

**La filosofía de la mente
en el pensamiento
de John Searle**

Ignacio Aguinalde Sáenz



**UNIVERSIDAD AUSTRAL
EDICIONES**

AUTOR

- **Aguinalde Sáenz, Ignacio.** Doctor en Filosofía y Licenciado en Psicología por la Pontificia Universidad Católica Argentina. Autor del estudio preliminar y traducción anotada del *Comentario de Tomás de Aquino al libro de Aristóteles sobre la generación y la corrupción. Los principios de la naturaleza y otros opúsculos cosmológicos* (2005) y del libro *Psicología* (2009).

Actualmente participa en proyectos de investigación promovidos por el Instituto de Filosofía de la Universidad Austral. Sus áreas de interés incluyen la filosofía de la naturaleza, la antropología filosófica, la epistemología y la psicología cognitiva. Se desempeña como profesor de Antropología Filosófica, Filosofía de las Ciencias y Ética en programas de grado y posgrado de la Facultad de Ciencias Biomédicas de la Universidad Austral.

Índice

Introducción	17
Aclaraciones preliminares	21
I. Biografía intelectual de John Searle	23
I.1. Los inicios en la Filosofía del lenguaje	23
I.2. La búsqueda de fundamentos en la Filosofía de la mente	32
I.3. La visión unificada de la realidad aplicada a la Filosofía social	38
II. Contexto histórico de la filosofía de la mente en Searle	45
II.1. El dualismo y los problemas de la filosofía de la mente	45
II.1.1. El dualismo en Descartes	45
a. Dualismo de substancias	45
b. Problemas del cartesianismo	49
II.1.2. El dualismo en la actualidad	64
a. Dualismo de substancias contemporáneo	64
b. Dualismo de propiedades	68
II.2. Posiciones teóricas materialistas	72
II.2.1. Noción general de materialismo	72
II.2.2. El conductismo	76
a. Conductismo metodológico	77
b. Conductismo lógico	81
c. Objeciones	82

II.2.3. El fisicalismo o teoría de la identidad entre mente y cerebro	87
a. La tesis fisicalista como reacción frente al conductismo	87
b. Objeciones	89
II.2.4. El materialismo eliminativo	99
a. El eliminativismo como versión extrema del materialismo	99
b. Objeciones	103
II.2.5. El funcionalismo, la Inteligencia Artificial fuerte y el cognitivismo	108
a. El funcionalismo	108
b. La inteligencia artificial fuerte	113
c. El cognitivismo	126
d. Objeciones	129
II.2.6. Teoría de la actitud intencional	139
a. La actitud o estrategia intencional	139
b. Objeciones	143
II.2.7. Valoración general de las teorías materialistas en la filosofía searlina	147
II.3. El emergentismo como antecedente inmediato de la posición de Searle	154
III. La crítica de la Inteligencia Artificial fuerte y el cognitivismo	165
III.1. La crítica de la Inteligencia Artificial fuerte	165
III.1.1. El argumento de la habitación china	165
a. Planteamiento general	165
b. Formulación del argumento	168
c. Estructura lógica de la demostración	179
III.1.2. Objeciones frente al argumento de la habitación china	182
a. Clasificación de las objeciones	182
b. Objeciones originales	183
1. Objeción del sistema total (Berkeley)	183
2. Objeción del robot (Yale)	186

3. Objeción del simulador cerebral (Berkeley y M. I. T.)	188
4. Objeción combinatoria (Berkeley y Stanford)	191
5. Objeción de las otras mentes (Yale) y de las muchas mansiones (Berkeley)	191
c. Críticas generales al argumento	193
d. Objeciones posteriores	194
1. Objeción de la analogía con el electromagnetismo (Paul y Patricia Churchland)	195
2. Objeción del conexionismo (Paul y Patricia Churchland)	199
3. Objeción de la imposibilidad de realizar el experimento (Daniel Dennett)	201
4. Objeción de la relación entre conciencia y computación (Jerry Fodor)	202
III.1.3. Conclusiones	207
III.2. La crítica del cognitivismo	217
III.2.1. El argumento del carácter relativo al observador de los procesos computacionales	217
a. Planteamiento general	217
b. Formulación del argumento	220
c. Corolarios	227
III.2.2. Procesos cognitivos y seguimiento de reglas	232
a. Procesamiento de información y reglas computacionales inconscientes	232
b. Propiedades del seguimiento consciente de reglas	235
III.2.3. Conclusiones	245
IV. La solución de Searle al problema mente-cuerpo: el naturalismo biológico	249
IV.1. Fundamentos epistémicos y ontológicos del naturalismo biológico	249
IV.2. El naturalismo biológico y la distinción de niveles de descripción de un sistema	259

