

LA CIENCIA COGNITIVA

INTRODUCCIÓN Y CLAVES PARA SU DEBATE FILOSÓFICO

Trabajo de Investigación dirigido por el Prof. Jaime Nubiola

Lino Iglesias Martínez

Octubre 2006

Universidad de Navarra

ÍNDICE

INTRODUCCIÓN.....	5
CAPÍTULO I. HACIA UNA DEFINICIÓN DE CIENCIA COGNITIVA	15
1.1 INTRODUCCIÓN	15
1.2 CAMPO DE ESTUDIO DE LA CIENCIA COGNITIVA	17
1.3 CARÁCTER CIENTÍFICO DE LA CIENCIA COGNITIVA	26
1.4 CARÁCTER INTERDISCIPLINAR DE LA CIENCIA COGNITIVA	33
1.5 RESUMEN.....	43
CAPÍTULO II. HISTORIA BREVE DE LA CIENCIA COGNITIVA.....	45
2.1 PRECURSORES.....	45
2.2 CIBERNÉTICA	53
2.3 COGNITIVISMO.....	62
2.4 CONEXIONISMO. TENDENCIAS ACTUALES	74
2.5 RESUMEN.....	85
CAPÍTULO III. LAS OTRAS DISCIPLINAS	87
3.1 PSICOLOGÍA	87
3.2 NEUROCIENCIA	97
3.3 LINGÜÍSTICA.....	102
3.4 ANTROPOLOGÍA	106
3.5 FILOSOFÍA.....	110
3.6 RESUMEN.....	114
CAPÍTULO IV. DOS TEORÍAS ACTUALES EN CIENCIA COGNITIVA	117
4.1 LA TEORÍA DE LA MENTE CORPOREIZADA DE LAKOFF Y JOHNSON.....	117
4.1.1 <i>Introducción</i>	117
4.1.2 <i>Fundamentos científicos e interdisciplinarios</i>	121
4.1.3 <i>La teoría de la mente corporeizada</i>	126
4.1.4 <i>Implicaciones filosóficas</i>	134
4.2 LOS “ESPACIOS CONCEPTUALES” DE PETER GÄRDENFORS.....	141
4.3 RESUMEN.....	150
CAPÍTULO V. CLAVES PARA EL DEBATE FILOSÓFICO DE LA CIENCIA COGNITIVA	151
5.1 INTRODUCCIÓN: FILOSOFÍA Y CIENCIA COGNITIVA.....	152
5.2 EL PROBLEMA MENTE-CUERPO	156

4	La Ciencia Cognitiva: introducción y claves para su debate filosófico	
	5.3 LA BÚSQUEDA DEL SENTIDO	163
	5.4 LÍNEAS FUTURAS DE INVESTIGACIÓN	167
	CONCLUSIONES	171
	BIBLIOGRAFÍA	179

INTRODUCCIÓN

La Ciencia Cognitiva surgió a mediados del siglo veinte en Estados Unidos con la pretensión de emprender el estudio del razonamiento humano de una manera empírica, científica e interdisciplinar, frente a los estudios sobre epistemología anteriores puramente filosóficos. Estas pretensiones estaban avaladas por experimentos en psicología que apuntaban a la necesidad del estudio de los mecanismos de funcionamiento interno del cerebro en contraposición a las tesis del conductismo imperante, y por la aparente similitud de los procesos lógicos y las emergentes computadoras con ciertos procesos neurológicos. Así, parecía que se tenía la base científica y las herramientas técnicas necesarias para acometer dicho estudio. Poco a poco, viendo la magnitud de la empresa y fruto de la necesidad de ampliar e incluso de modificar el marco explicativo, se fueron añadiendo nuevas disciplinas a la Ciencia Cognitiva, como la lingüística o la filosofía.

La Ciencia Cognitiva está todavía en un proceso de autodefinición y maduración. Dada la complejidad y el carácter abierto del objeto de su estudio, este proceso pasa por la adopción de fundamentos cada vez más firmes y completos que integren los procesos mentales como actividad biológica con el carácter social del pensamiento, y que incluyan a ambos en una perspectiva no determinista de la condición humana. Un estudio experimental lo más completo posible del razonamiento humano permite el acercamiento de la Ciencia Cognitiva a perspectivas filosóficas más amplias, al tener en cuenta los distintos aspectos humanos, tanto biológicos como sociales, que intervienen en el pensamiento. Estos estudios experimentales, aunque cuestionables por no mostrarse concluyentes, han acumulado sin embargo el suficiente número de evidencias como

para mostrar a la filosofía que ni el pensamiento se produce de cualquier forma, ni es independiente del medio en que vivimos.

Se establece así una peculiar relación entre Ciencia Cognitiva y filosofía. Aquélla intentando establecer principios científicos del pensamiento, que también serían aplicables a la filosofía. Ésta por su parte proporcionando teorías sobre las que construir la Ciencia Cognitiva y apuntando a sus limitaciones. Queda por ver lo que una podrá proporcionar a la otra, o mejor, lo que cada una aportará a la explicación del pensamiento humano. Pero si se admite, como se hace en este trabajo, que la razón va más allá de sí misma, pretender la desaparición de la reflexión filosófica en favor de las explicaciones de la Ciencia Cognitiva parece tan arriesgado como acotar a priori los resultados a los que será capaz de llegar la Ciencia Cognitiva.

Mi interés en esta materia tiene su origen en la realización de cursos en el programa de doctorado en Ciencia Cognitiva en la UNED, fundamentalmente el curso sobre la Teoría Cognitiva de la Metáfora, impartido por el profesor Eduardo Bustos. Me llamó la atención el descubrir que la Ciencia Cognitiva reclamaba para sí el estudio experimental del razonamiento humano y que construía teorías filosóficas intentando fundamentarlas en aspectos básicos de la experiencia humana. Vi entonces en la Ciencia Cognitiva un posible nexo entre filosofía y ciencia, nexo que siempre he echado de menos en los estudios sobre epistemología tradicionales. Este posible nexo requiere tanto una ciencia con fuertes bases filosóficas como una filosofía atenta y consistente con los resultados científicos. En la Universidad de Navarra, el profesor Jaime Nubiola me introdujo en la obra del pragmatista americano Charles S. Peirce, en la que descubrí una buena fuente de ideas donde buscar esta colaboración entre filosofía y ciencia.

El objetivo de este trabajo es doble: realizar una introducción a la Ciencia Cognitiva y a la vez presentar sus aspectos más controvertidos en cuanto a su validez y fundamentación como ciencia de todo el fenómeno del pensamiento. He organizado la introducción

en torno a tres temas. El primero de ellos es la definición de la Ciencia Cognitiva atendiendo a sus características principales: el pensamiento como su campo de estudio y su pretendido carácter científico e interdisciplinar. El segundo presenta una breve historia de la Ciencia Cognitiva. Considero que muchas respuestas a lo que es la Ciencia Cognitiva se encuentran en su historia, por lo que su estudio resulta importante. Esta historia consiste en la búsqueda de los distintos modelos “científicos” del pensamiento, principalmente a través de modelos matemáticos técnicamente replicables en algún tipo de máquina. Se presenta aquí también la historia de las distintas disciplinas que juegan algún papel en la Ciencia Cognitiva. Por último, se ilustra lo anteriormente expuesto mediante la presentación de dos teorías en Ciencia Cognitiva, mostrando sus supuestos principales y su estructuración en torno a ellos, su posible carácter científico e interdisciplinar, y la explicación que ofrecen del pensamiento.

A lo largo de todo el trabajo se van presentando los problemas a los que se enfrenta la Ciencia Cognitiva en su propósito de dar una explicación científica del pensamiento, lo que constituye el segundo objetivo del trabajo. Se podría decir que estas dificultades presentan una doble cara, una “técnica” y otra filosófica. En un primer análisis se presentan como dificultades “técnicas”: la dificultad en la definición de qué se entiende por pensamiento, los problemas de entendimiento entre las distintas disciplinas al no encontrar un lenguaje común, la dificultad en la especificación de los experimentos y teorías para considerarlos relevantes y científicos, la dificultad en la adopción de un paradigma estable, etc. Pero un análisis detallado revela que muchas de ellas tienen un fondo filosófico, ya que en el estudio del pensamiento trasluce la necesidad de tener en cuenta todos los aspectos de la persona sin partir de la asunción de que es posible un estudio completo desde la perspectiva y métodos de las ciencias naturales. Son las claves de esta discusión filosófica las que a mi entender apuntan a la raíz de los problemas y pueden arrojar más luz sobre la posibilidad de tener una ciencia de la cognición, por lo que se estudian con más detenimiento.

Esta estructura está reflejada en cinco capítulos. El primero trata de la definición de la Ciencia Cognitiva, el segundo y el tercero presentan su historia, el cuarto ilustra lo anteriormente expuesto mediante dos teorías, y el quinto presenta las claves para su discusión filosófica. Por último, se presentan las conclusiones del trabajo y la bibliografía.

El primer capítulo presenta la definición de Ciencia Cognitiva como el estudio científico e interdisciplinar del pensamiento, e introduce las opiniones respecto a la metodología que debe emplear, el campo de estudio sobre el que debe actuar y su carácter científico de dos autores, Paul Thagard y Barbara von Eckardt, que han intentado definir y delimitar el campo de actuación de la Ciencia Cognitiva. Se muestra en el capítulo cómo en la definición de Ciencia Cognitiva surgen los problemas de definición de su campo de estudio, no resultando fácil decidir qué queda dentro de la actividad cognitiva y qué no (por ejemplo, sería difícil situar a los sentimientos dentro de la actividad cognitiva), como tampoco es sencillo saber cuáles son los métodos más adecuados para el estudio de los distintos aspectos de la cognición (métodos de las ciencias naturales o ciencias sociales, etc.). Tampoco hay consenso en la comunidad científica en cuanto a su establecimiento como ciencia natural ni está claro que tenga un lenguaje propio fruto de una verdadera interdisciplinariedad. Se empieza a ver que los aspectos sociales y los que parecen más humanos como los sentimientos e intenciones son los que antes se resisten a ser explicados por la Ciencia Cognitiva.

El segundo capítulo es una historia breve de la Ciencia Cognitiva que parte de los intentos de formalización matemática del pensamiento y de los avances en la formalización de una expresión aritmética o lógica que supuso la máquina de Turing. Sin embargo, la consistencia de estos sistemas formales quedaba en entredicho desde los propios análisis matemáticos, como demostró Gödel, con lo que puso en serias dudas la empresa de formalización de todo el pensamiento. Al mismo tiempo comenzaban los estudios cerebrales y

su intento de formalización matemática por equipos de científicos de distintas disciplinas. Tuvieron lugar en Estados Unidos varios encuentros de diversos científicos en los que se plantearon gran parte de los temas vigentes hoy en día en Ciencia Cognitiva, y en estas mismas reuniones se empezaron a ver las dificultades de entendimiento entre distintas disciplinas al abordar el tema común de la cognición. Nuevamente, no estaba claro cómo se podrían formalizar los aspectos sociales y los más personales del pensamiento. Esto supuso un freno en las investigaciones cerebrales, pero la investigación en funciones computables y máquinas de Turing pronto ofreció resultados prácticos en los encuentros de Dartmouth, de la mano de Herbert Simon y Allen Newell, que presentaron una máquina capaz de demostrar teoremas de los *Principia* de Russell y Whitehead. El éxito de las funciones computables hizo que se considerase al pensamiento como la manipulación de símbolos basado en reglas, al igual que hacen las computadoras, en un periodo al que se denomina cognitivismo. Nuevamente, la euforia inicial decayó con el tiempo, al comprobar que muchos proyectos que se creía que se podrían abordar, como la realización de traductores y otros que realizaran tareas sencillas que hace el hombre a diario, resultaban proyectos imposibles dada la variabilidad y riqueza de la experiencia humana, con lo que la principal aplicación fue la realización de tareas muy restringidas y bien delimitadas, mediante lo que se llaman sistemas expertos. Posteriormente se fue volviendo a intentar el análisis y modelado del cerebro sin acudir a la máquina de Turing, analizando éste como un sistema interconectado de unidades que serían las neuronas, lo que se denomina conexionismo. Actualmente conviven cognitivismo y conexionismo, y a éstos se unen los intentos de incluir el mundo que rodea al sujeto mediante modelos dinámicos matemáticamente complejos, e incluso se aborda el modelado matemático de la genética para incluirlo así en el sistema, pero los resultados y el poder explicativo de estos modelos siguen siendo bastante restringidos.

En el tercer capítulo se presenta la recepción de la Ciencia Cognitiva en psicología, antropología, lingüística, neurociencia y filosofía, y se expone también las aportaciones a la misma Ciencia

Cognitiva de estas disciplinas en sus corrientes "cognitivas". Las corrientes "cognitivas" de estas disciplinas eligen alguno de los paradigmas de la Ciencia Cognitiva (conexionismo, cognitivismo) o una mezcla de varios, para aplicarlos en sus respectivos campos. A su vez, la Ciencia Cognitiva se nutre de estas aplicaciones realizadas bajo sus supuestos, para ofrecer pruebas de su validez como ciencia y ampliar su poder explicativo.

En el cuarto capítulo se ilustra lo anterior presentando dos ejemplos de teorías actuales en Ciencia Cognitiva: la teoría de la mente corpórea de George Lakoff y Mark Johnson y la teoría de espacios conceptuales de Peter Gärdenfors. Las teorías deben dar cuenta del pensamiento en su totalidad, y deben tener bases científicas e interdisciplinarias sólidas. Lakoff y Johnson dan ejemplos procedentes de la psicología y la neurociencia que según ellos así lo demuestra. Consideran que el pensamiento es fundamentalmente metafórico y que está basado en nuestras experiencias corporales. El hombre comprende a través de metáforas, que consisten en entender y experimentar una realidad (más abstracta) en términos de otra (más directamente conectada a la experiencia). Lakoff y Johnson no presentan un modelo matemático del pensamiento, pero refieren el funcionamiento de la metáfora al funcionamiento cerebral del modelo conexionista. Para Lakoff y Johnson las bases científicas del pensamiento son principalmente su dependencia de nuestra experiencia corporal, su desenvolvimiento en el mundo y la creación de conceptos, comportamientos, etc., sobre la base de una adaptación evolutiva como supervivencia de la especie, y el funcionamiento cerebral según los modelos conexionistas. Partiendo de estas premisas explican la filosofía como las distintas teorías que se han formado conforme a metáforas más o menos acertadas de nuestra experiencia corporal, y propugnan una filosofía acorde con su teoría corporal de la mente, que tenga al evolucionismo como motor de los actos, y al conexionismo como modelo físico del funcionamiento cerebral. A esta filosofía le dan el nombre de *experencialismo*. La verdad, la moral, la espiritualidad, etc., quedarían totalmente condicionados por estas bases "científicas", con las que deben ser consistentes. Lakoff y

Johnson admiten una variabilidad en la forma de conceptualizar del ser humano, ya que de no ser así las teorías filosóficas siempre habrían sido las “acertadas”. Cabe preguntarse entonces cómo se produce “científicamente” esta variabilidad, explicación que no ofrecen Lakoff y Johnson. Además, el hecho de que el modelo conexionista, el evolucionista o el corporal de la mente puedan explicar bastantes aspectos de la actividad cognitiva y de comportamientos humanos, no implica que sean aplicables a todo el pensamiento. Por ejemplo, no hay bases científicas para asegurar que el evolucionismo condicione toda la actividad cognitiva, aunque explique el curso de la vida animal, incluida la del hombre. Ni tampoco que el modelo conexionista sea el adecuado para explicar todos los procesos cerebrales, como ponen de manifiesto algunos científicos. Ambos son modelos lo suficientemente tentativos como para resultar muy arriesgado hacer descansar sobre ellos una explicación completa del pensamiento.

La teoría de espacios conceptuales de Gärdenfors explica la actividad cognitiva basándose en el hecho científico de, según él, la representación topológica en el cerebro de nuestra interacción con el mundo. Un problema importante en Ciencia Cognitiva es cómo pasar de las descripciones matemáticas del cerebro a nivel neuronal a los conceptos expresables simbólicamente. Una posible solución es distinguir "niveles" en la actividad cognitiva. Éstos podrían consistir en un primer nivel "neuronal" (conexionista), que interactúa directamente con el mundo, consistente en impulsos electroquímicos, un nivel de "enlace" entre la actividad neuronal y los significados o expresiones simbólicas, y un tercer nivel de expresiones y procesamiento simbólico (lenguaje, expresiones matemáticas, etc.). Gärdenfors asume esta estructura cognitiva de tres niveles y sitúa sus espacios conceptuales en el nivel intermedio, en el cual se daría la actividad cognitiva más relevante al combinar las “entradas” del nivel neurológico para producir las “salidas” con significado del nivel simbólico. En este nivel se darían fenómenos como por ejemplo la inducción o la creatividad. Como bases experimentales de su teoría ofrece los experimentos psicológicos que reflejan cómo se puede

realizar una modelización geométrica de las representaciones de la realidad, modelizaciones que son lo que llama espacios conceptuales. Gärdenfors piensa que solo tenemos acceso a las representaciones que nos hacemos de la realidad, y que éstas varían según las distintas soluciones adaptativo-evolutivas al entorno. Nuevamente, al igual que ocurre con la teoría de Lakoff y Johnson, resulta problemático aplicar su teoría a todo el pensamiento, tanto por la cuestionabilidad de los experimentos como por la asunción del enfoque evolucionista.

El quinto capítulo aborda el problema de por qué las teorías en Ciencia Cognitiva no son capaces de dar una explicación de todo el pensamiento humano y presenta dos posibles líneas de investigación para desarrollar en una tesis doctoral. El problema en Ciencia Cognitiva es más de tipo filosófico que de avance o madurez de las teorías científicas. Se trata de pensar si la actividad mental es reducible a mera actividad cerebral (problema mente–cuerpo) que se pudiese explicar mediante las ciencias naturales. Una explicación de este tipo elimina, entre otras cosas, la autonomía de la persona como ser que es más que sus partes, y deja poco margen para la explicación del carácter comunitario del ser humano. El fenómeno cognitivo más relevante en relación con la experiencia humana es la *autoconsciencia*, en donde el hombre se reconoce a sí mismo, y de esta manera se interesa por la verdad de lo que se le presenta, incluido él mismo. Pero la verdad del hombre no se encuentra en sus estados neuronales o en el evolucionismo, que jamás van a contestar a esta pregunta. Se propone como realidad cognitiva ineludible y más profundamente humana la teoría de Levinas: una conciencia no intencional que es llamada a amar al prójimo y que es apelada por él. La respuesta a la pregunta del sentido del hombre no es irracional, pero según esta propuesta exigiría la asunción del misterio de que el hombre se realice dándose, y de que esta entrega sea parte de una trascendencia de la que participa el hombre.

El trabajo se cierra con las principales conclusiones alcanzadas y el listado bibliográfico de las obras consultadas y expresamente citadas.

Finalmente, quiero agradecer a todos los que han estado cerca de mí y de una manera u otra son parte de este trabajo. Al profesor Eduardo Bustos, con quien descubrí la Ciencia Cognitiva y me animó a entrar en ella. De manera especial al profesor Jaime Nubiola, que guió desde el principio mis pasos en filosofía. Sus consejos, correcciones y paciente ayuda han hecho posible este trabajo. A mis padres, a mis hijos y especialmente a mi mujer Isabel por su amor y su paciencia.

