories and Human Kinds*. Cambridge: Cambridge University Press.
- Leavitt, John (1996), «Meaning and feeling in the anthropology of emotions». *American ethnologist* 23 (3): 514-539.
- Lemeire, Olivier (2015), «The Scientific Classification of Natural and Human Kinds». <https://lirias.kuleuven.be/retrieve/319407>.
- Loewenstein, George (2000), «Emotions in Economic Theory and Economic Behavior». *The American Economic Review* 90 (2): 426-32.
- Lutz, Catherine, y Geoffrey M. White (1986), «The anthropology of emotions». *Annual review of anthropology* 15 (1): 405-436.
- Mill, John Stuart (1872), *A System of Logic, Ratiocinative and Inductive : Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation : In Two Volumes*. London : Longmans, Green, Reader, and Dyer.
- Millard, Chris (2017), «Concepts, Diagnosis and the History of Medicine: Historicising Ian Hacking and Munchausen Syndrome». *Social History of Medicine* 30 (3): 567-89. <https://doi.org/10.1093/shm/hkw083>.

- Moscoso, Javier & Juan Manuel Zaragoza (2014), «Historia del bienestar. Desde la historia de las emociones a las políticas de la experiencia». *Cuadernos de Historia Contemporánea* 36: 73-89.
- Nussbaum, Martha C. (2001), *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.
- Nussbaum, Martha C. (2013), *Political Emotions: Why Love Matters for Justice*. Cambridge: The Belknap Press of Harvard University Press.
- Okasha, Samir (2002), «Darwinian Metaphysics: Species and the Question of Essentialism». *Synthese* 131 (2): 191-213.
- Panksepp, Jaak (2004), *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford University Press.
- Pigman, G. W. (1995), «Freud And The History Of Empathy». *Int. J. Psycho-Anal.* 76: 237-56.
- Prinz, Jesse J. (2004), *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Putnam, Hilary (1975), «The Meaning of “Meaning”». *Minnesota Studies in the Philosophy of Science* 7: 131-93.
- Reddy, William M. (2001), *The Navigation of Feeling. A Framework for the History of Emotions*. Cambridge & New York: Cambridge University Press.
- Reisenzein, Rainer (2015), «A Short History of Psychological Perspectives on Emotion». En Calvo, R., D’Mello, S., et. al., *The Oxford Handbook of Affective Computing*, Oxford: Oxford University Press.
- Sparti, Davide (2001), «Making up People. On some Looping Effects of the Human Kind - Institutional Reflexivity or Social Control?» *European Journal of Social Theory* 4: 331-49.
- Starkey, Charles (2008), «Classifying Emotions: Prospects for a Psychoevolutionary Approach». *Philosophical Psychology* 21 (6): 759-77. <https://doi.org/10.1080/09515080802513300>.
- Vytal, Katherine, y Stephan Hamann (2010), «Neuroimaging Support for Discrete Neural Correlates of Basic Emotions: A Voxel-Based Meta-Analysis». *Journal of Cognitive Neuroscience* 22 (12): 2864-85. <https://doi.org/10.1162/jocn.2009.21366>.
- Whewell, William (1847), *The Philosophy of the Inductive Sciences, Founded upon Their History*. London, J. W. Parker.
- Zaragoza, Juan Manuel (2013), «Historia de las emociones: una corriente historiográfica en expansión». *Asclepio* 65(1), DOI: 10.3989/asclepio.2013.12
- Zaragoza, Juan Manuel (2021), «A Change of Pace. The History of (Emotional) Experiences». *History of Psychology* 24 (2): 130-135. DOI: 10.1037/hop0000173.

RESEÑAS



BUTLER, Judith (2021). *Excitable Speech. A Politics of the Performative*. Nueva York: Routledge. 212 pp.

Anne Marie Smith ha registrado las disputas intelectuales a las que buscó responder Judith Butler con la publicación en 1997 de *Excitable Speech: A Politics of the Performative*. se trata de un libro que engancha directamente con las discusiones académicas en torno a la regulación del discurso público, el lenguaje de odio y la censura (Smith, 2001, 391). Sus páginas entran en diálogo con las controversias en teoría feminista entre la regulación estatal del discurso público, defendida por Catharine MacKinnon y Andrea Dworking, y la oposición a las políticas discursivas centradas en el Estado por parte de Carole Vance, Lisa Duggan, Nan Hunter y Gayle Rubin (Smith, 2001, 392). Smith resalta la similitud de los argumentos de MacKinnon y Dworking con aquellos que esgrimieron Mari Matsuda, Charles Lawrence III, Richard Delgado y Kimberlé Williams Crenshaw en *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment* (Smith, 2001, 396).

Excitable Speech trata de ser una respuesta a la instrumentalización de las reivindicaciones feministas contra la violencia de género en favor de un mayor control estatal del discurso público, de la representación y del activismo feminista y *queer*. La filósofa se dirige contra la equiparación en términos de performatividad soberana entre pornografía y discurso racista. La concepción soberana de la performatividad del habla cierra las posibilidades de agencia y de interpela-

ción que dieron lugar al propio sujeto del habla. Judith Butler establece en virtud de una vulnerabilidad constitutiva que la viabilidad del sujeto de habla está en continua producción y que es precisamente esta condición la que le permite hablar desde posiciones inesperadas y sorprender con usos inadecuados de las palabras que hieren.

El texto comienza con un prefacio donde la filósofa explica la importancia de las palabras que hieren y sitúa a larga vista su aportación. Prosigue con una introducción donde profundiza en la vulnerabilidad lingüística y despliega su crítica al performativo soberano y las posibilidades de agencia que inaugura en los siguientes capítulos. En el primero critica la supuesta excepcionalidad del acto de habla para dismantelar la distinción entre conducta y discurso. Comprender el acto de habla como conducta posibilita su restricción y castigo, motivo por el cual pasa a estudiar el tipo de performatividad que subyace a los distintos actos de habla y propuestas de control estatal. Se detiene así en el segundo capítulo en la posición de Catharine MacKinnon y en la polémica sobre la declaración pública de homosexualidad en el tercero. Finalmente, en el cuarto capítulo resitúa la cuestión de la censura en la estructura del lenguaje y da cuenta de las posibilidades de resignificación y recontextualización lingüísticas.

En esta nueva edición Judith Butler añade un prefacio donde posiciona sus reflexiones acerca del lenguaje de odio, la capacidad de

herir de las palabras y la ambivalencia con la que es invocada la libertad de expresión (Butler, 2021, XVI–XVII). Al prefacio le sigue una rica introducción en la que la filósofa piensa en profundidad sobre la condición de vulnerabilidad lingüística que afecta a la formación del sujeto de habla. Judith Butler conjuga la teoría de la interpelación de Louis Althusser y la teoría de los actos de habla de John Langshaw Austin para proponer que la capacidad de un sujeto para producir efectos con sus actos de habla depende fundamentalmente de dos aspectos (Butler, 2021, 26). El primero es el carácter ritual, temporal y reiterativo del acto de habla ilocutivo y el segundo es la instauración previa como sujeto de habla reconocible a través de la interpelación lingüística. Los cuerpos son susceptibles de ser traídos a la existencia social y lingüística en la medida en que están expuestos a la interpelación. La importancia de la introducción estriba en que delinea una serie de cuestiones que atravesarán el resto de las secciones: las formas específicas de violencia lingüística, la supervivencia en el lenguaje, la vulnerabilidad lingüística, la escena de la interpelación y de agencia.

El texto se compone de cuatro capítulos más. En el primero, «Burning Acts, Injurious Speech», Judith Butler critica la singularidad del acto de habla y la diferencia entre conducta y discurso. Judith Butler toma como referencia la polémica que desató en Estados Unidos en 1993 la declaración de inconstitucionalidad por parte de la Corte Suprema de una ordenanza de la ciudad de Saint Paul por la cual se podían condenar como delitos actos deliberados de discriminación por razón de raza, género, creencia o religión (Butler, 2021, 53). Seguidamente, aborda la comprensión de la pornografía como lenguaje de odio por parte de Catharine MacKinnon en *Only Words*. MacKinnon establece que la pornografía es un lenguaje de odio en cuanto

produce a las mujeres en relación de subordinación. Por el contrario, Butler arguye la dificultad de imaginar el campo visual de la pornografía bajo los términos de un sujeto soberano que al proferir da realidad a aquello que nombra (Butler, 2021, 69).

Con la problemática de la autoridad que corresponde al sujeto de habla la filósofa da inicio al segundo capítulo, «Sovereign Performatives». En él analiza con más detalle los argumentos de Catharine MacKinnon para exponer el retorno de formas soberanas de discurso y autoridad que su postura comporta (véase Disch, 1999, 555). MacKinnon defiende que la construcción de la mujer como subordinada en la pornografía es una construcción autorizada por el Estado y solicita en consecuencia que el Estado que intervenga a fin de construir el discurso pornográfico como una forma de injuria que ataca a la integridad y libertad de las mujeres (Butler, 2021, 72). Judith Butler se centra más bien en pensar qué tipo de performatividad subyace al planteamiento de MacKinnon. Según el planteamiento de MacKinnon las palabras y las imágenes constituyen hechos, es decir, producen el efecto que nombran o representan cuando tienen lugar. Concebir la imagen pornográfica y el discurso de odio en clave de un discurso ilocutivo conduce a la ficción de que quien lo ejerce realiza un performativo soberano que se opone a otro poder también soberano, aquel de la ley jurídica (Butler, 2021, 80). Esta aseveración anula la reversibilidad que caracteriza al discurso pornográfico y sexual por la cual el significado de una preferencia siempre excede su contenido locutivo. La preferencia siempre está en riesgo de significar de maneras que no fueron consideradas. Concretamente, Butler sostiene que el propio acto lingüístico de prohibir la sexualidad produce su reverso. Deviene un acto sexual de seducción y, por

tanto, perlocutivo. La pornografía conduce a admitir la prioridad de la dimensión retórica de la preferencia por sobre su función constatativa (Butler, 2021, 85).

Son las posibilidades de dar la vuelta y recontextualizar un discurso las que mueven a Judith Butler a pensar sobre la política del «Don't ask, don't tell» en el ejército estadounidense legislada por el gobierno de Bill Clinton (Smith, 2001, 391, 395). De este punto se encarga en el tercer capítulo, «Contagious Word: Paranoia and "Homosexuality" in the Military». La filósofa estudia el estatus performativo de la prohibición y regulación discursiva de la declaración pública de la condición de homosexualidad. La autora critica que se identifique el acto de habla que declara la propia homosexualidad con un acto homosexual. Salir del armario es un tipo de acto de habla muy preciso pero que no produce plenamente aquello que profiere, la homosexualidad. En palabras más precisas se trata de un acto de habla que vuelve a la homosexualidad discursiva pero no convierte al discurso en referencial (Butler, 2021, 125).

Esta distancia entre la performatividad del lenguaje y su referencialidad es la que permite a Butler distinguir entre el nombre y aquello que nombra. Este factor se une a las posibilidades de retorcer y recontextualizar nombres e insultos como el de *queer* para estudiar cómo funciona la censura y las posibilidades de agencia lingüística en el último capítulo, «Implicit Censorship and Discursive Agency». Si la censura es una limitación del discurso y el lenguaje de odio deja de ser visto como mero discurso para ser considerado legalmente una conducta punible, entonces la restricción del lenguaje de odio, la pornografía o la declaración pública de la homosexualidad podría no considerarse censura, pues se trataría de castigar una conducta. Judith Butler propone alejarse de la noción jurídica de censura para estudiar

cómo la regulación de la distinción entre discurso y conducta opera a favor de una forma de censura implícita (Butler, 2021, 128). La censura como fenómeno lingüístico opera en el plano de la forclusión constitutiva en la formación del sujeto de habla que diferencia entre el discurso admisible e inadmisibile. Tal forclusión produce como efecto al sujeto y a un régimen validado de discurso (Butler, 2021, 139). Cada vez que el sujeto de habla de la esfera de lo decible realiza una preferencia vuelve a invocar los términos de su forclusión que le produjeron y que aún le sostienen. Judith Butler admite que esta reflexión sobre la censura en el lenguaje como operación que es posibilitadora y límite mismo de la agencia no da una guía para decidir sobre las discusiones sobre la censura jurídica (Butler, 2021, 140). No obstante, sí permite entender que la mejor manera de oponerse a la censura es replantear las condiciones en las que tiene lugar la forclusión en la formación del sujeto de habla (Butler, 2021, 140). Si lo habitual es considerar varias prácticas de censura para después elegir entre aquellas en función de un conjunto de principios normativos, Butler alega que nuestras descripciones están normativamente estructuradas por adelantado a través de forclusiones que delimitan el ámbito de lo decible (Butler, 2021, 141). Con tal fin la filósofa recurre a dos pensadores, Pierre Bourdieu y Jacques Derrida. Pensar al sujeto en un contexto postsoberano no supone despejarlo o anular su capacidad de agencia, más bien, supone utilizarlo en un nuevo contexto. Este movimiento permite desplazar el firme sentido contextual que ese vocablo guarda y da cuenta de cómo una preferencia puede adquirir precisamente su fuerza de la ruptura con el contexto en el que se realiza. La filósofa vincula así las posibilidades de reinscripción, que diría Derrida, con la concepción del acto de habla como

rito de institución en el sentido de Bourdieu. (Butler, 2021, 145).

Excitable Speech es una obra que da cuenta de las tensas discusiones académicas en torno a la pornografía, el lenguaje de odio y la libertad de expresión. Judith Butler desarrolla sus argumentos resituando los problemas de la censura, la representación y la violencia en las propias condiciones lingüísticas que las hacen posibles de catorce años de su publicación y tras la derrota electoral de Donald Trump Routledge ha decidido incluir en su colección de clásicos la obra en la que por primera vez la filósofa estadounidense describe explícitamente la vulnerabilidad como condición constitutiva de la corporalidad. En el prefacio Judith Butler pone en relación con su nuevo contexto de publicación los elementos insuperables de las formas específicas de violencia y vulnerabilidad que el lenguaje comporta. Alude especialmente a cómo los discursos que expresan odio racial, misógino, transfobo, heterosexista y cissexista tienen una renovada licencia (Butler, 2021, xv). Usos discriminatorios del lenguaje como el racista, el misógino y el transfobo no buscan simplemente asestar un golpe, sino también menguar la agencia y degradar a las personas a las que se dirigen. La injusticia cometida no es sencillamente el sufrimiento no merecido, puesto que incluye la pérdida de derechos, de estima, de estatus y de libertades. Si el lenguaje discriminatorio es considerado simple ofensa por no suponer ningún daño corporal, esto conduce a ignorar el carácter corporal de los actos de habla y la exposición lingüística del cuerpo. Dirigirse a alguien para hablarle de cierto modo es actuar y tratarle de un modo específico y decididamente corporal. La exposición a la interpelación ajena, a un lenguaje que no es nuestro y a la circulación de palabras son elementos de una vulnerabilidad lingüística

de la cual no es posible desprenderse. Si alguien profiere «somos libres de hablar como queramos», siempre actúa dentro un campo lingüístico de restricción y libertad que es previo, con unas palabras que son prestadas y cuyo sujeto ha sido ya previamente constituido. La libertad de expresión siempre tiene lugar dentro del juego de restricciones y posibilidades que es un lenguaje dado y conviene pensarla en sus términos a fin de comprender mejor cómo puede realizarse (Butler, 2021, XVI).

Esta última aseveración es la que permite a Judith Butler discutir la distinción entre forma y contenido como argumento en defensa de una pretendida libertad de expresión (Butler, 2021, XVII). Si el contenido de un acto de habla es pensado al margen de la forma que lo vehicula y de sus efectos sólo habrá que atenerse al significado de las proposiciones, sin necesidad de atender al contexto de la preferencia. Pero si se admite que la forma lingüística que vehicula una preferencia y su contexto son inseparables de su significado habría que admitir que no se puede escindir el lenguaje discriminatorio del modo y contexto en que es proferido. Así sucede en los casos de abuso sexual en la universidad que Judith Butler menciona en su prefacio. Los actos de habla con carga sexual y de menosprecio hacia las y los estudiantes debilitan su capacidad para conocer sus propias habilidades. Tales comentarios dificultan la percepción sobre su valía y ponen trabas a sus derechos a la igualdad de trato, a trabajar y a estudiar en un ambiente que no sea hostil (Butler, 2021, XXII). La vulnerabilidad lingüística que sugiere Judith Butler da cuenta de un cierto grado de desposesión que tiene lugar en el lenguaje. Ser en el lenguaje significa ser susceptible a la herida (Burgos Díaz, 2006, 98), a la interpelación y a la destitución violenta como sujetos de habla. El lenguaje es algo que

se escapa de nuestro control y que impone un conjunto de condiciones que delimitan la supervivencia lingüística. Por otro lado, limitar el discurso público corre el riesgo de menospreciar que un cuerpo que llega a adquirir una existencia social puede significar en modos inesperados. Un cuerpo puede hablar desde el lugar de lo indecible citando inesperadamente una autoridad lingüística que no le corresponde y poner en cuestión los términos que hacen de un sujeto de habla viable y reconocible. *Excitable Speech* es una reflexión sobre la vulnerabilidad y la agencia lingüísticas que da continuidad a la tarea que Judith Butler dejó por escrito en *Deshacer el género: cesar de legislar para todas las vidas lo que es vivible tan sólo para unas pocas y no decretar para todas lo que es insoportable para algunas* (Butler 2004, 8).

Bibliografía

- Burgos, E. (2006), «Cuerpos que hablan», *Contrastes. Revista Internacional de Filosofía*, 11, 93-109.
- Butler, J. (2004), *Undoing Gender*, Londres, Routledge.
- Butler, J. (2021), *Excitable Speech: A Politics of the Performative*, Londres, Routledge.
- Disch, L. (1999), «Judith Butler and the Politics of the Performative», *Political Theory*, 27(4), 545-559.
- Smith, A. M. (2001). «Words that Matter: Butler's *Excitable Speech*», *Constellations*, 8(3), 390-399.

Javier Moscoso Cala
(Universidad de Málaga)

BRANDOM, Robert M. (2019). *A Spirit of Trust. A Reading of Hegel's Phenomenology*. Cambridge, Massachusetts: The Belnap Press of Harvard University Press, 768 pp.

A interpretação de Brandom sobre a *Fenomenologia do Espírito* de Hegel, era há muito tempo esperada, pois ele trabalhou durante muitos anos nessa pesquisa. Trata-se de uma contribuição orientadora para a filosofia contemporânea por um dos seus representantes mais conhecidos e influentes.

Nessa obra, Brandom apresenta uma nova reconstrução da aventura idealista da razão, elaborando o comentário e a interpretação semântico-pragmática da *Fenomenologia do Espírito* de Hegel. Ele combina a tradição analítica, continental e histórico-hermenêutica, ao mesmo tempo em que ele mostra como as formas do pensamento dominantes na filosofia contemporânea são desafiadas por Hegel.

A Spirit of Trust trata da profunda transformação na história humana que caracteriza a chegada da modernidade. Na sua *Crítica da Razão Pura*, Kant faz a distinção entre as coisas como elas são em si mesmas e como elas nos aparecem. Hegel discute a distinção de Kant, explicitando o que separa o mundo moderno do mundo antigo. No mundo antigo, as declarações normativas — juízos sobre o que deveria ser — eram tomadas como fatos objetivos. Esses juízos são frequentemente vistos como um dom de Deus. No mundo moderno, esses juízos são entendidos como juízos estabelecidos pelos sujeitos. Hegel tem uma visão que combina ambas as abordagens, que Brandom chama de “idealismo objetivo”: Existe uma realidade objetiva, a menos que

primeiro fazemos sentido à forma como pensamos sobre ela. Em certo sentido, o mundo é composto por pensamentos que compõem a matéria, o material da mente. Há em ambos os polos, do lado do mundo e do lado da mente (*mind*: consciência) pensamentos que estruturam esses lados e que compõem o “espírito” de acordo com a leitura de Brandom sobre Hegel.

Na opinião de Hegel, só nos tornamos sujeitos atuantes quando somos conhecidos e tomados como tal por outros sujeitos atuantes. Isso significa que o estatuto normativo de tais conceitos como obrigação, responsabilidade e autoridade é estabelecido por práticas sociais do reconhecimento mútuo. A modernidade é a era da dominação (*mastery*): a dominação do sujeito autônomo. Isso leva a uma concepção deficiente das normas, que está associada à alienação. A autonomia moral da autolegislação contém o problema de que a compulsão legal e a intelecção livre se contradizem. As normas apenas dadas a si mesmo perdem facilmente sua força vinculativa. Brandom argumenta que se nosso comportamento autoconfiante e reconhecido assume a forma do perdão, da generosidade e da confiança descritas por Hegel na transição da moralidade para o espírito absoluto da religião, podemos superar a época problemática da modernidade caracterizada pela alienação e entrar numa nova era, a era da pós-modernidade, que para ele é caracterizada por uma ética do mútuo perdão (*forgive*) e confiança (*trust*) e, portanto, da reconciliação (*reconciliation*).

Brandom é da opinião de que ainda hoje podemos aprender com Hegel. Na sua interpretação sistemática da *Fenomenologia do Espírito*, ele estabelece um diálogo fecundo com a metafísica, a ontologia social, a teoria atual da ação e a filosofia da linguagem. Brandom mostra, entre outras coisas, como

Hegel, muito antes de Wittgenstein, não apenas via claramente o problema fundamental de seguir uma regra, mas também propunha uma solução baseada na natureza social das normas. Assim, a visão passa da consciência moral subjetiva para normas, costumes, práticas e crenças comuns. A prática normativa é social, uma mútua relação ética da obrigação que é assumida pelos sujeitos.

Os conceitos-chave da interpretação de Brandom da *Fenomenologia do Espírito* (FE) são os seguintes: Negação e verdade; Conceitos empíricos e especulativos; Experiência; Luta e reconhecimento; Formas do idealismo; A vida ética pré-moderna e a alienação moderna; Uma ética pós-moderna do perdão, da confiança e da reconciliação.

O livro *Um espírito de confiança* é composto de três partes: Parte 1 (compreende 7 capítulos da p. 35-234) - Semântica e Epistemologia: Conhecer e representar o mundo objetivo. Nesta parte Brandom comenta os três primeiros capítulos da FE: Consciência (Certeza Sensível, Percepção, Força e Entendimento). Trata-se de uma leitura semântica do processo do conhecimento, ou seja, de uma epistemologia que usa a estratégia da semântica descendente.

Parte 2 (compreende o capítulo 8 a 12, p. 235-468) – Pragmática normativa: Reconhecimento e Metafísica expressiva da agência. Aqui Brandom faz a leitura do capítulo 4 da FE: Autoconsciência (a verdade da certeza de si mesmo), comentando a clássica figura da luta entre o “senhor e o servo” como a interdependência (autoridade e responsabilidade recíprocas) da atitude de reconhecimento e atribuição de estatutos normativos. O autor descreve a estrutura da autonomia e do reconhecimento das atitudes entre os agentes. Brandom interpreta o capítulo sobre a Razão como um novo modelo de ação, não mais

como um evento mental ou volitivo que simplesmente acontece, mas como um processo a ser produzido.

Parte 3 (compreende o capítulo 13 a 16, p.269-635) – Rememorando as eras do espírito: da ironia à confiança. Esta parte comenta o capítulo VI da FE: o espírito. Para Brandom a história das estruturas normativas do espírito compreende a eticidade imediata, a emergência da subjetividade, a alienação e a cultura. O ponto culminante da FE, na leitura de Brandom, é a estrutura de reconhecimento rememorativo da confissão e do perdão através da “confiança” (*Vertrauen*). A última lição do livro de Hegel é a seguinte: “Compreendermo-nos enquanto estrutura pós-moderna da normatividade, comunidade, e autoconsciência individual que é articulada por práticas de reconhecimento tem a forma da confiança. Nesta estrutura, reconhecimento mútuo tem a forma diacrônica de atitudes práticas de confissão e perdão. São práticas sociais de reconhecimento que instituem normas, assumindo a forma histórica magnânima de rememoração” (Id. p. 30-31). A estrutura normativa prática reconhecível da rememoração expressiva dá-se pelo perdão, confissão e confiança. Isto realiza a nova forma cognitiva e teórica da autoconsciência, uma nova forma prática magnânima de agência de autoconsciência. Na conclusão (p. 636 – 758), o autor reafirma sua leitura da FE: a articulação do reconhecimento e a rememoração como o caminho para a era da confiança.

O livro faz uma leitura semântico-pragmática da FE de Hegel. O tópico focal é o conteúdo e o uso dos conceitos, isto é, “as relações entre significado e uso, conteúdo conceitual e aplicação de conceitos (os tópicos, respectivamente, da semântica e da pragmática), entre normatividade e modalidade” (p. 4). A estratégia da semântica

descendente é expressar o uso e o conteúdo das práticas empíricas e conceituais que Hegel chama de “conceitos determinados”. Os conceitos categoriais tornam explícito o que está implícito no uso de qualquer conceito aplicado no juízo e nas ações (ver p. 5). A estratégia da semântica descendente não é a metodologia de Hegel, mas a de Brandom. Os metaconceitos ou metacategorias (consciência, autoconsciência, razão etc.) têm a função de tornar explícito o que é implícito no uso e no conteúdo empírico ordinário e nos conceitos práticos. Esse é o método semântico descendente, como um olhar através de binóculo e não de um monóculo, ou seja, isso articula nossa autoconsciência semântica e pragmática (ver p. 8).

O discurso com dimensão social constitui a normatividade e o reconhecimento. A atividade discursiva é a aplicação de conceitos os quais vinculam compromissos e responsabilidades práticas. A teoria pragmática normativa da atividade discursiva é a estratégia pragmática de compreender a semântica em termos pragmáticos (ver p. 10). Ou seja, Hegel constrói a dimensão social da normatividade enquanto reconhecimento, e a dimensão histórica do conteúdo conceitual em termos de racionalidade rememorativa (ver p. 11). Hegel explica as normas discursivas como produtos de práticas sociais, ou seja, a “Bildung”, a cultura é a nossa segunda natureza, construída em nossas interações sociais com os outros (ver p. 12).

A dimensão histórica do discurso expressa-se na racionalidade rememorativa, isto é, rememorar a ideia de normas implícitas em práticas sociais (costumes, usos, instituições). A substância, para Hegel, é a síntese do reconhecimento mútuo. Hegel usa três metaconceitos para explicitar esse reconhecimento substancial: O “Univer-

sal” (comunidades) se desenvolve a partir de relações de reconhecimento entre seres “Particulares” que se tornam uma auto-consciência normativa “Subjetiva”, adotando atitudes práticas de reconhecimento através do outro (ver p. 14).