IV.3. El concepto emergentista de lo mental y la superveniencia de macropropiedades en sistemas físicos	272
IV.4. Emergentismo y concepto de reducción	277
IV.5. Objeciones al naturalismo biológico	273
IV.6. Mente y causalidad	289
V. La intencionalidad y la conciencia	299
V.1. La intencionalidad	299
V.1.1. El problema de la intencionalidad: Antecedentes históricos	299
V.1.2. Naturaleza y estructura de los estados intencionales	304
a. Concepto de intencionalidad	304
b. Intencionalidad y lenguaje: el modelo de los actos de habla	307
c. La intencionalidad como representación de condiciones de satisfacción	313
V.1.3. La red y el trasfondo	319
V.2. La conciencia	327
V.2.1. El concepto de conciencia: aclaraciones preliminares	327
V.2.2. La naturaleza y estructura de la conciencia	331
a. Definición de conciencia	331
b. Características estructurales de la conciencia	333
V.2.3. El inconsciente y el principio de conexión	344
a. Definición y clasificación de los estados inconscientes	344
b. El principio de conexión	350
VI. El dinamismo mental	355
VI.1. La percepción	355
VI.1.1. Percepción y realismo	355
VI.1.2. Estructura intencional de la percepción	361
VI.2. La acción racional	364
VI.2.1. Estructura intencional de la acción	364

VI.2.2. Obrar humano y razones para actuar	371
VI.2.3. El problema del libre albedrío	374
VI.2.4. El sí mismo como fundamento de la acción racional	383
VII. Discusión crítica	389
VII.1. Presupuestos epistémicos, metafísicos y filosófico-naturales del pensamiento searliano	389
VII.1.1. Cientificismo	390
VII.1.2. Materialismo	393
VII.1.3. Mecanicismo	395
VII.1.4. La teoría hilemórfica como respuesta a los problemas de la filosofía de la mente	397
VII.2. El naturalismo biológico	401
VII.2.1. Equilibrio inestable entre materialismo y dualismo	401
VII.2.2. Inadecuación de las analogías	402
VII.2.3. Problema de la unidad	404
VII.2.4. Incapacidad de dar razón de las operaciones específicamente humanas	407
VII.3. Inconvenientes de la crítica de Searle al cognitivismo	409
VII.3.1. Forma, finalidad e información en la naturaleza	409
VII.3.2. Contradicción con el naturalismo biológico	415
VII.4. La intencionalidad y la conciencia	420
VII.4.1. Intencionalidad	420
VII.4.2. Conciencia	422
VII.5. El libre albedrío y el sí mismo	426
VII.5.1. El libre albedrío	426
VII.5.2. El sí mismo	428
VII.6. Valoración general	430
VIII. Conclusiones	433
Bibliografía	443

Resumen. Este libro ofrece un estudio sistemático y crítico de las contribuciones del filósofo estadounidense John Searle a los debates contemporáneos en el campo de la filosofía de la mente. En sus páginas aborda cuestiones centrales como la posibilidad de la inteligencia artificial en sentido fuerte, la naturaleza de la conciencia y el estatuto epistemológico de las ciencias cognitivas, e intenta brindar herramientas conceptuales para comprender uno de los problemas filosóficos más acuciantes de la actualidad: la relación entre la mente y el cuerpo. Dentro de la tradición analítica anglosajona, la obra de Searle representa un valioso aporte a la discusión sobre este problema, tanto por el modo original en que ha sabido desarrollar las intuiciones válidas de las diferentes corrientes como por su continua insistencia en lograr una cosmovisión filosófica que integre de un modo coherente los conocimientos fundamentales aportados por las ciencias empíricas. La obra reconstruye así el itinerario intelectual de Searle desde sus comienzos en la filosofía del lenguaje hasta su teoría de la mente, y culmina con una valoración de conjunto sobre su pensamiento a la luz de la filosofía aristotélico-tomista, en la que se busca destacar no solo los aportes de Searle, sino también sus limitaciones.