CAPÍTULO I

Hacia una definición de Ciencia Cognitiva

En este capítulo se define la Ciencia Cognitiva como el estudio científico e interdisciplinar del pensamiento. Esta definición se examina atendiendo a tres aspectos relacionados entre sí. Primero, el pensamiento como campo de estudio de la Ciencia Cognitiva, lo que se espera de sus teorías y los supuestos en los que se basan. Segundo, el carácter científico de sus teorías. Y finalmente, el carácter interdisciplinar de las mismas. El primero se explica fundamentalmente a través de las obras de Thagard y von Eckardt. A continuación se presentan distintas opiniones y dificultades respecto al establecimiento del carácter científico, interdisciplinar y unitario de la Ciencia Cognitiva, dada la complejidad y amplitud de sus objetivos y la ausencia de un paradigma estable y satisfactorio que la caracterice.

1.1 Introducción

La Ciencia Cognitiva es, o pretende ser, el estudio del pensamiento de forma científica e interdisciplinar. Esta definición enmarca de forma bastante amplia el campo de estudio y el modo de abordarlo. A continuación se presentan algunas definiciones más detalladas y se desarrollan los aspectos fundamentales que la definen: el concepto de “pensamiento” como campo de estudio, y la idea de que este estudio se realiza de forma científica e interdisciplinar.

En primer lugar, vale la pena prestar atención a algunas definiciones de Ciencia Cognitiva:

Es el análisis científico moderno del conocimiento en todas sus dimensiones¹.

Es el estudio interdisciplinar de la mente y de la inteligencia, abarcando la filosofía, la psicología, la inteligencia artificial, la neurociencia, la lingüística y la antropología².

Es la ciencia que responde a interrogantes epistemológicos de antigua data, en particular a los vinculados a la naturaleza del conocimiento, sus elementos componentes, sus fuentes, evolución y difusión³.

Es una disciplina creada a partir de una convergencia de intereses entre los que persiguen el estudio de la cognición desde diferentes puntos de vista. El aspecto crítico de la Ciencia Cognitiva es la búsqueda de la comprensión de la cognición, sea ésta real o abstracta, humana o mecánica. Su meta es comprender los principios de la conducta cognitiva e inteligente. Su esperanza es que ello nos permita una mejor comprensión de la mente humana, de la enseñanza y aprendizaje, de las habilidades mentales y el desarrollo de aparatos inteligentes que puedan aumentar las capacidades humanas de manera importante y constructiva⁴.

Es un campo que estudia la cognición extrayendo recursos de cierto número de disciplinas, incluyendo la psicología cognitiva, la

¹ Francisco J. Varela, *Conocer: las ciencias cognitivas, tendencias y perspectivas*, Gedisa, Barcelona, 1998, 1990, 11. Original en inglés: *Cognitive Science. A Cartography of Current Ideas*, 1988.

² Paul Thagard, "Cognitive Science", en *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/cognitive-science/>, 20 abril 2004, (8 junio 2004), párrafo 1.

³ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, Paidós, Barcelona, 1996, 1988, 21. Original en inglés: *The Mind's New Science. A History of the Cognitive Revolution*, Basic Books, 1985.

⁴ Donald A. Norman, "¿Qué es la ciencia cognitiva?", en *Perspectivas en Ciencia Cognitiva*, Paidós, Barcelona, 1987, 13. Original en inglés: *Perspectives on cognitive science*, Ablex Publishing Corporation. 1981.

inteligencia artificial, la lingüística, la filosofía, la neurociencia y la antropología cognitiva⁵.

1.2 Campo de estudio de la Ciencia Cognitiva

En 1977, en el primer número de la recién creada revista *Cognitive Science*, la primera publicación regular específicamente sobre Ciencia Cognitiva, Allan Collins⁶ explicaba el objetivo de la revista dando a la vez una definición de Ciencia Cognitiva con relación a los problemas que aborda y las herramientas que utiliza. Especificaba los problemas a tratar como la representación del conocimiento, la planificación y solución de problemas, la inferencia, y la comprensión del lenguaje, entre otros. Como técnicas a utilizar destacaban las relacionadas con la inteligencia artificial y con la psicología experimental de los últimos años. A la vez rechazaba la publicación de artículos específicos en psicología, inteligencia artificial o lingüística, en favor de aquellos en los que intervinieran varias disciplinas.

Se resumía así el propósito de la revista:

En resumen, la función de la revista es proporcionar un lugar para publicar nuevos tipos de análisis de ideas teóricas sobre representación y procesamiento cognitivo. Espero que la revista transmita el excitante ambiente creado por el giro de paradigma en el estudio de la cognición hecho posible por la síntesis de la inteligencia artificial, la psicología y la lingüística⁷.

Se hace hincapié por tanto, en cuanto al objeto de estudio, en el procesamiento cognitivo, el comportamiento inteligente y la

⁵ Barbara von Eckardt, *What is Cognitive Science?*, The MIT Press, Cambridge, Massachussets, 1996, 1993, 1.

⁶ Allan Collins, "Why Cognitive Science?", en *Cognitive Science*, vol. 1, 1977.

⁷ Allan Collins, "Why Cognitive Science?", en *Cognitive Science*, vol. 1, 1977.

representación del conocimiento, todos ellos tratados bajo el prisma del campo en el que se dan la mano las distintas disciplinas que componen la Ciencia Cognitiva, fundamentalmente la psicología y la inteligencia artificial. Se puede apreciar en los términos que definen el campo de estudio un marcado sabor computacional, y la tendencia a pretender explicarlos como procesos realizables en alguna máquina. Desde que se publicó este artículo, los modelos y metáforas aplicables al estudio de la cognición han ido cambiando. Principalmente, el paradigma de la mente como mecanismo de procesamiento de símbolos y la existencia de representaciones mentales ha sido puesto seriamente en duda y se han buscado modelos alternativos, como los modelos conexionistas y los sistemas dinámicos. Pero en la mayoría de los estudios se emplean modelos realizables en máquinas, como parte de una empresa científica capaz de dar una explicación del pensamiento, hacerlo “medible”, y así abordable científica y técnicamente. También han ido cambiando las disciplinas que componen la Ciencia Cognitiva, añadiéndose a las ya citadas la neurociencia, la antropología y la filosofía. Dentro de estas disciplinas, los trabajos que se consideran parte de la Ciencia Cognitiva son aquellos que son compatibles de alguna manera con esta concepción de la Ciencia Cognitiva como ciencia natural.

Queda entonces por intentar definir mejor su campo de estudio, que, si bien ahora es algo más específico que cuando se afirmaba que el campo de estudio era la cognición humana, aún suena confuso y desconocido: ¿en qué consiste el procesamiento humano, el comportamiento inteligente y la representación del conocimiento? ¿Es la forma adecuada de tratar la cognición mediante “procesamiento humano”, “comportamiento inteligente” y “representaciones”? ¿qué se espera que diga sobre ellos la Ciencia Cognitiva?, ¿qué preguntas debe responder la Ciencia Cognitiva?

Para abordar estas cuestiones presentaré algunas opiniones sobre lo que es la inteligencia y las teorías de Paul Thagard y Barbara von Eckardt sobre lo que debe explicar la Ciencia Cognitiva, y sobre cuáles son sus métodos y supuestos.

De forma general, se puede definir la inteligencia como la “capacidad de adaptarse, conformar y seleccionar nuestro entorno”⁸. Si la Ciencia Cognitiva va a estudiar esta capacidad, tendrá que conocer su naturaleza para así saber qué métodos aplica a su estudio. Es en este punto donde no existe acuerdo y surgen distintas definiciones de “inteligencia”, según hagan mayor énfasis en los distintos aspectos de esta “capacidad” de adaptación, conformación y selección del entorno. Por ejemplo, Howard Gardner propone que la inteligencia no es unitaria, sino que hay “diferentes inteligencias: lingüística, lógico-matemática, espacial, musical, dinámico-corpórea, interpersonal, intrapersonal y naturalista, Sternberg considera que hay tres aspectos en la inteligencia: pensamiento analítico, creativo y práctico, todos explicables mediante componentes de procesamiento de la información, y Goleman, Salovey y Mayer incluyen también los aspectos emocionales como parte de la inteligencia”⁹. De acuerdo con estas definiciones también el comportamiento animal se debe considerar, al menos parcialmente, inteligente, debido a su capacidad de adaptación y selección del entorno. Otro problema que surge al estudiar la inteligencia es si este estudio se debe hacer de forma “estática”, como una capacidad establecida en el adulto o se debe estudiar todo el proceso de desarrollo de la inteligencia desde la infancia.

Todos estos temas están presentes en la Ciencia Cognitiva y son parte del motivo de su heterogeneidad, debido a las distintas disciplinas que entran en juego dependiendo de los aspectos de la inteligencia que se estudien, y a los distintos paradigmas que se apliquen para su desarrollo. De esta forma se concibe lo estudiado en gran medida en relación con los métodos y modelos empleados. Un ejemplo de esta vinculación entre paradigma y campo de estudio se

⁸ Robert J. Sternberg. “Intelligence”, en Robert A. Wilson and Frank C. Kiel (eds.) *The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, The MIT Press, Cambridge, Massachussets, 2001, 1999, 409.

⁹ Robert J. Sternberg. “Intelligence”, *The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, 409.

tiene en las discusiones en 1956 en la conferencia de Dartmouth, una de las fechas considerada como importante respecto a los orígenes de la Ciencia Cognitiva, donde se adoptó como punto de partida la siguiente afirmación:

Cualquier aspecto del aprendizaje u otra característica de la inteligencia puede en principio ser descrita con precisión de tal forma que se puede construir una máquina que la simule¹⁰.

El significado de esta afirmación lo explica y amplía Crevier¹¹, denominando a este planteamiento “la hipótesis de símbolos físicos”, según la cual nuestras mentes operan sobre estructuras de símbolos que se corresponden con representaciones del mundo, pudiendo estas estructuras tomar forma en cualquier estructura física, sea el cerebro o la computadora. La inteligencia consistiría en la habilidad de procesar estas estructuras de símbolos. Esta hipótesis tenía un gran peso en la revista *Cognitive Science* cuando ésta se fundó.

Según Thagard¹², el propósito principal de la Ciencia Cognitiva es explicar *cómo* se piensa, estando la mayoría de los científicos cognitivos de acuerdo en que el conocimiento consiste en *representaciones* mentales. Las representaciones mentales serían las formas mentales que tenemos de lo que nos rodea, las *representaciones* del mundo. Por ejemplo, una situación sobre cómo decidir qué coche comprarnos podemos representarla como una estructura de razonamientos lógicos, como una imagen mental o como una adaptación por analogía de algo ya conocido antes, etc. Esta concepción se encuentra próxima a la hipótesis del sistema de

¹⁰ John McCarthy, Marvin L. Minsky, Nathaniel Rochester, Claude.E. Shannon, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. August 31, 1955, <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>, 1996 (8 junio 2004).

¹¹ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, Basic Books, Nueva York, 1993, 48.

¹² Paul Thagard, *Mind: Introduction to Cognitive Science*, The MIT Press, Cambridge, Massachussets, 1998, 1996, 3.

símbolos físicos, pero amplía los símbolos a lo que entendamos por representaciones, y el procesamiento computacional a lo que entendemos por procedimiento mental, conceptos estos más abiertos a los descubrimientos en psicología y neurología. Aun así, siempre se conserva como importante la capacidad práctica de simulación física de la teoría en una máquina. Lo que se debe esperar de una teoría de la representación mental según Thagard¹³, es:

1. Potencia representacional
2. Potencia computacional
 - a. Solución de problemas
 - i. Planificación
 - ii. Decisión
 - iii. Explicación
 - b. Aprendizaje
 - c. Lenguaje
3. Plausibilidad psicológica
4. Plausibilidad neurológica
5. Aplicabilidad práctica
 - a. Educación
 - b. Diseño
 - c. Sistemas inteligentes

Para Thagard, las principales formas de entender cómo se piensa o de modelar actualmente las representaciones mentales son la lógica, las reglas, los conceptos, las analogías, las imágenes y las conexiones neuronales. La lógica trata las representaciones como inferencias deductivas, las reglas son estructuras del tipo “si... entonces...”, los conceptos serían algo así como “esquemas mentales” que se corresponderían con situaciones típicas, las analogías son adaptaciones de situaciones ya conocidas, las imágenes se corresponderían con “dibujos mentales”, y las conexiones neurales

¹³ Paul Thagard, *Mind: Introduction to Cognitive Science*, 15.

son representaciones como patrones de actividad de una red de neuronas. Todas estas formas de entender las representaciones mentales se complementarían en lugar de competir entre sí, y se evaluarían conforme a los cinco criterios anteriormente expuestos. A continuación se explican estos criterios con algo más de detalle¹⁴.

La potencia representacional se refiere a la cantidad de información respecto al mundo que puede expresar un tipo de representación mental. Por ejemplo, la lógica consigue representar bien las deducciones pero no consigue formalizar todo el lenguaje, las analogías se aplican bien en la descripción de situaciones similares entre sí, y los conceptos sirven para organizar jerárquicamente el pensamiento y los casos como más o menos cercanos al estereotipo.

La potencia computacional de una teoría de la representación mental se puede evaluar en referencia a su viabilidad para resolver problemas, y tiene que describir cómo la gente debe razonar para conseguir sus objetivos. Esta resolución de problemas tiene tres aspectos: planear el resultado futuro desde la situación inicial, tomar decisiones, es decir, ver cómo elegir la mejor opción entre varias, y la explicación o comprensión de por qué algo ocurre. Por ejemplo, las reglas son muy válidas para explicar una cadena de decisión entre varias opciones, las imágenes sirven para la ordenación de variables en gráficos, y las analogías se utilizan al tomar decisiones legales sobre la base de la jurisprudencia. La potencia computacional debe explicar también cómo piensan las personas y debe dar cuenta del uso del lenguaje en al menos tres aspectos: la habilidad en la comprensión del lenguaje, la habilidad de expresarnos y la habilidad universal en los niños para aprender el lenguaje. Por ejemplo, el lenguaje se podría entender en algunos casos en términos de adquisición, modificación y aplicación de reglas, algunos aspectos de la generalización inductiva también pueden verse como aplicación de reglas, y también parece que el hombre posee algunos conceptos innatos estando otros organizados alrededor de éstos, etc.

¹⁴ Paul Thagard, *Mind: Introduction to Cognitive Science*, 15-18.

La plausibilidad psicológica de una teoría requiere que ésta explique los resultados cuantitativos en psicología respecto a cómo la gente piensa. Esto es así ya que la Ciencia Cognitiva tiene como meta entender la cognición *humana*, y por tanto no basta tan sólo con desarrollar modelos computacionales como hace la inteligencia artificial, sino que estos deben reflejar y estar orientados a la comprensión de la cognición humana. Por ejemplo, los conceptos estarían avalados por la teoría psicológica de prototipos y categorías, y las imágenes por los experimentos de rotación de figuras.

De la misma forma, dado que el pensamiento humano se lleva a cabo en el cerebro, una teoría de la representación mental debe al menos ser consistente con los resultados de los experimentos neurocientíficos (plausibilidad neurológica). En este caso, aun no se tiene constancia de esta consistencia en la mayoría de las teorías de representación mental, excepto en la de redes neurales, y solo en casos restringidos.

Finalmente, las teorías, además de contribuir a entender la mente, deben tener una aplicabilidad práctica, que también constituye una medida de su acierto. En propósitos educativos, la Ciencia Cognitiva debería ser capaz de incrementar nuestra comprensión del aprendizaje y sugerir métodos más eficaces. La aplicabilidad práctica de una teoría en cuanto al diseño se refiere a la aplicación de nuestros conocimientos sobre cómo piensa la gente, al diseño de sistemas que tengan esto en cuenta y nos ayuden en nuestras tareas, como por ejemplo, en el diseño de interfaces de ordenador. Las teorías también deberían contribuir a la creación de sistemas inteligentes y expertos tomando como base los modelos creados para la explicación del razonamiento humano.

Aun así, hay aspectos del conocimiento que, según reconoce Thagard, no es capaz de explicar esta concepción representacional de la Ciencia Cognitiva, en particular: las emociones, que afectan a los procesos cognitivos, la consciencia de nuestro propio ser y nuestros

pensamientos, la relación con el mundo, en el sentido de que el pensamiento no es algo privado sino que está “enganchado” al mundo, la imposibilidad de representar *todo* el conocimiento, y la forma con la que cambia el pensamiento con el tiempo y la capacidad de la razón de encontrar verdades que están más allá de ella si la consideramos una máquina.

Aunque Thagard es consciente de estas limitaciones de la teoría representacional y de que hay muchos ataques a esta teoría, entiende que, en lugar de sustituirla por una teoría no representacional, la forma de avanzar es ampliándola y completándola para que dé cuenta de lo que hasta ahora no es capaz de explicar.

Von Eckardt restringe el campo de estudio de la Ciencia Cognitiva a “la cognición típica y normal en el adulto”¹⁵, dejando así fuera del ámbito de la Ciencia Cognitiva todo estudio sobre ontogenia y desarrollo en el ser humano, estudios sobre animales o máquinas, y sobre aspectos emocionales, aunque reconoce que no hay acuerdo en la comunidad de científicos cognitivos sobre si deberían incluirse estos campos como parte de la Ciencia Cognitiva. Sobre este campo de estudio que especifica Eckardt pesarían unas asunciones que determinan la posibilidad de existencia de una ciencia. Estas asunciones consisten en asignar ciertas propiedades a la cognición:¹⁶ que consisten en estados que tienen contenido, es decir, que se corresponden con “algo” del mundo, que los estados son evaluables pragmáticamente y que tienen cierta coherencia de tal forma que se pueda construir una teoría unificada.

También pesarían unos supuestos metodológicos y de delimitación del campo en cuanto a la forma de abordar su estudio:¹⁷ que la cognición puede estudiarse enfocándose exclusivamente en el individuo, con lo que los factores sociales o culturales se explican en

¹⁵ Barbara von Eckardt, *What is Cognitive Science?*, 6.

¹⁶ Barbara von Eckardt, *What is Cognitive Science?*, 72.

¹⁷ Barbara von Eckardt, *What is Cognitive Science?*, 53-56.

cuanto a que están mediados por la representación individual, que las capacidades cognitivas humanas son lo suficientemente independientes de otros aspectos de la mente (por ejemplo la personalidad) y entre sí como para poder ser estudiadas por separado, que tiene sentido, a pesar de la variedad, distinguir entre una cognición “normal” y “anormal” y que se puede especificar un “típico” adulto cognoscente. También se asume que las preguntas sobre la cognición se pueden formular como preguntas sobre procesamiento de la información, que las respuestas a estas preguntas deben tener una base empírica y tener en cuenta los resultados de la neurociencia y que una teoría completa de la cognición no será posible sin la contribución de cada una de las disciplinas que la componen.

Von Eckardt reconoce que hay unos “retos para el futuro” de la Ciencia Cognitiva. Estos retos consistirían¹⁸ en desarrollar teorías que integren lo individual con lo social en la cognición, ya que ésta tiene lugar en contextos culturales y sociales, y por tanto no puede estudiarse completamente de forma aislada, también serían retos para el futuro desarrollar una teoría de la computación que tenga plausibilidad neurológica, esclarecer la naturaleza de las representaciones mentales, que es actualmente un misterio, y conseguir una estructura institucional que favorezca la investigación interdisciplinar en Ciencia Cognitiva.