A rememoração é a base da expressão, isto é, “a relação entre o que é implícito e o que é a expressão explícita daquilo” (Id. p. 18). Hegel compreende as fases da experiência prática e cognitiva como a emergência da explicitação rememorativa do que estava implícito. No lado cognitivo, o que está implícito é o mundo objetivo em si, no lado prático o que está implícito é a intenção do agente em si. Hegel explica o que está implícito como um processo se expressando, isto é, “o processo de tornar-se explícito”. A rememoração racional retrospectiva reconstrói o curso da experiência enquanto progride se expressando (ver. p. 18).

Brandom resume sua leitura: Conhecimento, reconhecimento, pragmatismo normativo e a historicidade do Espírito. “Rememoração é a chave para compreender a relação entre a semântica de Hegel (teoria do conteúdo conceitual) e a pragmática normativa (atitudes práticas que são o uso ou a aplicação de conceitos). Este é o pragmatismo como idealismo rememorativo conceitual” (Id. p. 19).

O novo tipo de autoconsciência teórica obtido pela rememoração fenomenológica é para tornar possível uma nova forma de prática normativa. “Quando o reconhecimento assume a forma magnânima de rememoração, é o *perdão*, a atitude que institui a normatividade como uma *confiança* totalmente autoconsciente” (Id. p. 19).

Após essa contextualização inicial sobre a obra de Brandom, cabe apontar alguns pontos críticos de sua obra *A Spirit of Trust*. Brandom tem uma compreensão

analiticamente reduzida de importantes conceitos básicos de Hegel. A negação determinada que tem três significados em Hegel – *Aufhebung* = superação/suprasunção: negar, conservar e elevar – é lida como incompatibilidade material (*material incompatibility*). Os múltiplos significados da mediação em Hegel (a mediação de algo por outro algo, a mediação de duas determinações por um terceiro mediador, a automediação de algo consigo mesmo, etc.) estão nivelados à capacidade de ligação (a bola pode ser vermelha e redonda) (*material consequence*). Enquanto em Hegel a negação é mediação e a mediação também é negação, em Brandom a mediação é o outro da negação. Quase não se fala da negação autorrelacionante e dos diferentes significados da imediatidade em Brandom. Com Wittgenstein Brandom diz que a negação é separação e a mediação é conexão. Assim, Brandom também escapa da lógica da negação determinada em Hegel, segundo a qual apenas a cisão negada resulta em reconciliação.

Para Hegel a rememoração tem a função de garantir a verdade das coisas finitas, enquanto Brandom enfatiza o contrário, isto é, a rememoração (*recollection*) é o reconhecimento dos erros e enganos cometidos na determinação do conceito das coisas no decurso da história, seguindo essa série: rememoração, perdão, reconciliação e confiança. A referência da verdade na rememoração em Hegel é contrastada com sua referência de erro em Brandom. O posterior reconhecimento e perdão de erros na determinação dos conceitos das coisas nos ensina a não condenar e rejeitar o que é contrário à verdade, mas a incluí-la no processo da verdade. As oposições não devem ser criticadas e eliminadas, mas atenuadas. Essa é a única maneira de construir confiança num mundo cheio de oposições

e contradições. Esta é precisamente a lição que deve ser tirada do processo normativo da experiência: devemos perdoar os erros morais e os enganos dos outros. Brandom apenas considera os conflitos de interesse como oposições morais, os quais ele apenas pretende mitigar.

Se a rememoração em Hegel tem o significado da suprassunção criticamente pensante da alienação social e política, no livro de Brandom ela tem o significado da contenção da alienação moral através de uma construção comunitária bem-sucedida e orientada por normas. Isso é uma consequência de sua má compreensão da lógica da negação determinada, que implica em dizer não a degradação moral, ao empobrecimento econômico e a violência política. A figura dialética da negação das formas repressivas e dilacerantes da sociedade permaneceu, portanto, estranha a ele. A medida para a vida “boa” ou “não fracassada” das pessoas é avaliada apenas em nível moral. A questão da organização social da produção da riqueza para uma vida materialmente boa nem sequer é posta. Por exemplo, na crise ecológica a ênfase deve ser colocada no domínio da tarefa de reproduzir a vida social no “processo metabólico” em relação de sustentabilidade com a natureza externa.

Para Brandom, espírito significa confiança recíproca e respeito mútuo. Seu ideal é a comunidade de pessoas moralmente boas e generosas. No entanto, continua a ser questionável se esse ideal é mais do que uma compensação fraca da concorrência capitalista.

Em Brandom pode-se observar o fenômeno de que na filosofia contemporânea,

a filosofia fundamental sobrevive apenas como ética normativa. O ser humano não é ser humano sem dever ser e obrigação moral. Sua crítica não passa da crítica moral à crítica social. No entanto, a alienação social que é vista como um não “deve ser” não tem o significado do “dever ser” moral, mas de um não querer viver assim. Ou seja, pode-se dizer: esse não querer é uma negação existencial, não uma mera negação moral. Brandom é um normativista ético. Mas a moral não é suficiente para alcançar as condições sociais objetivas.

Enfim, Brandom defende um Hegelismo analiticamente reduzido e neopragmáticamente inspirado, que por um lado encurta o conceito de espírito de Hegel pelo espírito absoluto da religião, por outro lado transforma o espírito objetivo em uma pragmática normativa (moralidade, eticidade). A razão suprapessoal e anônima em Hegel é reposta no processo intersubjetivo de reconhecimento de seres racionais. Nisso, há também pontos de conexão com Jürgen Habermas e a mais recente Teoria Crítica de Axel Honneth. O ideal da comunidade de pessoas moralmente boas carece, essencialmente, da falta da crítica da economia política do capitalismo com seus antagonismos sociais que apenas podem ser mantidos pela violência política.

Christian Iber

(Universidade livre de Berlim; Universidade de Friburgo),

Agemir Bavaresco (PUCRS, Brasil)

Danilo Vaz Costa (UNICAP)

NÚÑEZ, Amanda (2019), *Gilles Deleuze, una estética del espacio para una ontología menor*. Madrid, Arena Libros, 232 págs.

“Consideramos que Deleuze apuesta fielmente por una ontología, pero una *ontología menor* que genera una nueva *imagen* de lo que significa pensar (...). Deleuze parece ser más ontólogo, si cabe, a medida que la expresión “ontología” desaparece de sus escritos” (p. 44).

Con esta rotundidad y claridad se nos presenta el libro de la profesora de la UNED Amanda Núñez. Es una auténtica declaración de intenciones de lo que nos trae de novedad este texto sobre la ontología deleuziana. En esa cita aparecen las dos ideas fuerza que impulsan su análisis: por una parte la defensa de la existencia de una ontología explícita en Deleuze —y su caracterización como ‘menor’ absolutamente original— y, por otra la inspiración casi “genealógica” que alumbra el estudio de las condiciones de posibilidad de esa ontología en la obra del filósofo francés. Transita por los textos deleuzianos buscando cómo se han ido formando la multiplicidad de sentidos de los propios conceptos que escudriña. Y esto es especialmente decisivo en la obra deleuziana por la diversidad de fuentes de las que se nutre, o por la ‘personalización’ que hace de términos procedentes de la tradición filosófica.

Pero, hay que avisar que este libro no es una introducción a la filosofía de Gilles Deleuze o una monografía sobre él, sino una inmersión en el mundo deleuziano, es un pensar con él, desde él, trabajando con su gesto. Todo ello con un gran rigor en la escritura filosófica que no deja ningún cabo suelto, aunque sí deja siempre la puerta abierta a la propia experimentación. La lectura apasionada y apasionante que hace del autor tampoco le priva del sentido crítico

respecto a las posiciones que puedan tener otros filósofos sobre el pensamiento de Deleuze, con los que discute y debate, llegando a sus propias conclusiones.

El concepto “ontología”, que vehicula la reflexión del libro, siempre es presentado como problemático o, mejor, como fronterizo. No sólo en la obra de Deleuze —por sus distintos devenires y su progresiva desaparición— sino por el mismo destino que ha tenido en la filosofía y su crítica más contemporánea. El espacio en el que se mueve Deleuze es el del diálogo-debate con aquellos que la desvinculan de la Metafísica, como en Heidegger (con el que el autor mantiene un tira y afloja constante en toda su obra), como con aquellos que la reivindican. En este último grupo estaría la famosa discusión con Badiou con el que sostiene un enfrentamiento (más por parte de Badiou, que del propio Deleuze) sobre las posibilidades de una ontología de la diferencia que asuma la unicidad del ser, tal y como defiende Deleuze y que Badiou contradice en función de su ontología formalizada en la teoría matemática de conjuntos.

El proyecto ontológico “menor” deleuziano que se defiende en este texto conduce a una nueva ‘imagen del pensar’ y una nueva ‘materia del ser’ como partes esenciales en las que se derrama su sistema abierto de filosofía. Porque, en efecto, la única ontología que Deleuze se permite es aquella que coincide con la filosofía misma, siempre que esta se desprenda de sus condiciones, tal y como se han entendido en su tradición más canónica, y experimente —alterando— las nociones de espacio, tiempo o límite, ya que en esa alteración, en ese derrame, se juega la posibilidad de una nueva forma de pensar.

La forma en la que se aborda el tratamiento de la “imagen del pensar” en el autor francés no solo es una explicitación de la posición filosófica de Deleuze sino también de la propia autora de este libro. Y esa posición es una postura netamente filosófica, aunque, eso sí, en el sentido deleuziano del término: un pensamiento liberado, que abre vías a la exploración de nuevas significaciones, de nuevas formas de pensar que se alejan de las corrientes que la han convertido en un cliché escolástico, sin explorar su propia problemática.

Esto no significa que sea un pensamiento ensimismado en su propio discurrir; más bien lo que se reivindica es una manera de pensar “ingenuamente”, sin grandilocuencia, en tono menor y sobre todo, en relación con otras actividades (de ahí los textos de Deleuze sobre pintura, Bacon, literatura, Proust, el cine o los nuevos deportes, y las consecuencias que se pueden extraer de los nuevos escenarios que abren las ciencias al trastocar las concepciones del espacio/tiempo), pero sin hacer mixturas aberrantes o ilusorias, sino articulándolas de manera ontológica. Como muy bien se insiste en el libro, la ontología deleuziana no es del “ser” (“est”), sino del “y” (“et”), de la relación y la diferencia. En esta lectura de la realidad se niega la perspectiva estática de entes aislados y sometidos a la lógica de una jerarquización óptica respecto a un Ser Universal, y se defiende una visión donde hay relaciones y acciones que actúan de diferente manera en función del lugar en el que se encuentren actuando. De lo que se trata es de ver cómo las cosas, las acciones, los deseos, los pensamientos... se reparten en la extensión del ser de manera horizontal y sin que haya territorio dividido previamente: se ocupa temporalmente, y su límite es todo lo que puede ocupar, lo que la intensidad de su potencia, en relación con las otras, le permita.

En esta nueva imagen del pensar hay una filosofía que reniega de las alturas, del tono elevado —que por altivo solo es máscara y simulacro—, y de su condición de amante del saber que se arroga la potestad de adoc-trinar a los demás sobre la verdad de las cosas porque cree haber alcanzado la auténtica naturaleza de las mismas. Tal tradición solo ha generado ilusiones nefastas sobre el valor del saber filosófico: lo ha colocado en una posición de superioridad que, al final, lo termina expulsando del mundo: la ilusión de la filosofía como contemplación (del Bien supremo), como reflexión acerca de todo, o como orientadora de las relaciones inter-subjetivas. Deleuze advierte, con Foucault, que estas ilusiones nos hacen creer que la filosofía es un saber independiente y reactivo al poder, y con la capacidad de desvelar los secretos de éste y liberarnos de él. La vía que ensaya Deleuze, y que es la que se nos ofrece en este libro, es la de una filosofía que sabiendo los vínculos saber-poder, construye ‘artefactos’ para poder intervenir en lo que hay, pero en tono menor, sin dar por sentado nada de antemano, sin imponer esos útiles a los demás, solo transitando por la realidad produciendo máquinas que pueden funcionar o no, según las circunstancias. Amanda Núñez compara esta tarea con la de los operarios, “nada fuera de lo que pueda ser cualquier oficio y su tonalidad respectiva: el cualquiera” (p. 62) (el ‘cualquiera’ se refiere al singular, no al Universal abstracto intercambiable).

Para ensayar este pensar como experimentación, Deleuze mira de otra manera la historia de la filosofía. No “derriba” todo pensamiento anterior, ni a la tradición canónica, pero, sí va contra lo que lo que la autora llama las “zonas dogmáticas” (p. 143) de esos pensamientos. Además, se fija en otras corrientes que han sido tomadas menos en cuenta, corrientes subterráneas. Así, si en la

tradición más ortodoxa prima el presupuesto pre filosófico de una vinculación entre el saber y el bien, la dirección que toma el pensamiento de Deleuze quiere fijarse más en otra corriente, que sitúa en los Cínicos o en la Stoa, y que aportan otra ‘imagen’ del pensar que lo vinculan más al cuerpo (las ‘filosofías de la risa’, p. 46) y lo sitúan previo al logos, que no es el mito, sino la sensación, la sensibilidad como “atmósfera pre filosófica”, o ‘sensus communis’ imprescindible para desarrollar cualquier forma de pensamiento nuevo y articular una estética del espacio para una ontología materialista.

Sin embargo, el proceso no es solo gnoseológico, sino, como certeramente se expone en el libro, es también ontológico, pues la ontología minorada de Deleuze se ‘derrama’ tanto en una imagen del pensar como una materia del ser. Pero, el ser del que nos habla esta ontología no es ni el de lo inmovible, permanente, homogéneo y cerrado en sí mismo con el que venimos tratando con mayor o menor fortuna desde el platonismo; tampoco es el ser del movimiento, tal y como nos remite la dialéctica hegeliana y toda su cohorte de historicismos. Ambas tradiciones han demostrado su improductividad en lo real para el autor galo.

La materia del ser se acerca más al Cuerpo sin Órganos, potente imagen (que no metáfora) recogida de Artaud, que remite a la ausencia de una articulación estática que estableciera la identidad construida desde un origen jerárquico (Ser/entes). El Cuerpo sin Órganos realza, por encima de todo, la materialidad del ser, la negación de la trascendencia, y la apertura a la desviación de la normatividad. Por eso, insiste Núñez, en repetidas ocasiones del texto, el posicionamiento de la ontología menor como una ontología del “y”, donde no hay dominio de ningún orden: ni el del pensar (y ahí está toda la crítica que el libro hace a las interpre-

taciones idealistas de Deleuze, como la que va de las p. 95 - 101, aunque no es la única) ni el del ser. Lo que prima es una diferencia, pero hecha a-lógicamente, emancipada de la lógica dicotómica, para poder pensar de otra manera. No sólo para pensar las cosas de otra manera, sino para que pensar sea otra cosa: una actividad lúdica, impura, e inmanente.

¿Por qué esa exigencia de pensar de otra manera? Amanda Núñez nos acompaña a visitar a ese Deleuze que quiere una filosofía más apegada a la tierra, pues es ésta la que da que pensar, para asumir la complejidad de lo real, y expandir los límites de nuestro reflexionar. Nos obliga a alejarnos de la lógica de la re-presentación y las dicotomías excluyentes, a abandonar ese pensar que se maneja mejor con la analogía del ser —pues ello le permite conocer a partir de cánones ya establecidos— que con la univocidad del ser (entendida de forma no sustancialista, “el ser es voz que se dice” dice Deleuze en *Lógica del sentido*). Este ser unívoco coloca a todos en un mismo plano no jerárquico, afirmando no solo la multiplicidad de lo que hay, su diferencia radical inconmensurable, y su absoluto inmanentismo, sino también el continuo devenir no cronológico, la heterogénesis. De ahí, la misma dignidad para Dios que para la garrapata, porque constituyen una misma comunidad de ser, que no es identidad, pues lo que define la comunidad de ser es precisamente la diferencia de formas de individuación, la flexibilidad, los flujos y las potencialidades.

Esta filosofía tiene un sentido netamente político, y en él incide el libro cuando señala que si no queremos que la filosofía sea pura ideología, tendrá que actuar como “un agenciamiento del pensar que lucha abierta y constantemente contra la solidificación de las costumbres y las metáforas olvidadas convertidas en dogmas y leyes de la ciudad” (p.135). Pero si esto hace de la filosofía

una actividad, no la convierte en un mero recetario para activistas. Primero, porque no hay una relación causal y de dirección única entre teoría y praxis; ambos espacios se interfieren y enriquecen, se entorpecen y taponan, se alumbran y se oscurecen.

Y, en segundo lugar, porque la tarea principal de la filosofía es la de crear conceptos. Cuando en *Mil Mesetas*, Deleuze y Guattari insisten en la vinculación de la filosofía con los conceptos, señalan que los conceptos pueden ser tratados de muchas maneras. Y la que les interesa a ellos es la de que sean herramientas para circular por el mapa de lo real. Por eso, las palabras, sean creaciones propias, de otros pensadores, o tomadas de otros campos (la botánica: rizoma, la genética: heterogénesis, de la literatura, la química), son conceptos que abren a pensar la “no-filosofía de la filosofía”, es decir aquello no filosófico con lo que la máquina de la filosofía puede conectarse para que funcione. De ahí, su relación con otros artefactos literarios y artísticos por su potencialidad de transvaloración. Núñez señala que lo no-filosófico se dirige también a las mismas condiciones internas de la filosofía que no son exclusivamente conceptuales, sino que transitan entre la estética y la política. Desde aquí, el texto desarrolla una prolija y espléndida analítica del tiempo y el espacio, y, sobre todo, del bloque espacio/tiempo —cercana y lejana a la vez de la kantiana— que atraviesa todo el libro y es recogida en la “estética” que le da título y justifica la tesis defendida en el mismo.

La potencia de estos nuevos conceptos es dar cabida a la posibilidad de nuevos espacios donde habitar en este mundo nuestro, para poder cambiar algo. Aquí es donde encaja la dimensión política de esta “imagen del pensar” como pensar desde las minorías y ejercicio de desterritorialización. La mino-

ría es política en tanto que presenta la potencia (o virtualidad) de desviarse de la norma. La minoría no es una cuestión de cantidad, sino de norma y procesos de cambio; minorar es tener potencia de transformación, y en ese sentido todos somos minoría en algún momento, siempre y cuando estemos en disposición de asumir ese movimiento. Y por eso mismo, la ontología solo puede serlo si es menor, o minorada, es decir si deja de ser ontología del Ser y se pega a la inmanencia de lo matérico.

Amanda Núñez confirma esta forma de entender lo político en Deleuze cuando reitera el carácter experimental que infunde al pensamiento y a las prácticas personales como un salirse de la norma, y asume que las posiciones personales están vinculadas a lo colectivo en un espacio en el que el dentro y el afuera son indiscernibles. También advierte de que todo salirse de la norma no es “minoritario”, en el sentido en el que acabamos de referir; muchas salidas de norma no son más que confirmaciones de la misma. El alejamiento creativo de la norma es el que posibilita descubrir nuevos espacios menos jerárquicos, más intensos y menos sometidos al patrón, y que van transformándose en el mismo habitar. Lo minoritario es la potencia creativa de espacios nómadas. Ahí transita, de igual forma, el movimiento de la desterritorialización como proceso de decodificación de las prácticas sociales, cuyo riesgo es igualmente generar nuevas territorialidades sedentarias que arraiguen en los sujetos y los sometan a los mismos poderes, aunque sea en otros espacios.

Lo que hace atractivo el proyecto deleuziano, tan bien expuesto en este libro, es precisamente esa llamada a la creatividad y a la experimentación, pero, también en ello residen algunos de sus principales riesgos, precisamente porque vivimos en un mundo donde el “emprendimiento” y la

“creatividad” se cotizan al alza. De ello nos advierte la autora en una de las últimas páginas de este texto: “Minorizar y molecularizar nunca jamás llevan asociada una absolutización. Gracias a ello, por un lado, apelan a una creatividad que no es la de los

mercados de ideas que nos desbordan y, por otro lado, nos da herramientas para pensar y resistir a nuestro presente tan complejo y molecularizado por los poderes” (p. 210).

Rosa Jiménez Asensio

CORCHIA, L., MÜLLER-DOOHM, S. und OUTHWAITE, W. (Hg.). *Habermas global. Wirkungsgeschichte eines Werks*. Berlin: Suhrkamp, 2019, 894 pp.

Nos dice algo sobre la situación cultural de nuestra época —un tiempo de inundación informativa y de proliferación febril de congresos y revistas académicas— la publicación de un volumen de casi novecientas páginas sobre la recepción global de un autor vivo. Según los datos con los que los editores abren el prólogo, la bibliografía secundaria sobre Jürgen Habermas abarca no menos de ocho mil títulos, que un especialista ya ha recopilado en un informe de cientos de páginas. Su obra se ha traducido a decenas de lenguas, tal como refleja la última parte de este mismo volumen en una relación de títulos que, pese a ser incompleta, ocupa más de setenta páginas. No es sorprendente, dadas estas cifras, que para trazar la panorámica de la recepción global de Habermas haya sido necesario el trabajo de cuarenta autores procedentes de más de veinte países. Y la intención de este libro colectivo es mostrar las líneas por las que ha discurrido esa recepción, así como las principales controversias en torno a la obra de este filósofo en diferentes contextos nacionales y culturales.

Tras una larga introducción firmada por Stefan Müller-Doohm y Dorothee Zucca, la panorámica se organiza en nueve partes (subdivididas a su vez en varios capítulos) correspondientes a Alemania, Estados Unidos, Gran Bretaña, Escandinavia y Holanda,

Francia y Bélgica, Italia, España y Latinoamérica, Portugal y Brasil, y finalmente Asia. A fin de ofrecer una muestra representativa del contenido del libro, en lo que sigue nos atenderemos a algunos capítulos que nos han parecido especialmente interesantes: los que tratan la recepción de la obra de Habermas en España y Latinoamérica, los que se refieren a los países en los que la lectura de Habermas ha sido más fecunda (Estados Unidos) o más polémica (Francia), y por último el capítulo sobre China, uno de los contextos de recepción de la filosofía habermasiana políticamente más peculiares.

Juan Carlos Velasco firma el capítulo sobre la recepción de la obra de Habermas en nuestro país. Su estudio completa otras panorámicas publicadas hace décadas y ya un poco anticuadas, y presenta de un modo sintético tanto la historia de la introducción de Habermas en España como los temas a los que han prestado más atención los especialistas españoles. Velasco recuerda cómo Habermas se consolida desde finales de los años setenta del siglo XX como una figura de referencia no solo en el mundo académico, sino también en la sociedad de la Transición; un contexto en el que, por razones obvias, la sociedad española necesitaba actualizarse y apropiarse de las corrientes culturales y filosóficas de las democracias de su entorno. Pero de la expo-

sición de Velasco se desprende que en este caso la influencia política precedió incluso a la propiamente académica, puesto que Habermas ya era extraordinariamente famoso en España cuando la mayor parte de sus obras ni siquiera se había traducido. La victoria del PSOE en 1982 convirtió al filósofo alemán en uno de los referentes ideológicos de esa socialdemocracia postmarxista y europeísta con la que el nuevo Gobierno quería identificarse. Esta influencia de Habermas (un poco sorprendente, como señala Velasco, si se tiene en cuenta la complejidad y el carácter más bien técnico de sus escritos) quedó reflejada simbólicamente en el discurso que el filósofo pronunció en noviembre de 1984 nada menos que en el Congreso de los Diputados. En esa misma época las editoriales Taurus y Tecnos asumieron el compromiso de traducir sus libros al castellano (una labor que durante años desarrolló ejemplarmente Manuel Jiménez Redondo), y esto hizo posible que poco a poco los bueyes fueran poniéndose delante del carro, y que a la euforia política le siguiera una lectura más académica. Dicha lectura alcanzó su zenit en la última década del siglo XX, cuando el nombre de Habermas se mencionaba prácticamente en cualquier congreso, seminario o publicación filosófica española, ya fuera para elogiarlo, ya para criticarlo. Entre las principales líneas temáticas por las que ha discurrido la recepción de nuestro autor, Velasco trata la ética del discurso (desarrollada y criticada respectivamente por Adela Cortina y Javier Muguerza), la filosofía teórica (en la que destacan los trabajos de Cristina Lafont), los temas de filosofía social, política y jurídica (Francisco Colom, Juan Carlos Velasco) o la filosofía de la religión (José María Mardones, Juan Antonio Estrada).

El capítulo de Javier Aguirre y Eduardo Mendieta menciona algunas circunstancias peculiares que han condicionado la lectura

de Habermas en Latinoamérica. Entre ellas está el hecho de que unas pocas editoriales españolas (Taurus, Paidós, Tecnos) prácticamente monopolizaran durante décadas la traducción y publicación de los libros de Habermas. Los libros españoles son caros para buena parte del público universitario latinoamericano y también es costosa su distribución en la región, por lo que el monopolio editorial español ha resultado ser en este caso (y seguramente en otros) un obstáculo a la difusión del pensamiento de este filósofo. Otro impedimento ha sido la pervivencia, al menos hasta la década de 1990, de una izquierda marxista revolucionaria totalmente alejada de un enfoque político como el de Habermas, que sin dejar de ser de izquierdas tiene un carácter socialdemócrata y reformista. Solo la progresiva “desmilitarización” de la izquierda latinoamericana y la creciente disposición a solucionar los conflictos de clase en el marco del Estado democrático de derecho ha posibilitado que la filosofía política habermasiana haya adquirido presencia en un entorno académico y social que el propio filósofo, en una entrevista de 1984, interpretaba como muy ajeno a su obra. Desde entonces las cosas han cambiado, y libros como *Facticidad* y *validez* han tenido mucha influencia en la consolidación de procesos de constitucionalización o de reforma constitucional en los países latinoamericanos. Lo mismo sucede con la ética del discurso, que actualmente tiene una presencia académica notable, sobre todo en Argentina, gracias a la *Red Internacional de Ética del Discurso* fundada en 2011. No obstante, en Latinoamérica el pensamiento de Habermas ha sido objeto de una lectura atenta pero crítica, como lo refleja paradigmáticamente la figura del filósofo mexicano-argentino Enrique Dussel, partidario de una “filosofía de la liberación” que toma elementos de

la filosofía de Habermas y sobre todo de Apel, pero adaptándolos a las condiciones materiales de sociedades que mantienen una relación compleja con las tradiciones intelectuales procedentes de las potencias imperialistas (pasadas y presentes), y que están muy marcadas por grandes desigualdades socioeconómicas.