Palabras clave. John Searle, problema mente-cuerpo, inteligencia artificial, filosofía de la mente, filosofía aristotélico-tomista, dualismo, materialismo, naturalismo, conciencia, intencionalidad

Introducción

Todos estos avances y muchos más convergen para hacer del estudio de los procesos cognitivos una de las áreas más activas y apasionantes de la ciencia moderna. Es evidente que no se trata de una empresa que pueda confiarse a una sola disciplina, que pueda desarrollarse desde un punto de vista limitado o especializado.

George Armitage Miller,
*Center for Cognitive Studies Annual Report, 1960*¹

Desde mediados del siglo xx, la investigación de los procesos cerebrales y cognitivos ha ido experimentando cambios revolucionarios, que ofrecieron maneras inéditas de conceptualizarlos y estudiarlos. El increíble desarrollo que en la actualidad ha alcanzado el ámbito científico de la neurociencia, la lingüística, la psicología cognitiva y los sistemas computacionales relacionados con la llamada “inteligencia artificial” abre un panorama insospechado hasta ahora y vuelve a plantearle a la reflexión filosófica una serie de interrogantes fundamentales acerca de la naturaleza de la mente humana. En este sentido, una de las temáticas centrales de la investigación filosófica contemporánea es, sin duda, aquella que surge a raíz del diálogo entablado con las ciencias cognitivas; en otras palabras: la filosofía de la mente.

La actual filosofía de la mente se propone dilucidar la esencia de los fenómenos psicológicos y del espíritu humano haciendo hincapié en el análisis de los rasgos característicos de las capacidades y de los procesos cognitivos. De este modo, antiguos problemas han adquirido un vigor inusitado y temas nuevos —y, por ende, nuevos problemas— han venido

1. En Hirst, William (ed.), *The Making of Cognitive Science. Essays in Honor of George Armitage Miller*, Cambridge, Cambridge University Press, 1998, p. 94.

a ensanchar y enriquecer el área. Entre los nuevos problemas se encuentran, por ejemplo, la relación entre mente y cerebro a la luz de los últimos descubrimientos de la neurofisiología, el estatuto epistemológico de la psicología y la posibilidad de la inteligencia artificial en sentido fuerte. Pero estos temas no han hecho sino aumentar el interés por el planteamiento de ciertas cuestiones recurrentes en filosofía, como la naturaleza de los estados de conciencia, la explicación de la intencionalidad de los fenómenos mentales, el estatuto ontológico del espíritu humano o la noción de subjetividad.

El panorama de influencias históricas que han conducido a este desarrollo notable de la filosofía de la mente es muy amplio y se encuentra ligado a la gradual consolidación de las denominadas “ciencias cognitivas”. Howard Gardner, en su clásico estudio sobre la “revolución cognitiva”², distingue los diferentes aportes teóricos que sirvieron de base para el establecimiento de este ámbito interdisciplinar, mostrando cómo científicos que provenían de disciplinas diferentes efectuaron descubrimientos que tenían un efecto común: generar una manera novedosa de modelizar y de investigar los fenómenos cerebrales y del conocimiento. Los aportes teóricos en este campo fueron muy variados: los descubrimientos de Alan Turing sobre la descripción formal de una máquina que pudiera llevar a cabo todo cálculo concebible y su posterior elaboración de un test para discriminar las respuestas de un dispositivo adecuadamente programado de las respuestas de una persona; la teoría matemática de la información (Claude Shannon); la cibernética o teoría de los sistemas autorregulados (Norbert Wiener); los estudios neurológicos sobre el orden serial de la conducta (Karl Spencer Lashley); la arquitectura computacional (John von Neumann); la modelización lógica de las neuronas y las redes neurales (Warren McCulloch y Walter Pitts).

Tal cantidad de datos empíricos y de teorías provenientes de tan variadas disciplinas constituyó un fuerte estímulo para investigaciones filosóficas de gran relevancia, entre las cuales se destaca el pensamiento del filósofo norteamericano John Searle, objeto de este libro.