Como conclusión, puede decirse que en la definición del campo de estudio de la Ciencia Cognitiva están presentes los supuestos de los que se parte, o el modelo que tengamos más o menos a priori de la cognición. Se prescinde en cierta manera de hacer un estudio de la cognición como fenómeno natural por su complejidad y extensión, teniendo que dejar de un lado aspectos que probablemente no se puedan aislar o manejar como los que se consiguen implementar en máquinas (carácter social de la cognición, desarrollo de la cognición, efecto de las emociones, intencionalidad, etc.). En

¹⁸ Barbara von Eckardt, *What is Cognitive Science?*, 341-343.

definitiva, se buscan explicaciones que *funcionen*, en el sentido de poder manejarlas y simularlas. Así, se pasa de decir que la Ciencia Cognitiva es el estudio de la cognición, a decir que es el estudio de las *representaciones mentales* y su *procesamiento*, para después intentar encontrar la congruencia de este planteamiento con todos los aspectos de la cognición, que en gran medida son todos los aspectos del ser humano.

1.3 Carácter científico de la Ciencia Cognitiva

Uno de los aspectos principales de la Ciencia Cognitiva es su reivindicación del carácter científico en su estudio de la cognición, frente a los estudios anteriores en epistemología puramente filosóficos. La Ciencia Cognitiva basaría sus resultados en hechos contrastados experimentalmente y así crearía una auténtica ciencia aceptada por la comunidad científica. Esto contrastaría con los estudios anteriores en este campo, que eran meramente especulativos o tomaban experimentos aislados para corroborar suposiciones filosóficas hechas a priori, pero en los cuales en ningún caso estos experimentos constituían la base sobre la cual se edificaban las teorías epistemológicas. En palabras de Varela:

Por primera vez la ciencia (es decir, el conjunto de científicos que definen qué debe ser ciencia) reconoce plenamente la legitimidad de las investigaciones sobre el conocimiento mismo, en todos sus niveles, mas allá de los límites tradicionalmente impuestos por la psicología o la epistemología¹⁹.

Respecto a las características de esta disciplina como ciencia, von Eckardt opina que la Ciencia Cognitiva es una ciencia inmadura, y como tal viene caracterizada en términos de unos compromisos

¹⁹ Francisco J. Varela, *Conocer: las ciencias cognitivas, tendencias y perspectivas*, 13.

compartidos por los científicos cognitivos²⁰. Estos compromisos consisten en un conjunto de supuestos que especifican un dominio de actuación, un conjunto de preguntas básicas, un conjunto de supuestos que restringen las respuestas a estas preguntas y un conjunto de supuestos metodológicos. Algunos de estos supuestos se han presentado en la sección 1.1.

Según von Eckardt, lo que comparten los científicos cognitivos es una aproximación al estudio de la mente, más que un conjunto de teorías, explicaciones o leyes. Mientras una teoría es inmadura, las teorías que genera son tentativas. Von Eckardt defiende el carácter científico de este estudio apoyándose en que, según ella, ya hay filósofos de la ciencia que creen que la investigación científica tiene lugar dentro de un marco de compromisos como los que expone y que adoptar una serie de supuestos como hipótesis de trabajo general no implica tomarlos como verdades, con lo que se explicaría su falibilidad científica.

En esta defensa, von Eckardt está apelando a una noción de ciencia más como actividad científica que como conocimiento científico²¹. Esta concepción, además de tener en cuenta como criterio de evaluación de la ciencia su carácter explicativo de los fenómenos, hace énfasis en la idea de ciencia como una empresa humana, con sus agentes económicos y académicos, su capacidad innovadora y su aplicabilidad tecnológica. Presentando el carácter falible de la Ciencia Cognitiva, von Eckardt parece estar defendiéndola de las críticas que han recibido ciertas hipótesis de trabajo, como el modelo computacional de la mente.

Sin embargo no parece que esté tan claro que exista la unanimidad reclamada por Varela o von Eckardt respecto al reconocimiento de la legitimidad de sus investigaciones, ya que hay numerosas voces relevantes que se alzan en contra de la idea de la

²⁰ Barbara von Eckardt, *What is Cognitive Science?*, 5.

²¹ Cf. Javier Echeverría, *Filosofía de la Ciencia*, Akal, Madrid, 1995.

viabilidad del proyecto de la Ciencia Cognitiva. Algunas son opiniones en contra de poder establecer una ciencia de la cognición en los términos que plantea la Ciencia Cognitiva, y otras restringen tanto el alcance de sus resultados que la limitan seriamente en sus aspiraciones. Por ejemplo, Putnam y Cicouriel, en una crítica del modo de operar de la Ciencia Cognitiva, opinan que:

Creemos que como Newton en cierto sentido redujo el mundo físico al orden, algo parecido debe ser posible en psicología²².

La cuestión es: ¿es el estudio del cerebro la respuesta a lo que la mente es? Me parece que hay que ser cuidadoso sobre cómo se caracteriza eso²³.

Y en cuanto a los resultados obtenidos:

Hay logros valiosos en ingeniería, pero la reivindicación de haber hecho algo como un descubrimiento importante, o incluso una nueva aproximación en la forma de pensar sobre la mente y en psicología, me parece fraudulenta²⁴.

Las razones de estas objeciones proceden fundamentalmente de la creencia contraria a la posibilidad de tener una ciencia de la cognición en el sentido de una ciencia natural, donde además se puedan y tengan que simular las teorías en algún tipo de máquina para comprobar su viabilidad. También proceden de los pobres resultados obtenidos por la Ciencia Cognitiva durante sus años de vida en lo que se refiere al establecimiento de una teoría que dé cuenta de todos los

²² Hilary Putnam, "Against the new associationism", en Peter Baumgartner and Sabine Payr, *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 186.

²³ Aaron Cicouriel, "Cognition and cultural belief", en Peter Baumgartner and Sabine Payr, *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 53.

²⁴ Hilary Putnam, "Against the new associationism", en Peter Baumgartner and Sabine Payr, *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 183.

aspectos de la cognición, o al menos que pueda establecer las bases para ello.

Cabe esperar que para que a una disciplina cercana a las ciencias naturales, tal y como pretende serlo la Ciencia Cognitiva, se la reconozca como la disciplina autorizada en el estudio de todos los fenómenos cognitivos, deberían darse al menos las dos siguientes condiciones en la construcción de sus teorías: que estén basadas en la experimentación y que se vean consolidadas por los resultados obtenidos durante periodos lo suficientemente largos como para que generen confianza en la comunidad científica.

La Ciencia Cognitiva siempre contrasta y casi siempre basa sus teorías en resultados experimentales, pero un problema para la construcción de una teoría relevante de todo el fenómeno de la cognición es conseguir que los experimentos también sean relevantes. La dificultad está en definir los experimentos a realizar, como en psicología, donde la especificación de las condiciones de realización de un experimento muchas veces suponen todo un ejercicio teórico, y donde una vez realizado el experimento surgen serias dudas respecto a las inferencias que pueden obtenerse de los resultados obtenidos. Estas inferencias son difícilmente elevables a carácter de ley del pensamiento y muchas veces dependen de la perspectiva adoptada en la realización e interpretación del experimento. Se suceden los experimentos que refutan las conclusiones a las que se llegó anteriormente y así se echa de menos el necesario carácter acumulativo que hace que la ciencia avance con firmeza. En neurociencia se limitan los experimentos a un ámbito restringido de la actividad biológica y física del cerebro, pero las conclusiones psicológicas que den una visión algo más general del fenómeno cognitivo son todavía escasas y controvertidas. Gardner analiza la situación de la psicología y de la neurociencia:

La psicología es una disciplina central en todo estudio de la cognición; pero es también una disciplina difícil, en la cual los avances auténticos no se lograron sin esfuerzo. Casi cualquier

elemento concebible guarda relación con el desempeño del sujeto, y por ende, pocos son los temas vinculados con la naturaleza y el comportamiento de los seres humanos que pueden excluirse *a priori* en el laboratorio. En consecuencia, escoger un problema y descartar todos los que rivalizan con él se vuelve una labor particularmente perturbadora²⁵.

Es preciso abordar muchos de los problemas más importantes de la psicología desde una perspectiva molar, que a su vez entraña una perspectiva que vaya de lo general a lo particular; no obstante, ocurre que los métodos psicológicos más rigurosos no se adecuan con frecuencia a estas grandes cuestiones. Tal como yo lo veo, el desafío que hoy enfrenta la psicología consiste en unir íntimamente esta refinada panoplia metodológica con problemas e interrogantes de consecuencias inequívocas²⁶.

...cuando en la literatura psicológica aparece una demostración interesante, surgen decenas de otros experimentadores que dedican todo su celo metodológico a encontrar los puntos vulnerables del experimento; y a la postre, casi todos los experimentos muestran sus limitaciones y son en definitiva abandonados. La combinación de todas estas tendencias ha hecho que esta ciencia avanzara menos de lo que debería, y siga consistiendo en una serie de hallazgos notables pero aislados, en lugar de conformar una disciplina auténticamente acumulativa²⁷.

Los neurocientíficos dedican hoy la mayor parte de su tiempo al estudio minucioso de sistemas específicos de organismos específicos, en la esperanza de que estos sistemas puedan comprenderse por sí mismos y de que este conocimiento contribuya, a la postre, a los debates más amplios sobre las bases neurales de la cognición²⁸.

Estos problemas en la experimentación están relacionados con la dificultad de conseguir una teoría satisfactoria y estable, que genere

²⁵ Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 113.

²⁶ Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 115.

²⁷ Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 154.

²⁸ Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 304.

resultados durante periodos significativos. Si los experimentos no han resultado todo lo satisfactorios que se pretende, el salto a una teoría general de la cognición resulta arriesgado y un tanto inestable, de tal forma que sólo unos resultados prácticos podrían hacerla avanzar con paso seguro.

Otro problema en la realización de experimentos y en el establecimiento de una teoría, es que no se dispone de unas herramientas y métodos de trabajo comúnmente aceptados en el análisis del pensamiento, como pueden ser las matemáticas para describir los fenómenos físicos. Esto, junto con la vaguedad de los experimentos, hace que se realice ese salto teórico recurriendo a metáforas externas a la experimentación, como el modelo computacional de la mente y en alguna medida el conexionista. Respecto a la metáfora computacional de la mente, opina Searle:

Creo que va a estar presente durante un tiempo, porque no sabemos cómo funciona el cerebro... No sabemos cómo funciona, así que nos apresuramos a utilizar metáforas²⁹.

Surgen entonces teorías creadas a partir de modelos ajenos en gran medida a la experimentación, aunque luego se contraste si “encajan” con ésta. Estos modelos se toman de teorías manejables y realizables en máquinas. Como dice Searle³⁰, estos conceptos no son de la naturaleza, sino que son conceptos formados en base a una viabilidad tecnológica.

Por otro lado, la “prueba de fuego” de estas teorías, sus resultados prácticos, su contrastación con la realidad, no llega a ser muy positiva cuando se trata de explicar todos los aspectos de la cognición. Ejemplos de esta insuficiencia los encontramos en cualquier modelo, y en concreto en el computacional-representacional

²⁹ John R. Searle, “Ontology is the Question”, en Peter Baumgartner and Sabine Payr, *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 212.

³⁰ John R. Searle, “Ontology is the Question”, en Peter Baumgartner and Sabine Payr. *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 205.

mencionado antes, ante la explicación de la consciencia, la intencionalidad, la relación con el mundo o los sentimientos. Y en un ámbito más tecnológico, en el fracaso de la inteligencia artificial en simular la cognición humana³¹.

Después de ver los problemas para lograr una ciencia natural del pensamiento volvemos a la pregunta de si puede aplicarse un método científico para el estudio y explicación del pensamiento en todos sus aspectos, ya que podríamos encontrarnos con una disciplina muy “científica” pero que no fuese capaz de responder a los interrogantes que se le plantean. Se podría ver lo científico como la adecuación al campo de estudio en cuestión de los métodos y el lenguaje empleados. Por ejemplo, no sería científico el estudio de la poesía con herramientas matemáticas como no sería científico estudiar las leyes del movimiento de los cuerpos con un método discursivo.

El problema de la Ciencia Cognitiva es que su campo de estudio presenta todos los aspectos posibles, con leyes físicas que obedecer y con otros aspectos que parecen escapar al análisis empírico y formalizable. Por estos motivos, su carácter científico habría que buscarlo fundamentalmente en la ampliación del concepto de ciencia más allá de ciencia natural, en la adecuación del método empleado al aspecto de la cognición que se pretenda explicar y en el éxito en el uso del método y lenguaje interdisciplinar, al ser en el terreno donde se encuentran varias disciplinas donde se piensa que se puede encontrar quizá una explicación algo más completa de la cognición.

³¹ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 108-144.

1.4 Carácter interdisciplinar de la Ciencia Cognitiva

Otro aspecto central en Ciencia Cognitiva es su definición como un esfuerzo interdisciplinar. Esto quiere decir que se reconoce el valor de las distintas disciplinas que la componen en sus aportaciones a la explicación de la cognición, y se promueve su diálogo como forma de avanzar en la investigación. Pero si, como es el caso, se quiere tener un cuerpo completo de explicaciones, con un funcionamiento coherente, además de diálogo y colaboración entre sus miembros, o la aportación de cada uno de ellos a la resolución de un problema, lo que se podría llamar multidisciplinariedad, habrá que integrarlos para que formen una unidad, de tal forma que se pueda explicar, por ejemplo, cuándo y cómo cada uno de ellos debe entrar en acción ante un problema.

Esta integración entre disciplinas o miembros, en la metáfora del cuerpo de la Ciencia Cognitiva, deberá tenerlas en cuenta a todas y a la vez ser distinta de cada una de ellas, y deberá explicar cuál es el punto de encuentro entre las disciplinas y cómo y cuándo se debe pasar de una a otra al explicar los fenómenos cognitivos. Parece razonable pensar que esto deba hacerse mediante un lenguaje y métodos al menos en parte nuevos, ya que si estos métodos fuesen los de una de las disciplinas, en gran medida estaríamos referenciando el resto a ésta, que se constituiría en “vara de medir” de la validez de las explicaciones. En este caso, esta disciplina diría a las otras algo así como: “observad cómo adoptando mis métodos y suposiciones resolvéis los problemas de la cognición de vuestro campo, por lo que en última (y científica) instancia vuestras explicaciones se pueden expresar en mis términos”, algo que ocurre, por ejemplo con el materialismo eliminativo, al intentar reducir todas las explicaciones de los fenómenos cognitivos a la dinámica de las conexiones neurales, o con otros experimentos en psicología al situarlos como prueba de la validez de una hipótesis. Esto, como muestra Sperber, no es fácilmente aceptable al menos por algunos de los miembros de las otras disciplinas:

... Un equipo de eminentes psicólogos pasa años proporcionando evidencia experimental a favor de la opinión de que hay diferencias fundamentales en los modos de pensamiento de culturas diferentes. Mientras esta opinión va en contra de los prejuicios de la mayoría de los psicólogos, hace tiempo que la defienden los antropólogos, sin tener, sin embargo, el beneficio de la evidencia experimental. Se invita a nuestros psicólogos a que presenten su trabajo en una conferencia de antropología. La decepción es grande por ambos lados. Los antropólogos no son capaces de ver la relevancia de la evidencia experimental a favor de una tesis de la que se sienten seguros que ya ha sido ampliamente demostrada con datos etnográficos. Objetan la artificialidad de los experimentos realizados fuera de un contexto etnográfico. Mas aún, encuentran la concepción de cultura de los psicólogos demasiado simple, ejemplificándose en el hecho de que éstos hablan de las culturas occidental y asiática de forma general. Los psicólogos creen que los antropólogos son sencillamente incapaces de ver la importancia de la evidencia experimental, que critican la metodología experimental sin conocerla, y que no saben apreciar lo mucho que su trabajo podría contribuir a un intercambio fructífero entre psicólogos y antropólogos. Al final, la tesis en sí no se discute en absoluto.

De hecho, en la historia que acabo de contar, los psicólogos hicieron claramente el mayor de los esfuerzos para salir de su campo y producir un trabajo nuevo, pero más con la esperanza de tener un mensaje que compartir que uno que aceptar. Los antropólogos, por su parte, estaban deseando dar la bienvenida a los psicólogos, y esperaban que éstos se inclinasen ante la obvia superioridad de la antropología sobre la psicología en materia de modelos culturales de pensamiento.³²

Sin embargo, esto es en gran medida lo que ocurre, ya que dotar a la Ciencia Cognitiva de este carácter nuevo, y a la vez democrático y unitario es ciertamente difícil. Además, al concebir su validez según la posibilidad de verificación y manipulación empírica de sus teorías se está buscando su fundamentación en algún tipo de

³² Dan Sperber, “ Why Rethink Interdisciplinarity?”, en web: <http://www.interdisciplines.org/interdisciplinarity/papers/1> Abril 2003.

disciplina “dura”, en el sentido de disciplina experimentable y formalizable. Pero obtener una disciplina “dura” como base de la Ciencia Cognitiva, de la interacción de, por ejemplo, la filosofía o la antropología con la Inteligencia Artificial o la neurociencia sería toda una sorpresa, dada la disparidad de sus métodos y supuestos. Al final, lo que se hace es dar preponderancia a ciertas disciplinas como la Inteligencia Artificial y la psicología³³, o la neurociencia y la psicología, más recientemente, situando sus métodos y resultados en el centro de referencia de los trabajos en Ciencia Cognitiva. El resultado es que, en lugar de una integración interdisciplinar real, lo que se tiene son estudios en cada una de las disciplinas utilizando métodos de las reconocidas como centrales. En palabras de Sperber:

Casi veinte años después, ¿qué observamos? Las disciplinas no se han integrado (y, en casos como la filosofía o la antropología, solo están incluidas subdisciplinas en la empresa de la Ciencia Cognitiva), sino que cada disciplina ha tomado prestado de las otras conceptos, problemas, herramientas y criterios. Por dar tan solo un par de ejemplos, el modelado, inspirado por la Inteligencia Artificial, se usa cada vez más como herramienta en psicología y neurociencia, y, en términos generales, la existencia de una posibilidad clara de modelar una hipótesis dada se reconoce como un criterio para juzgar la aceptabilidad de una hipótesis en cualquiera de las Ciencias Cognitivas. Los problemas sobre el carácter y el papel de las representaciones, sacados a la luz primero por la Filosofía de la Mente, han llegado a ser tópicos en todas las Ciencias Cognitivas...³⁴

Como se ha visto anteriormente, la aceptación de supuestos y métodos en la comunidad de científicos cognitivos no es tan general como cabe esperar, con lo que en muchos casos los científicos de cada una de las disciplinas difícilmente aceptarán la mayor validez científica de los resultados de cualquier otra. Esto, junto a la escasez

³³ Christian D Schunn, Kevin Crowley, Takeshi Okada, “The Growth of Multidisciplinarity in the Cognitive Science Society”, en *Cognitive Science* 22(1), 107-130. 1998. En web:

<http://www.lrdc.pitt.edu/Schunn/research/papers/multidic.html>

³⁴ Dan Sperber, “Why Rethink Interdisciplinarity?”

de buenos resultados de los trabajos interdisciplinarios, se traduce en tribulaciones personales y cierta zozobra institucional a la hora de hacer frente al trabajo interdisciplinar. Sperber lo presenta como sigue:

La mayoría de las candidaturas para becas que tenemos que evaluar se han construido en base a una retórica interdisciplinar y describen futuras colaboraciones entre gente de distintas disciplinas, pero en la mayoría de los casos se hace así para cumplir los requisitos de la beca. El contenido científico real generalmente consiste en la yuxtaposición de proyectos monodisciplinarios junto con algún esfuerzo en articular su presentación. Unas pocas propuestas son genuinamente interdisciplinarias, pero a menudo son las peor elaboradas y las que con menor probabilidad mostrarán resultados claros. Y ahora nosotros tenemos que calificar dos propuestas: una propuesta verdaderamente buena cuyo carácter interdisciplinar es superficial y ad hoc, y otra simplemente decente, pero genuinamente interdisciplinar e innovadora. ¿Deberíamos preferir la primera esperando que, tal y como se dice que la fe se obtiene rezando, algo de verdadera interacción y pensamiento interdisciplinar aparezca en lo que era inicialmente un esfuerzo a medias y oportunista, o deberíamos apoyar la segunda propuesta y ver su carácter tentativo y confuso como el precio a pagar por abandonar los caminos bien establecidos? He conocido este dilema antes. Esta vez, voto la propuesta mejor que no es tan interdisciplinar, la cual creo que merece más la beca. Al tiempo, me pregunto, ¿qué clase de comedia es ésta, donde fingimos dotar de fondos a la investigación nueva e interdisciplinar, cuando, de hecho, en primer lugar, hay pocos fondos para la enseñanza y práctica interdisciplinar? ¿Con qué probabilidad podrán surgir propuestas interdisciplinarias sobresalientes en tales condiciones? Y, ¿no están la mayoría de mis colegas satisfechos con el estado actual de la cuestión, que permite que el negocio disciplinar siga adelante al precio barato de algo de retórica interdisciplinar?