Pero para comprender la propia evolución del pensamiento de Habermas, resulta especialmente interesante el capítulo de Eduardo Mendieta y Benjamin Randolph sobre Estados Unidos. A finales de los años sesenta Habermas comenzó a impartir clases regularmente, como profesor invitado, en universidades norteamericanas, y esto dio lugar en las décadas siguientes a cierta simbiosis. Desde esa misma época sus libros se traducen regularmente al inglés y se publican (como sucede en España) en editoriales no universitarias, lo que ha favorecido su difusión entre un público amplio, no especializado. Por otra parte, en el ámbito propiamente académico profesores como Thomas McCarthy, Richard J. Bernstein o el propio Rorty han leído a Habermas aproximando su pensamiento a las tradiciones filosóficas específicamente norteamericanas, en particular el pragmatismo. Y esta lectura ha influido a su vez en Habermas, que desde la década de 1970 ha incorporado cada vez más temas, debates y enfoques norteamericanos a una filosofía cuyas raíces se situaban originalmente en un suelo inconfundiblemente “continental”.

Interesante es también la recepción de Habermas en Francia. Dada la tormentosa e interminable polémica de este filósofo con los postestructuralistas franceses, el lector se asoma con curiosidad casi morbosa a este capítulo, firmado por Isabelle Aubert. Y lo primero que uno descubre es que, en la lectura francesa de Habermas, hay un *antes* y quizás incluso un *después* de dicha

polémica. Traducidas al francés en la década de 1970, las obras de Habermas tuvieron una acogida favorable entre quienes en ese momento buscaban en la Teoría Crítica frankfurtiana alguna pista para desarrollar un marxismo alternativo al de Althusser, pero también fueron seriamente estudiadas por un filósofo hermenéutico como Paul Ricoeur. En los años ochenta, sin embargo, Habermas encuentra en Francia una acogida bastante menos favorable. En este cambio resultó determinante la crítica sociológica de la pragmática universal habermasiana que Pierre Bourdieu desarrolló en el libro *¿Qué significa hablar?* (1982), pero sobre todo las críticas a Nietzsche, Heidegger y sus continuadores postestructuralistas que Habermas presentó parcialmente en un ciclo de conferencias impartido en el Collège de France en marzo de 1983, y que más tarde desarrolló en el libro *El discurso filosófico de la modernidad* (1985). La polémica se desencadenó en Europa y Estados Unidos, y solo remitió cuando, ya en la década de 1990, comenzó a declinar la influencia del postestructuralismo. A partir de esa época, los tópicos de la obra habermasiana que más han interesado en Francia han sido los mismos que en cualquier otro país: la teoría de la democracia deliberativa, la filosofía del derecho, las reflexiones sobre Europa y, más recientemente, la función de la religión en las sociedades actuales.

Los temas de filosofía política predominan también en la recepción de Habermas en China, que expone la sinóloga australiana Gloria Davies en uno de los capítulos más interesantes del libro. Aparte de algunos detalles curiosos (por ejemplo, que en China se conoce a Habermas como *Habeimasi* o también como *Ha shi*, es decir: “Señor Ha”), la lectura china de Habermas es, en general, muy sorprendente. Desde la década de 1990 existen traducciones de

sus principales libros, porque —de manera similar a lo que sucedió en los últimos años de la República Democrática Alemana— la China de Deng Xiaoping necesitaba referentes ideológicos inscritos en una tradición genéricamente marxista pero que pudieran representar una alternativa al pensamiento de Stalin y Mao. Sin embargo, el uso que se hace de esta filosofía en el mundo académico chino a menudo se orienta hacia un objetivo tan ajeno a su espíritu como es la legitimación del régimen comunista. Los intelectuales chinos aprecian, por ejemplo, la crítica habermasiana de la racionalidad instrumental dominante en las sociedades capitalistas occidentales, pero no extienden esa crítica a la burocracia comunista ni a las formas específicamente chinas del capitalismo. Del mismo modo, estudian en profundidad las teorías de Habermas acerca de la racionalidad comunicativa, la ética del discurso o la democracia deliberativa, pero no extraen de estas teorías sus evidentes consecuencias prácticas, es decir: no las aplican para elaborar un pensamiento que contribuya a una politización democrática de la sociedad civil ni a la reforma democrática de las instituciones chinas. Más bien al contrario, la democracia deliberativa se interpreta según un modelo “*top-down*” que reserva la deliberación a las élites sociales o administrativas del país, en una inesperada reedición de aquel paradójico republicanismo sin democracia (e incluso sin República) que defendía Kant para la Prusia de finales del siglo XVIII. Ahora bien, pocas cosas hay menos habermasianas que una interpretación de la democracia deliberativa que defiende formalmente la deliberación pero no la democracia. Como son conscientes de estas tensiones, los intérpretes chinos de Habermas argumentan que, después de todo, la construcción de la modernidad en este país asiático sigue vías particulares

que un intelectual occidental no es capaz de comprender. Una visión ésta, tanto de China como de Habermas, que contrasta con la lectura que hacen de esta filosofía los intelectuales disidentes, quienes desde hace décadas ven en Habermas un referente para la defensa de los derechos humanos contra el discurso político oficial.

De la lectura de *Habermas global* cabe extraer algunas reflexiones. Una de ellas se refiere a los posibles destinatarios de este libro. Muchos de los datos que contiene sobre la recepción de Habermas en diferentes regiones del mundo son muy interesantes, pero en mi opinión no está del todo claro a quién puede ir dirigido el libro en su conjunto. A los especialistas pueden interesarles los detalles de la recepción del pensamiento habermasiano en el contexto cultural o nacional propio y en países que, como Estados Unidos o Francia, tienen tradiciones filosóficas que han influido de uno u otro modo en el pensamiento del filósofo. En cambio, difícilmente podrá interesar al especialista español la acogida de Habermas en, digamos, Bélgica o Noruega, como tampoco al especialista noruego le interesará especialmente la lectura española, italiana o japonesa de este filósofo. Cabría pensar que, pese a todo, el libro puede resultar útil como una especie de manual de consulta, puesto que reúne en *un* solo volumen las exposiciones de la recepción de la obra de Habermas en distintos países. Ahora bien, el libro solo podría cumplir eficazmente esa función si se hubiera publicado en inglés. Al haberse publicado en alemán, se reduce mucho el número de sus potenciales lectores y, curiosamente, se dificulta su proyección “global”, la única que tendría sentido para un volumen como éste.

Una segunda reflexión se refiere a la imagen predominante de la filosofía de Habermas que se desprende de la lectura de este

libro. Son impresionantes los datos con los que los editores abren el volumen, y que hemos mencionado también en las primeras líneas de esta reseña. Pero esos datos no son, por sí mismos, un indicador suficiente de que la obra de este autor vaya a sobrevivir a su contexto histórico y político natural (que se sitúa entre las últimas décadas del siglo XX y nuestro propio tiempo), ni aseguran que Habermas esté teniendo una influencia profunda en el pensamiento filosófico actual. El lector de *Habermas global* tiene más bien la impresión de que la recepción de este filósofo tiende a limitarse actualmente a su filosofía política, y en concreto a la teoría de la democracia, al tiempo que se olvidan otras vetas de sus textos, muy ricas pero más profundas, más complejas, y quizás no tan cómodamente accesibles como lo son las ideas políticas. Me refiero a la teoría de la racionalidad, a la filosofía del lenguaje, a la ética del discurso o a los análisis, a menudo magistrales, que Habermas ha hecho de otros filósofos y que pueden abrir vías fecundas para la lectura e interpretación de éstos. Sería una lástima que la vasta producción habermasiana quedase reducida a algunas ideas más bien inofensivas

sobre la Unión Europea o sobre la conveniencia de la deliberación en la política de las sociedades contemporáneas.

Pero quizás no está todo perdido para la recepción de esta filosofía en todo su potencial. A finales de 2019, ya con noventa años y en una pirueta que ha dejado asombrada y puede que incluso un poco descolocada a la comunidad habermasiana “global”, Habermas publicó una voluminosa obra en dos tomos titulada *Auch eine Geschichte der Philosophie* (Berlín: Suhrkamp, 2019). En ella repasa la historia de la filosofía occidental tomando como hilo conductor la relación entre fe y saber. Aún está pendiente la recepción de este inesperado y exigente fruto tardío, pero sería deseable que, sean cuales sean sus aportaciones, este último Habermas lograra sacudir algunas inercias de sus propios partidarios y reorientase el estudio de su obra hacia problemas más netamente filosóficos y hacia los estratos más fundamentales, y hoy cada más olvidados, de su filosofía.

José Luis López de Lizaga
(Universidad de Zaragoza)

HEGEL, G. W. F. (2019). *Lecciones sobre la filosofía del espíritu subjetivo II. Antropología*. Traducción Alberto Ciria y Juan J. Padial, notas y estudio preliminar por Juan J. Padial, Sevilla: Thémata, 540 pp.

En septiembre de 2019 vio la luz *Leciones del espíritu subjetivo II*, el segundo volumen traducido de los cursos impartidos por G.W.F. Hegel en la Universidad de Berlín en la década de 1820. En este segundo volumen se recogen los apuntes de Heinrich Gustav Hotho en 1822, Karl Gustav von Griesheim en 1825, y de Stolzenberg entre 1827 y 1828. Los tres fueron alumnos de Hegel y asistieron a sus clases, anotando

las explicaciones pertinentes que el filósofo alemán daba acerca de la *Antropología* —la antropología es en el sistema hegeliano la primera ciencia del espíritu, es decir, lo que en el paradigma de la actualidad psicológica denominamos «preconsciente»—.

Para poder desarrollar correctamente la descripción de esta nueva obra, primero trataremos de mostrar la relevancia de tener una traducción al español de esta obra.

Luego pasaremos a la exposición de los traductores y a comentar el estudio preliminar y las notas. Finalmente expondremos el propio contenido de la obra centrándonos en las partes novedosas en las que este volumen puede ser de extraordinaria relevancia.

Este volumen es una obra que cualquier investigador e interesado en la filosofía de Hegel debería utilizar por varios motivos. El primer motivo es que este conjunto de lecciones solamente estaba disponible hasta ahora en alemán y en inglés. Por eso, que se haya elaborado una traducción al español ya es un gran avance para los países de habla hispana. Además, la traducción ha sido realizada por Alberto Ciria y Juan José Padiá, lo que constituye una garantía asegurada. Alberto Ciria, el reconocido académico español, ha traducido en anteriores ocasiones obras de Fichte, Heidegger y Lauth. Un experto en la traducción del alemán-español de su nivel ha decidido en esta ocasión traducir a Hegel, uno de los filósofos alemanes más potentes de toda la historia del pensamiento occidental. Por otra parte, encontramos la figura de Juan José Padiá, partícipe no solo en la traducción, sino que ha elaborado en este volumen el estudio preliminar y las notas. Padiá es profesor de la Universidad de Málaga y un reconocido experto español en la filosofía de Hegel, siendo además uno de los miembros más activos del Grupo de investigación sobre el Idealismo alemán.

Habiendo presentado a los dos encargados de haber traído las *Lecciones del espíritu subjetivo II* al lenguaje hispano, huelga decir que el estudio preliminar de la obra es bastante detallado. Este estudio permite al lector interesado en la filosofía del espíritu subjetivo de Hegel sumergirse en las ideas clave para la correcta comprensión de lo que expone este volumen. Así,

la nota preliminar presenta al inicio una breve exposición comparada con lecciones impartidas anteriormente por el propio filósofo para que el lector tenga un bagaje de la obra de Hegel suficiente para entender en qué momento se dictan estas lecciones y su contenido.

Posteriormente pasa a exponer con la terminología más actual las ideas de Hegel, de manera en que el lector pueda sentirse cómodo en su lectura y entienda a qué se refiere exactamente cada concepto. Es a destacar la enmarcación sobre las posibles críticas y similitudes hacia otros autores en el propio estudio preliminar. Véase como ejemplo la rigurosidad con la que Padiá expone al lector los puntos de unión de Hegel con Kant: «Según Hegel, es un acierto de Kant el no pensar al yo como objeto, y denunciar con ello los paralogismos», (p. 19). A su vez, el estudio preliminar también servirá de contextualización, pues Padiá guía al lector incluso en la terminología científica y filosófica de la época: «Esto plantearía todas las dificultades desde la metafísica del alma, la Pneumatología» (p. 23). Finalmente, la nota preliminar dedica la mitad de su extensión a exponer correctamente el concepto de «antropología» en la obra de Hegel y su relación con la exposición que realiza en la *Enciclopedia de las ciencias filosóficas*. De este modo expone las tres partes del alma: el alma natural; el alma que siente; el alma efectivamente real. Estas tres nociones son explicadas con una precisión elogiable por parte de Padiá.

Aparte de la labor de traducción y el estudio preliminar, son muy a destacar las notas. Todas las notas tienen el objetivo de clarificar y contextualizar. De esta manera, encontramos que notas como la 372 (p. 521) explicita las diferentes acepciones de la palabra «sentir» para poder introducir

al lector hispano de una manera precisa en un tema tan complejo como el «alma sintiente». Incluso asume posibles correcciones en los apuntes de los tres alumnos de Hegel. Por ejemplo, la nota 384, en la que encontramos que «Parece que hay un error en la transcripción. Hegel debe estar refiriéndose a Jenofonte y no a Jenófanes. Sobre Jenofonte, véase la nota 18» (p. 522). Es decir, el volumen no solo se limita a ser una buena traducción acompañada de un estudio preliminar riguroso, sino que las propias notas son un material rico para trabajar por parte de los académicos, a la par que constituye una guía altamente recomendable para el lector interesado y no especializado.

En conclusión, se podría decir que el estudio preliminar y las notas del presente volumen constituyen no solo un claro ejemplo de la maestría en Hegel por parte de Padial, sino que permiten acercar la obra al lector común sin perder la calidad y rigurosidad de la filosofía. Conviene decir que, debido a su trayectoria, Ciria y Padial, son el símbolo de la seguridad y rigurosidad con la que tratar una traducción. Al haberse encargado estos dos académicos españoles no hay duda de la fiabilidad de la traducción, lo que tiene como consecuencia el afianzamiento con el que se puede leer y citar el texto, pues una mala traducción puede incluso llegar a dar una visión filosófica totalmente distinta de la que el propio autor quería dar. Sin embargo, en este caso la fiabilidad es máxima.

El contenido de la obra es muy destacable. En ella podemos observar cómo el propio Hegel va madurando su pensamiento paulatinamente en las diferentes clases que impartió y que viene recogidas en este nuevo volumen. Como se sabe, Hegel redactó su *Enciclopedia de las ciencias filosóficas* para sus clases en la

Universidad de Berlín. Esto quiere decir que lo que expone en la *Enciclopedia* era desarrollado en sus clases. Por esta razón la traducción de los apuntes completos de su alumnado es, en definitiva, un elemento decisivo para poder comprender verdaderamente al autor. Esto nos permite seguir el transcurso de la propia *Enciclopedia* en sus tres ediciones. Los apuntes de 1822 y los de 1825 se corresponden con la primera edición de la *Enciclopedia*, en 1817. Sin embargo, hay cambio entre ellos, razón por la que aparecería la segunda edición de la *Enciclopedia* en 1827. Por este motivo gracias a *Lecciones sobre la filosofía del espíritu subjetivo II* podemos observar el proceso de maduración del pensamiento de Hegel sin limitarnos a comparar la primera y la segunda edición, sino atendiendo al proceso de reflexión que se estaban llevando a cabo por parte de Hegel en sus clases. Finalmente, los apuntes del periodo de 1827-1828 ya atienden a la segunda edición de la *Enciclopedia*.

Sumergiéndonos en el propio contenido del volumen, Hegel tiene ciertas temáticas sobre las que, como se puede observar gracias a estos apuntes, reflexiona durante años. Podremos a continuación dos ejemplos: las edades del ser humano, y el desarrollo del propio concepto de *antropología*. Yendo al primer punto, el ser humano tendría de acuerdo con las lecciones de 1822 un primer periodo desde su niñez a su adolescencia. El discernimiento entre ambas etapas viene marcado por los momentos vitales que se manifiestan en la actitud. Así, recogen los apuntes de 1822: «En el muchacho, la oposición contra el mundo aún no tiene verdad, sino que es solo representada. El muchacho juega y aprende. La culminación de la edad juvenil es el adolescente. Aquí comienza la seriedad, la oposición del sujeto contra un mundo que aparece

en su aspereza» (p. 117). Incluso Hegel trata brevemente sobre el carácter marcado de la juventud frente al periodo maduro del adulto: «Así pues, el hombre adulto tiene la relación correcta con el mundo como es. La juventud, por el contrario, es revolucionaria» (p. 118). Finalmente, se puede encontrar en esta misma línea que el anciano es el hombre que ha perdido el interés, y es justamente esto lo que le hace envejecer. No de manera biológica, sino espiritualmente por haberse acostumbrado al mundo. El anciano es aquel al que ya nada le sorprende. Por eso Hegel señala que: «Esta habituación constituye el tránsito a la ancianidad (...). Es decir, es la pérdida de intereses lo que hace envejecer al hombre. Para él no hay nada nuevo, todo es recuerdo, lo individual ya no tiene para él ningún interés particular, se atiene sólo a lo general» (p. 119). Todas estas explicaciones sobre las edades del hombre pueden tener una enorme conexión con la propia vida de Hegel. Y sin duda, si se realizase este tipo de estudio, *Lecciones sobre la filosofía del espíritu subjetivo II* es lo que el cónclave a la bóveda.

Incluso en las concepciones sobre cómo Hegel iba considerando la *antropología* este volumen es iluminador. Por ejemplo, en 1822 no encontramos una definición clara al inicio de los apuntes sobre qué es la antropología. Directamente trata el alma. En contraste, en el semestre de verano de 1825, encontramos en los apuntes de Griesheim una concepción muy amplia de la antropología. Tras exponer la relación del gesto con la comunicación, comenta Hegel

que «todas estas cosas que atañen a la relación del espíritu con el cuerpo y del cuerpo con el espíritu entran aquí, en la antropología» (p. 179). Esto es un claro avance en la profundidad del pensamiento de Hegel. El lector puede contemplar cómo se va refinando. Finalmente, en los apuntes de 1827 y 1828, encontramos que la *antropología* es ni más ni menos que lo primero que Hegel define. Y es prácticamente seguro que el curso comenzase así, a diferencia del resto, porque los apuntes de Stolzenberg coinciden en el comienzo con los de Walter como nos sugieren sutilmente las notas a pie de página. Aquí encontramos la *antropología* ya muy bien definida y asentada: «En la Antropología el espíritu se encuentra en su determinación [o condición] natural. Esta condición natural, [el alma], es fundamento o base para el todo» (p. 339). Por este motivo, debe tomarse en consideración la evolución del pensamiento de Hegel que está refrendado por estas lecciones.

A modo de conclusión, se debe decir que es un volumen de una calidad excelente, y que, además, es iluminador para cualquier investigador. En esta reseña solamente se han tratado de manera abreviada dos ejemplos posibles, pero conceptos tan sumamente relevantes como el alma en su individualidad, el alma como totalidad y el alma efectivamente real son otros de los muchos temas que pueden investigarse gracias a este nuevo volumen.

Andrés Ortigosa
(Universidad de Sevilla)

CAVALLETTI, A. (2019). *Vértigo. La tentación de la identidad*. Traducción de María Teresa D'Meza. Buenos Aires: Adriana Hidalgo editora.

Andrea Cavalletti presenta un texto que busca poner en evidencia la forma de la subjetividad vertiginosa de la máquina securitaria-biopolítica en curso, desde el cine y la fenomenología. Si bien retoma la temática de la sugestión y la seguridad que había elaborado en textos anteriores, se trata de un giro en la obra del autor, ya que el ensayo está centrado en el pensamiento de Husserl y la discusión de sus textos en la tradición fenomenológica en autores como Robert Klein, Stein, Sartre, Heidegger, Merleau-Ponty, Melandri, Derrida, etc. y no en las ideas de Furio Jesi o Walter Benjamin, aunque su impronta está presente. Husserl comparte el protagonismo del libro con la película *Vértigo* de Hitchcock, a través de la cual busca establecer paralelos entre la subjetividad y el dispositivo del cine, así como el carácter técnico y artificial de la ontología y la política. El libro está formado por un breve proemio, 6 capítulos, y un epílogo.

El sucinto proemio establece la problemática del vértigo como técnica gubernamental desde Vigilius Haufniensis, Hobbes, Antonio Genovesi. La amenaza de muerte violenta propia del poder de vida y de muerte, de la soberanía moderna, podía resumirse en que morimos por nuestra seguridad. La palabra vértigo nombra en De Giuliani en 1790 la tendencia “a perseguir lo que más se teme o a realizar el gesto que no se debería realizar a ceder resistiendo o a llamar acción a la pasividad” (p. 10). Es así que aquello que logra exponer la crítica de los valores nietzscheana es al ser humano como vértigo, en la que sin dudas este ensayo se inscribe.

El primer capítulo llamado “Efecto-*vértigo*” expone la signatura del vértigo entre el cine, su prehistoria, lo social, y el análisis moderno de tradición kantiana, filosófico y medicinal, así como las más diversas definiciones de vértigo, desde el psicoanálisis, lo habitual y no la voluntad, la embriaguez, la danza, la crítica al socialismo o incluso el capital financiero según Marx.

Quien lleva al vértigo a un estatuto de “impureza trascendental” es Marcus Herz, a quien retorna a lo largo del libro como alternativa a Husserl. Según Ludwig Ernst Borowski, a pesar de que Herz le dedicó su investigación sobre el vértigo (*Versuch über den Schwindel*), Kant afirmando que era “inmune al vértigo” lo devolvió de inmediato; actuar que según Peter Fenves se habría debido a una *ambivalence toward the Jew-Palestinians* (p. 31). La idea propia de una medicina filosófica de Herz es que para cada persona y en cada situación se da una medida de *Weile*, una medida interna del desarrollo de las representaciones, dependiente del orden del alma y la economía corpórea. Es así que el vértigo es un aceleración de la *Weile* hasta la confusión, produciendo una excentricidad del tiempo y del espacio, un “vértigo del tiempo” en palabras de Purkinje, poniendo en tensión así a la estructura a priori kantiana, llevando la psicología al reino de la estética trascendental, “una impureza de lo trascendental” (p. 36). Clave para la articulación de la enorme variedad de conceptos y problemas es el vínculo que establece Cavalletti entre el cine como *délire d'une machine* y la *machine sociale* (p. 58) caracterizada por Simone Weil (*Réflexions*

sur les causes de la liberté et l'oppression sociale, 1934). Es así que Weil escribió que es “una máquina que fábrica inconsciencia, estupidez, corrupción, apatía y, sobre todo, vértigo” (p. 130), incluso “arrastrado a la ruina” según Roger Callois.

El capítulo 2 “No estamos aquí” es sin dudas el capítulo más extenso del libro, y está dedicado a la noción husserliana de *habitus-Habe*, a la *Einfühlung*, y sobre todo, a lo inapropiable que retoma tácitamente la discusión con la fenomenología de su maestro Agamben en *L'uso dei corpi*. Husserl en *Ideen II* y *Cartesianische Meditationen* propone que “el ego puro, polo del flujo de conciencia, se apropia del mundo en otros términos mientras se forma como sujeto concreto (o mónada) contrayendo tanto pasiva como activamente los *habitus*, que luego se vuelven apreciables en el plano empírico como disposiciones y determinaciones del carácter.” (pp. 74-75)

Cavalletti señala críticamente que el concepto de *habitus*, proveniente de la *héxis* aristotélica “es pues consignado y reducido al principio posesivo de la identidad: la temática husserliana vincula el haber con el ser, transforma a uno en otro, restringe el horizonte temporal de lo posible a la relación entre la determinación y la libertad, entre la pura elección y la constricción, es decir, a las propiedades permanentes del sujeto personal, centro de todas las apropiaciones.” (p. 76). Eventualmente, el hábito es interpretado por Cavalletti a partir de lo inapropiable, la intersubjetividad y la comunidad, como por ejemplo en el caso de Fink, a partir de una intersubjetividad en ausencia, a través de la historia, “excéntrica y vertiginosa” (p. 157).

El siguiente capítulo titulado “Hábito, máscara”, comienza con un análisis de la escena en la que Scotti contempla a Madeleine, o más bien Judy haciéndose pasar

por Madeleine en el museo, observando el cuadro de Carlotta Valdés; el ramo de flores y el peinado de Carlotta son idénticos a los de Madeleine; “*se ve a sí misma con los ojos de la muerta*”. (p. 159); un juego de máscaras entre la vida y la muerte. Luego continúa problematizando la escisión a partir de la fenomenología, la escisión del ego trascendental, que Sartre había intentado solucionar a partir de que el ego sea la muerte de la conciencia, deviniendo así una figura pasiva; que Adorno había problematizado sosteniendo la diferencia entre el ego absoluto y el ego fáctico. Retoma la perspectiva deleuziana según la cual Hitchcock con sus “imágenes mentales” da cuenta de un vértigo, de una duración ya no subjetiva ni interior, sino impersonal, dando así con la fórmula del desinterés más puro, separando el vértigo y la identidad del sujeto. Eventualmente, profundiza en la crítica de Derrida y de Melandri hacia la identificación entre la alteridad y lo originario, así como la negación y obliteración de la diferencia, para proponer con Foucault que la tensión concreta es aquella entre la actualidad que nos limita y constituye, y la descripción del archivo que nos permitiría dar cuenta de las discontinuidades y las diferencias, dejando a la fenomenología en favor de la arqueología.

El capítulo 4 “Un singular arrebatado” profundiza la idea de vértigo como “aquí que es un allí” desde Jankélévitch, así como la muerte del otro como condición de la intersubjetividad y la necrofilia de Scottie. El ser humano es para Jankélévitch un aprendiz de brujo, dueño de sus decisiones pero de consecuencias incalculables, irreversibles y vertiginosas: “Un segundo para decir sí, toda una vida para arrepentirse” (p. 210). Tal inestabilidad y escisión otorga un carácter onírico a la vivencia desde la actualidad de las vivencias. Aún más el vér-

tigo como descentramiento y vacío es previo al individuo en cuanto muerto: “antes es la intersubjetividad que constituye al sujeto como sujeto personal e idéntico a sí, es la muerte del otro que presenta el yo a sí mismo, y sólo en la muerte el otro se presenta al yo, participa en su constitución: lo que Fink llama la “habitualidad del idéntico ser-yo” se da exactamente con la muerte del otro” (p. 219). Mary Pickford lloró al verse por primera vez en la pantalla, al no saber ser Mary Pickford, aquella identidad de la cual podían dar fe millones. Vértigo cinematográfico que anula la identidad propia, Mary Pickford se ve desde otra, así como Husserl al confundir el maniquí con un *alter ego* se había reconocido en el análogo sin vida, ya muerto.