Nuestra intención es presentar aquí un estudio detallado de la filosofía de la mente de Searle a fin de lograr una valoración adecuada de sus

2. Cf. Gardner, Howard, *La nueva ciencia de la mente. Historia de la revolución cognitiva*, Barcelona, Paidós, 1987.

aportes a la comprensión del espíritu humano. Para ello, procuraremos, en primer término, ofrecer una visión panorámica de la biografía intelectual de Searle junto con el contexto histórico en el que se desarrolló su filosofía de la mente. En el capítulo I trazaremos el itinerario intelectual que lo llevó desde su interés inicial por los problemas del lenguaje al redescubrimiento de la importancia de las características distintivas de los fenómenos mentales para establecer una fundamentación tanto de los procesos lingüísticos como de las instituciones sociales humanas en general.

Este itinerario de Searle se enmarca dentro de una serie de antecedentes remotos y más inmediatos de los problemas que él intenta resolver y que abordaremos en el capítulo II. Con relación al contexto remoto, analizaremos como antecedentes el dualismo, tanto en Descartes como en la actualidad, y, por otra parte, tomaremos en consideración las principales posturas materialistas que derivaron más tarde en la “revolución cognitiva”, con las cuales Searle entra en diálogo al establecer las líneas generales de su postura filosófica. De este modo, el conductismo, la teoría de la identidad entre mente y cerebro, el materialismo eliminativo, el funcionalismo, la tesis sobre la posibilidad de la inteligencia artificial “fuerte”, el cognitivismo, la teoría de la actitud intencional y, por último, el emergentismo serán las principales posiciones sistemáticas que analizaremos para comprender las respuestas de Searle a las grandes cuestiones planteadas por las ciencias cognitivas.

Una vez determinados estos puntos, el capítulo III abordará de lleno la doctrina searlina de la mente, donde se examinarán ante todo las críticas de Searle a los defensores de la inteligencia artificial en sentido fuerte, analizando el famoso experimento mental de la “habitación china”. En este experimento, Searle señala la imposibilidad de que una máquina entienda verdaderamente los significados de los símbolos que manipula como resultado exclusivo de la instanciación de un programa computacional. Searle distingue entre los defensores de la IA débil —que se contentan con afirmar la utilidad de los modelos computacionales para el estudio de la inteligencia— y los defensores de la IA fuerte —que sostienen que un ordenador, con un programa adecuado, posee propiamente estados cognitivos—. En contra de los defensores de esta última tesis, propuso su célebre experimento mental en el que argumenta que un ordenador se limita a manipular símbolos a partir de unas instrucciones.

Este funcionamiento, meramente sintáctico, no permite afirmar que el ordenador sea capaz de comprender sus actos. La discusión del argumento de la habitación china, junto con el estudio de las principales objeciones de Searle a la corriente cognitivista, permitirá, a su vez, establecer los ejes sobre los que él plantea su solución emergentista al problema de la relación entre la mente y el cuerpo.

El capítulo IV intentará, entonces, trazar la tesis emergentista de lo mental —tal como Searle la entiende en su relación con la superveniencia de macropropiedades en los sistemas físicos—, tesis que Searle propone bajo la denominación de *naturalismo biológico*. Para Searle, su postura no constituye en absoluto una forma de reduccionismo, dado que, a su entender, es posible distinguir entre mente y cerebro: la mente no se reduce a los mecanismos neurológicos ni se identifica con ellos, sino que es una propiedad emergente causada por el cerebro. Sin embargo, del hecho de que los procesos mentales no sean reductibles a meros procesos biológicos no se sigue, a su juicio, ninguna forma de dualismo, sino un cierto tipo de emergentismo en el que se considera que determinadas propiedades —entre ellas, la mente— no pueden meramente reducirse a las actividades físicas que las causan. El cerebro es *causa* de la mente, pero esta no se reduce a aquel. Con esta teoría, Searle se propone superar las deficiencias que encuentra tanto en el dualismo como en las posturas materialistas —principalmente en el funcionalismo—, en la medida en que no permiten dar una explicación satisfactoria de fenómenos propios de la vida mental humana como la intencionalidad, la conciencia, la subjetividad, la causalidad mental, la dinámica de la percepción o la libertad de la acción racional, fenómenos que consideraremos a la luz del pensamiento searliano en los capítulos V y VI, respectivamente.

Por último, el capítulo VII estará dedicado a la realización de una discusión crítica general de la filosofía de la mente searliana, en la que tendremos en cuenta sus presupuestos de fondo, tanto epistemológicos como metafísicos y filosófico-naturales, para luego arribar a una valoración de conjunto sobre su antropología desde una perspectiva aristotélico-tomista, en la que buscaremos destacar no solo los aportes de Searle, sino también sus limitaciones.