... Es fascinante, y en cierto modo descorazonador, observar cómo, semana tras semana, año tras año, se expresan las mismas discrepancias en los mismos términos, entre disciplinas y a veces incluso dentro de cada disciplina, como si no se pudiesen superar nunca las afiliaciones disciplinares y teóricas. Pero esto es sólo la

mitad de la historia. Algunas personas vienen durante un tiempo y lo dejan para siempre, sintiendo que es una pérdida de tiempo. Otros han permanecido durante años y han desarrollado una concepción clara y detallada del trabajo realizado en otras disciplinas, y, en su propio trabajo, abordan asuntos verdaderamente interdisciplinarios extrayendo recursos, aunque a veces a la defensiva, de distintas disciplinas. Algunos de los estudiantes del programa piensan y trabajan realmente “entre” disciplinas, aunque provengan de las ciencias sociales o de la psicología. Así que todos los de nosotros que participamos en este programa, tanto como miembros permanentes o visitantes regulares, tenemos tanto un sentimiento de frustración – ¿no podría esto funcionar mejor, avanzar más rápidamente y dejar de una vez por todas los malentendidos iniciales?- como un sentimiento de consecución de objetivos – aunque no tan ni tan bien desarrollado como nos gustaría, algo nuevo y relevante está emergiendo que no podría haberse forjado en un contexto disciplinar.

... Como les he dicho a unos cuantos estudiantes que querían trabajar en el tipo de interdisciplinariedad que defiendo, elegir un tema de investigación interdisciplinar en el doctorado conlleva serios riesgos en el desarrollo profesional... Veo aquí un círculo vicioso: postponer el trabajo interdisciplinar para cuando un investigador está bien establecido significa que esa investigación se lleva a cabo como una actividad secundaria, con más buena voluntad que verdadera competencia, y que, por tanto, será mucho más difícil para un estudiante encontrar la supervisión adecuada en un área interdisciplinar que en una disciplinar. En términos generales, esto significa que se desalienta a dedicar la inventiva y creatividad de los estudiantes al trabajo interdisciplinar, haciendo este trabajo más lento, e intelectualmente y en la práctica menos atractivo ...³⁵

Una muestra de la dificultad en conseguir la interdisciplinariedad es que muy a menudo se habla de “Ciencias Cognitivas”, en lugar de “Ciencia Cognitiva”, haciendo de esta forma referencia a las aportaciones de las distintas disciplinas a la

³⁵ Dan Sperber, “Why rethink interdisciplinarity?”

explicación de la cognición, pero sin entrar en la cuestión de si forman una ciencia unificada.

Así, por ejemplo, la lingüística aportaría datos sobre la estructura del pensamiento a través del estudio del lenguaje, la antropología a través del estudio de la organización y comportamiento de los individuos en sociedades, la neurociencia a través del estudio de los mecanismos neurológicos al producirse el pensamiento, la psicología a través del análisis y experimentación del comportamiento humano en distintas situaciones, la inteligencia artificial a través de la creación de modelos matemáticos realizados en máquinas que funcionen de la misma forma que el razonamiento humano y así puedan predecir como éste lo haría en nuevas situaciones, y la filosofía, aunque de manera controvertida respecto a si debe ser parte de la Ciencia Cognitiva, ayudando a definir las cuestiones fundamentales y el marco en el que se integren los trabajos en distintos campos.

Pero no toda la lingüística o la psicología, por ejemplo, se considera “oficialmente” una Ciencia Cognitiva. En este reconocimiento, además de que la disciplina en cuestión se ocupe de la cognición, se incluye también el que utilice los métodos y supuestos anteriormente presentados como centrales en Ciencia Cognitiva, como por ejemplo, la simulación por ordenador o la aplicación del paradigma del procesamiento simbólico.

Se tiene entonces una segunda denominación muy extendida para las Ciencias Cognitivas: Antropología Cognitiva, Lingüística Cognitiva, Psicología Cognitiva, Neurociencia Cognitiva e Inteligencia Artificial. La filosofía se resiste a partir de ningún supuesto previo en sus análisis, aunque en algunos casos cuando se ocupa de cuestiones cognitivas se le llame Filosofía de la Mente, y la Inteligencia Artificial no cambia de denominación. Estas disciplinas harían su aportación al estudio de la cognición y a la vez a su propio campo de estudio empleando los supuestos, métodos y resultados de la Ciencia Cognitiva que prevalezcan en cada momento. Esto supone que

heredan los problemas de fundamentación implícitos en estos supuestos y métodos. En ocasiones se llega a clasificaciones y terminologías que pueden resultar poco esclarecedoras, dada la rapidez con la que cambian dichos supuestos y a menudo la disparidad entre ellos. Por ejemplo, el interpretivismo de Clifford Geertz algunas veces se clasifica como dentro de la Antropología Cognitiva³⁶, clasificación con la que él mismo no parece estar muy de acuerdo:

La cultura, ese documento activo, es pues pública, lo mismo que un guiño burlesco o una correría para apoderarse de ovejas... Esto podrá parecer una verdad evidente, pero hay numerosas maneras de oscurecerla. Una de ellas es imaginar que la cultura es una realidad “superorgánica”, conclusa en sí misma, con fuerzas y fines propios; esto es reificar la cultura. Otra manera es pretender que la cultura consiste en el craso esquema de la conducta que observamos en los individuos de alguna comunidad identificable; esto es reducirla: Pero aunque estas dos confusiones todavía subsisten e indudablemente subsistirán siempre, la fuente principal del embrollo teórico que presenta la antropología contemporánea es una concepción que se desarrolló como reacción a esas dos posturas y que ahora está ampliamente sostenida; me refiero a la concepción, para citar a Ward Goodenough, quizá su principal expositor, según la cual “la cultura (está situada) en el entendimiento y en el corazón de los hombres”.

Designada de varias maneras, etnociencia, análisis componencial o antropología cognitiva (una terminología fluctuante que refleja profunda incertidumbre), esta escuela de pensamiento sostiene que la cultura está compuesta de estructuras psicológicas mediante las cuales los individuos o grupos de individuos guían su conducta.³⁷

³⁶ Merwin S. Garbarino, *Sociocultural Theory in Anthropology. A Short History*. Waveland Press, Inc 1983, Holt, Rinehart and Winston, Prospect Heights, Illinois, 1977, 83.

³⁷ Clifford Geertz, *La interpretación de las culturas*, Gedisa, Barcelona, 1990, 1987. 24-25. Original en inglés: *The Interpretation of Cultures*, Basic Books, 1977.

Se presentan a continuación algunas definiciones de cada una de estas nuevas disciplinas o los objetivos de sus publicaciones más representativas:

La Antropología Cognitiva es un subcampo unificado de la antropología cultural cuyo principal propósito es entender y describir cómo la gente concibe y experimenta el mundo dentro de sociedades. La definición de cultura que guía la investigación en Antropología Cognitiva mantiene que la cultura es un sistema cognitivo idealizado – un sistema de conocimientos, creencias y valores – que existe en las mentes de los miembros de la sociedad. La cultura es el equipamiento mental que los miembros de la sociedad utilizan al orientar, interactuar, discutir, definir, categorizar e interpretar el comportamiento social en su sociedad.³⁸

Los antropólogos cognitivos ven la antropología como una ciencia formal. Mantienen que la cultura se compone de reglas lógicas basadas en ideas que están en la mente y a las que se puede acceder (Moore 1999:5). Hay cuatro categorías básicas en antropología cognitiva: semántica, estructuras de conocimiento, modelos y sistemas, y análisis del discurso (D'Andrade 1995:1). Los estudios semánticos de los sistemas de terminologías constituyeron las bases de la primera antropología cognitiva. Los métodos analíticos y etnográficos desarrollados en estos estudios proporcionaron los fundamentos de la etnociencia (también llamada nueva etnografía). En contraste con algunas de las primeras aproximaciones antropológicas a la cultura, ésta no estaba considerada como un conjunto de fenómenos materiales, sino como organizaciones cognitivas de fenómenos materiales (Tyler 1969:3).³⁹

La Lingüística Cognitiva es una escuela de lingüística y Ciencia Cognitiva cuyo propósito es proporcionar explicaciones del lenguaje que se integren con los conocimientos actuales de la mente humana. El principio general de la Lingüística Cognitiva es que el uso del

³⁸ Ronald W. Casson, "Cognitive Anthropology", *MITECS*, 120-121.

³⁹ Sharon N. Solomon, "A Brief History of Cognitive Anthropology", en <http://www.geocities.com/xerexes/coganth.html>, 2000.

lenguaje debe ser explicado haciendo referencia a los procesos mentales subyacentes.

Cognitive Linguistics presenta un fórum para la investigación lingüística de todos los tipos sobre la interacción entre el lenguaje y la cognición. La Lingüística Cognitiva se concentra en el lenguaje como instrumento para organizar, procesar y expresar información, y la revista está dedicada a la investigación de alto nivel en temas como:

- Las características estructurales de la categorización del lenguaje natural (como la prototypicalidad, los modelos cognitivos, las metáforas y las imágenes).
- Los principios funcionales de la organización lingüística (como la iconicidad).
- La interfaz conceptual entre sintaxis y semántica.
- La relación entre lenguaje y pensamiento, incluyendo asuntos como la universalidad y la especificidad del lenguaje.
- El trasfondo experiencial del uso del lenguaje, incluyendo el trasfondo cultural, el contexto del discurso, y el entorno psicológico del desarrollo lingüístico.⁴⁰

Cognitive Psychology se ocupa de los avances en el estudio de la memoria, el procesamiento del lenguaje, la percepción, la solución de problemas y el pensamiento. La revista presenta trabajos originales experimentales y teóricos, artículos sobre metodología y revisiones críticas. *Cognitive Psychology* está especializada en artículos que tienen un gran impacto en la teoría cognitiva y/o que proporcionan nuevos avances teóricos.

Sus áreas de investigación incluyen:

- Inteligencia Artificial
- Psicología del desarrollo
- Lingüística
- Neurofisiología
- Psicología social ⁴¹

⁴⁰ Cognitive Linguistics Journal, en www.degruyter.de/rs/384_386_ENU_h.htm

Hacia finales de los ochenta (...), espoleados en parte por los nuevos y excitantes resultados que surgen del análisis de varias formas de déficits cerebrales y de la técnica de imágenes de tomografía por emisión de positrones (PET), los neurocientíficos y los científicos cognitivos comenzaron a colaborar y a integrar sus metodologías desarrollando un análisis continuo de las bases cerebrales de los procesos cognitivos. El psicólogo George Miller y el neurobiólogo Michael Gazzaniga acuñaron el término *neurociencia cognitiva* para designar la investigación colaborativa que integra las herramientas de los psicólogos sobre la conducta con las técnicas que revelan el funcionamiento cerebral para determinar cómo el cerebro realiza el procesamiento de información que genera el comportamiento. Hoy los científicos neurocognitivos estudian regularmente tanto los procesos psicológicos como la actividad neural, y desde los años noventa la neurociencia cognitiva se ha mostrado como una de las áreas de estudio científico más excitantes y con mayor desarrollo.⁴²

La *neurociencia cognitiva* es un campo que ha surgido recientemente debido a la disponibilidad de nuevas formas de medir y formar imágenes de la actividad del cerebro durante la actividad cognitiva. Los diseñadores de modelos de redes están enfocando su atención en el cerebro, adecuando las arquitecturas y tareas realizadas por las redes al conocimiento sobre áreas del cerebro particulares obtenidos no solo de las neuroimágenes, sino también métodos tradicionales como estudios de lesiones, potenciales de respuesta evocadas (ERP) y registro de células en animales.⁴³

Es difícil saber si es necesaria la interdisciplinariedad para avanzar en el estudio de la cognición o basta, al menos de momento,

⁴¹ Cognitive Psychology Journal, en <http://authors.elsevier.com/JournalDetail.html?PubID=622807&Precis=&popup=>

⁴² William Bechtel, Peter Mandik y Jennifer Mundale, "Philosophy Meets the Neurosciences". En William Bechtel, Peter Mandik, Jennifer Mundale y Robert S. Stufflebeam, *Philosophy and the neurosciences. A reader*. Blackwell Publishers Ltd, Malden, Massachussets, 2001, 6.

⁴³ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, Blackwell Publishers Ltd, 2002, Malden, Massachussets, 1991, 15-16.

con las aportaciones multidisciplinares. Lo que tanto en un caso como en otro está claro es que hay una necesidad de fundamentación, de supuestos comunes bien establecidos y relevantes, para ir avanzando. La dificultad está en encontrar estos supuestos comunes que sirvan para avanzar en todos los aspectos de la cognición, y sin los cuales tanto la Ciencia Cognitiva como las Ciencias Cognitivas perderán poder explicativo.

1.5 Resumen

En este capítulo se ha visto cómo, ligado al campo de estudio que se establece para la Ciencia Cognitiva, se encuentra la concepción y los supuestos que se tengan sobre el pensamiento, al ver cómo la mayoría de éstos tiene que ver con la idea de poder simular la cognición en algún tipo de máquina. Se han visto las dificultades que tienen estos supuestos para dar una explicación de los aspectos de la cognición más difícilmente formalizables, como el carácter social del pensamiento. También se ha visto cómo la Ciencia Cognitiva busca su carácter científico como ciencia formal y natural intentando basar sus teorías en la experimentación, y cómo hay una gran dificultad en establecer ese vínculo y realizar experimentos claros y relevantes, por lo que en las teorías a menudo se recurre a metáforas externas a los experimentos. Esto, junto a la escasez de resultados, hace que existan dudas respecto a su aceptación como ciencia. Por último se ha visto que, a pesar de que la Ciencia Cognitiva se define como una empresa interdisciplinar, todavía no se ha conseguido una integración real de disciplinas, prevaleciendo los métodos y supuestos de algunas de ellas en la mayor parte de los trabajos.

En el siguiente capítulo se verá cómo llegaron a adoptarse estos supuestos y su influencia en las teorías y disciplinas que componen la Ciencia Cognitiva.

CAPÍTULO II

Historia breve de la Ciencia Cognitiva

En este capítulo se presenta el desarrollo histórico de la Ciencia Cognitiva como intento de encontrar una explicación científica verificable, y en cierto modo realizable tecnológicamente, del pensamiento. Se parte del intento de formalización de las leyes del pensamiento por Boole en el siglo diecinueve y se hace un recorrido hasta la actualidad. Se presta atención a los supuestos centrales en cada periodo y a cómo han ido cambiando a medida que los anteriores resultaban insuficientes.

2.1 Precursores

Los intentos de formalización del pensamiento son bastante antiguos. En el siglo trece, Ramón Llull combinaba los razonamientos al azar mediante un mecanismo oriental llamado “zairja”. En el siglo diecisiete, Leibniz intentó asignar un número a cada concepto para luego resolver cualquier cuestión mediante la manipulación de esos números. Nunca consiguió su objetivo y se dio cuenta de la dificultad de su empresa. Uno de los mayores obstáculos era la interconexión de todos los conceptos: “No hay un término tan absoluto ni independiente que no contenga alguna relación y del cual un análisis detallado no lleve a otras cosas o incluso a todo el resto de cosas”⁴⁴.

⁴⁴ Emilienne Naert, “Leibnitz”, en Denis Huisman, ed. *Dictionnaire des Philosophes*, Presses Universitaires de France, París, 1984, en Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 15.

En 1854, en sus *Investigaciones de las Leyes del Pensamiento, en las que se basa la teoría matemática de la lógica y de la probabilidad*, Boole intentó realizar una formalización algebraica de las leyes del pensamiento mediante la teoría de conjuntos al asociar propiedades a los objetos y agruparlos en clases según la presencia o ausencia de cierta propiedad. Posteriormente asignó símbolos a las operaciones necesarias y suficientes para combinar los elementos de esas clases para así tener una notación rápida en su manejo. Por ejemplo, la clase de los elementos pertenecientes a dos clases A y B vendría dada por la unión de ambas, $A \cup B$, la clase de los elementos de A y B con cierta propiedad común a ambas clases sería su intersección $A \cap B$, y la clase de los elementos no contenidos, por ejemplo, en la clase A, como $\neg A$ (complemento). Así, la clase de los elementos con cierta propiedad común a A y B y que no estén en C, se expresaría como $(A \cap B) \cap (\neg C)$. Observó las propiedades de estas operaciones entre sí (asociativa, distributiva, existencia de complemento y existencia de operador unitario para la unión y la intersección) para obtener una formalización de operaciones y por tanto su manejo mecánico. Esta forma de agrupar los objetos en categorías y de manejarlos eran para Boole aspectos fundamentales del pensamiento humano⁴⁵.

Es razonable pensar que si la lógica expresa las leyes del pensamiento, como indica toda la tradición lógica, y la teoría matemática de Boole también las representa, debe existir algún tipo de equivalencia entre ambas. Sustituyendo el concepto de conjunto por el de proposiciones lógicas que pueden ser verdaderas o falsas, siendo equivalente el valor verdadero al conjunto universal (todos los elementos que cumplen lo afirmado) y el valor falso al conjunto vacío (no hay ningún elemento que cumpla lo afirmado) y sustituyendo los operadores unión, intersección y complemento por las operaciones lógicas “O”, “Y” y “NO” se tiene que se pueden combinar dichas

⁴⁵ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 16.

proposiciones como conjuntos. Por ejemplo, la frase “Pedro nació en 1980” (A) Y “Pedro murió en 1975” (B) sería equivalente a $A \cap B$. Si “Pedro nació en 1980” es verdadero (conjunto universal) y “Pedro murió en 1975” es falso (conjunto vacío) entonces la intersección de A y B es el conjunto vacío. O lo que es lo mismo, la proposición “Pedro nació en 1980 y murió en 1975” es falsa. Si además, se sustituye “verdadero” y “falso” por los valores 1 y 0, y las operaciones lógicas O, Y, por los operadores + y \times de la aritmética y NO como la conmutación entre 0 y 1, y se observan, como en teoría de conjuntos, las propiedades de estos operadores, se tendrá una formalización algebraica y manipulable aritméticamente (álgebra de Boole) de los razonamientos lógicos. En el ejemplo, “Pedro nació en 1980” y “Pedro murió en 1975” sería igual a $1 \times 0 = 0$, o lo que es lo mismo, la afirmación es falsa.