El capítulo 5 “Abismo” está dedicado al *Schwindel* (vértigo) heideggeriano de la pregunta por el ser, al abismo (*Kluft*) entre el ser y el ente, entre el tiempo y el espacio, entre el moribundo (*Sterben*) y el muerto (*exitus*); además, recuperando las notas críticas de Husserl a *Sein und Zeit*, así como las cercanías entre ambos, y las diversas lecturas posibles, la épocal de Arendt, aquella anárquica de Reiner Schürmann, como la fascista de la paratranscendencia de Oskar Becker.

El capítulo 6 “Superficie” establece el carácter técnico-artístico de la subjetividad, la política, y el cine, es decir, estudia la superficie sin profundidad, la máscara sin rostro, teniendo siempre en cuenta el lazo entre técnica y política la tesis benjaminiana: “la tendencia justa en la política es una tendencia justa de la técnica” (p. 278). El vértigo es presentado ahora como un existencial, y así, *Sein und Zeit* es releído como autoexhibición técnica del ser-ahí vertiginoso: “Cada acto del sujeto idéntico a sí deriva del miedo a la muerte, procede de la conclusión que teme y persigue, es el

acto sugerido por la muerte y expuesto a la muerte. [...] Esta relación es en efecto la relación con el otro, y hace comparecer al moribundo en el “aquí” que es un “allí” y, antes aún, desde el “no-aquí” que habita míticamente en el “allí” del muerto, ligando la muerte (de ahora) con la vida (de antes) como ausencia de la vida, y haciendo así del “allí” un “aquí”. Esta corresponde a la técnica violenta que dispone de los seres, produce y domina todas las cosas y la “mundidad” misma del mundo.” (p. 266)

Es la idea de superficie absoluta la que interrumpe y excede a la subjetividad vertiginosa. Retomando al teórico del arte Sergio Bettini, quien reconoció en Roma y Venecia un colorismo puramente óptico, desustancializado, ni lejano ni cercano. Es así que es imposible la subjetividad vertiginosa del *Dasein*, que habita a partir del *Ent-fernung*, del des-alejamiento, de un “allí” a partir del cual comprender su “aquí”. El ensayo filosófico es otra estrategia contra el encantamiento mítico, contra la sugestión vertiginosa: “Quizás en este punto podemos definir el ensayo filosófico, más que como una forma, como una teoría de la maquinación. Esta traiciona la efectividad, impide que el vértigo actúe en nosotros constituyéndonos como sujetos, a fin de que lo que está en nosotros sea, en cambio, hecho por nosotros.” (pp. 283-284)

El epílogo presenta una relectura de la dialéctica hegeliana del amo y del esclavo desde la perspectiva desarrollada a lo largo del libro, en el marco del “hedonismo político” hobbesiano siguiendo a Leo Strauss, que reduce el placer a la ausencia de dolor. Sugiere finalmente, con Furio Jesi y Nietzsche, que la única técnica que no disfraz el “no-hay” del poder que amenaza de muerte al esclavo, es aquella artística que no teme al ojo capaz de ver su pequeño fraude.

En suma, se trata de un ensayo sumamente erudito que tiene como logro tematizar la subjetividad como vértigo, enlazando el problema desde la estética del cine, la fenomenología, y la máquina securitaria. Quizás haya una alternativa benjaminiana más al terror vertiginoso de la amenaza soberana. Tal vez el alegorista del *Origen del Trauerspiel* alemán —que va de ruina

en ruina y que tiene a la muerte como su límite, pero que niega la nada en ellos dando así un salto a la resurrección— sea una estrategia por indagar dentro de la crítica de la soberanía.

Juan Cruz Aponiuk
(Universidad de Buenos Aires)

GONZÁLEZ BLANCO, Azucena (ed.) (2019). *Política y literatura. Nuevas perspectivas teóricas*. Berlín: De Gruyter.

Esta obra colectiva editada por Azucena González Blanco analiza las relaciones entre literatura y política desde nuevas perspectivas contemporáneas. El volumen se abre con el texto “Política de la ficción” de Rancière, autor de referencia del *giro político* de la literatura, y a cuyo pensamiento se dedica la primera parte del libro. En segundo lugar se analizan distintas estéticas del desacuerdo y finalmente se problematiza la política del teatro.

En “Política de la ficción” Rancière propone una definición del concepto de ficción no opuesto al de realidad, sino como una “estructura de racionalidad”, es decir, como un “modo de relación que construye formas de coexistencia, sucesión y encadenamiento” que “confiere a estas formas la modalidad de lo posible, de lo real o de lo necesario”. En consecuencia, cuando se produce un cierto sentido de realidad se pone en obra la operación ficcional: toda experiencia de realidad comporta previamente un trabajo de la ficción. Por tanto, así como los novelistas hacen uso de la ficción para crear mundos, también “la acción política que identifica situaciones y señala sus actores [...] hace uso de ficciones”. Rancière llama a las distintas formas en las que las ficciones articu-

lan lo perceptible “régimenes ficcionales” y su estudio comparado compete a la política ficcional. Los presupuestos que subyacen bajo este planteamiento son los mismos que le permitieron el alumbramiento de sus tesis sobre la política de la literatura: el rechazo de la concepción aristotélica de mimesis y la sustitución de la *veritas* clásica por una *aletheia* que, al adoptar el principio de igualdad como a priori, toma un carácter político (González Blanco, 2019, 742-743).

Desde este planteamiento aborda dos modelos de interpretación política de la ficción, los propuestos por Lukács y por Auerbach. En el primer caso, Rancière observa la convergencia entre la crítica marxista de Lukács al naturalismo de Zola y la de los críticos reaccionarios de su tiempo. Para el filósofo húngaro, Zola simplemente describe cuadros, mientras que en el naturalismo correctamente interpretado por Tolstoi y Balzac se asiste a la narración de un drama en el que un héroe se enfrenta a las relaciones sociales existentes. Por su parte, los críticos reaccionarios acusaban a Zola de propiciar la invasión de la democracia que otorgaba igual valor a cualquier individuo. Rancière sostiene que esto se debe a que ambas críticas comparten la idea de ficción de rai-

gambre aristotélica: la poesía es drama y es cuestión de hombres activos. Esta implica un régimen de legitimidad ficcional fundado sobre una jerarquía de tiempos —los que viven en el mundo de la acción y pueden proyectarse fines, y los que despliegan su vida en el espacio de la reproducción de la vida cotidiana— que entraña a su vez una jerarquía de formas de vida —hombres activos frente a hombres pasivos—. En cambio, Rancière cree que Auerbach sí entrevé la potencia de la cotidianidad en su lectura de *Al faro* de Woolf. En esta novela, el crítico ve que la fragmentación subjetiva provoca la revalorización del instante cualquiera, de la riqueza de la vida ordinaria.

No obstante, ninguno atiende a la que para el autor es la gran enseñanza de las novelas realistas: la falta de cohesión entre el saber y la acción, que a su juicio implica la conquista política de la experiencia igualitaria del tiempo y, con ella, de las jerarquías de formas de vida. Rancière cree que en esta lección de la literatura se encuentra su núcleo emancipador: los intentos de seres anónimos de romper su inserción en un tiempo ilegítimo para vivir otra vida. En este sentido, dice Rancière, la literatura es “una ficción de igualdad” (Rancière, 2011, 123)

Judith Revel, en su capítulo “La invención y el *déjà-là* del mundo”, realiza un estudio de dos casos en los que la literatura tuvo una función reveladora para la filosofía, no un instrumento a su servicio. La autora rastrea este “trabajo de la literatura” en las obras de Foucault y Merleau-Ponty. Desde este punto de vista, Revel sostiene que en la investigación sobre Raimond Russel de los años 60, Foucault descubre la posibilidad de un lenguaje subversivo con respecto al orden del discurso pero que se produce desde su interior, lo que constituye uno de los principales motores de su obra. Aquí encuentra la posibilidad de una

práctica intransitiva de la libertad desde el interior de la configuración histórica —o de las relaciones de poder—. Por su parte, Merleau-Ponty se plantea la cuestión de cómo es posible la creación en el interior de la historia acumulada de las cosas dichas en “Le langage indirect et les voix du silence”. Aquí afirma que la novedad del lenguaje literario —o de la práctica artística— no se encuentra en los aspectos radicalmente nuevos, sino en la experimentación de nuevas formas de relación; se trata de confrontar la obra de Mallarmé con otras para hacerla brillar de nuevo. El arte en general y la literatura en particular —concluye Revel— enseñan que la creación se produce desde el interior de las palabras y estructuras disponibles, lo que a su juicio tiene grandes consecuencias políticas: una nueva comprensión de la revolución como “desequilibrio creador”.

En “Parresía y disidencia: veridicción como política de la literatura”, Blanco González aborda la herencia que la literatura moderna recibe de la parresía, especialmente sus consecuencias políticas en tanto que discurso disidente, a través de las obras de Foucault y Rancière. Frente a la poética ficcional y representativa, se trata de un arte específico de escritura que entra en conflicto con la estabilidad del cuerpo social gracias a su carácter contradictorio: una palabra sin fundamento que, por ello mismo, pone bajo sospecha toda fundamentación pretendida. Esto le permite a la autora señalar también las diferencias entre ambos autores: mientras que para Foucault la verdad del lenguaje tiene que ver con un poder-decir con consecuencias performativas, para Rancière, la verdad de la literatura es un puro “dar a ver” al servicio del principio de igualdad de una democracia radical.

Miguel Corella dedica su capítulo “Metáforas de la política” a un análisis de las imágenes y las metáforas utilizadas para

iluminar los nuevos movimientos políticos y, en especial, el concepto de multitud, en aras de dilucidar sus implicaciones filosóficas y políticas. El autor subraya la necesidad de no otorgarles un papel ontológico, ya que se corre el peligro de introducir elementos teleológicos o evolucionistas que encubrirían el vector voluntarista que no puede soslayarse: la decisión de actuar.

A partir de una lectura crítica de las ideas nucleares de la filosofía de Rancière, Barroso Fernández propone en “El humanismo como condición de la igualdad. A propósito de La distinción entre política y policía en Rancière” la necesidad de recuperar los ideales de humanidad para hacer posible una política emancipadora. A su juicio, el hecho de que la emancipación consista en la realización del principio de igualdad, que para Rancière implica la eliminación de la diferencia entre los que saben y los que no, expulsa a la verdad de la política. ¿Cómo es posible entonces —piensa el autor— encontrar un criterio desde el cual modificar las representaciones de la naturaleza humana o los hábitos a ella aparejados? Frente a una propuesta que conduce al nihilismo y a la incapacidad de tornarse hegemónica, Barroso defiende la necesidad de recurrir a ideales de humanidad para que el principio de igualdad no se esfume; ideas que no nieguen el carácter acontecimental de la realidad, sino que lo acompañen, y que requieran el trabajo activo de la multitud. De aquí, finalmente, su apuesta por un populismo virtuoso que exprese el encuentro entre el pueblo y los intelectuales en sentido gramsciano.

De la Higuera estudia en “Negatividad y experiencia del pensamiento” el anclaje del pensamiento de Rancière en el marco de la constatación metafísica contemporánea según la cual el mundo no es más que su propia interpretación. La filosofía se con-

vierte en la práctica de la negatividad que denuncia la arbitrariedad última de todo orden, lo que opera una transformación en la cuestión política-literatura. En la literatura se asiste al conflicto con lo presente visible que suspende las formas individuales y que inventa formas pre-individuales, poniendo bajo sospecha cualquier relación del sí mismo con una identidad determinada, lo que implica un alcance ontológico. En cambio, la experiencia política es la de la creación de un sujeto colectivo nuevo pero en el orden de lo visible. Por tanto, a la vez que la ausencia de lo incondicionado posibilita la transformación del mundo mediante su interpretación infinita, la política adquiere un carácter literario: solo la ficción puede unir lo dado y lo no-dado.

El texto de Juan Carlos Rodríguez “Tres estallidos en el horizonte literario de la modernidad y la posmodernidad” está dedicado a la denuncia de tres contradicciones de la ideología capitalista que el autor resume en la idea “yo-soy-libre porque he nacido libre por naturaleza”. En primer lugar, en relación con el “misterio de la poesía”, el intento de Jakobson de proporcionar una definición de la poesía cifrada en el “en sí” del lenguaje se encuentra inserta en el mismo paradigma, en la medida en que esta funciona como la plasmación de la pureza de un alma de ese yo-soy-libre más puro. La segunda contradicción toma cuerpo en *Tristram Shandy* de Laurence Sterne que ataca el corazón de ese “yo libre” lockeano, puesto que demuestra que nacemos en un mundo que nos precede y que se nos impone. Finalmente, la tercera contradicción literaria y artística es la invisibilidad de la mujer, mostrada magistralmente por Hamilton o Godard, visible como cuerpo desnudo disponible pero invisible como persona.

Miguel Ángel García devuelve la crítica de Rancière a Althusser a la luz de las últi-

mas publicaciones del autor en “¿Política y literatura? La lección de Althusser”. En opinión del autor, la política de la literatura que propone Rancière sí es un “discurso del orden en el léxico de la subversión”, en la medida en que la política es definida más allá de la lucha de clases. Además, bajo la supuesta transformación del reparto de lo sensible se encuentra la ideología burguesa del sujeto que reclama el conocimiento de sus derechos individuales.

En “Renuncia y proliferación: sobre los cuerpos inéditos de un libro de aforismos” Erika Martínez propone un estudio del aforismo contemporáneo como “pensamiento de lo impensado”. Frente a la creciente “institucionalización de lo políticamente correcto”, el aforismo muestra la posibilidad incesante del desacuerdo a través de una relación de discordancia con la verdad. Fraccionamiento que se redobra en la incapacidad de obtener un conocimiento completo de un libro de aforismos y que se multiplica si tenemos en cuenta que habitualmente han sido compuestos y ordenados por editores. En esta medida, es un lugar privilegiado para el estudio de la relación de heterogéneos, o la “fraternidad de la metáfora” que diría Rancière, y de la función de autor.

Fischer-Lichte plantea el “entretejimiento cultural” como matriz de comprensión que permita comprender las nuevas relaciones entre culturas escénicas en la actualidad en el capítulo titulado “Entretejimiento de culturas escénicas: re-pensando el «teatro intercultural»”. En opinión de la autora, ni “intercultural” ni “poscolonial” son operativos para describir la situación inédita de la actualidad, debido a los presupuestos occidentalistas en el primer caso y binarios o raciales en el segundo. En el entretejimiento, las hebras se unen formando un hilo y muchos hilos se juntan para formar

un tejido, sin que sea reconocible el origen ni necesaria la formación final de un todo. Este proceso dinámico y siempre abierto permite atisbar la dimensión utópica del arte: puede dar a conocer nuevas formas de vida hasta el momento desconocidas y potencialmente posibles en una sociedad globalizada.

Rábade Villar recupera en “Un análisis del giro afectivo en el teatro contemporáneo” la crítica del giro afectivo (y su convergencia con el político) del teatro en el “El espectador emancipado” de Rancière. Como es sabido, el autor francés rechaza la comprensión platónica de la visión como pura pasividad y critica la mimesis como operación que fundamenta el quehacer artístico teorizada por Aristóteles. Y desde aquí, a diferencia de la mayor parte de los estudios teatrales actuales sobre la performance, subraya la necesidad de la mediación para no caer en la fantasía de la presencia inmediata. La “inmediatez ética” impide la subjetivación política, que solo es posible gracias a un trabajo de “traducción poética”; exclusivamente desde la divergencia entre el sentimiento y la expresión es posible la igualdad. Esto se debe a que, en el fondo, el paradigma mimético descansa sobre “la ley de las causas y los efectos” según la cual los afectos se identifican con efectos que la obra busca producir. La emancipación no es posible ni en la mediación representativa ni en la inmediatez ética, solo en creaciones que nos impongan la necesidad de ser traducidas. El sentir, defiende la autora con Rancière, no es un punto de llegada sino de partida. De esta forma, la razón de que la estética del autor se fundamente en el percibir es que la observación siempre es volitiva, el *contemplan hace*.

En cambio, en “La estética teatral contemporánea como fórmula posmoderna de revolución: política y posdramaticidad”, López Silva discute las tesis de Rancière sobre la performance recurriendo a los pre-

supuestos del teatro posdramático. En este, la mirada o la escucha no están vinculadas con la pasividad, ni el espectador es considerado desigual. El espectador y el artista tienen roles distintos en la comunicación, pero el primero es activo en tanto que su presencia en la creación siempre es activa. El problema, a juicio de la autora, es que Rancière interpreta la relación con el espectador en la performance en términos del paradigma dramático, creyendo que la transformación solo ha afectado al código. No obstante, la ruptura con la mimesis ha roto la distancia entre el teatro y el público en la medida en que este es libre de interpretar y debe organizar su propio proceso de comprensión. Además, la participación del espectador condiciona la propia obra artística, por lo que no está relegado al mero aprendizaje de lo que la performance enseña. En definitiva, para la autora, el posdrama debería ser considerado

como una fórmula democratizadora capaz de subjetivación política.

El público encontrará en este libro plural un estudio profundo de las relaciones entre la política y la literatura a través de una discusión intensa y exhaustiva con las últimas perspectivas del pensamiento contemporáneo.

Bibliografía

- González Blanco, A. (2019). «Política de La Ficción / Ficción de La Política En Jacques Rancière», *Signa*, 28, pp. 733–46, doi:<https://doi.org/10.5944/signa.vol28.2019.25079>.
- Rancière, J. (2011). *Política de la literatura*. Buenos Aires: Zorzal.

Beltrán Jiménez Villar
(Universidad de Granada)

WITTGENSTEIN, Ludwig (2019). *Correspondencia con Rudolf Koder*. Traducción de Venancio Andreu Baldó e Isabel Gamero Cabrera. Madrid: Editorial Ápeiron.

La apuesta de Ápeiron Ediciones esta vez, se centra en el autor vienés Ludwig Wittgenstein. Este filósofo analítico, después de arrojar luz sobre los límites del lenguaje, la lógica, filosofía de las matemáticas y filosofía de la psicología, entre otras temáticas de las que se hace cargo directamente, y otras de las que se puede extender del pensamiento del austriaco. Esta vez la apuesta de la editorial se centra en matizar el filosofar wittgensteiniano en un tema tan problemático de pensar para la filosofía como es el de la «música», que indudablemente ocupaba un papel nuclear en la vida del autor. Este análisis, más que atomizar o taxonomizar, intenta matizar, ayuda a comprender mejor el pensamiento del vienés, su forma de vida y su

relación personal. Los recursos tratados para la elaboración de este trabajo, se localizan centralmente en el material recogido de la correspondencia que tiene durante un largo periodo de la vida de Ludwig con su amigo Rudolf Koder (a pesar de las interrupciones a causa de las guerras y demás dificultades). Esta publicación traducida al español por Venancio Andreu Baldó e Isabel Gamero Cabrera es muy útil para profundizar e introducirte en la concepción musical, la vida y la familia de los Wittgenstein. La estructura del manuscrito comienza con un prólogo y una breve introducción sobre el filósofo matemático y su amigo Rudolf Koder, continúa con la correspondencia entre ambos y se adentra en una segunda parte del libro que

contiene unos ensayos y anexos escritos por Martin Alber, titulados: “Josef Labor y la familia Wittgenstein” y “Sobre los aspectos de lo musical en la vida y la obra de Ludwig Wittgenstein”. En una tercera y última parte del libro nos aporta más información específica como el informe editorial, bibliografía e índice cronológico de las cartas.

Ludwig Wittgenstein tocaba el clarinete, aunque no tenía una formación académica sí que practicaba y era muy meticuloso y lo tocaba muy en serio, es decir, no entendía la música como algo decorativo, sino como un alimento vital, al que en ocasiones le molestaba si creía que no se la trataba con el suficiente respeto. Tocaba y “se esforzaba hasta conseguir un sonido lo más hermoso posible [...] tocaba el pasaje muchas veces seguidas, y seguía en todo caso hasta que él estuviera satisfecho con el sonido del clarinete”,¹ como apunta Martin Alber en la introducción sobre Rudolf Koder previa a la correspondencia. Nuestro autor, también era muy conocido por ir silbando las orquestas que tenía en mente. Solía comparar y hacer analogías de la vida con la música observando sus «aires de familia» y parentescos conceptuales resaltando la infabilidad epistémica de ambas, como se puede constatar en muchos escritos menores recopilados como en Zettel u otras correspondencias.

Su amigo Rudolf Koder, durante toda su vida ejerció de profesor de primaria y secundaria en varias escuelas, llegando a ser director dos años antes de su jubilación de la Escuela Primaria de Klosterneuburg, especializándose en la educación secundaria y en asignaturas como: la música, el alemán y las matemáticas. Koder compartía con Ludwig su pasión por la música, gustos y era

muy cercano a la familia Wittgenstein, sobre todo a su hermana Helene Wittgenstein con la que también compartía gustos musicales y tocaba frecuentemente el piano a cuatro manos. Entre los gustos del vienés, juicios y críticas tratan autores como Brahms, Bruckner, Schubert, Schumann, Mendelssohn, Beethoven, Mozart, Haydn, Wagner, Mahler y Richard Strauss, entre otros. En el intercambio epistolar, se observa cómo trabajan sus apreciaciones sobre los conciertos a los que asistían, las veladas, o sus escuchas personales con el gramófono. Entre multitud de referencias, Ludwig comenta que quizá su favorita es —la última y la penúltima parte de *Las cofradías de David*, “*Wie aus weiter ferne*, de Robert Schuman—, o la 7^o de Bruckner (la dulzura que siempre tuvo), son muy de su agrado como unos cuartetos de Haydn muy bien interpretados, o la gran expresividad de la *Gran Misa* de Bach. Se puede rastrear en correspondencias muy cortas cómo comparan autores y nuevas interpretaciones y matices que aportan diferentes directores, dando argumentos de quién era más destructor o creador, hasta filosofar a base de concepciones literarias, estéticas y críticas de la teoría musical.

Estos detalles analíticos que hacen en pocas líneas entre preocupaciones existenciales en un periodo de entre guerras, con conversaciones de interés que tienen sobre sus familias, enfermedades, el trabajo y muertes cercanas, se hacen muy enriquecedoras para observar la importancia de la música en sus vidas.

La familia Wittgenstein estaba inmersa en el arte de su tiempo, como, por ejemplo, su hermana Helene, que cantaba, tocaba el piano y sobre todo ejerció de pintora, Paul Wittgenstein, que fue pianista, o el mismo Ludwig como filósofo, entre los demás hermanos. En otro ejemplo, se puede apreciar en el patrimonio de la familia, que tenían bajo su

1 Ludwig Wittgenstein, *Correspondencia con Rudolf Koder*, traducción Venancio Andreu Baldó e Isabel Gamero Cabrera. Editorial Ápeiron Ediciones, Madrid (España) 2019. Pág. 10.

propiedad obras, como gran parte de la obra de Rodín. Por otro lado, formaba parte de la cotidianidad de los Wittgenstein organizar veladas donde iban a tocar reconocidos y no tan reconocidos del mundo musical, que se reunían por su “afinidad espiritual en torno a la música, un círculo común de conocidos y amigos. Las familias austriacas, aristocráticas y de la gran burguesía, que Brahms frecuentaba, ofrecían una y otra vez un espacio íntimo para escuchar el sonido de la música de cámara en una atmósfera que le fuera propicia”². En estas celebraciones se encontraban autores como al que hace referencia el primer ensayo de la obra “Josef Labor y la familia Wittgenstein”. En esta parte de la investigación, presentan al músico que se quedó ciego a causa de la viruela con 3 años. Labor de origen austriaco, cosechó sus logros como pianista, organista y compositor (1842-1924), en el ensayo también, muestran las influencias de Labor y su relación con los Wittgenstein, aparte de visibilizar otras valoraciones de la familia y amigos cercanos sobre el músico, extraídas de otras correspondencias con Ludwig, como, por ejemplo, críticas que constan en cartas de Hermine Wittgenstein o Maurice Dury, entre otras.

Josef Labor y “el vínculo que lo une a la familia, donde gozaba de gran consideración, se traza a través de la amistad con el violinista Joseph Joachim y, a través de este con Johannes Brahms (1831, Kittsee/Köpcsény – 1907, Berlín), este era pariente por parte de la madre y primo de Fanny Figdor (1813-1890), quién se casó en 1839 con el abuelo de Ludwig, Herman Christian Wittgenstein (1802-1878)”³. En este mismo ensayo, recuerdan la presencia de Labor, entre otros, como compositor de Paul Wittgenstein, que perdió una mano en la guerra y aun así, continuó tocando

el piano, a pesar de los juicios críticos de algunas de sus hermanas, como refleja la obra. Como dato curioso Josef Labor muere el día del cumpleaños de Ludwig Wittgenstein, el 26 de abril de 1924.

En el apartado “Sobre los aspectos de lo musical en la vida y obra de Ludwig Wittgenstein”, la investigación comienza a profundizar, proponiendo a la música y sus compositores como objeto filosófico. En esta segunda parte del libro, en los ensayos filosóficos se encuentran muchas reseñas de los autores de la escena musical antes mencionados, los nuevos protagonistas que se ocupan de esclarecer el contexto histórico-musical alemán y vienés. Estos teóricos de la música como: Hugo Riemann, Eduard Hanslick, Zimmermann, Edgar Istel, Walter Niemann, Ernst Kurth, Halm, Guido Adler, también se cita a Engelmaan, entre otros, enriquecen la publicación con juicios y comparaciones, entre las que también se encuentran las observaciones de nuestro autor Ludwig Wittgenstein.

Se pueden apreciar entre otras reflexiones ontológicas y gnoseológicas sobre la estética filosófica como: la magia irónica y dramática de las óperas de Wagner o el límite de lo bello en Mozart según Grillparzer. U otras, en las Wittgenstein afirma, por ejemplo, que Mozart esconde una clave para las reflexiones sobre el entendimiento y que para nuestro filósofo: Beethoven es el genio schopenhaueriano. Wittgenstein anoche en su diario el 1.3.1931: Beethoven “es realista del todo, es decir, su música es completamente verdadera; quiero decir: ve la vida tal como es y entonces la realiza. Es completamente religión, y no es en absoluto poesía religiosa”⁴. También comenta una fuerte crítica sobre Brahms y Bruckner. De Brahms afirma, que es demasiado abstracto,

2 *Ibid.*, pág 139.