La formalización de Boole solo puede aplicarse a un número muy limitado de expresiones del lenguaje. Por ejemplo, las proposiciones “Algún niño nació en 1980” y “algún niño nació en 1975” son ambas verdaderas y por tanto según el álgebra de Boole “algún niño nació en 1980 y en 1975” también sería verdadera ($1 \times 1 = 1$), cuando se ve que es claramente falsa. En 1879, Frege desarrolló un formalismo en el que expresar proposiciones más complejas que tuviesen en cuenta situaciones más generales o abstractas llamado lógica de predicados, en el cual se asignaba o “predicaba” un valor de verdad (verdadero o falso) a cierta proposición, y creó un cálculo algebraico con ese formalismo. Aplicó este cálculo en su programa logicista de reducción de la aritmética al formalismo lógico. Sin embargo, como el mismo Frege reconoció, el simbolismo y su manejo no estaban totalmente desligados del lenguaje, ya que el valor asignado a la expresión y su condición de verdad seguía estando en el uso que se daba al lenguaje de antemano. Esto se puso de manifiesto en las dificultades técnicas de estos sistemas a la hora de probar su “consistencia” o “coherencia” y su completitud. La consistencia de un sistema garantiza que todo lo que se pueda probar dentro de él sea verdadero. Es decir, que no haya contradicciones dentro del sistema.

La prueba de consistencia es, como dice Garrido⁴⁶, algo tan vital para un sistema deductivo como el certificado de normalidad psiquiátrica para un ciudadano respetable. La completitud de un sistema indica que todas las proposiciones verdaderas que pueden expresarse en el sistema son formalmente deducibles de los axiomas. En 1902, Russell le indicó a Frege la presencia de “paradojas” o contradicciones dentro de su sistema, con lo que no se conseguía la perseguida formalización del pensamiento.

En 1910, Russell y Whitehead publicaron los *Principia Mathematica*, donde se intentaba reducir las matemáticas a la lógica mostrando cómo todas las matemáticas se siguen de premisas puramente lógicas y usan conceptos solamente definibles en términos lógicos. Esta obra tuvo éxito en la formación de un sistema lógico consistente que resolviese las paradojas del sistema de Frege y estableció demostraciones lógicas de muchas partes de las matemáticas. Sin embargo, lo realizó introduciendo razonamientos que muy dudosamente se pueden considerar dentro de la lógica, como el axioma de infinitud, que indica la existencia de al menos un conjunto infinito a partir del cual se pueden formar otros conjuntos infinitos. Esta introducción de elementos no lógicos en las demostraciones generaba cierta inquietud respecto a la validez de las demostraciones de consistencia y completitud de los sistemas lógicos.

Al mismo tiempo, Hilbert, trabajando en la axiomatización de toda la matemática y en la demostración de la consistencia de los sistemas formales, y viendo los problemas con los que se encontraron Frege, Russell y él mismo, trató de construir pruebas “absolutas” de consistencia, en las que no se necesitase utilizar premisas y métodos de otro sistema que el que se quiere probar. Esto garantizaría el poder formalizar el pensamiento al no haber siempre un elemento “exterior”, no formalizable, que escapase al sistema y al que se tuviese que

⁴⁶ Manuel Garrido, “El *ego cogito* del siglo XX”, en Ernest Nagel y James R. Newman, *El Teorema de Gödel*, Tecnos, Madrid, 1999, 2000, 7. Original en inglés: *Gödel's Proof*, New York University Press, 1958.

apelar. Como primer paso se requería la completa formalización del sistema deductivo mediante la extracción del significado de las expresiones del sistema, a las que se debe considerar signos vacíos. Después se derivarían los teoremas como una estricta transformación, siguiendo un número finito de reglas bien definidas, de un conjunto de signos en otros.

Los signos que servirían para este propósito serían los números naturales finitos, y las reglas de transformación serían las utilizadas en la aritmética. Hilbert consideraba que había una rama privilegiada de la matemática, la teoría de los números naturales, que estaba fundamentada en bases puramente intuitivas de signos concretos, no mediadas por la lógica. Estos números y sus relaciones (operaciones) existirían como experiencia inmediata no reducible a ningún otro elemento. La teoría desarrollada así estaría libre de contradicciones, simplemente porque sus proposiciones no tendrían estructura lógica. En palabras de Hilbert se tenían que utilizar:

Objetos discretos extralógicos, que existen intuitivamente como experiencia inmediata anterior a todo pensamiento. Si las inferencias lógicas deben ser ciertas, entonces se debe poder examinar estos objetos en todas sus partes, y su presentación, su diferencia, su sucesión (como los objetos mismos) debe existir para nosotros de forma inmediata, intuitiva, como algo que no se puede reducir a otra cosa.⁴⁷

Al utilizar la aritmética de los números naturales finitos como instrumento de formalización del sistema lógico se estaba reduciendo el sistema a parte de la aritmética y a la vez se utilizaba a ésta de forma “metamatemática” para probar la consistencia del sistema, por lo que los mismos métodos de prueba quedaban acotados al sistema, evitándose así la continua referencia a “algo” exterior al sistema, tal y como se pretendía evitar. La importancia del programa de Hilbert

⁴⁷ Richard Zach, “Hilbert’s Program”, en *Stanford Encyclopedia of Philosophy*, capítulo 1.3.

estaba en que si se conseguía la prueba absoluta de consistencia de los sistemas lógicos por estos medios finitistas se estaría en camino de conseguir una formalización completa de al menos ciertas partes del pensamiento bastante complejas. Una vez formalizado un sistema lógico que lo representase adecuadamente, no cabrían más interpretaciones ni elementos nuevos que añadir, ya que todo lo que se pudiese decir estaría ya dentro del sistema y sería deducible en él.

En 1931, en un artículo con el título “Sobre las proposiciones formalmente indecidibles de los *Principia Mathematica* y sistemas conexos”, Gödel demostró, mediante métodos aritméticos finitistas, que había proposiciones válidas (bien formadas) en el sistema de los *Principia Mathematica* y en general, en todo sistema axiomático aritmetizable en el que toda expresión que se pudiera probar fuese correcta en cuanto a su contenido, que no se podían deducir de sus axiomas y teoremas. O lo que es lo mismo, demostró la incompletitud de dichos sistemas. También demostró mediante los mismos métodos la imposibilidad de demostrar “desde dentro del sistema” (es decir, mediante métodos aritméticos finitistas) la ausencia de contradicciones o consistencia de dichos sistemas.

El desarrollo de las matemáticas en dirección a una mayor exactitud ha llevado – como es bien sabido – a la formalización de gran parte de ella, de manera que se pueden llevar a cabo las pruebas de acuerdo a unas pocas reglas mecánicas. Los sistemas formales más completos realizados son, por una parte, el sistema de los *Principia Mathematica* y, por otra parte, el sistema axiomático de la teoría de conjuntos de Zermelo-Fraenkel (ampliada más tarde por J. v. Neumann). Estos dos sistemas son tan extensos que todos los métodos de prueba utilizados actualmente han sido formalizados en ellos, es decir, se han reducido a unos pocos axiomas y reglas de inferencia. Se podría por tanto inferir a primera vista que estos axiomas y reglas de inferencia son también suficientes para decidir *todas* las cuestiones matemáticas que se puedan expresar formalmente de alguna manera en los sistemas en cuestión. Se muestra más abajo que este no es el caso, y que en los dos sistemas mencionados hay de hecho problemas relativamente simples en la teoría de números enteros que no se pueden derivar de los

axiomas. Esta situación no se debe en algún aspecto a la naturaleza especial de estos sistemas, sino que es válida para una clase bastante amplia de sistemas formales, incluyendo, en particular, a todos los que surgen de la adición de un número finito de axiomas a los dos sistemas mencionados, suponiendo que por ellos no se pueda probar ninguna proposición falsa...⁴⁸

La importancia de este descubrimiento estaba en que al demostrar la incompletitud de los sistemas axiomáticos se ponía fin al programa logicista, que trataba de expresar todas las matemáticas mediante axiomas y teoremas lógicos. A la vez, se ponía fin al programa formalista de Hilbert, que intentaba dar una prueba de consistencia absoluta de un sistema axiomático mediante métodos aritméticos finitistas, ya que utilizando estos mismos métodos se demostró la imposibilidad de probar la consistencia del sistema. Continuando con la metáfora de Garrido, podríamos decir que un sistema no puede demostrar la ausencia de contradicciones en él, igual que ningún ciudadano puede expedir su propio certificado de normalidad psiquiátrica. También se demostró, al encontrar una expresión aritmética bien formada no deducible dentro de la aritmética axiomatizada, que no era posible dar cuenta de toda la aritmética formalizándola, ya que se podían encontrar expresiones (verdades) aritméticas bien formadas que no se podían deducir mediante el sistema. Puede pensarse que esto se corregiría añadiendo la nueva expresión como un axioma, pero el sistema resultante sufriría el mismo defecto. No solo eso, sino que nunca se sabría, desde “dentro del sistema” si éste está libre de contradicciones (si es consistente).

Considerando las matemáticas como una parte del conocimiento humano, y viendo que es imposible formalizarlas axiomáticamente, se deduce que el pensamiento humano no es

⁴⁸ Kurt Gödel, “On Formally Undecidable Propositions of Principia Mathematica and Related Systems I”, en web: <http://home.ddc.net/ygg/etext/godel/godel3.htm>, 1998. Original en alemán: “Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme, I”, en *Monatshefte für Mathematik und Physik*, 38, 1931, 173-198.

formalizable en un sistema axiomático (y en el que las reglas de deducción que se empleen sean las de la lógica clásica), ya que es capaz de “captar” verdades, expresiones válidas, que no se pueden derivar de dichas formalizaciones.

En 1936, abordando también el problema de Hilbert de encontrar un método general de prueba de la validez de un sistema matemático o lógico (*Entscheidungsproblem*) mediante un método intuitivamente inmediato y desligado de toda interpretación, Alan Turing consideró este método intuitivo como los movimientos mecánicos simples de escritura y borrado de números naturales representados en código binario (“1” y “0”), y como el cambio de posición en un papel de la longitud necesaria (finita o infinita) conforme a unas reglas fijas establecidas (algoritmo). A este conjunto lo denominó “máquina computadora”⁴⁹. Demostró, utilizando el mismo tipo de expresiones que Gödel y los métodos que se acaban de describir, que hay ciertas expresiones o fórmulas válidas que no se pueden obtener desde “dentro” del sistema, lo que significaba llegar de otra forma a las mismas conclusiones que Gödel.

Al mismo tiempo, Alan Turing mostraba un tratamiento mecánico finito a seguir para el cálculo, evaluación y representación de una expresión matemática o lógica. A estas funciones las llamó *computables*. Las funciones (o expresiones) computables son aquellas a las que se puede llegar en un número finito de pasos. De esta forma sentó las bases para construir una máquina real a partir de su máquina computadora. Considerando una máquina computadora general como la que evalúa todas las, a su vez, posibles evaluaciones de un sistema se llega al concepto actual de computadora, donde cada programa que utiliza se identifica con una máquina computadora de Turing (o máquina de Turing). En el fondo de todas estas demostraciones está la

⁴⁹ Alan M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem”, cap 2, Definitions. En web <http://www.abelard.org/turpap2/tp2-ie.asp>, 2001. Original en *Proceedings of the London Mathematical Society*, Series 2, 42, 1936, 230-265. Errata en Series 2, 43, 1937, 544-546.

posibilidad de representar mediante un número natural cualquier expresión, y cualquier expresión de cualquier expresión, y así sucesivamente. Si se llega al resultado deseado en un número finito de pasos, la función es computable.

Gödel aplicó su demostración a sistemas axiomáticos y lógicos expresables mediante la aritmética de los números naturales. Cabe preguntarse si hay sistemas de pensamiento no axiomáticos, o axiomáticos con una lógica distinta, con “otra forma de deducir”, en los que estos resultados no sean válidos, y por tanto, en los que estuviese abierta la posibilidad de la formalización del pensamiento humano. Este es el camino que siguen las investigaciones respecto a lo que se llama “máquinas no de Turing” o “lógica paraconsistente”⁵⁰. Uno de los problemas a la hora de realizar estos sistemas en una máquina radica en la dificultad de automatizar reglas de decisión que admitan como válidas a la vez, por ejemplo, una expresión y su negación. Por otra parte, intuitivamente resulta difícil encajar la idea de “sistema”, sea axiomático o no y con una lógica u otra, en las características más humanas del pensamiento, como pueden ser la creatividad o la toma de decisiones aparentemente contradictorias.

2.2 Cibernética

Durante la misma década de los años treinta en la que Turing y Gödel presentaron sus trabajos, tenían lugar en la Escuela de Medicina de Harvard unas reuniones informales, dirigidas por el fisiólogo Arturo Rosenblueth, para discutir trabajos sobre metodología científica. La mayoría de los asistentes eran científicos jóvenes de la Escuela, aunque se incorporaron también físicos y el matemático y profesor del Instituto de Tecnología de Massachussets Norbert Wiener, este último debido a la convicción de Rosenblueth y del

⁵⁰ Graham Priest, Koji Tanaka, “Paraconsistent Logic”, en *Stanford Encyclopedia of Philosophy*.

mismo Wiener de la necesidad de la revisión crítica de los temas matemáticos. Ambos estaban convencidos de la importancia para el avance científico del trabajo interdisciplinar en materias que quedan en la frontera de varias disciplinas. Como expone Wiener:

Durante muchos años habíamos compartido el Dr. Rosenblueth y yo la convicción de que los campos más fructíferos para el desarrollo de las ciencias eran aquellos que habían quedado desatendidos como tierra de nadie entre las distintas especialidades establecidas. Desde tiempos de Leibniz no ha existido probablemente nadie con semejante dominio sobre la actividad intelectual de una época. Desde entonces la competencia científica se ha ido restringiendo cada vez más a los especialistas, circunscritos a unos campos con progresiva tendencia a estrecharse. Hace un siglo no había un Leibniz, pero había un Gauss, un Faraday y un Darwin. Actualmente pueden contarse con los dedos de la mano los científicos que no sean exclusivamente matemáticos, físicos o biólogos. Puede haber topólogos, acústicos o coleopteristas que dominen la jerga de su especialidad y conozcan toda la literatura de su especialidad y sus ramificaciones, pero la mayoría de las veces considerarán cualquier otra disciplina como algo propio de un colega con despacho tres puertas más allá en el pasillo, y creerán que cualquier interés por su parte hacia el tema es una injustificable violación de una exclusiva.

Durante años hemos soñado con una institución de científicos independientes que trabajasen en equipo en una de esas áreas científicas, no como subalternos de un ejecutivo, sino voluntariamente incorporados, por necesidad espiritual, y que entiendan ese campo de trabajo como algo unitario en donde se comparte la capacidad de conocimientos.⁵¹

Durante los años veinte, y también movido por una “necesidad espiritual”, en este caso, la comprensión de cómo se produce el

⁵¹ Norbert Wiener, *Cibernética o el control y comunicación en animales y máquinas*, Tusquets Editores, Barcelona, 24-25. Original en inglés: *Cybernetics or Control and Communication in the Animal and the Machine*, The Massachusetts Institute of Technology, 1948, 1961.

conocimiento matemático⁵², en los términos de los *Principia* de Whitehead y Russell, Warren S. McCulloch cursó los estudios de psicología, donde conoció el conductismo, el psicoanálisis y el introspeccionismo (ver sección 3.1). Insatisfecho con las respuestas que la psicología ofrecía a su pregunta, estudió medicina y posteriormente neurología. Trabajó durante los años treinta en hospitales psiquiátricos junto a Eilhard von Domarus, quien impulsó la idea de que el habla de los esquizofrénicos no obedece las reglas de la lógica normales, y estudió el campo de la naturaleza lógica de la mente y su relación con la neurofisiología y la demencia. Realizó experimentos sobre el funcionamiento del córtex cerebral en chimpancés mediante la aplicación de impulsos eléctricos, llegando a ser una autoridad reconocida en este campo. Pero su proyecto vital de investigación fue el encontrar una correspondencia entre la lógica, como integradora también de las matemáticas en el programa de Russell, y las redes de neuronas del córtex. Si se pudiese establecer formalmente esta correspondencia, se podría considerar el funcionamiento de las redes neurales como el de una gran máquina lógica o como el de una computadora. Los experimentos sobre la organización funcional del sistema nervioso serían un paso adelante hacia una epistemología experimental, y a través de la combinación del rigor de la lógica formal y de experimentos neurofisiológicos se podría presumiblemente saber científicamente cómo conocemos los números y mucho más, y expresar este conocimiento formalmente.⁵³

McCulloch vio la similitud entre las operaciones lógicas expresadas con los valores “verdadero” o “falso” y las conexiones neuronales en el caso de que éstas transmitan o no la corriente a partir de un umbral, pero carecía de los conocimientos lógicos y matemáticos para establecer formalmente esta similitud. En 1941

⁵² Warren S. McCulloch, “What is a Number, that a Man may know It, and a Man, that he May know a Number?”, en Warren S. McCulloch: *Embodiments of Mind*, The MIT Press, Cambridge, Massachussets, 1970, 1965, 2.

⁵³ Steve Joshua Heims, *The Cybernetics Group 1946-1953. Constructing a Social Science for Postwar America*, The MIT Press, Cambridge, Massachussets, 1993, 1991, 35.

aceptó un puesto de director de investigación en el departamento de psiquiatría de la Escuela de Medicina de la Universidad de Illinois, y en esta universidad conoció a Walter Pitts⁵⁴, un adolescente con un talento matemático prodigioso que aportó los conocimientos y la originalidad necesaria para la formalización de las teorías de redes neuronales.

Mientras tanto Wiener, Rosenblueth y el ingeniero Julian Bigelow, basándose en los trabajos anteriores de Walter Cannon (1929) sobre autorregulación de parámetros en los seres vivos (homeostasis), y en los propios estudios que estaban llevando a cabo Wiener y Bigelow sobre predicción curvilínea del vuelo en artillería antiaérea, llegaron a establecer una equivalencia entre ambos sistemas en donde los resultados y las causas se determinan mutuamente de manera circular para comportarse de tal forma que se llegue a un objetivo final o “propósito”. Este objetivo final sería, en el caso de los seres vivos, la regulación de un determinado parámetro (temperatura, presión sanguínea, etc.) y en el caso de las máquinas minimizar cierta función (como la función error entre la trayectoria descrita por el avión y la prevista en la máquina antiaérea). La originalidad de estos trabajos estaba principalmente en la consecución de la formalización matemática de todo el proceso, lo cual no se había conseguido antes, mediante la predicción de series temporales con ayuda del cálculo de variaciones, en donde los cambios obtenidos en la variable a regular (efecto), actuaban a la vez como entrada (causa) a tomar en cuenta para su regulación. Esta formulación matemática hacía posible pasar de hablar intuitivamente de la similitud de organismos vivos y máquinas a hablar acerca de la equivalencia formal entre el sistema biológico y el sistema mecánico.

En 1942 tuvo lugar una reunión organizada por la fundación Macy, fundación que tenía como objetivo apoyar los programas para la formación de los profesionales de la salud y mejorar la

⁵⁴ Para más detalles, aunque siempre escasos, sobre este personaje sorprendente puede verse Heims: *The Cybernetics Group*, 40-48.

representación de las minorías en dichas profesiones, con el título “Inhibición Cerebral”, que tenía como temas el reflejo condicionado y la hipnosis. A esta reunión acudieron Rosenblueth, Lorente de Nó, fisiólogo discípulo de Cajal, Gregory Bateson, antropólogo y conocido de Frank Fremont-Smith, que era el organizador del encuentro y director médico de la fundación, Margaret Mead, antropóloga y esposa de Bateson, McCulloch, Lawrence K. Frank, psicólogo social y colaborador de la Fundación, y el psicoanalista Lawrence Kubie. Saliéndose de los temas oficiales a tratar, Rosenblueth presentó los resultados de su trabajo con Wiener y Bigelow bajo el título “Conducta, Propósito y Teleología”⁵⁵, en el que se establecía un paralelismo entre ciertos tipos de máquinas y organismos vivos. La mayoría de los asistentes vieron abierto un gran campo de estudio en la presentación de Rosenblueth. McCulloch instó a Fremont-Smith a continuar las reuniones y éste aceptó, aunque se postpusieron para después de la guerra.

El año siguiente McCulloch y Pitts publicaron sus trabajos en un artículo con el título: “Un cálculo lógico de las ideas inmanentes en la actividad nerviosa”⁵⁶. En este trabajo se conseguía, bajo ciertos supuestos de funcionamiento del sistema nervioso, formalizar la actividad nerviosa como una máquina de Turing y dar una explicación de la memoria como lazos cerrados de potenciales eléctricos, basándose en los trabajos sobre circuitos reverberantes cerebrales de Lorente de Nó.

Tanto el trabajo de Wiener y Bigelow, como el de McCulloch y Pitts, mostraban el paralelismo entre máquina y persona, presididos por la formalización matemática de estos “sistemas”. Esto constituía, por un lado, una antropomorfización de la máquina, al comenzar a hablar de propósitos y fines en ella, aunque estos propósitos y fines

⁵⁵ Arturo Rosenblueth, Norbert Wiener, Julian Bigelow. “Behavior, Purpose and Teleology”, en *Philosophy of Science*, Vol. 10, No. 1 (Jan., 1943), 18-24.

⁵⁶ Warren S. McCulloch and Walter H. Pitts, “A Logical Calculus of the Ideas Inmanent in Nervous Activity”, en *Bulletin of Mathematical Biophysics*, Chicago, University of Chicago Press, 1945, 89-93.

estuviesen implícitos en la estructura física de la máquina y no tuviesen ningún contenido semántico. Y por otro, y quizás en mayor grado, una mecanización de lo humano⁵⁷, un reduccionismo orgánico, al interpretar los procesos fisiológicos, neurológicos y en último término cognitivos, como sistemas descriptibles matemáticamente. McCulloch, que estaba interesado en la formulación matemática de los mecanismos cognitivos, lo expresa como la reducción de la epistemología a una ciencia experimental, con la ayuda de la física, química y las nuevas teorías de la comunicación:

Aunque mi interés siempre ha sido reducir la epistemología a una ciencia experimental, el grueso de mis publicaciones se ha ocupado de la física y química del cerebro y mente de los animales. Debido a mi interés en la organización funcional del sistema nervioso, este conocimiento era tan necesario como insuficiente en problemas de comunicación en los hombres y en las máquinas.⁵⁸

Tanto Rosenblueth y McCulloch, como Wiener y otros matemáticos como John von Neumann, se conocían con anterioridad a este encuentro sobre “Inhibición Cerebral”, lo cual cerraba el círculo de intereses interdisciplinarios entre matemáticos, fisiólogos y neurólogos. Von Neumann era un matemático de prestigio reconocido, brillante y polifacético, con contribuciones al desarrollo de las computadoras, de la bomba atómica, de las teorías matemáticas de juegos y de las metamatemáticas. Sus intereses eran muy variados y poseía una gran erudición en diversos campos, estando entre ellos el que estaban desarrollando Rosenblueth, Wiener, McCulloch y Pitts. Así, en 1943 y 1944 von Neumann y Wiener organizaron una reunión de todos los principales interesados en estos temas en Princeton. Wiener, Pitts, von Neumann, McCulloch y Lorente de Nó, entre otros, acudieron a estas reuniones.

⁵⁷ Jean-Pierre Dupuy, *The Mecanization of the Mind. On the Origins of Cognitive Science*, Princeton University Press, Princeton, Nueva Jersey, 51. Original en francés: *Aux origines des sciences cognitives*, La Découverte, Paris, 1994.

⁵⁸ Warren S. McCulloch, “What is a Number, that a Man may know It, and a Man, that he May know a Number?”, en Warren S. McCulloch: *Embodiments of Mind*, 1.

Se generó un gran entusiasmo entre ellos y en 1945 Wiener intentó fundar en el MIT un laboratorio para investigar y analizar experimentalmente estos conceptos, para el que contrataría a von Neumann como gran director del proyecto, pero finalmente éste se decidió por el proyecto de construcción del primer ordenador de propósito general, el EDVAC, en Princeton, quedando las reuniones apalabradas anteriormente por McCulloch y Fremont-Smith para después de la guerra, las llamadas “Conferencias Macy”, como únicos centros de discusión futuros.

En 1946 tuvo lugar la primera de las diez conferencias Macy con el título “Mecanismos de Realimentación y Sistemas Causales Circulares en Biología y en Ciencias Sociales”. A partir de la sexta conferencia se denominaron “conferencias sobre Cibernética”, tomando el término acuñado por Wiener para nombrar su teoría de la retroalimentación en seres vivos y máquinas. Este término proviene del griego “timonel”, como buen ejemplo de sistema autorregulado formado por el timonel y la nave, ya que el timonel corrige el rumbo en función de sus variaciones respecto al objetivo. Acudieron von Neumann, Wiener, McCulloch, Pitts, Rosenblueth, Bigelow, Lorente de Nó, Bateson, Frank, Fremont Smith, el psicólogo Eric Klüver y el neurofisiólogo Ralph W. Gerard, entre otros. Von Neumann presentó el concepto de máquina de propósito general y realizó comparaciones con la estructura del sistema nervioso, tal y como lo describía Lorente de Nó, en términos eléctricos, y éste habló de las neuronas como elementos de un autómata corporal. Rosenblueth, McCulloch y Pitts presentaron las teorías de sus artículos anteriormente citados y Wiener hizo énfasis en la noción de mensaje o información, en el sentido de variación de energía entre un valor anterior y posterior, como el determinante del comportamiento de los mecanismos y organismos autorregulados. Posteriormente Bateson comentó la posibilidad de aplicar estos conceptos en ciencias sociales, en las que echaba de menos una teoría sólida, y von Neumann presentó las posibles aplicaciones de sus teorías matemáticas a la explicación del comportamiento económico.

A partir de estas presentaciones se generó un debate respecto a la adecuación a la realidad del modelo “todo o nada” de transmisión neuronal y sobre cómo, suponiendo que este modelo fuese cierto, era posible que el cerebro resolviese situaciones imposibles para la máquina de Turing, tal y como demostró Gödel. Klüver expuso el problema de la “gestalt” como percepción del todo más allá de la adición de cada parte y cómo la neurología aún no era capaz de explicarlo. Gerard habló sobre la posibilidad del comportamiento analógico del cerebro como opuesto a las teorías digitales presentadas y Frank destacó la necesidad de tener un lenguaje más general que el de cada una de las disciplinas, para la clarificación y refinamiento de los conceptos que se estaban presentando.

En las siguientes conferencias se sucedieron los debates sobre los temas de la primera conferencia y se incluyeron algunos nuevos. En 1948, en el congreso de Hixon y en la conferencia Macy de 1949, von Neumann abordó el tema de la complejidad en autómatas celulares, mostrando cómo el cerebro es más lento que el ordenador pero dispone de muchas más conexiones en paralelo y más complejas, y cómo surge el problema de identificar el tamaño y organización que debe tener para poder dar cuenta de determinada propiedad de la mente. En la misma conferencia el psiquiatra Lawrence Kubie mencionó la necesidad de incluir al observador en la definición del sistema. Se sucedieron las disputas entre psicoanalistas y neurólogos, los primeros no concediendo capacidad explicativa a la neurología y los segundos acusando de poco científicos a los psicoanalistas. Dado el impacto en los medios de comunicación, que presentaron el libro de Wiener “Cibernética”, como toda una revolución, Gerard advirtió que se estaban exagerando las posibilidades de la cibernética, cuando las teorías que se exponían en las reuniones aún no están contrastadas. En 1951, después de la publicación del artículo de Turing “Maquinaria

Computacional e Inteligencia”⁵⁹, se abordó el tema de si la información debe contener “significado”, o debe ser un mero hecho físico, de la misma forma que se hablaba de “experiencia subjetiva” frente a “mecanismos neuronales”. En la misma conferencia de 1951 surgió el debate de si las computadoras son capaces de realizar inferencias inductivas si emplean capacidades probabilísticas, y en 1952 Bigelow alertó de la extrapolación sin justificación suficiente de los resultados matemáticos a las ciencias sociales.⁶⁰

A medida que avanzaban las conferencias se sucedían las bajas significativas, como las de Wiener y von Neumann, y los intereses de los participantes eran más divergentes. Finalmente, en la última en 1953, McCulloch presentó el resultado de las conferencias:

Einstein definió una vez la verdad como el consenso obtenido al tomar en consideración observaciones, sus relaciones, y las relaciones de los observadores... Desafortunadamente para nosotros, nuestros datos no se podrían definir tan sencillamente. Se han obtenido por métodos extremadamente distintos, por observadores condicionados por formaciones y talentos distintos y relacionados entre sí solo por una torre de Babel de jerga y terminología técnica. Nuestro acuerdo más notable es que hemos aprendido a conocernos un poco mejor, y a “luchar deportivamente con lo puesto”. Eso suena democrático, o mejor, anarquista, como me lo han recordado ustedes en dos ocasiones. Aparte de las tautologías de la teoría, y de la autoridad del acceso único a un hecho en cuestión por medio de la observación personal, nuestro consenso nunca ha sido unánime. Y aunque sí lo hubiera sido, no veo la razón por la que Dios tuviese que haber estado de acuerdo con nosotros.⁶¹

⁵⁹ Alan Turing, “Computing Machinery and Intelligence”, en *Mind*, Vol. 59, octubre 1950, 433-60. Disponible en web: <http://crl.ucsd.edu/~elman/Courses/cog202/Papers/turing-cmi.html>

⁶⁰ Una exposición cronológica de los temas tratados en cada una de las conferencias Macy puede encontrarse en la web de la Sociedad Americana para la Cibernética: <http://www.asc-cybernetics.org/>

⁶¹ Steve Joshua Heims, *The Cybernetics Group 1946-1953. Constructing a Social Science for Postwar America*, 277.

La etapa cibernética supuso un verdadero encuentro interdisciplinar en temas de gran interés entre científicos de primera línea. Pero eran más unas conferencias informales que un proyecto de investigación. Se trataba de exponer temas a debatir en lugar de desarrollar un proyecto como le hubiese gustado a Wiener en el MIT. Aun así la gran riqueza intelectual y variedad de los contenidos tratados hacen que estos temas estén presentes en la actualidad como temas abiertos de discusión y como fuentes de ideas nuevas. El impulso de las conferencias Macy continuó en forma de programas de investigación en redes neuronales hasta finales de los años cincuenta, hasta que un nuevo enfoque en el estudio de la cognición llegó a ser la corriente dominante.

2.3 Cognitivismo

En 1949, el fisiólogo Donald Hebb postuló que las conexiones cerebrales cambian a medida que aprendemos distintas tareas, y que estructuras neurales específicas y nuevas dan cuenta del conocimiento. Según esta teoría, mediante la activación repetida de una neurona por otra a través de una sinapsis particular se incrementaba su conductividad, haciendo que las posteriores activaciones de esta neurona fueran más probables ante el mismo estímulo, con lo que se inducía la formación de caminos neuronales conectados más estrechamente, lo que equivalía a decir que la tarea quedaba “aprendida”.

Marvin Minsky, un matemático con interés en neurología y psicología de la Universidad de Harvard, estaba llegando al mismo tiempo a conclusiones similares a las de Hebb. En 1951, Minsky intentó construir una red neural con dispositivos electrónicos, pero el proyecto fue un fracaso debido a las dificultades de ingeniería para reforzar las conexiones y hacerla “aprender”. Siguió esta línea de investigación hasta 1954, cuando se convenció de que una red de este

tipo lo suficientemente grande para poder razonar requeriría miles de millones de neuronas. Ante la imposibilidad de realizar tal red, buscó otras formas de construir una máquina que realizara ciertas tareas. En 1955 conoció a Ray Solomonoff, un matemático que estaba trabajando en la especificación formal de una máquina con capacidad de aprendizaje. Esto sorprendió a Minsky, quien a partir de ese momento seguiría esa línea de investigación:

Quedé tan impresionado que decidí que esto era mucho más productivo que el sistema de red neural, en el que se construye una pieza de *hardware* y esperas que haga lo correcto. Con este nuevo enfoque, intentabas realizar teorías de qué tipo de inferencias querías hacer, y luego te preguntabas, “¿Cómo haría una máquina que hiciese esto?”. Era una línea de pensamiento distinta.⁶²

En 1950, el artículo de Turing, “Computing Machinery and Intelligence”, reforzaba la idea de crear máquinas inteligentes. En éste mostraba su convicción de que las computadoras (o máquinas de Turing) podrían algún día ser inteligentes, y tomaba como prueba de la inteligencia el que la máquina fuese capaz de engañar a un entrevistador humano en un juego de imitación. Para este propósito, Turing predijo que la máquina no necesitaría una cantidad ilimitada de memoria, sino que bastaría una cantidad que estaría disponible en unos años.

Creo que aproximadamente dentro de cincuenta años será posible programar computadoras con una capacidad de almacenaje de alrededor de 10^9 , para hacerlas realizar el juego de imitación tan bien que el interrogador medio no tendrá más del 70 por ciento de probabilidades de realizar la identificación correcta después de cinco minutos de interrogatorio⁶³.

⁶² Marvin Minsky, “Entrevista con Crevier, 13 de mayo de 1991”, en Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 37.

⁶³ Alan Turing, *Computing Machinery and Intelligence*. *Mind*, Vol. 59, octubre 1950, 433-60, en web:

<http://crl.ucsd.edu/~elman/Courses/cog202/Papers/turing-cmi.html>

Al mismo tiempo, los matemáticos John McCarthy y Claude Shannon exploraban las posibilidades de los autómatas finitos de von Neumann como dispositivos inteligentes. McCarthy se dio cuenta de que la inteligencia y el comportamiento de una máquina de Turing en cierto modo no se parecían al comportamiento humano, ya que pequeños cambios producían comportamientos muy distintos en la máquina y pequeños cambios de comportamiento requerían grandes cambios en la estructura de la máquina. Sin embargo el trabajo práctico con ordenadores en 1955 en IBM le convenció de que se podían construir máquinas inteligentes. McCarthy conoció a Nathaniel Rochester en 1955, quien estaba trabajando en simular redes neurales en el ordenador IBM 704, y éste se interesó por su enfoque y el de Minsky respecto a la construcción de una máquina inteligente.

En 1955 McCarthy, Minsky, Shannon y Rochester, con la financiación de la fundación Rockefeller, convocaron unos encuentros durante el verano de 1956 en Dartmouth College (Hanover, New Hampshire), con el objetivo de impulsar la investigación en lo que denominaron “Inteligencia Artificial”.

Proponemos que diez personas durante dos meses en el verano de 1956 en Dartmouth College en Hanover, New Hampshire, profundicen en el estudio de la inteligencia artificial. Este estudio debe proceder sobre la conjetura de que cualquier aspecto del aprendizaje o cualquier otro rasgo de la inteligencia puede en principio describirse de forma tan precisa que se puede construir una máquina que lo simule. Se intentará averiguar cómo realizar máquinas que utilicen el lenguaje, construyan abstracciones y conceptos, solucionen algunos tipos de problemas reservados de momento a los hombres, y se mejoren a sí mismas. Pensamos que se puede avanzar significativamente en uno o más de estos problemas si un grupo de científicos convenientemente seleccionado trabaja en ellos durante un verano.

Los siguientes son algunos aspectos del problema en inteligencia artificial:

1. Computadoras Automáticas

Si una máquina puede realizar una tarea, entonces se puede programar una calculadora para que simule la máquina. Las velocidades y capacidades de memoria de las computadoras actuales pueden ser insuficientes para simular muchas de las funciones de más alto nivel del cerebro humano, pero el mayor obstáculo no es la carencia de capacidad, sino nuestras limitaciones para escribir programas que obtengan el mayor rendimiento de lo que tenemos.

2. Cómo se puede Programar una Computadora para que use un Lenguaje

Se puede especular que gran parte del pensamiento humano consiste en manipular palabras de acuerdo a reglas de razonamiento y reglas de conjetura. Desde este punto de vista, realizar una generalización consiste en admitir una palabra nueva y algunas reglas por medio de las cuales las frases que contengan la palabra implican y son a su vez implicadas por otras. Esta idea nunca se ha formulado con precisión ni se ha trabajado en ningún ejemplo.

3. Redes Neuronales

Cómo se puede organizar un conjunto de (hipotéticas) neuronas para que formen conceptos. Se ha realizado un considerable trabajo experimental y teórico por parte de Uttley, Rashevsky y su grupo, Farley y Clark, Pitts y McCulloch, Minsky, Rochester y Holland, y otros. Se han obtenido resultados parciales pero el problema requiere más trabajo teórico.

4. Teoría de la Dimensión del Cálculo

Si se nos da un problema bien especificado (aquél para el cual es posible probar mecánicamente si la respuesta que se propone es o no una respuesta válida) una forma de solucionarlo es probar todas las soluciones posibles una a una. Este método no es eficaz, y para

excluirlo se debe tener un criterio de eficacia del cálculo. Cierta consideración del problema mostrará que para obtener una medida de la eficacia de un cálculo es necesario tener disponible un método de medida de la complejidad de los dispositivos de cálculo, lo cual a su vez puede realizarse si se tiene una teoría de la complejidad de funciones. Shannon, y también McCarthy, han obtenido algunos resultados parciales sobre este problema.

5. Autoperfeccionamiento

Probablemente una verdadera máquina inteligente llevará a cabo actividades que podríamos describirlas como autoperfeccionamiento. Se han propuesto algunas líneas de trabajo para esto que merecen más estudio. Parece probable que se pueda estudiar también esta cuestión de forma abstracta.

6. Abstracciones

Se pueden definir claramente y menos claramente ciertos tipos de “abstracción”. Parece que valdría la pena intentar una clasificación y descripción de métodos realizables por máquinas de realizar abstracciones a partir de datos sensoriales y de otro tipo.