3 *Ibid.*, pág 133.

4 *Ibid.*, pág 156.

se nota que compone con la pluma, mientras que Bruckner es un compositor que compone más con el oído, un sonido de orquesta pleno, a saber, donde Brahms se desliza, Bruckner no evita cantos de esquinas. Este último era tratado como el músico de Dios, católico y de la alta Austria, discípulo de los místicos alemanes. El periodista Hanslick reconocido antagonista de Bruckner, afirma, que cómo siendo un hombre tan tierno y pacífico, se torna un anarquista en el momento de la composición, mientras que el wagneriano Niemann apunta, que Bruckner es el músico de más grande sentimiento del tiempo moderno. Por otro lado, también se ocupan de autores como Mendelssohn, al que califica nuestro autor: que más que una cima es una llanura elevada, “si se quisiera caracterizar la música de Mendelssohn, se podría hacer diciendo que no habría ninguna música suya difícil de entender”⁵, aunque no lo tenía en muy alta estima. Mendelssohn le parece a Ludwig como demasiado propenso a amoldarse. Sobre Mahler del que escribe que hay que entender mucho de música, de su historia, de su evolución para entenderlo, apunta: «es simplemente *Kitsch*, en el arte como lo perverso». A el viernes, Mahler le parece insoportable, en el análisis de la obra se puede leer una fuerte crítica: “me gustaría decirle siempre: es que esto solo lo has escuchado de los otros, es que esto no te pertenece nada realmente”⁶, esto se encuentra en una anotación en el Diario del 6.5.1931 según la investigación de Martin Alber.

La viveza de la música, tal cómo se nos muestra experiencialmente en un hipersubjetivismo, nos obliga a hablar de ella en un lenguaje poético. Lo poco que podemos explicitar de esta a través del logos nos hace topar una y otra vez con los «límites del len-

guaje», y por esta razón, en nuestro intento por decir o teorizar, falsificamos e utilizamos nuestras mejores metáforas, retorciendo el lenguaje hasta el límite de sus sentidos, intentando de este modo, verbalizar y traducir al otro, en nuestro fracaso, poco más que su escritura o su propia forma. Aquí se confronta la cuestión que trata más profundamente Wittgenstein en su obra: *Lecciones sobre estética* (1938) tal como indica Martin Alber: «Uno de los puntos más interesantes, que va unido a la cuestión del no-poder-describir, es que es indescriptible la impresión que engendra un determinado verso, o un par de compases»⁷. La peculiaridad de la sensación musical, lo intraducible, es esa puerta epistémica que nos queda vetada. A saber, no tenemos acceso a este tipo de conocimiento subjetivo y cualitativo a través del lenguaje: nuestra herramienta material para construir objetividad. Este se encuentra desbordado porque no encuentra ni anverso, ni reverso, ni secreto muro, ni oscuro puente, ni secreto centro, como diría Jorge Luis Borges, sobre el que operar (mensurar) o calibrar sus esquemas gramaticales. Sin brújula y sin poder mediar, todo juicio sobre el alma o la psique queda carente de sentido al no haber posibilidad de un ingreso en la realidad tal como ve y la siente el otro.

Se constata de este modo, una crisis de racionalidad, no entre lenguajes inconmensurables, sino —que no hay lenguaje racional viable ni condiciones de posibilidad previas para poder disponer de un material de referencia—. Entonces, sobre qué muestra ciega hemos de sobreponernos, qué nuevo paradigma nos abrirá nuevos datos de validación analítica que nos ofrezca la transferibilidad de esta energía que es nuclearmente vital y evidente ante nuestros ojos, y que, aun así, nos desafía con su naturaleza intrínseca-

5 *Ibid.*, pág 157.

6 *Ibid.*, pág 165.

7 *Ibid.*, pág 188.

mente humana y anti-teórica. La resolución nos llevaría a la disolución de este ensayo, la pregunta sería, el esfuerzo intelectual qué sacrificaría. Es decir, cómo llevar el logos a hipérbole, cómo hacer de la ausencia, presencia, este es el fundamento del misterio, este es el pilar de esta iglesia. Esto hace una referencia directa al *Tractatus Lógico Philosophicus*, la primera y única obra de nuestro autor en vida, en el que enuncia en su séptima y última tesis: 7. De lo que no se puede hablar es mejor callar, que a su vez, está relacionada con una premisa de la tesis sexta, 6.522 Lo inexpresable, ciertamente, existe. Se muestra, es lo místico.

En la correspondencia se piensa en la aspiración antropológica del ser humano por estudiarse así mismo: ¿qué clase de elevación moral nos impulsa? Según Krennek, el destino de la música lleva consigo el extrañamiento del Yo con el mundo, la desautomatización con la realidad, verse desde una nueva mirada. En la obra, también hace mención al constructo del pensamiento musical y lo límites de su comprensión, es

decir, este objeto completamente holístico qué tipo de verdad nos revela. Según Ludwig Wittgenstein, la verdad, si existe, tiene que ser cultural, si hubiese alguien lo suficientemente inteligente como para pronunciarla, la verdad sonaría a toda la humanidad como completamente paradójica: “Y el compositor que la siente en sí debe enfrentarse, con su sentimiento, a todo lo expresado ahora, y debe también aparecer como absurdo, estúpido, según nuestros criterios. Pero no atractivamente absurdo (pues eso es lo que en el fondo sí se correspondería con la concepción de nuestros criterios culturales, sino trivial. Labor es un ejemplo de ello allí donde ha creado algo significativo”⁸.

¡Ni una sola nota no sentida!
Hermine a Ludwig Wittgenstein,
[finales de enero de 1922].

Jeremy Mederos
(Universidad Complutense de Madrid)

⁸ *Ibid.*, pág 175.

LUMBRERAS SANCHO, Sara (2020). *Respuestas al Transhumanismo. Cuerpo, autenticidad y sentido*. Madrid: Digital Reasons, 202 pp.

En los últimos años no han dejado de aparecer publicaciones sobre lo que, desde la perspectiva del historiador y filósofo Fukuyama (2002) —aunque también la del pensador de izquierdas Slavoj Zizek—, bien pudiera ser “la idea más peligrosa del mundo” y la que representa la mayor amenaza para el bienestar de la humanidad. A saber: el Transhumanismo, aquel movimiento intelectual y político que busca la superación de lo humano a través de la ciencia y la tecnología. Estas publicaciones han adoptado

distintas formas. Así, podemos encontrarnos con: artículos académicos, Trabajos de Fin de Grado y Trabajos de Fin de Máster, e, igualmente, con libros colectivos y de un único autor. Buena parte de éstas han venido de la mano de académicos pertenecientes a las Humanidades —en general— y la Filosofía —en particular—, aunque también podemos hallar estudios de científicos más o menos “reflexivos” (aquellos que, más allá de su propio ámbito de estudio, se hacen preguntas generales que escapan a lo puramente

empírico y verificable) provenientes de las ciencias duras, blandas y de la salud.

Otra clasificación de todas estas publicaciones recientes, cruzada con la anterior e igualmente importante, es aquella en la que se distingue entre las que han tratado un tema específico del Transhumanismo (bien desde una óptica más analítica o una axiológica y poliética) y las que han decidido abordarlo en su generalidad, desde una perspectiva sinóptica y global. Dentro de este último grupo hemos de destacar, en lo referente a publicaciones en forma de libro de un único autor y dentro del ámbito nacional, el ya famoso libro de Diéguez (2017), *Transhumanismo*, así como, más allá de España, *La revolución transhumanista. Cómo la tecnomedicina y la uberización del mundo van a transformar nuestras vidas* de Ferry (2017). En él también se hallaría el libro que aquí reseñamos: *Respuestas al Transhumanismo (Cuerpo, autenticidad y sentido)*, editado por Digital Reasons, un sello incipiente que pretende favorecer la publicación de libros sobre temas de pensamiento contemporáneo. En este caso, sin embargo, su autora no es una filósofa nominal, como sí lo son Diéguez y Ferry, sino que es ingeniera: se trata de Sara Lumbreras Sancho, doctora en Ingeniería Industrial, profesora titular de la escuela de ingeniería ICAI y miembro del consejo de la Cátedra de Ciencia, Tecnología y Religión (Universidad Pontificia de Comillas).

Que la autora no sea filósofa de formación no implica, empero, ni que ella no lo sea *avant la lettre* (al fin y al cabo ha publicado numerosos artículos de epistemología social y bioética) ni que el libro no tenga un claro componente filosófico. Por el contrario, éste es una apuesta por una Tercera Cultura que trataría de desdibujar la frontera entre ciencias y humanidades (p. 8). Es, si se quiere, el camino inverso adoptado por Diéguez y Ferry: mientras que éstos son filósofos de

marcado carácter naturalista (filósofos que piensan en continuidad con las ciencias y las tecnologías), Sara Lumbreras es una ingeniera que tiene una mirada amplia que le lleva a pensar en torno a cuestiones filosóficas y ético-políticas.

En este sentido, *Respuestas al Transhumanismo* tiene elementos en común con aquellos libros que tratan de contornear una panorámica general del susodicho movimiento, de entre los que destacan los ya mencionados. En él, como en los otros, nos encontraremos con la exposición de: los antecedentes sociológicos del Transhumanismo y la gestación del movimiento transhumanista; la historia del Transhumanismo tal y como es concebida por sus propios intelectuales (especialmente por Nick Bostrom), que es prácticamente indistinguible de la propia historia de la humanidad; las predicciones más alocadas y pintorescas de sus autores, desde la Singularidad hasta la promesa de inmortalidad, pasando por el advenimiento de la creación de una Inteligencia Artificial Fuerte o general y el argumentario principal a favor de la adopción de medidas como las que proclaman los autores transhumanistas, siendo de especial importancia aquellos argumentos en los que se apela a la naturaleza humana, a la inevitabilidad de los procesos de innovación y desarrollo y a la obligación moral de mejorarnos. Todo ello, además, viene acompañado de un balance crítico por parte de la autora, algo que también se hace en Diéguez (2017) y Ferry (2017). El libro, por tanto, no es meramente descriptivo, sino que entra a valorar la calidad de los argumentos presentados por el Transhumanismo desde un punto de vista tanto teórico como práctico, adoptando una serie de compromisos que habremos de entrar a valorar en las líneas que siguen. Cabe decir que esto es algo que la profesora Lumbreras decide hacer de manera intermitente, sin distinguir de forma clara y

nítida, muchas veces, entre la exposición y la valoración, algo que, aun siendo una cuestión menor, en algunos puntos puede llegar a ser algo confuso.

Pese a que haya una clara afinidad temática y de contenido con respecto a otras introducciones al Transhumanismo (como las que hemos mencionado), creemos, no obstante, que hay muchos aspectos originales dentro del libro que pueden aportar nuevas herramientas tanto de crítica como de análisis. Sin irse muy lejos, ya dentro de la propia introducción al Transhumanismo, ubicada en el capítulo 1, es digno de mención el hecho de que se relacione la utopía subyacente a este movimiento explícitamente con las cuestiones del fin del trabajo (donde, como no podía ser de otra manera, se trae fundamentalmente a colación a Jeremy Rifkin) (p. 21), por una parte, y con el debate o la oposición entre malthusianos y cornucopianos (p. 29) en lo relativo a los límites del crecimiento y la superpoblación, por otra. Además, dentro de la historia del movimiento se aportan datos sobre algunos de los grandes nombres que se acogen bajo el paraguas del Transhumanismo (FM-2030, Natasha Vita-More, Max More...) que no habían sido señalados hasta ahora en los más relevantes libros generales acerca de este movimiento (pp. 13-14). Más aún, fuera de esta historia, en el Apéndice (p. 156), podremos leer una entrevista a Natasha Vita-More por parte de Lumbreras que nos ayudará a hacernos una idea del discurso y la retórica de la parte más visible y con más poder del movimiento.

Otro de los principales aspectos que conviene subrayar de cara a la originalidad e innovación del contenido del libro es lo bien informado que está en líneas generales. Cabe aclarar que, frente a otros textos (como, muy especialmente, el de Antonio Diéguez), *Respuestas al Transhumanismo* se centra más en las propuestas transhumanistas cibernéticas

(aquellas que están más próximas a las ciencias de la computación y la Inteligencia Artificial) que en las biomédicas (aquellas que, por el contrario, apuestan por la superación de lo humano a través de mejoras genéticas, protésicas y hormonales). Es en este aspecto en el que la bibliografía seleccionada nos parece indudablemente pertinente y enriquecedora, así como el uso que se hace de ella.

La profesora Lumbreras hace un gran repaso de la historia de la Inteligencia Artificial y nos informa de sus grandes avances, pero también de las limitaciones de un campo como éste con un estilo claro y accesible a los no expertos en la materia. Igualmente se hace eco de (y dialoga con) muchos de los filósofos que han venido discutiendo problemas teóricos en torno a la mente humana y su posible (o no) replicación artificial. En este sentido, hacen acto de presencia pensadores como Andy Clark, David Chalmers (éste, junto con el anterior, en lo relativo a la teoría de la mente extendida, y de forma individual en lo relativo a la famosa “hipótesis zombie”) o Antonio Damasio (en lo relativo a la relación entre razón y emoción). Quizá sí se eche algo más en falta, empero, la presencia de John Searle (al que solo se le menciona de forma lateral en un único sitio) (p. 120), Daniel Dennett y Douglas Hofstadter, sobre todo si se tiene en cuenta que todos ellos han participado en la difusión de los últimos resultados en Inteligencia Artificial y que también han pensado a partir de éstos, tal y como hace Sara Lumbreras en varios pasajes del texto que aquí reseñamos. Más aún, muchos de los argumentos de corte teórico por parte de la autora en lo tocante a las limitaciones de los avances en Inteligencia Artificial y ciencias de la computación son afines a algunas de las posturas que han venido manteniendo estos tres autores a lo largo de los últimos años. La incapacidad que tienen incluso las máquinas más avan-

zadas para comprender aquello en lo que son competentes, por ejemplo, la encontramos en el último gran libro de Dennett (2017), *De las Bacterias a Bach*. Asimismo, el énfasis en la subjetividad, en el punto de vista de la primera persona del que carecen (y probablemente carecerán siempre) las máquinas, está en directa relación con las posturas de John Searle, algo que puede comprobarse si se atiende, entre otros muchos lugares, a *El misterio de la conciencia* (Searle, 2000). Por no hablar, por supuesto, de la diferencia entre simulación y duplicación implicada en algunas de las respuestas críticas que la autora da —desde un punto de vista teórico— al transhumanismo cibernético.

Las objeciones más importantes en este libro, sin embargo, no son de tipo teórico, sino antropológico y práctico. Esto es así porque éstas son a las que dedica más páginas dentro del texto y las que, en buena medida, permiten su articulación. Conviene subrayar, pues, cuál es a grandes rasgos la crítica efectuada por la profesora Sara Lumbreras así como el trasfondo de la misma.

Desde el punto de vista antropológico, la denuncia de la profesora Lumbreras es directa y clara: el Transhumanismo peca de manejar una imagen del ser humano que es reduccionista (p. 84) y que, además, es dualista en lo que respecta a la distinción entre mente y cerebro (o mente y cuerpo, si se prefiere). Dicho dualismo, huelga aclarar, no sería horizontal (tal y como suele ocurrir con todos los grandes dualismos abiertos en la Modernidad), sino jerárquico: bajo éste, se primaría todo lo que tuviese que ver con la mente y la cognición (más bien habría que decir “información”) por encima de todo lo relacionado con lo material y corpóreo. A esta imagen del ser humano contraponen una en la que se le comprende “como una única entidad dotada de varias dimensiones entrelazadas: biológica, intelectual y espiritual” (p. 84), donde el papel

tanto del lenguaje como de la sociabilidad son fundamentales. La forma que tiene de contornear esta imagen es mediante el empleo de la mejor literatura en psicología, antropología y biología. Todo ello acaba convergiendo, además, en la abierta defensa de lo que Sara Lumbreras ha venido a llamar “una espiritualidad encarnada” (p. 87).

Desde el punto de vista ético-político, solo inseparable conceptualmente del anterior, el gran pecado del Transhumanismo es, sin embargo, su exacerbado y ciego tecnooptimismo, cuyas raíces pueden ser rastreadas en el pasado (p. 18) y que, esencialmente, presupone que todos los problemas del ser humano (tanto individuales como colectivos) pueden ser resueltos mediante mejoras tecnológicas. En palabras de Lumbreras: “El tecnooptimismo asume que todos los problemas se solucionarán por el camino, porque la tecnología resolverá también los problemas sociales” (p. 147). Frente a ello, la autora argumenta que buena parte de las mejoras que más nos deberían importar (como las morales o personales, pero también, cabría añadir, las políticas) no vendrían propiciadas por avances tecnológicos, sino por la forma de relacionarnos con nosotros mismos y con los que nos rodea. El cuidado —en su sentido más amplio: de uno mismo y de los demás—, la agencia (epistémica y moral) e incluso cierto sentido de la trascendencia, son aspectos primordiales de una buena vida que no pueden ser resueltos simplemente con la potenciación o la innovación de las máquinas que ya tenemos. Esto último, además, cabría apostillar a la argumentación de la autora, es netamente incompatible con un mundo finito con recursos finitos como en el que vivimos, algo a lo que la propia Sara Lumbreras es sensible, tal y como se podrá ver al interior del apartado 1.4. (p. 37). Hubiese sido

deseable, en este sentido, que se hubiese engarzado en su libro estas dos cuestiones: por una parte, lo concerniente a la buena vida y a la mejora moral y política, y, por otra, la cuestión de los límites del crecimiento y las consecuencias catastróficas de vivir en un estado de *hybris* permanente.

Sea como fuere, creemos que *Respuestas al Transhumanismo* es de obligada lectura no solo si se quiere estar bien informado de las radicales propuestas de este movimiento, sino también provisto de buenos argumentos para no quedar cegados ante sus promesas desorbitadas y asumirlo críticamente. Creemos que el libro consigue su objetivo principal en lo que respecta a este último punto: a saber, responder al Transhumanismo sin caer preso de su tecnooptimismo e hiper-humanismo — por una parte— ni hacerlo desde una postura reaccionaria, catastrofista ni tecnófoba —por otra—. Una empresa que, pese a lo que pudiera parecer a primera vista, no es tan fácil de acometer.

Referencias

- Dennett, D. (2017). *De las bacterias a Bach*. Barcelona: Pasado y Presente.
- Diéguez, A. (2017). *Transhumanismo*. España: Gedisa.
- Ferry, L. (2017). *La revolución transhumanista. Cómo la tecnomedicina y la uberrización del mundo van a transformar nuestras vidas*. España: Alianza.
- Fukuyama, F. (2002). *Our posthuman future: consequences of the biotechnology revolution*, Nueva York: Farrar Straus & Giroux.
- Searle, J. R. (2000). *El misterio de la conciencia*. España: Paidós.

Adrián Santamaría Pérez
(Universidad Autónoma de Madrid)
José María Santamaría García
(Universidad de Alcalá de Henares)

HUSSERL, Edmund (2019). *Textos Breves (1887-1936)*. Coord.: Agustín Serrano de Haro y Antonio Zirión Quijano. Salamanca: Sígueme, 716 pp.

Textos breves es un verdadero hito para la literatura fenomenológica traducida al español y muy en concreto para la obra de Edmund Husserl, que disfruta de un presente dulce de novísimas traducciones y reediciones. Quizás sea esta la más deslumbrante, sobre todo si nos atenemos a su complejidad técnica y editorial, su voluntad abarcadora de toda la vida creativa de su autor, una exquisita, experta y meditada selección de textos y una vocación plenamente consciente de estar aportando un material imprescindible para comprender a fondo y en detalle la trayectoria, bastante menos sinuosa de lo

que en principio pudiera parecer, del pensador cuya obra determinó la dirección de la filosofía en el siglo XX. No sería una exageración decir que esta reunión de *Textos breves*, elaborada por un nutrido grupo de especialistas de ambos lados del Atlántico a lo largo de una dilatada cantidad de años, roza la importancia de las obras canónicas del fundador de la fenomenología.

Es conocida la severísima exigencia que Husserl se marcaba en cada una de sus publicaciones. Esto se constata en la absoluta falta de proporcionalidad entre sus libros aparecidos en vida (apenas seis) y los

43 volúmenes de sus obras completas, conocidas popularmente por el estilizado nombre de *Husserliana*, integrada en buena parte por la desbordante cantidad de manuscritos que dejó como legado —más de 50.000 páginas escritas en una intrincada estenografía—. De esta colección se nutren fundamentalmente estos *Textos Breves*, pero en una selección muy particular. En *Husserliana*, como es natural, se incluyen contenidos de muy distinta procedencia y elaboración: desde papeles de trabajo en los que Husserl dejaba correr la reflexión a modo de tentativa y prueba, en los que casi se le puede “ver” pensar, hasta textos definidos y acabados que contienen ya una síntesis y un sentido propio por sí mismos. Muchos de estos últimos proceden de artículos que publicó en revistas de investigación o de discursos y conferencias pronunciadas en foros públicos ante audiencias muy variadas. Son textos de este tipo, no ya vuelos del pensamiento que rastrea caminos posibles, sino auténticas piezas científicas en las que Husserl se posiciona filosóficamente y ofrece resultados bajo su firma y responsabilidad, los que son mayoría en la selección ofrecida en estos *Textos breves*.

La minuciosa y cuidada selección de los editores, el español Agustín Serrano de Haro y el mexicano Antonio Zirián Quijano, pone a disposición de la comunidad filosófica de habla hispana una serie de textos que Husserl dio a conocer públicamente —o preparó para ello, ya que alguno de ellos se guardó en un cajón—, lo que sin duda les añade un valor muy singular. Podríamos hablar de este libro, quizá, como de una gran obra paralela del padre de la fenomenología que vendría a completar enfoques, iluminar ángulos, definir contornos, ofrecer contextos e incluso afirmar alternativas a lo que entregó en sus obras más conocidas. El conjunto integra artículos de investigación —procedentes

sobre todo de la época anterior, podríamos decir que preparatoria, a *Investigaciones lógicas*—, discursos públicos —con especial mención al de la toma de posesión de su cátedra en la Universidad de Friburgo—, conferencias en Londres, Ámsterdam y Praga, en las que fundamentalmente presentaba el método fenomenológico a un público ávido por conocer esa nueva filosofía de la boca de su fundador; lecciones de carácter más abierto —como los cursos sobre Fichte destinados a combatientes de la Primera Guerra Mundial—, ponencias de temática concreta —con la obra de Kant, la relación entre fenomenología y antropología o “el presente de la filosofía” como motivos rectores—, una reseña de 1903 en la que contesta con no poca vehemencia a una crítica a sus *Investigaciones lógicas* y el famoso proyecto de prólogo, escrito en 1913 pero finalmente descartado, para la segunda edición de esta última obra. Solo tres de los dieciocho escritos proceden propiamente de manuscritos de investigación, aunque hay que resaltar la extraordinaria naturaleza de estos documentos. Uno de ellos recoge dos manuscritos de investigación sobre la constitución del presente vivo y del cuerpo propio que fueron publicados por Alfred Schutz y autorizados por el propio Husserl; mientras que los otros son dos clásicos del pensamiento husserliano conocidos bajo los títulos de *La tierra no se mueve* (1934) y *El origen de la geometría* (1936).

Si estos dieciocho textos recorren medio siglo de incansable búsqueda filosófica, la traducción parece querer emular la epopeya original: más de una década y trece traductores después de iniciada la tarea, la obra vio la luz. En este sentido, hay que destacar la paciente, meticulosa y precisa edición llevada a cabo por Serrano de Haro y Zirián Quijano, auténticos pulidores y abrillantadores de términos, expresiones, referencias y contextos

históricos. A su labor infatigable y vocacional debemos que las obras de Husserl, otras anteriormente y ahora en concreto estos *Textos breves*, hayan recibido el tratamiento que se merece: dignísimo, rigurosísimo y cuidado hasta el límite. Y no por ello carente de un pulso crítico también sobresaliente, preñado de agudeza y perspicacia. Ellos mismos han llevado a cabo la titánica tarea de modificar las versiones originales de los otros once traductores, “revisando ambos cada uno de los textos y revisando luego mutuamente sus revisiones”, confiesan en el prefacio de la obra. Esta ardua labor ha logrado unificar y homogeneizar, con amplitud de miras, mano izquierda y flexibilidad, la siempre difícil tarea de traducir a Husserl, lo que ha evitado de forma casi general el recurso a encorchetar el término alemán.

A ellos se debe asimismo un logro de inestimable mérito añadido a la traducción y la coordinación de las traducciones: la elaboración de una introducción para cada uno de los ensayos del libro (aparte de la introducción general a la obra). Estos mini-prólogos ayudan al lector a contextualizar el texto en cuestión, tanto histórica como filosóficamente. Pero no solo explican su procedencia —si es un discurso, un artículo, un manuscrito, así como el año en que se escribió—, sino que también se entresacan sus logros teóricos más decisivos —lo que justifica su inclusión en el volumen—, se descubren conexiones con las obras maestras de Husserl —lo que se percibe muy bien en los artículos previos a 1900, donde ya se adivinan algunos de los desarrollos fundamentales de *Investigaciones lógicas*—, se practican análisis a modo de cata sobre conceptos clave —como la problemática traducción de “Vorstellung” como “representación”— y se discuten polémicas clásicas de fenomenología.

También se penetra en la intrahistoria de los textos, algunos de los cuales están mar-

cados por curiosos avatares o amplificadas polémicas. Por ejemplo, *Sobre la tarea presente de la filosofía* debía haber sido leído por Jan Patočka en un congreso en Praga en 1934, lo que nunca llegó a suceder, ya que Husserl pidió que se le devolviera el ensayo que había enviado al no estar conforme con su resultado final. También se incluye aquí *Fenomenología y antropología*, texto procedente de la multitudinaria conferencia pronunciada por Husserl en Berlín en 1931, que sorprendentemente supuso su primer viaje por Alemania en calidad de conferenciante. Este evento, presenciado en vivo por un joven Xavier Zubiri, sería recordado varias décadas después por Martin Heidegger, que lo calificó en su famosa entrevista a *Der Spiegel* como un “ajuste de cuentas” contra él y contra Scheler en un “ambiente de palacio de deportes” (ajuste que habría que poner doblemente entre comillas, una vez leída la limpia y rigurosa exposición de un posicionamiento teórico que impugna la posibilidad de establecer la antropología como filosofía primera).