7. Aleatoriedad y Creatividad

Una conjetura bastante atractiva y sin embargo claramente incompleta es que la diferencia entre el pensamiento creativo y el pensamiento competente no imaginativo radica en la presencia de cierta aleatoriedad. La intuición debe guiar a la aleatoriedad para que ésta sea eficaz. En otras palabras, la intuición educada o las corazonadas incluyen la aleatoriedad controlada en lo que de otra forma sería un pensar ordenado.

Además de los problemas anteriormente expuestos, hemos pedido a los participantes que describan sus trabajos futuros.⁶⁴

La intención en Dartmouth era eminentemente práctica, se trataba de desarrollar la construcción de máquinas inteligentes, principalmente máquinas de Turing. La forma dominante de abordar el problema era mediante la especificación del programa o instrucciones a dar a la máquina, aunque se seguía mencionando la simulación de redes neuronales como línea de investigación, como herencia de las investigaciones “cibernéticas”. Tienen especial interés los puntos 2, “Cómo se puede programar una computadora para usar un lenguaje” y 6, “Abstracciones”. En estos puntos se empieza a introducir la idea del pensamiento humano como formación de estructuras de símbolos y su manejo de acuerdo a reglas, y a las computadoras como dispositivos capaces de realizar estas mismas funciones.

En Dartmouth se trazaron las principales líneas de lo que sería la investigación en inteligencia artificial durante los siguientes veinte años. Oliver Selfridge introdujo la idea de su “Pandemonium”, una máquina con distintos “agentes” especializados en una tarea, que colaboran en el funcionamiento global. Solomonoff indicó la conveniencia de centrar la investigación en lo que aparentemente eran las tareas más “sencillas” del pensamiento, como el reconocimiento de imágenes, aunque luego su dificultad se revelaría como un serio obstáculo para el desarrollo de la inteligencia artificial. Trenchard More presentó sus trabajos de prueba de teoremas por deducción natural y Arthur Samuel sus investigaciones respecto a cómo enseñar a un ordenador a jugar “checkers”⁶⁵.

⁶⁴ John McCarthy, Marvin L. Minsky, Nathaniel Rochester, Claude.E. Shannon, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. August 31, 1955, <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>, 1996 (8 junio 2004).

⁶⁵ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, Basic Books, 40-41.

Sin embargo, el trabajo más impactante lo presentó el científico social matemático Herbert Simon. Simon estudió en Chicago ciencias políticas y allí conoció el trabajo en lógica de Carnap y en biología matemática de Raschevsky. Estaba interesado en establecer con rigor matemático una ciencia social, y más en concreto, una ciencia del comportamiento humano de toma de decisiones en las organizaciones. A partir de ahí, sus investigaciones le llevaron a ver la inteligencia como un proceso de búsqueda entre un número limitado de opciones. Simon presentó en Dartmouth el “Logic Theorist”, un trabajo en colaboración con Allen Newell y Clifford Shaw consistente en un sistema de deducción de teoremas lógicos de los *Principia* de Whitehead y Russell, basado en la búsqueda selectiva de los pasos a dar en la deducción aplicando criterios heurísticos. La idea era que si todos los teoremas se siguen de los axiomas y las reglas de deducción, entonces aplicando todas las combinaciones posibles se podrían deducir todos los teoremas. Sin embargo, el número de combinaciones era demasiado elevado como para poder realizarlas en un sistema real. El “Logic Theorist” elegía las combinaciones según criterios heurísticos de semejanza entre el teorema a demostrar y los axiomas del sistema, y comenzaba un procedimiento de prueba y error y vuelta a empezar durante un número limitado de intentos hasta llegar o no al objetivo.

El método es heurístico, ya que emplea indicaciones basadas en las características del sistema a probar para limitar el rango de búsqueda. No enumera sistemáticamente todas las pruebas. Este uso de indicaciones representa un gran ahorro en la búsqueda, pero tiene el defecto de que puede que de hecho no se encuentre la prueba. La prueba de un tanteo es empírica: ¿funciona?⁶⁶

Se hacía funcionar a este sistema en una computadora mediante un programa en el lenguaje adecuado. Frente al resto de

⁶⁶ Allen Newell and Herbert A. Simon, *The Logic Theory Machine, a Complex Information Processing System*, 15 Junio 1956, The Rand Corporation, P-868, 36. en web: <http://shelf1.library.cmu.edu/IMLS/MindModels/logictheorymachine.pdf>

trabajos presentados en Dartmouth, que exponían criterios sobre líneas de investigación a seguir, el trabajo de Simon era un mecanismo que funcionaba (consiguió demostrar 38 de los 52 teoremas del capítulo 2 de los *Principia*). Hacía realidad la hipótesis de que cualquier aspecto de la inteligencia, en este caso la inteligencia entendida como búsqueda entre distintas opciones, podía implementarse en una máquina.

A partir de Dartmouth, Simon y Newell continuaron simulando en ordenadores la forma de pensar de las personas, mientras que la mayoría de los investigadores centraron sus esfuerzos en realizar máquinas inteligentes, pero no necesariamente siguiendo el modelo de razonamiento humano. Simon y Newell se dieron cuenta de que las personas no razonaban como en el “Logic Theorist”. En 1957 idearon el “General Problem Solver”. Este programa se basaba en la idea de retroalimentación de Wiener pero aplicada de forma más abstracta y general. Se trataba de detectar las diferencias entre lo deseado y lo obtenido y aplicar correcciones al resultado, pero para un espectro amplio de tipos de variaciones, como al que reaccionan los humanos, al contrario que las máquinas, que lo hacen en un dominio mucho más restringido. Si además las máquinas almacenan lo que han probado, funcione esto o no, van adquiriendo más conocimiento del que tenían cuando se los programara.

Aunque Minsky no triunfó en su intento de creación de modelos más cercanos a la estructura cerebral, otros continuaron intentándolo. Frank Rosenblatt presentó su intento, un sistema llamado “Perceptron”⁶⁷, en 1958. El “Perceptron” era una capa de un solo nivel de neuronas de McCulloch-Pitts situada entre sensores y unidades de activación. Los sensores recibían los impulsos del entorno exterior y enviaban una señal proporcional a su excitación a las neuronas. Éstas recogían esos impulsos, les asignaban cierto “peso” o relevancia, y los sumaban, para pasado un umbral, excitar las unidades

⁶⁷ Frank Rosenblatt, *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*, en *Psychological Review*, 65, 1958, 368-408.

de activación. Por ejemplo, en un animal, los sentidos podían ser los sensores y los músculos las unidades de activación. La aportación principal de Rosenblatt fue un método para ajustar los valores en los pesos para hacer que el “Perceptron” aprendiera a reconocer ciertos estímulos. Primero realizó las conexiones entre sensores y neuronas al azar, en un intento por simular el cerebro de un bebé. Tomando como unidades de activación la iluminación de unas lámparas, esto implicaba que al principio se iluminaban al azar, pero pasado un periodo de “aprendizaje” se iluminarían las lámparas adecuadas. En este periodo de aprendizaje Rosenblatt ajustaba los pesos de las neuronas. Cuando una luz no se iluminaba y debería haberlo hecho, incrementaba los pesos de su correspondiente neurona, y cuando se iluminaba la luz correcta, no modificaba los pesos. En 1962 presentó su regla de aprendizaje, y afirmó la superioridad del enfoque neural frente al simbólico en el estudio de la inteligencia. Aun así, no acompañó su presentación de argumentos científicos bien estructurados⁶⁸. Otros investigadores siguieron su ejemplo y la investigación en redes neurales artificiales tuvo cierto auge durante los años sesenta. En 1969 Minsky y Papert⁶⁹, demostraron que había patrones que el Perceptrón no podía reconocer, como distinguir entre “1” y “0”, aunque Rosenblatt había afirmado que era posible que aprendiese a reconocer cualquier patrón. Estas operaciones sí eran realizables por las computadoras (operación XOR entre dos bits). Esto, junto con la autoridad de los autores del libro y la muerte de Rosenblatt al poco tiempo, hizo que la investigación en redes neurales se quedase estancada durante quince años.

⁶⁸ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 105.

⁶⁹ Marvin Minsky y Seymour Papert, *Perceptrons. An Introduction to Computational Geometry*. Cambridge, Mass., The MIT Press, 1969, 1988. Citado en Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 105.

En 1969, Simon⁷⁰ había afirmado que tanto la computadora como la mente humana deberían considerarse “sistemas simbólicos”. Este mismo paradigma estaba presente en otras disciplinas que se ocupaban de la cognición humana. Como en lingüística en el trabajo de Fodor y Katz (*The Structure of language*, 1964), en psicología el trabajo de Ulrich Neisser (*Cognitive Psychology*, 1967), y en antropología el de Stephen Tyler (*Cognitive Anthropology*, 1969), entre otros. Con la confluencia de estas aproximaciones al estudio del conocimiento, comenzó a crearse un sentimiento de que un campo nuevo estaba emergiendo. La Fundación Sloan aportó 20 millones de dólares para la fundación e investigación en un nuevo campo interdisciplinar sobre estudios cognitivos. En 1977 se fundó la revista *Cognitive Science* y en 1979 la Sociedad con el mismo nombre. Desde el mismo comienzo resultó difícil crear un campo de estudio homogéneo, ya que junto con la denominación de la nueva disciplina y su intento de definición surgieron las tensiones acerca de lo que era este campo de estudio, quién lo comprendía correctamente y qué rumbo debía seguir, persistiendo esas tensiones hoy en día⁷¹. El paradigma dominante en esta etapa que comenzó en Dartmouth en 1956, consistente en considerar la cognición como manipulación de símbolos conforme a unas reglas, se denominó posteriormente “cognitivismo”.

Durante los años sesenta y setenta se fue avanzando en la programación de ordenadores para la realización de distintas tareas. Seymour Papert creó “LOGO” en 1967, un programa de aprendizaje para niños, y junto con Minsky creó “Blocks Micro World”, un programa para hacer que los ordenadores aprendan agrupando y viendo bloques sencillos. Terry Winograd creó “SHRDLU” en 1968-69, un programa que extendía el aprendizaje de los ordenadores a las letras. Joseph Weizenbaum, a mediados de los sesenta, creó “ELIZA”, un programa que comprendía y hablaba inglés para situaciones

⁷⁰ Herbert A. Simon, *The Sciences of the Artificial*, The MIT Press, Cambridge, Massachussets, 1996, 1981, 1969, 21.

⁷¹ Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 53.

sencillas. En 1965-75 surgieron los primeros sistemas expertos, que restringían el ámbito de aplicación de los ordenadores a una tarea formalmente bien definida. Entre 1965 y 1975 Edward Feigenbaum desarrolló “DENDRAL” un sistema experto de formulación química. En 1972 Edward Shortliffe creó “MYCIN”, un programa de diagnóstico de enfermedades de la sangre infecciosas. A estos primeros sistemas expertos han seguido muchos otros con aplicaciones en distintos campos industriales.

Pero las predicciones optimistas, que esperaban mucho más de la Inteligencia Artificial que la creación de sistemas expertos, sufrieron un revés a principios de los setenta. Las primeras carencias de la Inteligencia Artificial se vieron en el campo de la traducción automática y en la realización de tareas que no estuviesen restringidas a campos muy bien delimitados. Uno de los principales financiadores de proyectos en Inteligencia Artificial, la Agencia de Proyectos de Investigación Avanzados del Departamento de Defensa (DARPA), comenzó a restringir las ayudas. El problema subyacente en casi todos los casos, y que persiste hoy en día, era que formalizar el conocimiento cotidiano de cualquier situación, por sencilla que parezca, resultaba prácticamente imposible, dado que las posibles alternativas a manejar en la vida real parecen ser infinitas⁷². Aún así, y como respuesta al proyecto de creación de ordenadores de quinta generación en Japón, en 1984 Douglas Lenat abordó un ambicioso proyecto de Inteligencia Artificial en Estados Unidos, con los siguientes resultados:

Me gustaría presentar un conjunto de métodos de razonamiento sorprendentemente compacto, potente y elegante que formen un conjunto de primeros principios que explican la creatividad, el humor y el razonamiento del sentido común, una especie de “Ecuaciones de Maxwell” del pensamiento. Me gustaría mucho presentar esos métodos de razonamiento, pero, por desgracia, no creo que existan. Así que, en lugar de ello, este trabajo les informará sobre Cyc, el

⁷² Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 118.

proyecto de la enorme base de datos de conocimiento en la que hemos estado trabajando en MCC durante los últimos cuatro años.⁷³

Puede decirse como resumen, que después de la etapa cibernética, una nueva corriente con un enfoque distinto respecto a la cognición fue la predominante durante los años sesenta y setenta. Esta corriente tenía que ver fundamentalmente con el desarrollo de la potencia de las computadoras y la identificación de la cognición humana con su funcionamiento, considerando la manipulación de la información o mensajes no como un hecho físico sin contenido semántico constitutivo de la estructura de un sistema autorregulado, como en la fase cibernética, sino como manipulación de “símbolos” a los cuales se les atribuye contenido o significado semántico. Los símbolos y sus estructuras serían la “representación” del sistema del mundo exterior. A esta hipótesis se la denomina “hipótesis del sistema de símbolos físicos”, y a la corriente en Ciencia Cognitiva que se basa en esta hipótesis, “cognitivismo”. Fue en esta época cuando se creó la Ciencia Cognitiva como disciplina con nombre propio, a través de la creación de una revista y una sociedad. Aunque el cognitivismo es aún importante en la actualidad, dada la ausencia de alternativas satisfactorias en algunos aspectos de la cognición, los pobres resultados en la simulación y posterior explicación de la cognición humana hicieron que a finales de los setenta ya empezara a pensarse que había que volver la mirada hacia el funcionamiento del cerebro, y a la posibilidad de realizar no sólo su simulación, sino su efectiva replicación, en algún tipo de máquina.

⁷³ Douglas B. Lenat y R. B. Guha, “The World According to CYC”, MCC Technical Report no. ACA-AI-300-88, septiembre 1988, págs. 2-3. Citado en Daniel Crevier, *AI: the tumultuous history of the search for artificial intelligence*, Basic Books, 1993, 240. Para más información sobre CYC puede verse la tesis de Mark Telford Georges: *The Problem of Storing Common Sense in Artificial Intelligence. Context in CYC*. Universidad de Navarra. 2000, en web: <http://www.unav.es/gep/TesisDoctorales/MarkGTesis.pdf>

2.4 Conexionismo. Tendencias actuales

En los encuentros de 1956 en Dartmouth se prestó mucha atención a la creación de modelos artificiales de pensamiento basados en la manipulación de símbolos por máquinas de Turing, en las que se codificaba la tarea a realizar para llevarla a cabo mediante las necesarias iteraciones en pasos sucesivos, en detrimento de los modelos que simulaban la actividad neural. Sin embargo, a finales de los setenta ya se alzaban voces en contra del modelo computacional. La insatisfacción con los resultados generales obtenidos hizo que se volviera la mirada hacia los modelos de inspiración neuronal. En el enfoque neural se ponía de relieve la aparente no-adequación del procesamiento computacional a algunos procesos cerebrales. Como ejemplo puede citarse la lentitud de las conexiones neurales en comparación con las operaciones lógicas realizadas en los ordenadores, lo que hacía pensar en la necesidad de tener un procesamiento en paralelo de la información en lugar de en pasos sucesivos. O el hecho de que el cerebro no parezca tener una unidad controladora central, sino que las funciones cerebrales aparezcan distribuidas en distintas zonas, y que la eliminación de zonas importantes no haga que el cerebro deje de funcionar por completo, como podría ocurrir en un ordenador. A la vez que se presentaban las dificultades del modelo computacional en serie, se creaban modelos de inspiración neural que superaban las limitaciones del Perceptron de Rosenblatt apuntadas por Minsky y Papert. Todo esto hizo que este tipo de modelos, llamados modelos conexionistas, resurgieran en los años setenta.

Durante el final de los años setenta se realizaron varios trabajos sobre modelado conexionista de distintos aspectos de la cognición, tanto de investigadores que no habían abandonado el enfoque neural (Anderson, Kohonen, Hinton, Willshaw), como de otros atraídos por este nuevo campo (Sejnowski, Feldman)⁷⁴. En

⁷⁴ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 13.

1981, David Rumelhart, un matemático que estaba trabajando en modelos computacionales sobre el proceso de lectura, influido por el trabajo de Reddy en Carnegie Mellon sobre el sistema de reconocimiento del habla llamado HERSAY, en el que la idea de computación cooperativa era central⁷⁵, conoció en la Universidad del Sur de California en San Diego a James McClelland, un psicólogo que trabajaba en experimentos psicológicos sobre tiempos de reacción ante estímulos. Juntos acordaron crear un modelo de los procesos cognitivos con mayor plausibilidad neurológica que los existentes hasta la fecha. A esta iniciativa se unió Geoffrey Hinton, que estaba trabajando en temas de memoria asociativa. Los tres crearon en la citada universidad el *PDP Research Group* (grupo de investigación en procesamiento distribuido en paralelo). Los resultados de sus investigaciones se publicaron en 1986, en la que se puede considerar la obra más representativa del nuevo enfoque conexionista, bajo el título de *PDP: Explorations in the Microstructure of Cognition*.⁷⁶

Siguiendo este trabajo, un sistema cognitivo conexionista consta de los siguientes elementos: un conjunto de unidades de procesamiento, un estado o valor de activación de dichas unidades, una función de salida para cada unidad, un patrón de conexión entre las unidades, una regla de propagación para propagar patrones de actividad a través de la red de conexiones, una regla de activación para combinar las entradas que afectan a una unidad con el estado actual de esa unidad, de forma que produzca un nuevo nivel de activación para la unidad, una regla de aprendizaje por la cual los patrones de conexión son modificados por la experiencia y un ambiente dentro del cual debe operar el sistema⁷⁷.

⁷⁵ David E. Rumelhart, "From Searching to Seeing", en Peter Baumgartner and Sabine Payr. *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, 191.

⁷⁶ David E. Rumelhart, James L. McClelland and the PDP Group: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, The MIT Press, Cambridge, Massachussets, 1986.

⁷⁷ David E. Rumelhart, James L. McClelland and the PDP Group: *Introducción al procesamiento distribuido en paralelo*, Alianza Editorial, Madrid, 1992, 81-82.