En este sentido, hay en estas mini-introducciones numerosas referencias a hitos biográficos de Husserl, que junto a peripecias vitales y lances variados, tanto suyos como del movimiento fenomenológico, hacen que el libro tenga también interés en el plano de la crónica biográfica-histórica. Junto a estas introducciones y el prólogo general a la obra, sí se echa de menos un epílogo general de los editores que viniera a cerrar el volumen con una palabra sintética que acompañara al lector en sus últimas páginas. Después de un trayecto —sin duda apasionante y enriquecedor, pero también arduo y fatigoso— en compañía, en el que los editores van casi llevando de la mano al lector, no hubiera estado de más una palabra final para cerrar tan intenso recorrido.

Como ya se ha señalado, los dieciocho ensayos de este Husserl poco conocido abar-

can cinco decenios de trabajo, los que van de su primera publicación en 1887, un trabajo sobre el concepto de número que le sirvió de habilitación a la docencia en la Universidad de Halle, hasta el célebre *El origen de la geometría*, un manuscrito de 1936 que alcanzó gran notoriedad tras el estudio sobre el mismo realizado por Jacques Derrida en los años sesenta. Una amplia mayoría de escritos, en concreto doce, se enmarcan en la época trascendental, que en la obra se abre con gran pertinencia y simbolismo con el citado prólogo de 1913 a *Investigaciones lógicas*. Solo dos ensayos pertenecen al período de 1900 a 1913, uno de las cuales es el conocido e influyente artículo *La filosofía, ciencia rigurosa*, en la traducción de Miguel García-Baró.

Especialmente interesantes son los cuatro textos del Husserl pre-fenomenológico, que van de 1887 a 1894, un período de gestación y conformación de su pensamiento más original donde ya se intuyen y se dejan adivinar algunos de los núcleos teóricos básicos de *Investigaciones lógicas*. Todavía en la órbita de la escuela de Brentano, aquí brilla el Husserl matemático y lógico en toda su definición y potencia. Eran los años circundantes al primero de sus grandes libros, *Filosofía de la aritmética* (1890); de hecho, el primero de los ensayos está incorporado casi punto por punto a los cuatro primeros capítulos de esta obra. La gran preocupación de sus inicios era llevar a cabo una sólida filosofía de la matemática que liberara a esta ciencia de su crisis de fundamentación, así como de abrir el camino hacia una teoría del conocimiento sostenida en cimientos firmes. Es significativo, como ya advierte Serrano de Haro, la cantidad de temas y cuestiones de la fenomenología posterior que están presentes en estos cuatro escritos iniciales: la pregunta ontológica por la “noción de algo en general”, las dimensiones pasiva y activa de la

intencionalidad, la teoría de todos y partes, la prefiguración de categorías clave como las de cumplimiento intuitivo o intención vacía, la ampliación del término intuición a las representaciones de la fantasía... También hay que destacar profundas diferencias con el Husserl propiamente fenomenológico, como la teoría psicológica representacionista de marcada influencia brentaniana que alimenta las páginas de los *Estudios psicológicos ordenados a la lógica elemental*, en la que lo intuido se identifica con el contenido presente realmente en la conciencia. La proyección de estos textos no solo alcanza hasta su referencia editorial inmediatamente posterior, *Investigaciones lógicas*. Es de admirar que en las indagaciones en el nivel prelógico, en el que obligatoriamente tiene su origen todo concepto aritmético o lógico, ya se están adelantando postreros desarrollos del programa de la fenomenología genética.

Después de una reseña a un libro crítico con *Investigaciones lógicas* en la que trata de restituir el sentido de su obra y denunciar las malinterpretaciones (1903) y de la ya citada y conocidísima confrontación con Dilthey en *Filosofía, ciencia rigurosa* (1911), *Textos breves* ofrece una de las piezas más distinguidas de la colección: los dos fragmentos para el proyecto de un prólogo a la segunda edición de *Investigaciones lógicas*, fechado en el año 1913. El texto nunca llegó a anteceder a la magna obra —ni en 1913, cuando se reeditó el primer volumen y las cinco primeras investigaciones; ni tampoco en 1921, cuando se reeditó la *Sexta investigación*—, pero ocupó y preocupó mucho a Husserl, que en una fecha tan extemporánea para su propósito inicial como 1924 modificó y amplió los borradores del proyecto. El texto en el que Husserl quiere explicar una obra incomprendida por la crítica es, al mismo tiempo, la oportunidad de limar aristas con la novísima

perspectiva trascendental alumbrada en las mismas fechas en el primer libro de *Ideas*. El voltaje de la exposición alcanza máximos en el segundo de los fragmentos, donde Husserl rebate la teoría del objeto de Meinong, se confronta (amplísimamente) con las malinterpretaciones del libro por parte de Wundt, critica a los que quieren identificar la fenomenología con una reductiva filosofía del lenguaje y descubre los hilos invisibles que unen las *Investigaciones lógicas* con el pensamiento de Lotze y Bolzano.

La trayectoria de Husserl se puede interpretar como un intento sostenido y continuado de presentar la fenomenología de la mejor manera posible. *Textos breves* lo confirma. En la obra se pueden encontrar diferentes conferencias que el filósofo de origen moravo pronunció en distintos puntos de Europa para presentar el método fenomenológico y su campo de estudio. En las cuatro conferencias de Londres, en junio de 1922, Husserl va a radicalizar la vía cartesiana del *ego cogito* en su defensa de la fenomenología como filosofía primera, así como descubrirá los motivos existenciales y éticos que se esconden en los fundamentos últimos de su método. La conferencia de Ámsterdam, en la primavera de 1928, está en cambio muy influenciada por la entrada “Fenomenología” que Husserl escribió para la *Enciclopedia Británica*, donde se establece el acceso a la fenomenología trascendental a través de la psicología pura. Por último, las conferencias en Praga de 1935 suponen el último intento de Husserl por desplegar de una manera innovadora las virtualidades de su método y contienen el germen de lo que luego se convertiría en su último gran libro y uno de los más reconocidos: *La crisis de las ciencias europeas y la fenomenología trascendental*.

Antes de terminar, merecería la pena destacar uno de los textos procedentes de manuscritos de investigación. Se redactó en mayo

de 1934, se ha conocido popularmente como *La tierra no se mueve* y fue presentado por el propio Husserl como “una inversión de la teoría copernicana”. Se trata de una extraordinaria meditación fenomenológica, única en el pensamiento de Husserl, sobre el lugar que ocupa la Tierra, el suelo que pisamos, en nuestro experimentar el mundo. En el texto se destaca esta asombrosa condición del suelo, en cierto sentido tan sorprendente como la del mismo cuerpo propio, por la que, de un lado, se ofrece de manera absolutamente singular respecto al resto de cuerpos percibidos; mientras que, por el otro, dificulta el que se le otorgue el estatus de cuerpo físico. Para la ocasión, Agustín Serrano de Haro ha llevado a cabo una profunda revisión de la traducción que él mismo realizó en 1995 para la Universidad Complutense de Madrid.

Concluyendo, estos *Textos breves* se confirman sin duda alguna como una de las obras esenciales de Husserl en español. Un voluminoso libro de 716 páginas, a pesar de la “brevedad” que anuncia en la portada, que viene a constatar la buena salud editorial de Husserl en el momento actual. A los *Textos breves* hay que sumar la reciente reedición de los *Problemas fundamentales de la fenomenología* —descatalogada desde 1994— y la publicación de sus lecciones de ética de los años 20 (y pronto llegarán las lecciones de ética de 1908 a 1914). Es preciso saludar con alegría este creciente interés por la fenomenología, lo que es sinónimo del interés por la filosofía en letras mayúsculas, aquella que se sigue preguntando por las grandes cuestiones de la existencia, por la verdad, por el bien, las cuestiones que han hecho de la disciplina filosófica uno de los factores humanizadores y liberadores clave de nuestra historia.

Jesús Miguel Marcos del Cano
Instituto de Filosofía (CSIC)

HINDRICHS, Gunnar (2020). *La autonomía del sonido. Una filosofía de la música*. Traducción de Pedro García-Baró Huarte. Salamanca: Ediciones Sígueme.

Gunnar Hindrichs ha escrito diversas obras como *Das Absolute und das Subjekt. Untersuchungen zum Verhältnis von Metaphysik* (2011), *Philosophie der Revolution* (2017) y *Zur kritischen Theorie* (2020). La reciente traducción publicada por la editorial Sígueme es la versión castellana de *Die Autonomie des Klangs. Eine Philosophie der Musik*, escrito en 2014. El libro está estructurado en prólogo, introducción, seis capítulos y dos índices (nombres y materias).

El prólogo explicita la intención del autor: proponer una ontología de la música, que no es producto de la razón teórica ni práctica, sino de la razón estética. Se establece que la obra de arte es el referente central de esta ontología. El ente estético obra de arte musical va a ser desplegado en una serie de categorías y conceptos elementales.

La introducción comienza con una interpretación de la famosa frase de Sócrates sobre la filosofía como “la mayor de las músicas” (*Fedón* 61a). Desde allí surgen una serie de consideraciones: la relación entre filosofía y música en cuanto ambas tienen por referente lo bello, el trasfondo pitagórico —con la mención de la música de las esferas— que implica un embelesamiento por la verdad y, finalmente, una forma de vida marcada por la religión. Esas ideas propias del pensar antiguo son puestas en contrapunto con las ideas modernas de la filosofía de la música y el pensamiento actual que han perdido aquel orden unitario de la Antigüedad. El pensamiento moderno se caracteriza por la diferenciación de las razones en teórica, práctica y estética, expuestas paradigmáticamente en las tres críticas kantianas. Estos distintos ámbitos tienen sus propios procedimientos y contenidos, responden a ámbitos diferentes del ente.

La música deja de ser una figura teórica que lo abarca todo, para configurarse dentro del ámbito de la razón estética. La concepción clásica observaba a la obra de arte dentro del orden general del cosmos y sometida a sus leyes, pero la modernidad estableció la autonomía de la obra, que obedece a sus propias leyes. El concepto de obra musical, surgido en el siglo XVI, se convierte en el eje sobre el que el autor desarrolla su teoría como principio de la música europea en tanto tipo ideal (concepto tomado de M. Weber). La obra es articulada en su constitución espiritual, esto es, racional, en tanto organizada en un sistema de notas que le otorgan sentido. La filosofía de la música debe esforzarse por hallar la constitución fundamental del ente que objetiva las relaciones estéticamente determinadas entre los sonidos. Se convierte así en ontología de la obra de arte musical. El autor revisa críticamente las posiciones de varios estudiosos analíticos de la ontología musical que no establecen un punto de vista propiamente estético.

El capítulo primero titulado “El material musical” inicia el análisis de la ontología de la obra musical. Destaquemos la forma de escritura de los capítulos: están estructurados en párrafos. Hindrichs comienza por establecer las cualidades propias de la obra musical. Una primera es que está hecha. Esto lo lleva a analizar la concepción aristotélica sobre producción (*poiesis*) y acción (*praxis*) y la lectura de Tomás de Aquino, vinculadas mediante el concepto de trabajo que la tradición hegeliano-marxista entiende como objetivación. De aquí surge que las obras de arte requieren de un material en tanto son productos del trabajo. El material musical, primera característica ontológica de la obra,

es un material receptivo al espíritu. Si bien tiene ya una preformación, a la vez necesita del trabajo elaborativo que lo realice. El autor aborda luego la tesis de Adorno sobre la tendencia del material musical. La realización de una obra tiene lugar dentro de un espacio de posibilidades históricas, en las cuales el material padece un estrechamiento y/o ensanchamiento, vinculado a su vez a la continuidad-discontinuidad en el trabajo compositivo. Aparecen aquí otras dos ideas de Adorno. Una implica comprender el material como espíritu sedimentado, subjetividad objetivada o razón objetivada, de modo que resulta una tangente entre música y sociedad. La segunda idea es la de forma material: un trabajo sobre el material. De este modo, el concepto de material musical —señal distintiva de las vanguardias que fue luego ridiculizado por la posmodernidad— queda establecido como la primera caracterización ontológica de la obra de arte musical.

El segundo capítulo se dedica al sonido musical. En la reflexión sobre el ente obra de arte musical encontramos dos puntos de vista: la pregunta sobre el “qué” (la forma) y el “de qué” (el material), estrechamente vinculados. Este capítulo indaga sobre las categorías de la forma material. La pregunta por el qué es la obra de arte musical tiene una primera respuesta: es algo que suena. El sonido musical queda establecido como primera categoría. Entonces, debe analizarse qué designa un sonido y qué lo califica de musical. Hindrichs se detiene en diversas discusiones modernas-contemporáneas sobre el sonido musical y extramusical —la música concreta, los estudios de P. Schaeffer, los futuristas italianos, la música electrónica, la música formal de I. Xenakis y J. Cage—. Señala dos aspectos que constituyen al sonido musical: puede diferenciarse y puede ser reconocido. Critica las concepciones que no mantienen la autonomía del sonido musical (Strawson, Evans, Scruton),

en tanto establecen la diferenciación entre lo musical y lo no musical desde consideraciones extramusicales. La razón estética procede de otro modo: caracteriza el sonido musical exponiendo su sentido propio, es decir, lo entiende en su autonomía. Entonces, una primera caracterización lo define como ente audible identificable procedente de un sistema de notas que configura su sistema de reglas autónomo. Desarrolla luego los tipos de sonido musical de acuerdo a Lachenmann: sonidos de cadencia, de color, de fluctuación, de textura y de estructura, en contraposición al tradicional esquema de melodía, armonía y ritmo. Estos cinco tipos configuran la fisonomía del sonido.

El tercer capítulo estudia el tiempo musical en tanto los diferentes tipos de sonidos representan diversos ordenamientos temporales. El autor revisa la habitual descripción de la música como arte temporal, ya formulada por Lessing y Leibniz. Su particular condición temporal llevó a Leonardo da Vinci a cuestionar el carácter de obra de la música en tanto efímera. Hindrichs recurre a Schelling a fin de establecer el concepto de tiempo como inherente a la música misma. No se trata de que la música se despliega en el tiempo, sino que es configuradora de tiempo y de este modo permanece en su ser. La necesidad propia del orden de sucesiones en la música es una necesidad estética. Una primera caracterización del sistema de reglas autónomas de la obra musical indica que se trata de un sistema para la configuración del tiempo, de modo que es necesario desplegar la categoría de tiempo musical: una medición musical del tiempo en tanto la música tiene el tiempo dentro de sí. Se destacan tres formas de medición temporal en música: *mensura* (duración de las notas), *tactus* (compás) y *tempo* (movimiento, aire, por ej.: *andante*, *allegro*, etc.). Esta característica de no ser mensurable sino de contener dentro de sí el tiempo, propia del

tiempo musical, es analizada desde la noción de magnitud intensiva de Kant. Esto significa que es una unidad que conlleva un fluir de una continuidad gradual en su aproximación infinita a la propia negación. Entendido así cumple el rol de razón suficiente desde la cual devienen otras magnitudes extensivas: notas, compases y tempos. El autor analiza luego las formas históricas que ha tomado la medida musical: entendida como proceso en devenir en la música entre los siglos XVI-comienzos del XX, como imagen de la eternidad en la música medieval y, finalmente, como forma momento en la música desde 1950 —inversión del devenir—. Al tener al tiempo en sí, la música es libertad frente al dominio del tiempo.

El cuarto capítulo trata sobre el espacio musical. Esta categoría se introduce en función de los argumentos de Hauptmann, quien la considera necesaria pues el tiempo no permite articular la noción de unidad simultánea de la obra, un orden de coexistencias, rasgo que hace a la definición del espacio en cuanto tal según Leibniz. Por otra parte, esta categoría resulta necesaria para la identificación del sonido musical. En tanto cumple estos dos requisitos el espacio musical configura la segunda caracterización del sistema de reglas de la ontología musical. El espacio musical no es un espacio —extramusical— en el que la música suena, sino que se refiere al espacio que ella tiene en sí y es, del mismo modo que el tiempo musical, una magnitud intensiva que genera magnitudes extensivas. Las relaciones entre notas se valen de una noción espacial: el intervalo. El espacio es fundamental en las conceptualizaciones de Schönberg y el serialismo, pero desde la Edad Media tiene un lugar destacado en la reflexión musical en Aureliano de Reome, en el anónimo *Musica Enchiriadis*, en Guido de Arezzo. Hindrichs analiza las tres dimensiones del espacio: anchura (yuxtaposición de

sonidos y direccionalidad), altura (superposición de sonidos) y profundidad (sonidos en su anteposición). Añade luego las diagonales en el espacio, que abundan, por ejemplo, en el contrapunto, y la densidad, es decir, si el espacio está más lleno o más vacío.

El siguiente capítulo se titula “El sentido musical”. Un ente tiene sentido cuando es comprensible tal y como es en su manera de ser. Así, si el qué es de la obra se ha definido como sonido musical constituido por el sistema autónomo de reglas espacio-temporales, el sentido consiste en ser comprensible en la estructura de “algo en cuanto algo”, en lo que el autor toma elementos de Aristóteles. Esto quiere decir, en primer lugar, que los sonidos musicales adoptan una función dentro de la obra. Esta puede caracterizarse por su “ser en otro” (según la expresión escolástica *in alio esse*): el sentido del sonido viene determinado por su relación con otros sonidos y por el entramado relacional en el que se encuentra. Los sonidos resultan de este modo ser consecuentes dentro de la lógica musical. Esta lógica no es una analítica, sino una tópica, en el sentido de que se parece a los argumentos retóricos. La teoría de Riemann entiende los acordes como las funciones de la exposición de una tonalidad y Hindrichs, empleando asimismo otras teorizaciones, presenta esta funcionalidad en términos de armonía tonal a la cadencia. Este desarrollo cadencial genera una resistencia del oyente. Todo esto configura la necesidad estética de los sonidos musicales. Finalmente, desarrolla en este capítulo la relación con los sentidos extramusicales, empleando los cuatro modos de interpretar la Escritura: literal (funcional), alegórico (lo que encomienda creer), topológico (lo que propone en relación con la existencia práctica) y anagógico (lo que hace esperar).

El sexto y último capítulo aborda el pensamiento musical que da cuenta de la validez

estética de la obra. Esta no es una categoría más de la obra musical, pero depende de la relación entre forma y material. El autor analiza la música como un pensar no conceptual. Con Schönberg, comprende que esto significa, en primer lugar, que el pensamiento es método que conduce del desequilibrio al equilibrio, es la construcción metódica de la tonalidad de una pieza. Recurre a los conceptos de Adorno de coherencia (tensión entre forma y material) y homeostasis (reequilibrio de las tensiones). El pensamiento musical es la operación organizativa llevada a cabo en la obra, posee su ser en la forma material de los sonidos, es la formación de la figura sonora misma. La validez estética responde al ser veritativo en los términos de “se da el caso de que”, pero esto no señala un estado de cosas más allá de sí mismo, sino que la obra de arte musical será la que será en la historia de sus interpretaciones. Este es el ser veritativo de la música. El fundamento de todos los sonidos con sentido recae en un sin sentido. En esto el autor sigue a Adorno en la idea de la música como figura del nombre divino. Aunque la obra de arte esté para sus oyentes, no es solo para ellos, sino que se les escapa y permanece siempre distinta de

la historia de sus interpretaciones. Por tanto, contra Adorno, el ser de la obra musical no es devenir, sino que está en devenir.

El libro en su conjunto es una obra compleja. El autor despliega los conceptos recurriendo a filósofos y a teóricos de la música, entrelazando los razonamientos, aportando ilustraciones de la historia de la música, de modo de facilitar al lector una comprensión de los argumentos. No hemos referido en este comentario la totalidad de conceptos de los que se sirve el autor, dado que son múltiples, sino aquellos que caracterizan de manera general la estructura del texto. Ese doble juego entre pensadores filosóficos y críticos de teoría musical es propio de la filosofía moderna en adelante, claro contrapunto de la referencia socrática a la filosofía como la más grande música recordada en la introducción. El libro muestra cómo filosofía y música, separadas en diversas realizaciones de la razón, van deviniendo interlocutoras.

*Laura Carolina Durán
(Instituto de Filología Clásica,
Facultad de Filosofía y Letras,
Universidad de Buenos Aires).*

RIVERA, A. y MATEO, E. (eds.) (2020). *Las narrativas del terrorismo. Cómo contamos, cómo transmitimos, cómo entendemos*. Vitoria, Madrid: Fundación Fernando Buesa y Ediciones Catarata.

La palabra “narrar” proviene del latín *narrare*, derivada a su vez del adjetivo *gnarus*: el que conoce¹; por el contrario, el ignorante, término que comparte la misma raíz, sería aquél que no puede o no quiere narrar: el que desconoce porque nunca supo o bien porque ha olvidado. Por ello el

terrorismo —que ha aniquilado personas y silenciado voces— demanda la *narración*, entendida en tres sentidos interconectados: el compromiso moral con las víctimas, para que no se *ignore* el daño infringido; el compromiso con la sociedad, para que no se repita la violencia y no haya más víctimas y victimarios; y, por último, el compromiso intelectual: abordar lo que sucedió, por qué y cómo comprenderlo.

1 *Diccionario Etimológico Castellano en Línea*, disponible en: <http://etimologias.dechile.net/?narrar> (última consulta: 18.07.2020).

Esta forma de conocimiento —polifónico e interdisciplinar— es la premisa de la que parte el presente libro: *Las narrativas del terrorismo. Cómo contamos, cómo transmitimos, cómo entendemos*. Sus dieciséis capítulos comparten, implícita o explícitamente, los mismos interrogantes: ¿puede permitirse una sociedad ignorar —no narrar— su pasado? Y si se narra, ¿desde dónde hacerlo y por qué? Son preguntas desde luego abiertas, pero no por ello abstractas o relativistas. No son abstractas porque se sitúan sobre todo en la sociedad vasca contemporánea, pero también en el resto del Estado: cómo se ha vivido y sufrido el terrorismo y cómo debería narrarse ahora, si bien las reflexiones éticas y políticas que contiene el volumen son extensibles, por fundamentales, también a otros contextos. Y no son relativistas pues tales interrogantes sólo tienen sentido si están basados en ciertas ideas; ideas sin las cuales no puede existir narración posible pues, como se apunta en el prólogo, “la realidad no se ve en el vacío”.

Dichas ideas podrían sintetizarse en cuatro conceptos clave que, de un modo u otro, atraviesan todos los capítulos del libro: la memoria, pues cualquier narración debe ir de la mano de la conciencia de un pasado colectivo, aunque haya sido, o precisamente por ello, *dolorosamente* compartido; el reconocimiento del daño, pues la convivencia y la posibilidad del perdón exigen de justicia reparadora; la afirmación de la dignidad humana; y la democracia como marco normativo irrebable.

Así, el volumen se abre con un interrogante: *¿Dónde está(n)? La pregunta ineludible y las respuestas del relato*, donde Patxi Lanceros reflexiona, a partir de la poética del Génesis y de la secularización de la Historia, sobre la dimensión moral de todo asesinato. A continuación, en *Narrativas y memoria colectiva: entender cómo*

narramos para entender cómo recordamos, César López Rodríguez expone algunas claves de la memoria individual y la memoria en y del grupo, sus esquemas narrativos y las posibles consecuencias para un presente compartido; por ejemplo, dónde situar las emociones (orgullo, culpa...) en los relatos. En consonancia con ello, María Teresa López de la Vieja reflexiona en el tercer capítulo, *Entender el daño*, sobre los conceptos de comprensión, perdón y justicia, y los matices complejos y necesarios que deben existir entre ellos. El relato de José Ibarrola, en *El estigma familiar*, constituye un ejemplo, precisamente, del alcance directo e indirecto y las dimensiones del daño, tanto desde una perspectiva individual como colectiva. Ahora bien, ¿cómo y dónde, en qué disciplinas y medios, expresar dichas narrativas? El libro aborda tres campos necesarios y conectados: el amplio espacio de lo audiovisual, la representación literaria o gráfica y la transmisión en el ámbito educativo.

Elena López Aguirre analiza en *¿Quosque Tandem, Vasconia?* las ideologías que se filtraron en la música que escuchó toda una generación de la sociedad vasca, las consecuencias de ello, y también los ejemplos de resistencia. Y no sólo la música, sino también el cine: Eterio Ortega Santillana examina en *Cine y terrorismo* las razones de la ausencia y la progresiva presencia en la representación audiovisual del fenómeno terrorista y las claves de su propio trabajo cinematográfico al respecto. De forma complementaria puede leerse el capítulo de Roncesvalles Labiano Juangarcía, *Las víctimas de ETA en el cine y la narrativa literaria*, que recorre la historia de la representación de las víctimas en el cine y la literatura desde el tardofranquismo hasta la actualidad. Y del cine a la televisión: En *Una Historia televisada: ETA y sus víctimas en la pequeña pantalla*, Santiago de Pablo, Virginia López de Maturana y David

Mota Zurdo llevan a cabo un detallado análisis empírico que analiza las razones de la escasez de las producciones televisivas en torno a ETA hasta el cambio siglo. Pero la representación del fenómeno no se agota en las pantallas. El teatro ofrece asimismo una mirada complementaria y distinta; como señala María San Miguel Santos en *El teatro documental y la ruptura de canon*, “el hecho teatral llega al Otro por un canal diferente: el sensitivo, emocional y visceral” y, por ello, tal vez sea la disciplina artística más *radical* en la representación de un fenómeno tan doloroso como el terrorismo. En suma, las narraciones audiovisuales y performativas del daño permiten ampliar sus posibilidades de comprensión, el reconocimiento de las víctimas y las formas de reparación.

Lo mismo, mediante una vía paralela, es lo que ofrece la literatura. Vicente Carrión Arregui revisa en *Literatura y violencia política* las novelas que han abordado, aun de forma oblicua, el terrorismo y las claves psicológicas que hacen de la narración literaria un ámbito de vivencia y comprensión tan necesario. Porque es capaz de romper “el mar de hielo” que a veces inunda las sociedades, tal como señala Luisa Etxenike Urbistondo en el capítulo *Escribir con el terrorismo*, que tiene como objeto examinar las profundas razones de la literatura contra la violencia y el fanatismo. Decía Amos Oz que lo que caracteriza a un fanático es su falta de sentido del humor. Pues bien, este bloque temático se cierra con la contribución de Jesús Zulet Izura: *Humor gráfico contra los terrorismos*, donde expresa, por medio de palabras e imágenes, el poder comunicativo del dibujo y el humor.

¿Y cómo transmitir todo ello a las siguientes generaciones? Los cuatro últimos capítulos abordan la explicación del terrorismo en el ámbito educativo: *La presencia de las víctimas en las aulas vascas*, de Jonan

Fernández Erdocia, examina el programa Adi-adian desarrollado por el gobierno vasco y basado en la transmisión del testimonio de las víctimas al alumnado. En continuidad con ello, Raúl López Romo explica en *La educación ante el terrorismo*, los antecedentes de la propuesta Adi-Adian, el enfoque pedagógico de la misma y algunos resultados reveladores; y Fernando Muga Villate analiza en *Aportación de las víctimas de la violencia política en los centros educativos del País Vasco*, la experiencia entre víctimas y alumnado, su importancia en el cumplimiento de los compromisos educativos, y algunos apuntes para su mejora. Por último, María Lozano Alía expande el horizonte de tal propuesta en *Las víctimas en las aulas en el contexto europeo*, donde reflexiona sobre el programa del testimonio de víctimas en algunos centros educativos europeos, a partir de los conceptos de resiliencia y de evitación del daño. Se trata, en última instancia, de educar en un suelo moral común, nacional o internacional, basado en el respeto por los derechos humanos.

En definitiva, narrar el terrorismo en el momento en que ocurría, según Elena López Aguirre, era como relatar un naufragio mientras te estás ahogando. Ha tenido que transcurrir el tiempo suficiente para que sea posible *narrar, comprender y cuestionar* lo sucedido. El presente libro es un ejemplo de las múltiples formas y perspectivas en que tales acciones pueden llevarse a cabo.

Isabel Roldán Gómez
(Universidad de Salamanca)²

2 Este trabajo se inscribe en el marco de dos proyectos de investigación: Bioética y Ciudadanía VI. Ética Traslacional. MINECO. PID2019-105935RB-I00 (2020-2021), y El desván de la razón: cultivo de las pasiones, identidades éticas y sociedades digitales. MINECO. FFI2017-82535-P

PASCAL, Blas (2020). *Escritos espirituales*. Edición de Alicia Villar. Madrid: Tecnos. 242 pp.

Alicia Villar es una reconocida especialista en diversos campos. Desde luego en el de la Ilustración, sobre el que, entre otros estudios, nos ofreció su magnífica edición y comentario a los escritos de Voltaire y Rousseau, *En torno al mal y la desdicha* (Alianza, Madrid, 1995). También en Unamuno, respecto al cual no solo editó su trabajo inédito *Mi confesión* (Sígueme, Salamanca, 2011; 2ª ed. aumentada en texto y notas, 2015), sino una amplia selección de sus *Escritos sobre ciencia y cientificismo* (Tecnos, Madrid, 2017). Y asimismo en Pascal, autor al que ahora hemos de referirnos, sobre el que tiene ya una dilatada e importante obra. A su temprano *Pascal: ciencia y creencia* (Cincel, Madrid, 1987), siguió la excelente edición de sus *Obras* (Gredos, Madrid, 2012), con un amplísimo y valioso estudio introductorio. Y ahora nos ofrece estos *Escritos espirituales*, seleccionados, en buena parte traducidos y prologados por ella misma.

Esos *Escritos espirituales*, menos conocidos, no por ello tienen menor interés y, además de su intrínseco valor, sirven para poner en relación la evolución espiritual de Pascal con su desarrollo científico, su elaboración filosófica (tal como se expresa ante todo en los *Pensamientos*) o su participación en las disputas teológico-religiosas de su tiempo, según manifiestan las *Provinciales*. Con el fin de seguir mejor esa evolución, Alicia Villar ordena los escritos cronológicamente, conforme a lo que se suele conocer como “período mundano” y primera (en torno a 1646) y segunda (en 1654) conversiones religiosas. Y, apoyándose en la investigación más solvente y reputada (Gouhier, Lafuma, Mesnard quizá

ante todo, como ella misma reconoce en diversos pasajes; también entre nosotros Aranguren o Albiac, aunque no los señalemos a todos), nos introduce, con claridad y pertinencia, en esos escritos. Alicia Villar destaca la exigencia y radicalidad que Pascal requiere, como más tarde hará Kierkegaard, para la vida cristiana, lo que no exime de la alegría que su espiritualidad cristocéntrica también destaca (de ahí que cuestione, en cierto modo al menos, la visión que de Pascal ofreció Kolakowski en *Dios no nos debe nada. Un breve comentario sobre la religión de Pascal y el espíritu del jansenismo*, Herder, Barcelona, 1996).

Sea de ello lo que fuere, los escritos seleccionados incluyen, en primer lugar, tres cartas (a su hermana Gilberta, otra a la propia Gilberta, pero escrita junto a su hermana Jacqueline, y a su cuñado el sr. Périer, a la muerte del padre de Pascal), que revelan importantes y hondos acentos de esa evolución espiritual.

Pero tal espiritualidad se manifiesta especialmente en el *Memorial*, experiencia de carácter místico en el momento de su segunda conversión, de la que Pascal nos dejó ese testimonio, al que solo se pudo acceder tras su muerte, pues lo llevaba oculto en el dobladillo de su chaqueta, y que ha sido objeto de estudio por parte de diversos especialistas, que subrayan su hondura y densidad.

Quizá el menos conocido de esos escritos espirituales sea el *Resumen de la vida de Jesucristo*, descubierto en 1845, y en el que Pascal, siguiendo una antigua tradición de los Padres de la Iglesia y otros antecedentes (como, dentro de la corriente jansenista, el del propio Jansenio o el de Antoine

Arnauld), intenta concordar los diferentes Evangelios, acompañando las paráfrasis de los textos con interpretaciones en las que trata de destacar el valor dramático de las secuencias destinadas a favorecer la meditación.

Al *Misterio de Jesús*, incorporado habitualmente en los *Pensamientos* (aunque no fue su destino inicial; ahí se encuentra la famosa expresión: “Consuélate, no me buscarías si no me hubieras encontrado”), siguen las *Cartas a Charlotte de Roannez*, sobre la que ejerció el papel de director espiritual. En ellas se manifiesta cómo la lucha y el esfuerzo por la conversión se ven acompañados por la alegría, por lo que, se observa en buen estilo pascaliano, “debemos trabajar insistentemente en conservar nuestra alegría que modera nuestro temor y conservar ese temor que modera nuestra alegría”, en la esperanza del cumplimiento de la promesa y de una alegría plena.

La *Comparación de los cristianos de los primeros tiempos con los de hoy*, puede llamar un tanto a engaño, pues, pese al título, la comparación se establece tan solo desde la perspectiva de la significación del bautismo, que, al haberse convertido en práctica habitual realizarlo al poco de nacer, trastoca la condición del cristiano, dado que ya no se recibe instrucción para el mismo ni implica asunción de lo que comporta.

Sobre la conversión del pecador, en donde se manifiesta a veces un fuerte agustinismo, ilumina algunos fragmentos de los *Pensamientos*, como el de la “apuesta”, y subraya la vanidad de perseguir la felicidad

duradera en lo que se aniquilará, aunque el propio Pascal no abandonó del todo, pese a ello, sus investigaciones científicas, quizá siguiendo el consejo paulino de “vivir en el mundo como si no viviéramos en él” (1 Cor, 7, 29-30).

En fin, la *Oración para pedir a Dios el buen uso de las enfermedades* revela las muchas que Pascal sufrió, pero también el intento de que las mismas no arruinen el ánimo, en el convencimiento de que, si la oración no puede obtener la curación misma, sí puede lograr un buen uso de la enfermedad, que revertirá incluso sobre el cuerpo, conforme a lo que ya su hermana Jacqueline había sostenido.

Mas, antes de concluir el libro, Alicia Villar, además de una cronología e indicaciones bibliográficas pertinentes y precisas, ha incluido una bien pensada selección de algunos de los *Pensamientos*, agrupados en torno a determinados temas cardinales. Selección que nos ayuda a vincular la espiritualidad de Pascal con núcleos centrales de sus planteamientos filosóficos, y que estimo muy acertada. No es el menor de los méritos de esta preciosa y cuidada edición de los *Escritos espirituales* de Pascal, cuyas reflexiones —por más que podamos disentir en ciertos momentos de ellas— siempre tienen la ventaja de ayudarnos a salir de la costra de la banalidad y enfrentarnos, sobrecogidos y admirados, a nuestra humana condición.

Carlos Gómez
(UNED)

BRENTANO, Franz C. (2019). *Vermischte Schriften*. Sämtliche veröffentlichte Schriften, Band IX. Mit einer Einleitung von D. FISETTE. Herausgegeben von T. BINDER und A. CHRUDZIMSKI. Berlin/Boston: Walter de Gruyter.

Desde 2008 Thomas Binder y Arkadiusz Chrudzimski están contribuyendo a desagraviar la memoria de Franz C. Brentano, el cual debía de ser uno de los pocos filósofos renombrados de habla alemana que no habían sido honrados en el siglo XX con una colección de obras completas. Este proyecto es algo más modesto, pues los editores se conforman con recoger las obras publicadas por Brentano durante su vida. Se omite así una ingente cantidad de material manuscrito que dejó sin publicar y que ha ido viendo la luz el siglo pasado: algunas de esas ediciones de inéditos son excelentes —como las de Hedwig, George o Baumgartner—, pero las más antiguas —como las realizadas por Kastil y Meyer-Hillebrand— siguieron procedimientos muy discutibles.

En estas *Sämtliche veröffentlichte Schriften* han aparecido ya casi todas las obras de Brentano. Los editores las han distribuido en cuatro grupos: escritos sobre psicología, sobre ética y estética, sobre Aristóteles y otras publicaciones. En total, el proyecto consta de nueve volúmenes ocupados cada uno de ellos por uno de los grandes libros, si bien los opúsculos de menos páginas publicados de manera independiente (como los que tratan de estética o sobre Aristóteles) son reunidos en un solo volumen. Aún han de aparecer tres tomos sobre Aristóteles y el último que estaría dedicado a textos breves no propiamente filosóficos.

El volumen que aquí reseñamos es el noveno, el primero que recoge obritas aparecidas en distintos medios. Viene prece-

didada de una introducción de Denis Fissette, que precisamente ha estado trabajando estos años con Guillaume Fréchette en una magnífica recopilación en que son traducidas al francés estas obras y algunas más: Brentano, F. (2018-2021), *Essais et conférences*, Paris: Vrin.

Los escritos, de variada naturaleza, están impresos en orden cronológico, abarcando desde el inicio de su carrera en 1867 hasta cerca del final de su vida, 1908. Podemos encontrar, en primer lugar, varios que tratan sobre la naturaleza de la labor filosófica, entre los que se pueden incluir las tesis latinas de habilitación, y los tratados “Las razones del desaliento en filosofía” y “El porvenir de la filosofía” que son bien conocidos para el lector español merced a la traducción de X. Zubiri en 1936 y la reciente reedición de Juan Miguel Palacios en 2011 para la editorial Encuentro. A estos trabajos, se debe vincular el tratado “Mis últimos deseos para Austria” de 1894-1895, asociado con su propio itinerario personal en la Universidad de Viena.

Otras obras se ocupan de pensar sobre la naturaleza de la historia de la filosofía misma, entre las cuales destaca “Las cuatro fases de la historia de la filosofía”, también traducida por Zubiri en 1936 dentro de la colección titulada *El porvenir de la filosofía*, Madrid: Revista de Occidente. A ella se suma “Qué clase de filósofo hace época”, sobre Plotino, y “Augusto Comte y la filosofía positiva”, muy importante para comprender la formación del pensamiento personal de Brentano. Asimismo, se encuentra el artículo “Tomás de Aquino”

que yo mismo traduje y completé con notas en 2016: *Espíritu*, 65, pp. 539-555.

El tercer grupo de obras son más coyunturales. Entre ellas destaca “El ateísmo y la ciencia”, publicada anónimamente en 1873 como réplica a un artículo también anónimo aparecido poco antes. Este escrito había sido identificado ya por Tiefensee, E. (1998), *Philosophie und Religion bei Franz Brentano (1838-1917)*, Marburg: Francke, p. 80. Se trata de una defensa del teísmo contra la manida acusación de mantener una presunta alianza con la ignorancia y el oscurantismo. Al lector español le agrada encontrar citado por Brentano, en traducción alemana, nuestro soberbio *Soneto a Cristo crucificado*, de autor anónimo (p. 57). A esta obra se unen una reseña sobre un libro acerca del conocimiento de Dios en santo Tomás, una réplica a una reseña de Horwicz en que criticaba su *Psicología desde el punto de vista empírico*, un comentario sobre un libro de Miklosisch sobre la naturaleza del juicio y un artículo acerca del control del Estado de las cátedras universitarias: “Sobre la investigación sin presupuestos”.

Yo habría impreso también aquí el primer artículo publicado por Brentano (1867), acerca de la *Historia de las ciencias eclesiásticas*, el cual habría constituido muy buen acompañamiento para las páginas sobre santo Tomás y para *Las cuatro fases de la historia de la filosofía*. De

hecho, así procedieron Fisette y Fréchette en la edición francesa antes mencionada. No obstante, Binder ha preferido dejarlo para el último volumen por no considerarlo un escrito estrictamente filosófico.

El libro está tan bien presentado como los anteriores, aunque aparecen algunos errores editoriales menores, como las erratas en el encabezado de las pp. xxxix y xli. Algo más perjudicial lo constituye el que en la introducción de Fisette no correspondan los números de página citados con la edición de las obras: por ejemplo, en p. xxvi se cita p. 30 pero en realidad se refiere a p. 28; en p. xxxi se menciona p. 253 pero el texto citado aparece en pp. 240-241, etc.

En todo caso, es estupendo disponer en un solo volumen de estos escritos de Brentano que andaban dispersos en distintas publicaciones, algunas de muy difícil acceso. La introducción de Fisette que precede a los escritos resulta muy útil e ilustrativa y Binder proporciona toda la información requerida acerca de la primera publicación de estos opúsculos. Seguiremos, pues, con agradecida atención, la publicación de los volúmenes por aparecer en esta colección.

David Torrijos Castrillejo
(Universidad Eclesiástica San Dámaso.
dtorrijos@sandamaso.es)
(<https://orcid.org/0000-0003-2005-5634>)

CRETU, Ana-Maria; MASSIMI, Michela (ed.) (2020). *Knowledge from a human point of view*, Springer.

Knowledge from a human point of view, editado por Ana-Maria Crețu y Michela Massimi y publicado por la editorial Springer, es una obra colectiva resultado del esfuerzo que desde hace algunos años están llevando a cabo filósofas y filósofos de la Universidad de Edimburgo por reivindicar la importancia de la noción de *punto de vista* o del *perspectivismo*¹. Fruto de un esfuerzo análogo han surgido también en nuestro país obras referencia como *Puntos de vista* (2013), editado por Manuel Liz y publicada en la editorial Laertes, *Temporal points of view* (2015), editado por Margarita Vázquez y Manuel Liz o *Formal Approach to the Metaphysics Perspectives* (2018) de Juan J. Colomina-Almiñana, ambas publicadas en Springer. Además, desde Finlandia Antti Hautamäki ha publicado en Springer *Viewpoint Relativism: A New Approach to Epistemological Relativism based on the Concept of Points of View* (2020). Estas investigaciones recientes intentan rellenar un espacio que, quitando las aportaciones de Leibniz, Wittgenstein, Ortega y Gasset o William James, entre otros, sigue incomprensiblemente vacío a pesar de su potencial filosófico. Si bien la obra que se reseña aquí parte de este contexto general, su particularidad reside en que sus análisis sobre el perspectivismo se desarrollan en el marco de la *filosofía de la ciencia*, aportando contenido novedoso y complementando con análisis históricos y epistemológicos algunas de las líneas ya trazadas

en *Understanding Perspectivism. Scientific Challenges and Methodological Prospects* (2019), otro libro colectivo también editado por Michela Massimi y Casey D. McCoy y publicado en Routledge, Aunque son obras distintas, se recomienda acercarse a ellas por orden cronológico, ya que facilitaría una mejor comprensión de sus planteamientos.

El perspectivismo ha surgido recientemente en la filosofía de la ciencia como una nueva variante en el debate entre el realismo científico y el anti-realismo. Presenta nuestro conocimiento científico como algo histórica y culturalmente situado. Como la propia Massimi comenta en la introducción, dicho conocimiento es el resultado de toda una serie de praxis modernas, teorías científicas, técnicas experimentales y recursos conceptuales inherentes a “perspectivas científicas” específicas que nosotros, como sujetos epistémicos situados históricamente, podemos ocupar. Por ejemplo, cuando nos referimos a la “perspectiva newtoniana” estamos de algún modo especificando el punto de vista particular desde el cual se hacen típicamente las afirmaciones de conocimiento. Desde este punto de vista, la filosofía de la ciencia debe plantearse algunas cuestiones interesantes: ¿es posible un conocimiento del mundo tal como es (en oposición al conocimiento del mundo visto a través de nuestras perspectivas)? ¿Hay una concepción absoluta de la realidad? como diría Bernard Williams o ¿en qué medida afecta el perspectivismo a la noción misma de conocimiento (en cuanto a la verdadera creencia justificada, según el punto de vista tradicional) si la

1 Trabajo realizado en el marco del proyecto de investigación “Personal Perspectives”: Concepts and Applications, FFI2018-098254-B100, Ministerio de Economía y Competitividad, Gobierno de España

justificación y la verdad son dependientes de una perspectiva?

Esta interesante obra está dividida en dos partes y cuenta con nueve ensayos. La primera de las partes gira en torno a ideas fundamentales de pensadores clásicos como Kant, Nietzsche, Putnam o los pragmatistas norteamericanos. La segunda parte aborda cuestiones relativas a la relación entre el perspectivismo, el relativismo y la noción de puntos de vista, y las implicaciones del perspectivismo en los debates epistemológicos sobre la normatividad epistémica y los fundamentos del conocimiento humano. En general, la investigación trata de arrojar luz sobre el origen de la idea del conocimiento desde el punto de vista humano en la historia de la filosofía y qué papel juega la idea en los debates contemporáneos sobre epistemología. En el capítulo 1, *Attempting to Exit the Human Perspective: A Priori Experimentation in Kant's Critique of Pure Reason*, Rachel Zuckert tratar de esclarecer a través de Kant lo que identifica como una tensión inherente a la idea misma del conocimiento desde un punto de vista humano. Kant sostiene que sólo se puede obtener el conocimiento del mundo desde la perspectiva humana. Sin embargo, el reconocimiento de este hecho requiere que uno sea capaz de salir de la perspectiva humana y reconocer la existencia de otras perspectivas, lo cual Kant parece negar. En el capítulo 2, *Nietzsche's epistemic perespectivism*, Steven D. Hales, analiza dos tipos o niveles de perspectivismo en Nietzsche. Por un lado, uno que abarca una teoría epistémica de primer orden que toma la verdad como perspectiva y, por otro lado, un perspectivismo metodológico de segundo orden, dirigido a mejorar la "comprensión". La "epistemología positiva" de Nietzsche que defiende aquí Hales responde tanto a los críticos que perciben

a Nietzsche como un escéptico como a los que lo han interpretado como un pragmatista. El capítulo 3, *Pluralism and Perspectivism in the American Pragmatist Tradition*, Matthew J. Brown muestra el impacto de la noción de conocimiento situado desde un punto de vista humano en los trabajos de los pragmatistas norteamericanos. Brown plantea un panorama general a través de una extensa literatura y analiza la metafísica pluralista en la tradición pragmatista. Su investigación revela ciertas deficiencias del perspectivismo, como un potencial colapso hacia el relativismo, o un estrecho enfoque eurocéntrico en la ciencia. Brown sugiere que estas deficiencias pueden superarse si los perspectivistas estuvieran dispuestos a integrar ciertas lecciones sobre la verdad, la realidad y la pluralidad de los pragmatistas norteamericanos. En el capítulo 4, *Hilary Putnam on Perspectivism and Naturalism*, Mario De Caro, critica la concepción de Putnam sobre la realidad y el conocimiento, y en particular su visión sobre la posibilidad de sostener un tipo de realismo. De Caro comienza su estudio con las opiniones de Putnam sobre el fisicalismo y su crítica al realismo metafísico. Seguidamente, aborda el realismo interno de Putnam, que, según De Caro, se inspiró en parte en Kant, Peirce y Dummett, y fue motivado por un renovado esfuerzo por responder al realismo metafísico. Finalmente, Putnam abandonó el realismo interno en detrimento del "naturalismo liberal", una variación que De Caro identifica como afín al realismo perspectivista de Michela Massimi. En el capítulo 5, *Scientific Perspectives, Feminist Standpoints, and Non-Silly Relativism*, Natalie Alana Ashton observa algunas de las problemáticas planteadas en los otros capítulos desde la teoría del punto de vista feminista contemporáneo. Argumenta que tanto el perspectivismo como la teoría del

punto de vista feminista tienen mucho que aprender del relativismo, así como el uno del otro. Además, identifica elementos del relativismo inherente al perspectivismo y a la teoría del punto de vista de Ronald Giere. Para Ashton, hay una versión inocua del relativismo que puede beneficiar a ambos puntos de vista. En el capítulo 6, *Perspectives, Questions, and Epistemic Value*, Kareem Khalifa y Jared Millson muestran un punto de vista que denominan “monismo de verdad inquisitiva”, según el cual no sólo las verdaderas creencias tienen un valor epistémico, sino también las verdaderas respuestas a las preguntas pertinentes. Según Khalifa y Millson, es la perspectiva del investigador o investigadora la que determina qué preguntas son pertinentes, dicha perspectiva también abarca sus intereses, su función social y sus supuestos de fondo. La principal motivación de Khalifa y Millson en la búsqueda de un monismo inquisitivo y verdadero es, en sus propias palabras, explicar “la complejidad de las empresas de valor epistémico que caracterizan el esfuerzo científico”. En el capítulo 7, *Perspectivalism About Knowledge and Error*, Nick Treanor aborda una discusión de una visión generalizada sobre la normatividad epistémica que toma la verdad como una norma clave para las creencias. En este punto de vista, también compartido por Alvin Goldman o Ernesto Sosa, saber es creer la verdad y evitar el error. Treanor señala los problemas de esta concepción de la normatividad epistémica y sugiere una forma diferente de pensar sobre la normatividad epistémica. En el capítulo 8, *Virtue Perspectivism, Externalism, and Epistemic Circularity*, Adam Carter también hace referencia a Sosa, considerando al “perspectivismo virtuoso” como una postura lógica epistemológica de dos niveles, en la que la fiabilidad del conocimiento

animal de primer orden requiere un ascenso al conocimiento reflexivo o perspectivo de segundo orden. El objetivo de Carter en este capítulo es desentrañar las críticas y defender el perspectivismo de la virtud de Sosa de las objeciones basadas en la circularidad que se ha suscitado. Finalmente, Barry Stroud en el capítulo 9, *Knowledge from a Human Point of View*, reanaliza la noción misma de “conocimiento desde el punto de vista humano”. Investiga las formas en que los seres humanos llegan a conocer y lo que significa para nosotros llegar a conocer algo, Stroud también se hace cargo de los desafíos escépticos que existen en torno a la posibilidad de conocimiento y las preocupaciones generales sobre relación entre el conocimiento y la verdad en un realismo perspectivista. Afirma que ocupar un “punto de vista humano” es estar plenamente comprometido con la comunidad de concededores humanos y estar comprometido con el ser del mundo de la manera que es ampliamente conocido.

En definitiva, *Knowledge from a human point of view*, es una obra tremendamente interesante y sugestiva. A pesar de ser conjunta mantiene un tono homogéneo basado en la claridad de su escritura. Respecto al contenido, su virtud reside precisamente en todo lo contrario; en una clara pluralidad respecto a las maneras de abordar el impacto del perspectivismo en las cuestiones epistemológicas. Dado que su contenido se relaciona de múltiples maneras con algunas tendencias desarrolladas en nuestra sociedad actual —por ejemplo, un relativismo respecto de la verdad o la peligrosa propensión de elevar a *episteme* la opinión—, me atrevería a decir que estamos ante una de esas lecturas obligatorias.

Abraham Hernández
(Universidad de La Laguna)

ARAGÜÉS, Juan Manuel (2020): *De idiotas a koinotas. Para una política de la multitud*. Madrid: Arena Libros.

En su último libro, Juan Manuel Aragüés retoma, para profundizarla, una reflexión que ya había empezado a perfilar en su *Deseo de multitud* (2019) sobre la urgencia de construir un sujeto político antagonista que sea capaz de hacer frente a esta nueva configuración —depredadora— del capitalismo que lleva por nombre «neoliberalismo». Levantando acta de la «insuficiencia planificada» de la democracia parlamentaria para poner freno a la «pulsión suicida» de este nuevo régimen de acumulación, cuya voracidad ya no conoce límites, especialmente —como ha mostrado Piketty en *El capital en el siglo XXI* (2013)— desde la disolución del socialismo burocrático y autoritario de Estado, y cuyo modo de funcionamiento se puede entender a la luz de este proceso tematizado por Marx en *El Capital* bajo el rótulo de «subsunción real del trabajo en el capital», se le antoja a Aragüés que la lucha contra el capitalismo neoliberal y su lógica de apropiación privativa de los comunes consiste ante todo en un «combate por la democracia»: pues se trata de (re)conquistar la soberanía popular, de devolver al pueblo su voz y palabra; en suma, de restituir a la ciudadanía este poder de decidir y deliberar colectivamente sobre el futuro que le ha sido arrebatado por todo un dispositivo jurídico, político y mediático (tratados, forma-partido, medios de comunicación, etc.) que, al operar como codificador ideológico y orden institucional de lo que Laval y Dardot, en *La nueva razón del mundo* (2013), han denominado la «racionalidad neoliberal», parece haber sembrado por doquier el «silencio plebeyo» y convertido así al ciudadano de a pie en una suerte de «Tersites [del] siglo XXI».

Sin embargo, acontecimientos como la Primavera Árabe, *Nuit Debout*, *Occupy*

Wall Street o el propio 15-M (con su célebre eslogan: «¡No nos representan!») están aquí para recordarnos que el deseo de «democracia real ya», lejos de haberse esfumado, sigue latiendo en el seno de la ciudadanía y que resulta preciso «materializar el acontecimiento», el anhelo de transformación social, si lo que se pretende es que la democracia, de meramente formal, pase a ser plenamente efectiva; o lo que, dicho con la terminología spinoziana, viene a ser lo mismo: que la ciudadanía, de *plebs* o *vulgus* (muchedumbre), pase a convertirse en *multitudo* (multitud). Esto implica reconocer ante todo que democracia y conflicto, lejos de ser antinómicos, son más bien términos coextensivos: pues frente al giro deliberativo de la filosofía política, encarnado por autores como Rawls o Habermas, los cuales tienden a reducir la democracia a un mero proceso de pacificación de los conflictos, a promover correlativamente una cierta concepción del consenso (como acuerdo razonado entre voluntades libres) y por ende a reconducir las «falacias del liberalismo» que le hacen el juego al capitalismo postfordista, cabe destacar el carácter profundamente agonístico de la misma y comprender por consiguiente que el conflicto, como tiene a bien señalar Lazzeri en *Conflits et démocratie. Quel nouvel espace public?* (2010), no designa otra cosa que no sea la «antinomía constitutiva de la democracia». Una democracia a la que, por otra parte, sería un craso error achacar —como desgraciadamente hace Badiou en tono platónico— una serie de características (infantilización, embrutecimiento, egoísmo, etc.) que son más bien las que le confiere el capitalismo en «su forma consumista neoliberal». Para Aragüés —que se vale aquí

de las puntualizaciones de Sousa Santos—, antes de abandonar el concepto de democracia se trata más bien de resignificarlo, ya que en última instancia el objetivo es engendrar, no una alternativa a la democracia, sino una «democracia alternativa».

Ahora bien, la construcción de una democracia alternativa, que dé cabida a voces discordantes, restituya a la ciudadanía su correspondiente soberanía y se inscriba en una estrategia de defensa de lo común, es a su vez indisoluble —insiste Aragüés— de la construcción de un sujeto político alternativo que sea lo suficientemente «plural, abierto y múltiple» como para poder aunar en su seno las diferentes reivindicaciones de las que dichas voces son portadoras. Punto en el que nuestro autor se ve conducido a polemizar, por un lado, con los promotores del populismo, y por otro, con los críticos de las políticas de la diversidad (feminismo, ecologismo, postcolonialismo, etc.), es decir, con quienes entienden dichas políticas como una «trampa» insidiosamente tendida por el neoliberalismo para acabar con el marxismo. Con los promotores del populismo, porque si bien Aragüés comparte con ellos la idea de desarrollar una política lo más transversal posible, no por ello deja de considerar que el de «populismo» es un «concepto tóxico» por cuanto recuerda a formas de poder no democráticas, además de carecer de una amplia base social que se identifique con él. Y con los críticos de las políticas de la diversidad, porque la lucha económica atinente a la «posición de clase», por muy importante que sea, no es en absoluto la única lucha que el sujeto colectivo antagonista deba emprender y con la que haya de identificarse exclusivamente: pues si no cabe duda de que existe cierta tendencia a «glorificar la diferencia por la diferencia» —tendencia que, al fragmentar las luchas, contribuye indiscutiblemente a mermar la expresión política

unitaria de la multitud—, resta que no se da solución a este problema embarcándose en la estrategia inversa consistente en reificar identidades bien circunscritas. En efecto, como bien recuerda Aragüés: «las identidades [sociológicas] no implican posición política»; y agrega acto seguido: «ser obrero no implica ser revolucionario».

Si el combate por una democracia alternativa, como sostiene Aragüés, pasa pues por la articulación de una «política de la multitud», cuya potencia —y, por tanto, cuyo derecho a decidir sobre su futuro (como bien recoge la fórmula spinoziana: *tantum iuris quantum potentiae*)— será tanto mayor cuanto más amplio sea su espectro, entonces entendemos por qué nuestro autor considera que la trampa a evitar, de cara a dicha empresa, no estriba tanto en los discursos (postmodernos y «deleuznables») de la diferencia como en los (dogmáticos) de la identidad. Unos discursos que, al pecar —como muestra Aragüés— de un «idealismo conceptual» no confesado, terminan las más de las veces por degradarse en eslóganes identitarios y esencialistas. En efecto, del intento por caracterizar al sujeto colectivo antagonista mediante una combinación fija de rasgos distintivos (compartir espacio en la fábrica, tener las manos manchadas de hollín, etc.) se deriva, a buen seguro, la marginación, cuando no el acallamiento de las diversas reivindicaciones y luchas «minoritarias» (contra el patriarcado, el racismo, etc.) que, sin ser necesaria o estrictamente económicas, no dejan por ello de recorrer el espacio social y político. Piénsese al respecto, y para no citar más que un ejemplo, en la errada actitud del Partido comunista francés quien, en Mayo del 68, se negó inicialmente a secundar las revueltas estudiantiles por considerarlas propias de «burgueses» o «pequeñoburgueses». Al contrario —sostiene Aragüés— participa de la construcción

del sujeto colectivo antagonista quien, no en virtud de su identidad o posición sociológica, sino *en y por* su concreta práctica política, se rebela contra el poder constituido. Así, dice nuestro autor: «clase es todo aquel, aquella, que se une en la lucha contra el capital, independientemente de que sea estudiante, autónoma o trabajadora de fábrica, mujer es toda aquella, aquel, que se enfrenta al patriarcado, minoría étnica, todos aquellos que luchan contra el racismo». Y de concluir: «multitud, o pueblo, es el nombre que podemos otorgar a la coordinación de esas diversas luchas».

Comprender, como hace Aragüés en un gesto que recuerda a lo teorizado por Deleuze y Guattari en *Mil mesetas* (1980), que el sujeto colectivo antagonista, la multitud, consiste en la articulación y convergencia de una multiplicidad de luchas debe llevarnos a entender, en las antípodas de todo posicionamiento identitario y esencialista, que este mismo sujeto político es forzosamente «plural en su constitución». De modo que «el problema, tal como arguye nuestro autor, no es la diversidad, la diferencia, sino la forma cómo la entendemos y gestionamos»; lo cual implica reconocer y asumir de una vez por todas que de esta pluralidad originaria de la multitud se derivan necesariamente contradicciones con las que hace falta saber componer —de ahí la importancia crucial que Aragüés ya atribuía en su *Deseo de multitud* a los actos de escucha y traducción con el objetivo de minimizarlas—, y que la búsqueda a todo costa del consenso unánime, por loable que sea, no desemboca sino las más de las veces en una «eternización» de los debates que tiende a mermar la vivacidad del movimiento. Si el discurso de la identidad constituye, pues, la verdadera trampa que obstaculiza la construcción de esta subjetividad política alternativa, es porque no hace otra cosa que

no sea recaer en esta «política idiota (de *idion*, particular, propio)» con todo lo que esto implica en términos de sectarismos y dogmatismos, como dio —y continúa buenamente dando— muestras de ello cierto marxismo ortodoxo, mecanicista y economicista (cuya forma más acabada fue sin duda ese proyecto de «enterrar a Marx» que llevaba por nombre «estalinismo»).

Para poner dique a la tentación de subsumir las diversas reivindicaciones en la sola lucha por la posición de clase, Aragüés se vale de lo mejor de la tradición materialista con el propósito de formular un diagnóstico «realista» tanto del mundo social y político al que se ambiciona transformar, cuanto del sujeto que será llamado a hacer las veces de agente transformador de dicho mundo, la multitud. En efecto, «construir el sujeto político desde la conciencia de la diferencia de sus elementos constituyentes es el empeño primero que ha de abordar toda política materialista»; o, como diría Althusser, el primer paso que hace falta dar para «no contarse cuentos» (acerca de pretendidas identidades sociológicas que esperarían pasivamente a ser activadas) y para no caer preso, por consiguiente, del *wishfull thinking*. A este fin, Aragüés se vale de la ontología deleuziana para recalcar que lo primero en el orden del pensamiento, la acción y el ser no es la identidad sino la diferencia, pues como dice Deleuze en *Diferencia y repetición* (1968): «la diferencia está detrás de todo, pero no hay nada detrás de la diferencia». Lo que se traduce en el plano de la realidad social por el hecho de que, pese a la acción homogeneizante de las estructuras vigentes, no deja de haber una multiplicidad de individuos plurales, diferentes y con perspectivas distintas. Concepción del sujeto o individuo que ya se puede encontrar —arguye nuestro autor— en los primeros escritos de Marx por cuanto éste, lejos de

hacer de la posición de clase la esencia del individuo, recalca más bien, como en sus *Tesis sobre Feuerbach*, que es el «conjunto de las relaciones sociales» lo que determina dicha esencia. En este sentido, la construcción del sujeto colectivo antagonista se debe llevar a cabo desde la conciencia de que su «materia prima», por así decirlo, son individuos *sociales*, es decir, individuos cuya conciencia está determinada, no por su sola posición de clase, sino por la totalidad de la vida *social* en la que se despliega su existencia — totalidad abierta y fragmentaria de la que, para expresarlo con la terminología deleuziana, dichos individuos no son sino *pliegues* diferenciados, irreducibles por lo tanto a una pretendida identidad previamente dada.

Al proceder de este modo, Aragüés se inscribe ciertamente en una estrategia de descentramiento del marxismo al que ya contribuyeron, para no citar más que ellos, autores como Lukács con su obra póstuma *Marx, ontología del ser social* (1972), Sartre con su *Crítica de la razón dialéctica* (1960) o Fanon con *Los condenados de la tierra* (1961), cuando levantaban acta de la aparición de nuevas reivindicaciones y luchas (como el feminismo, el ecologismo o el anticolonialismo). En pocas palabras, de lo que se trata para Aragüés es, como bien señala José Luis Rodríguez en el prefacio a la presente obra, de operar un «desencapsulamiento de la identidad», en la que ciertos sectores intentan encerrar al sujeto político, para así poder allanar el camino a una «recepción de la(s) diferencia(s)»; vale decir, para convertir la multitud en una suerte de plataforma de encuentro, interseccionalidad y coordinación de todas las luchas en contra del orden establecido. Henos aquí, pues, con un claro llamamiento a desarrollar una estrategia política de transversalidad y de encuentro de las diferencias en pos de lo

común; vale decir, una «política koinota» (de *koinon*, lo común) que, según Aragüés, es la única capaz de empoderar a la multitud en su lucha por desgastar y acabar, si no con el capitalismo, al menos sí con su configuración neoliberal. Plural en su constitución, la multitud ha de ser —lo habremos entendido— «unitaria en su expresión»; lo cual apela forzosamente la puesta en marcha de nuevas formas organizativas, como bien reza el lema del que se hace eco nuestro autor: *estonoesunpartido*.

En efecto, si el propósito del sujeto colectivo antagonista, especialmente bajo el impulso de acontecimientos como los antes citados, es hacer frente al neoliberalismo y revertir el secuestro de la democracia orquestado por los dispositivos jurídico, político y mediático afines, entonces —señala Aragüés— no le será posible a la multitud «materializar el acontecimiento» sin «desterritorializar la organización» que debe respaldarla. Desterritorialización organizativa que, sin implicar la supresión de todos los procedimientos de corte representativo existentes (vuelto necesarios algunos por la complejidad misma de la comunidad koinota), pasa por la inyección de una buena dosis de *horizontalismo* (sorteo, revocabilidad, control y rotación de los cargos), cuyo cometido es, por así decirlo, doble: por un lado, se trata de mitigar, cuando no de impedir lo que Sartre ha denominado la «institucionalización», es decir, este momento en el que el grupo se olvida de sus objetivos iniciales para tomarse a sí mismo como fin último; o, si se prefiere la terminología negriana, este momento en el que el «poder constituyente» se degrada en «poder constituido» (con todo lo que esto implica en términos de reactivación de la forma-partido, de restauración de una visión idiota y de desvanecimiento del acontecimiento). Por otro lado, y como corolario de lo que

precede, se trata de empoderar el deseo (el *conatus*, diría Spinoza) así como los afectos comunes que impulsan a la multitud hacia la transformación social, es decir, los resortes pasionales a defecto de los cuales la política entendida como «trabajo de la imaginación volcada hacia el futuro», o «transformación de lo imposible en posible», deja de ser operativa. Porque de lo que se trata en última instancia es de delinear el futuro, por definición abierto e incierto. Este *porvenir*, o *por-ser*, como lo llaman respectivamente Negri y Castoriadis, acerca del cual la teoría, instrumento de comprensión del presente (lo

que es) y del pasado (lo que fue) de la sociedad que se ambiciona transformar, se revela empero ineficaz a la hora de decirnos de qué estará hecho (lo que todavía no es). *Ars imaginandi*, *ars affectandi*: éstas son, pues, las dos alas constitutivas de esta «política de la multitud», o política *koinota*, que tiene por cometido delinear y quizá —¿quién sabe?— materializar este «fantasma» que, tal como decían Marx y Engels en *El manifiesto comunista*, «recorre Europa».

Julien Canavera
(Universidad de Zaragoza)

ASSMANN, J., CACCIARI, M., DUQUE, F., PONS, J. y VELA, A. (2020). *La palabra que falta. El dios indecible de Moisés y Aarón, de Schönberg*. Madrid: Ediciones CBA.

Cinco pensadores de la talla de Jan Assmann, Massimo Cacciari, Félix Duque, Jordi Pons y Juan Ángel Vela, (los dos últimos también prestigiosos musicólogos) reflexionan en *La palabra que falta* sobre la conocida ópera del dodecafónico Arnold Schönberg, *Moisés y Aarón*. Escrita entre mayo de 1930 y marzo de 1932, cuando el antisemitismo era ya un fenómeno extendido e innegable, dicha ópera forma parte de una serie de obras que abordaron la figura de Moisés en esos años, entre las que destacan *Moisés y la religión monoteísta* (1939) de Sigmund Freud y las novelas bíblicas de Thomas Mann. Schönberg, reconvertido al judaísmo tras un periodo de confesión luterana, aborda en los dos actos de esta inacabada pieza el complejo diálogo teológico de los textos sagrados hebreos.

Dos personajes configuran la tensión entre la idea monoteísta del dios invisible y aquellos elementos politeístas del judaísmo, como los holocaustos, todavía presentes en su origen. Moisés es el inspirado, con-

vencido de la existencia de un dios invisible, irrepresentable e inefable, que otorga la nueva Ley. Pero la palabra le falta para convencer a un pueblo que no termina de digerir su innovación ni de fiarse de un dios al que no puede representar. Sus “dificultades de habla”, que son mencionadas en la Biblia y Freud atribuye a su origen no judío, son representadas por Schönberg magistralmente a través del dodecafonismo, en una forma de discurso semi-cantado muy distinto a la armónica y acompasada voz de Aarón. Éste último posee lo que su hermano carece: la persuasión, la palabra fácil, los milagros y las imágenes. Aarón sabe que el pueblo necesita algún tipo de prueba en el mundo de lo sensible; una constatación, en definitiva, del *amor* entre el pueblo y su dios, en lugar del puro intelectualismo de la Idea y un rigorismo ético sin recompensas palpables.

La tensión de la obra explota al final del segundo acto. El pueblo impaciente, hambriento y dudoso del dios invisible de

un Moisés que no aparece, urge a Aarón a construir algo que ver, que adorar. Éste forja el becerro de oro y se suceden festejos orgiásticos y sacrificios de animales. Cuando desciende Moisés de la montaña, recrimina a su hermano el pecado de idolatría, pero la lógica de Aarón es implacable: “también son imágenes” las tablas de la Ley, ella misma es “un milagro” (Cacciari.) Moisés rompe entonces las tablas y queda sólo con su idea inexpresable e incomprendida, exclamando: “¡oh tú, palabra que me faltas!”.

La obra no tuvo tercera parte, aunque en esta recopilación de ponencias, editadas y ampliadas para la ocasión, se especula mucho con un final a partir de anotaciones y apuntes de Schönberg. Los autores reflexionan sobre viejos y nuevos problemas a partir de la ópera, muchos de ellos de actualidad manifiesta. De entre todos ellos cabe enumerar algunos:

1) El egiptólogo Jan Assmann incide en los *peligros políticos* tanto del pensamiento solitario de Moisés, representante del rigor de la Idea de la Modernidad y el desapego hacia la realidad material, como de la divinización del poder del nacionalsocialismo, encarnado en el etnoteísmo de Aarón, con su palabra seductora y vacía. Imposible no pensar aquí en las tensiones entre política y filosofía, o en los movimientos nacional-populistas de hoy en día y en la forma en que utilizan las imágenes. Una pregunta que subyace es : ¿debe el líder meramente complacer los deseos de la masa (previa toma de temperatura a través de encuestas, etc.) o intentar reformarla a través

de exigencias éticas probablemente impopulares?

- 2) Por otro lado, el libro aborda el problema de los límites del lenguaje, tan tratado por pensadores de la época como Ludwig Wittgenstein, Karl Krauss, Walter Benjamin, Robert Musil y por supuesto Arnold Schönberg. Ante lo místico, ¿cabe el silencio o la palabra? La melodía, en el sentido Schopenhaueriano, no queda excluida como respuesta. Tampoco ese coro que Schönberg utiliza para representar lo divino, fiel a la pluralidad del primer nombre de dios en la Tora: “Elohim”. Pero, ¿cómo *suenan* Dios, y quién es capaz de hablar en su misma tonalidad? Jordi Pons, Massimo Cacciari y Juan Ángel Vela del Campo profundizan en la cuestión con rigor e imaginación, combinando el análisis musical con sugerentes reflexiones filosóficas.
- 3) Finalmente queda la cuestión de la ley, *palabra* huérfana y contradictoria, siempre interpretable, antinatural, eternamente desajustada, ¿quién la cumple, si es que es posible hacerlo? ¿no toma ella el papel de las imágenes? Y, en definitiva, ¿qué es, y a que propósitos sirve? De ello se ocupa Félix Duque, cuyas reflexiones sobre las figuras de Pablo y Moisés cierran el libro, dejando al lector con ganas de más.

Sin duda, una lectura muy recomendable para seguir pensando el clásico de Schönberg.

Gonzalo Pérez Santonja
(Universidad Autónoma de Madrid)

NORMAS DE PUBLICACIÓN

La finalidad de *Daimon - Revista Internacional de Filosofía* es publicar trabajos de investigación en filosofía. *Daimon* es, desde 2001, una publicación cuatrimestral. Algunos de los números son monográficos y otros no. Los números monográficos son anunciados con antelación suficiente (al menos un año) mediante la correspondiente *llamada para aportaciones (call for papers)*, en la que se anuncia el tema del monográfico y el nombre de la persona encargada de coordinarlo. En el caso de que un monográfico no reciba originales suficientes para completar el volumen (actualmente tenemos fijado un límite en torno a las doscientas páginas), se completará con una sección de artículos variados.

Formato de los originales: Véase en <https://revistas.um.es/daimon/about/submissions>

El texto de los artículos y de notas críticas que sea enviado para revisión NO debe contener datos personales del autor o autores, ni en el propio texto, ni en las propiedades del archivo informático, ni en las citas bibliográficas (en este último caso, cada cita de trabajos del autor ha de ser sustituida por la palabra "Autor" y el año de la publicación referida).

Las citas bibliográficas han de hacerse, o bien a pie de página, o bien mediante paréntesis insertos en el texto.

En el primer caso, las referencias han de redactarse como en los ejemplo que siguientes:

- *Libro*: Savater, F. (1999), *Las preguntas de la vida*, Barcelona: Ariel.
- *Capítulo de libro*: Camps, V. (1999), «Universalidad y mundialización», en: M. Cruz y G. Vattimo (eds.): *Pensar en el siglo*, Madrid: Taurus, pp. 61-85.
- *Artículo*: Campillo, A. (2009), «La biopolítica en España», *Daimon*, nº 48, pp. 223-232.

Derechos de autor:

Las obras que se publican en esta revista están sujetas a los siguientes términos:

1. El Servicio de Publicaciones de la Universidad de Murcia (la editorial) conserva los derechos patrimoniales (copyright) de las obras publicadas, y favorece y permite la reutilización de las mismas bajo la licencia de uso indicada en el punto 2.

© Servicio de Publicaciones, Universidad de Murcia, 2011

2. Las obras se publican en la edición electrónica de la revista bajo una licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 España (texto legal). Se pueden copiar, usar, difundir, transmitir y exponer públicamente, siempre que: i) se cite la autoría y la fuente original de su publicación (revista, editorial y URL de la obra); ii) no se usen para fines comerciales; iii) se mencione la existencia y especificaciones de esta licencia de uso.



3. Condiciones de auto-archivo. Se permite y se anima a los autores a difundir electrónicamente las versiones pre-print (versión antes de ser evaluada) y/o post-print (versión evaluada y aceptada para su publicación) de sus obras antes de su publicación, ya que favorece su circulación y difusión más temprana y con ello un posible aumento en su citación y alcance entre la comunidad académica.

Procedimiento: Véase en <http://revistas.um.es/index.php/daimon/about/submissions>

Daimon. Revista Internacional de Filosofía

Publicación cuatrimestral. Número 84. Septiembre-Diciembre 2021

Monográfico sobre Expressing Hatred: The Political Dimension of Expressives
(Editores: Eduardo Pérez Navarro y María J. Frápolli Sanz)

Introducción

Presentación del número monográfico sobre Expressing Hatred: The Political Dimension of Expressives. <i>Eduardo Pérez-Navarro y María J. Frápolli</i>	5
---	---

Artículos

Exactly why are slurs wrong? <i>Thaddeus Metz</i>	13
The resistant effect of slurs: A nonpropositional, presuppositional account. <i>Alba Moreno Zurita y Eduardo Pérez-Navarro</i>	31
Paving the road to hell: The Spanish word <i>menas</i> as a case study. <i>David Bordonaba-Plou y José Ramón Torices</i>	47
Provocative insinuations. <i>Álvaro Domínguez-Armas y Andres Soria-Ruiz</i>	63
Not all speakers are equal: Harm and conversational standing. <i>Claudia Picazo</i>	81
Exactly, what do you mean? <i>Manuel Almagro y Neftalí Villanueva</i>	97
I am large, I contain multitudes: Epistemic pragmatism, testimonial injustice and positive intersectionalism. <i>María J. Frápolli y Llanos Navarro</i>	115
Las cualidades secundarias y el valor epistémico de la experiencia estética. <i>Nieves Acedo</i>	131
Una concepción pragmatista sobre las reglas. <i>Pedro Antonio García Jorge</i>	147
Enactivismo y valoración. Cómo superar la querrela entre teorías somáticas y cognitivas de las emociones. <i>Andrea F. Melamed</i>	163
La autoridad de la primera persona como criterio de determinación y reconocimiento de la identidad de género. <i>Jorge Suárez Muñoz</i> ..	183
Investigación filosófica y emociones: la ontología histórica de Ian Hacking como alternativa. <i>Juan Manuel Zaragoza Bernal</i>	199

Reseñas

BUTLER, Judith (2021). <i>Excitable Speech. A Politics of the Performative</i> . Nueva York: Routledge. 212 pp. (<i>Javier Moscoso Cala</i>).....	219
BRANDOM, Robert M. (2019). <i>A Spirit of Trust. A Reading of Hegel's Phenomenology</i> . Cambridge, Massachussets: The Belnap Press of Harvard University Press, 768 pp. (<i>Christian Iber; Agemir Bavaresco; Danilo Vaz Costa</i>).....	223
NÚÑEZ, Amanda (2019). <i>Gilles Deleuze, una estética del espacio para una ontología menor</i> . Madrid, Arena Libros, 232 págs. (<i>Rosa Jiménez Asensio</i>).....	228
CORCHIA, L., MÜLLER-DOOHM, S. und OUTHWAITE, W. (Hg.). <i>Habermas global. Wirkungsgeschichte eines Werks</i> . Berlin: Suhrkamp, 2019, 894 pp. (<i>José Luis López de Lizaga</i>).....	232
HEGEL, G. W. F. (2019). <i>Lecciones sobre la filosofía del espíritu subjetivo II. Antropología</i> . Traducción Alberto Ciria y Juan J. Padiá, notas y estudio preliminar por Juan J. Padiá, Sevilla: Thémata, 540 pp. (<i>Andrés Ortigosa</i>).....	236
CAVALLETTI, A. (2019). <i>Vértigo. La tentación de la identidad</i> . Traducción de María Teresa D'Meza. Buenos Aires: Adriana Hidalgo editora. (<i>Juan Cruz Aponiuk</i>).....	240
GONZÁLEZ BLANCO, Azucena (ed.) (2019). <i>Política y literatura. Nuevas perspectivas teóricas</i> . Berlín: De Gruyter. (<i>Beltrán Jiménez Villar</i>)	243
WITTGENSTEIN, Ludwig (2019). <i>Correspondencia con Rudolf Koder</i> . Traducción de Venancio Andreu Baldó e Isabel Gamero Cabrera. Madrid: Editorial Apeiron (<i>Jeremy Mederos</i>).....	247
LUMBRERAS SANCHO, Sara (2020). <i>Respuestas al Transhumanismo. Cuerpo, autenticidad y sentido</i> . Madrid: Digital Reasons, 202 pp. (<i>Adrián Santamaría Pérez y José María Santamaría García</i>).....	251
HUSSERL, Edmund (2019). <i>Textos Breves (1887-1936)</i> . Coord.: Agustín Serrano de Haro y Antonio Ziriñ Quijano. Salamanca: Sígueme, 716 pp. (<i>Jesús Miguel Marcos del Cano</i>).....	255
HINDRICH, Gunnar (2020). <i>La autonomía del sonido. Una filosofía de la música</i> . Traducción de Pedro García-Baró Huarte. Salamanca: Ediciones Sígueme. (<i>Laura Carolina Durán</i>).....	260
RIVERA, A. y MATEO, E. (eds.) (2020). <i>Las narrativas del terrorismo. Cómo contamos, cómo transmitimos, cómo entendemos</i> . Vitoria, Madrid: Fundación Fernando Buesa y Ediciones Catarata. (<i>Isabel Roldán Gómez</i>).....	263
PASCAL, Blas (2020). <i>Escritos espirituales</i> . Edición de Alicia Villar. Madrid: Tecnos. 242 pp. (<i>Carlos Gómez</i>).....	266
BRENTANO, Franz C. (2019). <i>Vermischte Schriften. Sämtliche veröffentlichte Schriften, Band IX</i> . Mit einer Einleitung von D. Fiset. Herausgegeben von T. Binder und A. Chrudzimski. Berlin/Boston: Walter de Gruyter. (<i>David Torrijos Castrillejo</i>).....	268
ÇRETU, Ana-Maria; MASSIMI, Michela (ed.) (2020). <i>Knowledge from a human point of view</i> , Springer. (<i>Abraham Hernández</i>).....	270
ARAGÜÉS, Juan Manuel (2020): <i>De idiotas a koinotas. Para una política de la multitud</i> . Madrid: Arena Libros. (<i>Julien Canavera</i>).....	273
ASSMANN, J., CACCIARI, M., DUQUE, F., PONS, J. y VELA, A. (2020). <i>La palabra que falta. El dios indecible de Moisés y Aarón, de Schönberg</i> . Madrid: Ediciones CBA. (<i>Gonzalo Pérez Santonja</i>).....	277