Las unidades de procesamiento serían las equivalentes a las neuronas cerebrales. El trabajo de cada unidad es recibir la entrada de sus vecinas y, en función de las entradas recibidas, computar un valor de salida, que envía a las unidades vecinas. El sistema es paralelo en el sentido de que muchas unidades pueden desarrollar su actividad al mismo tiempo. Puede haber tres tipos de unidades: de entrada, de salida y ocultas. Un sistema siempre tiene unidades de entrada y de salida, pero puede no tener unidades ocultas. Las unidades de entrada reciben las entradas de fuentes externas al sistema, las unidades de salida envían señales hacia fuera del sistema, y las unidades ocultas son aquellas cuyas entradas y salidas están dentro del sistema y no son accesibles desde el exterior. El estado o valor de activación es el conjunto de números reales que representa la activación de todas las unidades de procesamiento, donde cada número representa la activación de una de las unidades. Puede verse el estado de activación total como lo que el sistema representa en cierto momento, y el procesamiento en el sistema como la evolución, a través del tiempo, del estado de activación global. La función de salida en cada momento de cada unidad determina las operaciones a realizar en su valor de activación actual para producir un valor de salida. El patrón de conexión entre las unidades es la forma en que éstas están conectadas entre sí, pudiendo cada unidad recibir conexiones y a la vez conectar su salida a muchas otras unidades. El patrón de conexión especifica además los valores o pesos de las conexiones. El peso es un número positivo si es excitatorio y negativo si es inhibitorio, y su valor absoluto indica la “fuerza de la conexión”. Los tipos de conexiones se consideran iguales si se pueden sumar algebraicamente para obtener su resultado en la entrada de la unidad en cuestión. El patrón de conexión determina en gran medida lo que el sistema conoce, cómo responderá a cualquier entrada arbitraria, y lo que cada unidad representa dentro del sistema. La regla de propagación calcula el valor de las entradas netas a las unidades mediante la combinación adecuada

Original en inglés: *Parallel Distributed Processing*, Massachusetts Institute of Technology, 1986.

del patrón de conexión del sistema y los valores de salida de las unidades que están directamente conectadas a éstas y las excitan o inhiben. La regla de activación toma las entradas netas de cada tipo que afectan a una unidad particular y las combina con los otros tipos de entradas y con el estado actual de la unidad para producir un nuevo estado de activación. La regla de aprendizaje es una función que modifica los pesos del patrón de conexión para obtener las salidas deseadas teniendo en cuenta, en cada momento, las salidas obtenidas frente a las entradas del sistema. Por último, el conjunto de entradas del sistema a lo largo del tiempo constituye el ambiente en el que opera el sistema.

En definitiva, se trata de un sistema de unidades interconectadas que recibe estímulos del exterior, los transmite y modifica a través de sus unidades hasta enviar otros estímulos al exterior, y al cual se guía en su comportamiento, mediante la variación de los valores de sus conexiones, hasta obtener el valor de salida deseado. Según la complejidad de cada uno de estos componentes, se tendrá un tipo u otro de sistema conexionista. Por ejemplo, siguiendo cada uno de los componentes, el Perceptron de Rosenblatt es un sistema sin unidades ocultas, con solo unidades de salida y de entrada, con tan solo dos valores de activación, 1 y 0, con una función de salida de cada unidad que es igual a su valor de activación, donde las unidades de entrada influyen en las de salida pero no al revés y ni las unidades de entrada ni las de salida están interconectadas entre sí, la regla de propagación es la suma ponderada por los pesos de las salidas de las unidades y es igual a la regla de activación de la unidad que recibe las conexiones, la regla de aprendizaje es un tipo de regla de Hebb, y el ambiente presenta sus estímulos simultáneamente a todas las unidades de entrada. Esto podría ser mucho más complicado, con varias unidades ocultas, múltiples interconexiones e influencias entre los distintos tipos de unidades, y con funciones de salida, reglas de propagación, activación y aprendizaje muy complejas.

El funcionamiento de estos sistemas es como sigue: mediante la presentación sucesiva de estímulos en las unidades de entradas se

obtiene en cada momento una salida, que se evalúa junto con la regla de aprendizaje para modificar el patrón de conexión. Posteriormente se vuelve a presentar la misma entrada, a evaluar la salida y a volver a modificar el patrón de conexión. Este proceso se repite cuantas veces sea necesario hasta obtener la salida deseada. En cada uno de estos pasos se espera a que el sistema se estabilice (produzca una salida invariable) y entonces se toma la salida. Estos sistemas han mostrado tener propiedades interesantes y sorprendentes, como la de seguir comportándose adecuadamente ante daños en su estructura, han conseguido calcular funciones lógicas y reconocer patrones espaciales, y se han creado sistemas para leer inglés (NETtalk),⁷⁸ y sistemas para realizar ciertas tareas, como ir a un restaurante, o comprar algo, en ámbitos bien definidos y restringidos (DISCERN).⁷⁹

Pero también hay problemas sin resolver por los sistemas conexionistas, unos específicos y otros compartidos con los sistemas computacionales en serie. Uno de estos problemas es el elevado número de veces que hay que presentar la misma entrada hasta que el sistema la “aprenda”, lo que cuestiona su validez como modelo del cerebro, ya que esto no encaja en muchos casos con la rapidez con la que se da el aprendizaje humano. Otro problema es la elección del patrón de conexión si se quiere simular el funcionamiento cerebral, ya que parece ser que cada neurona puede estar conectada a miles de neuronas. Esto, junto a la elección del número de unidades, que en muchos casos puede ser demasiado alto, hace que la realización de un sistema que simule el conocimiento humano más allá de un campo muy restringido sea difícilmente abordable, tanto matemáticamente como en su realización práctica. En este sentido se preguntaba Crevier en 1992 por el futuro del conexionismo:

Uno pensaría, sin embargo, que las dificultades pasadas de los sistemas expertos moderarían las pretensiones de otras tecnologías

⁷⁸ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 77-78.

⁷⁹ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 200-234.

respecto a realizar máquinas inteligentes. Sin embargo, mientras escribo esto, la puesta en marcha de una red neural parece una réplica exacta, una década después, del drama de los sistemas expertos. La prensa comenta la nueva tecnología, y las fuerzas armadas anuncian su financiación en 400 millones de dólares.⁸⁰

En los sistemas conexionistas, la computación puede verse como una interacción física del exterior con el sistema una vez que éste ha llegado a un estado de relajación, o lo que es lo mismo, una vez que ha “aprendido” a realizar una tarea. Esto difiere de la idea cognitivista del pensamiento como computación consistente en la manipulación de estructuras de símbolos, como representaciones del mundo, conforme a reglas. Los símbolos desaparecen como unidades sobre las que actuar ya que en la mayoría de los casos cada unidad no representa un símbolo y las representaciones y las reglas se pueden reinterpretar como el patrón de actividad general de la red. Patrón que “emerge” de su estructura, si bien es verdad que el patrón de conexión y la regla de aprendizaje vienen determinadas por el diseñador. Si nos olvidamos de esto último, que no es poco, diríamos que en un sistema conexionista el conocimiento no está enunciado explícitamente por símbolos sino que se encuentra en las conexiones de la red. Esto no anularía el enfoque basado en sistemas de símbolos, sino que lo reinterpretaría. Según Rumelhart y McClelland:

Hemos observado que el carácter cooperativo del pensamiento distribuido en paralelo a menudo nos ha permitido explicar conductas que previamente se habían atribuido a la aplicación de reglas del pensamiento. Esto nos ha conducido, a veces, a afirmar que una conducta que se atiene a leyes no necesariamente esta *regida por reglas*. En este caso hemos de distinguir entre *reglas* y *regularidades*. La bola que se desplaza y el planeta que orbita muestran regularidades en su comportamiento, pero ninguno de ellos está aplicando una regla. Hemos demostrado la potencia de este enfoque en nuestro trabajo previo sobre la percepción de palabras (McClelland y Rumelhart, 1981, y Rumelhart y McClelland, 1982) y sobre el aprendizaje de la

⁸⁰ Daniel Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, 216.

morfología del inglés (Rumelhart y McClelland, 1986). En estos casos hemos podido mostrar cómo la aparente aplicación de reglas podía *emerger* de las interacciones entre simples unidades de procesamiento, en lugar de hacerlo a partir de la aplicación de reglas de nivel superior⁸¹.

La perspectiva básica de este libro es que muchos de los constructos incluidos en las descripciones de macronivel, como los esquemas, los prototipos, las reglas, las producciones, etc., pueden considerarse emergentes a partir de las interacciones de la microestructura de los modelos distribuidos... Consideramos las macroteorías como aproximaciones a la microestructura subyacente que pretende captar el modelo distribuido presentado en nuestro texto. Como tales aproximaciones, a menudo resultan útiles, pero en algunas situaciones ocurre que el examen de la microestructura da lugar a una comprensión más profunda... No obstante, también podría ocurrir que algunos fenómenos fueran demasiado complejos para ser fácilmente representados como modelos PDP. Si estos fenómenos tienen lugar en un marco temporal en el que un modelo macroestructural fuera una aproximación adecuada, no hay razón alguna por la que no debiera aplicarse el modelo macroestructural. En consecuencia, pensamos que los conceptos de símbolo y procesamiento de símbolos pueden resultar muy útiles. En ocasiones, tales modelos pueden ofrecer la explicación más sencilla⁸².

Se propugna así una cooperación entre la interpretación simbólica del pensamiento y la conexionista, eligiendo la que sea más adecuada en cada momento. Según Rumelhart y McClelland hay cierta autonomía entre ambos niveles, y aunque el nivel simbólico “emerja” de las conexiones, hay propiedades que sólo se pueden entender y explicar en ese nivel, ya que el todo es diferente a la suma de las partes, al haber interacciones no lineales⁸³.

⁸¹ David E. Rumelhart, James L. McClelland and the PDP Group, *Introducción al procesamiento distribuido en paralelo*, 153.

⁸² David E. Rumelhart, James L. McClelland and the PDP Group, *Introducción al procesamiento distribuido en paralelo*, 157-158.

⁸³ David E. Rumelhart, James L. McClelland and the PDP Group, *Introducción al procesamiento distribuido en paralelo*, 159-160.

Es importante la distinción entre interacciones lineales y no lineales. Las interacciones lineales implican que la salida de un sistema sea la suma ponderada de sus entradas. Estos sistemas asocian un patrón de entrada con un patrón de salida, pero no se da “casi nada nuevo” en ellos y, si se muestra su comportamiento en el tiempo, se ve que no sufre cambios bruscos. Las interacciones no lineales tienen lugar cuando las salidas son una función más complicada de las entradas, implicando en algún momento el producto de éstas. Las funciones que representan el comportamiento de estos sistemas tienen una gran dificultad matemática en la gran mayoría de los casos. Son estos sistemas los que tienen propiedades marcadamente “emergentes”, ya que puede cambiar su comportamiento radicalmente dependiendo del valor de las entradas o de un ligero ajuste de los parámetros de conexión de sus unidades. El conexionismo considera que éste es el tipo de interacción neuronal y la forma en que se generan los pensamientos.

Muy próxima a este enfoque se encuentra la tendencia actual llamada “aproximación dinámica a la cognición”⁸⁴. En ésta se analiza el comportamiento de los sistemas a lo largo del tiempo, atendiendo a una o varias variables que se quieran estudiar. Estas variables se representan gráficamente frente al tiempo y se estudian sus propiedades. En algunas ocasiones los sistemas evolucionan hacia puntos estables llamados *atractores*, en otras, presentan trayectorias inestables que a veces dependen fuertemente de los parámetros iniciales, lo que se llaman trayectorias *caóticas*.⁸⁵ Muchos de estos sistemas son bien conocidos en física. La aproximación dinámica a la cognición interpreta el pensamiento como el paso en el tiempo de unos atractores a otros.

⁸⁴ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 235.

⁸⁵ William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 240.

Si en este estudio se toma una de las variables a estudiar del exterior, para ver como interactúa con el sistema, se tendrá un sistema acoplado entre el cerebro y el exterior. Este es el enfoque actual conocido como “cognición situada o corporeizada” (*situated or embodied cognition*). Algunos autores ven una posible cooperación entre el simbolismo y el enfoque dinámico:

Quizás, entonces, lo que se necesite sea un tipo de computacionalismo mecánico en el que los detalles del flujo de información sean bit a bit tan importantes como la dinámica a gran escala, y en el que algunos rasgos dinámicos locales tengan una función también de procesamiento de la información. Aquí, entonces, se encuentra una forma en la que los análisis dinámico y computacional pueden ir de la mano. El análisis dinámico puede ayudar a identificar los procesos físicos que actúan como los vehículos del contenido representacional.⁸⁶

Para otros autores, la noción de “representación” no tiene validez, ya que el sistema no “representa” nada, sino que son sus propiedades físicas y su interacción mecánica con el medio, lo que determina su comportamiento:

No hay nada intrínsecamente representacional en este proceso dinámico hasta que interviene el observador. Es el observador el que infiere lo que los patrones de actividad observados representan, para así explicarse a él mismo los resultados.⁸⁷

Un paso más que se puede dar en el modelado del cerebro en la búsqueda de una explicación del pensamiento es la caracterización de éste como un sistema autónomo, sin reglas de aprendizaje “exteriores

⁸⁶ Andy Clark, *Mindware. An Introduction to the Philosophy of Cognitive Science*. Oxford University Press, Oxford, 2001, 135.

⁸⁷ W. J. Freeman y C.A. Skarda, "Representations: Who needs them?", 1990, en J.L. McGaugh, N.M. Weinberger, y G. Lynch (eds), *Brain Organization and Memory: Cells, Systems, and Circuits*, Oxford University Press, Oxford, 375-380. Citado en William Bechtel y Adele Abrahamsen, *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, 271.

a él” o fijadas por otro agente, sin ajustes exteriores del patrón de conexión. Se tendría en este caso un sistema adaptativo-evolutivo, en el que su misma estructura y su adaptación al medio y evolución explicasen su comportamiento.

Esto es lo que trata de hacer el campo de investigación llamado *Vida Artificial*, entender los sistemas y procesos biológicos y los mecanismos evolutivos para después simularlos en robots electrónicos o en ordenadores. Generalmente esto implica la realización de algún tipo de los llamados *algoritmos genéticos*, que son simulaciones por medio de cadenas de bits, de una versión simplificada de los mecanismos de la selección natural en los genes. Así, un cromosoma se corresponde con una cadena de bits (genes), sobre los que se opera según criterios de selección natural, como la mutación, recombinación y la adaptación al medio. Mitchell indica el apoyo que tiene, a su juicio, este enfoque a la hora de explicar los fenómenos cognitivos:

Hoy en día muchos investigadores en Inteligencia Artificial creen que las “reglas” que subyacen a la inteligencia son demasiado complejas como para poder codificarlas mediante una aproximación de “arriba abajo”. Al contrario creen que la mejor forma es mediante un paradigma de “abajo arriba”, en el que los humanos solo escriben reglas muy simples, y los comportamientos complejos como la inteligencia emergen de la aplicación masiva en paralelo de estas reglas sencillas. El conexionismo es un ejemplo de esta filosofía; la computación evolutiva es otra. En el conexionismo las reglas son típicamente simples umbrales “neurales”, extensión de activaciones y fortalecimiento o debilitación de conexiones. El comportamiento emergente se asocia al reconocimiento de patrones y al aprendizaje sofisticado. En la computación evolutiva las reglas se corresponden típicamente con la “selección natural”, con las variaciones motivadas por recombinación y/o mutación. El comportamiento emergente se asocia al diseño de soluciones de alta calidad a problemas difíciles y a la adaptación de estas soluciones a un entorno cambiante.⁸⁸

⁸⁸ Melanie Mitchell, *An Introduction to Genetic Algorithms*, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1998, 1996, 4.

Al igual que ocurre con el cognitivismo y el conexionismo, lo cierto es que a día de hoy los algoritmos genéticos encuentran su aplicabilidad en la realización de sistemas expertos, en lugar de en la posible explicación de fenómenos cognitivos. Lo que sí hace este nuevo campo de la Vida Artificial es abrir un posible nexo entre biología, computación y cognición, y por supuesto, levantar toda una serie de cuestiones polémicas. Andy Clark señala, desde una posición bastante materialista como ocurre en gran parte de la Ciencia Cognitiva, algunas de éstas:

El trabajo en Vida Artificial también genera preguntas fundamentales que conciernen a la idea misma de vida y a la relación entre la mente y la vida. Sobre la idea misma de vida, el desafío es directo y simple: ¿podría instanciarse la vida (más allá de simplemente modelarla) en medios artificiales como robots o ecosistemas basados en ordenador?...

Un debate, por tanto, concierne a la efectiva definición de la vida, y quizás de distintas propiedades como la autorreplicación. En esta línea, Bedau (1996, p.338) indica la urgencia por definir la vida como “adaptación flexible” – la capacidad de responder apropiadamente, de un modo indefinido de formas, a una variedad (desde la perspectiva del organismo) impredecible de contingencias. Tal definición [al contrario que, por ejemplo, la centrada en la metabolización de materia en energía – véase Schrödinger (1969) y Boden (1999)] permite claramente que hechos y procesos que ocurren en dispositivos electrónicos y otros medios cuenten como instancias de vida propiamente dicha. Otros autores se centran sin embargo en otras propiedades y características, como la autopoiesis (los sistemas autopoéticos crean y mantienen sus propias fronteras, dentro de las cuales, interacciones circulares complejas mantienen la producción continua de los elementos químicos y materiales esenciales – véase Varela, Maturana, y Uribe, 1974), autocatálisis (conjuntos de elementos – químicos o computacionales – que catalizan sus propios productos a partir de recursos a su alcance – véase Kauffman, 1995), autorreproducción, genética y metabolización (Crick, 1981), etc. Una posibilidad muy real –también mencionada por Bedau (1996) – es que la “vida” sea un conjunto de conceptos, implicando múltiples

características típicas, ninguna de las cuales es necesaria individualmente para considerar que el sistema está vivo, y de las cuales múltiples subconjuntos diferentes podrían considerarse suficiente.

También hay un debate sobre las relaciones entre la vida y la mente. Una forma de resistir a la preocupación de que estos sistemas sencillos, similares a los vivos, nos digan poco respecto a los fenómenos realmente cognitivos es mantener que la vida y la mente comparten características organizativas profundas y que el proyecto de comprender la mente forma un continuo con el proyecto de entender la vida. Esta posición la expresa bien Godfrey-Smith (1996^a, p. 320) en su descripción de la tesis de “continuidad fuerte”: “la vida y la mente tienen un patrón abstracto común o conjunto de propiedades organizativas básicas. Las propiedades funcionales características de la mente son una versión enriquecida de las propiedades funcionales que son fundamentales para la vida en general. La mente es literalmente *como* la vida”.⁸⁹

En esta sección se han presentado el conexionismo y el enfoque dinámico y corporeizado de la cognición, junto con el campo de la Vida Artificial, como las tendencias más recientes en Ciencia Cognitiva. Esto no quiere decir sin embargo que el cognitivismo haya desaparecido. Como se ha visto en muchos casos se propugna una colaboración de teorías cognitivistas y conexionistas. En la sección 4.2 se verá un ejemplo de teoría que intenta aunar ambos enfoques.

2.5 Resumen

Desde la sección 2.1 hasta esta última se ha mostrado la Ciencia Cognitiva como el intento de modelar matemáticamente el pensamiento, sea mediante funciones recursivas como en el cognitivismo, o mediante ecuaciones lineales y no lineales como en el

⁸⁹ Andy Clark, *Mindware. An Introduction to the Philosophy of Cognitive Science*. 117-118.

conexionismo. En definitiva, se trata de tener modelos que poder manejar para observar y predecir su comportamiento y aplicarlo a la cognición humana. Pero ya se indicó en el capítulo 1 que la Ciencia Cognitiva se presenta a sí misma como un esfuerzo interdisciplinar formado también por otras disciplinas, siendo éstas la psicología, la neurociencia, la lingüística, la antropología y la filosofía. También se mostraba cómo estas disciplinas aplicaban los métodos de los paradigmas de la Ciencia Cognitiva, como el cognitivismo y el conexionismo, y los aplicaban a sus propias disciplinas, a la vez que hacían su aportación particular en la explicación del conocimiento.

En el próximo capítulo se presentará la relación histórica entre Ciencia Cognitiva y psicología, neurociencia, lingüística, antropología y filosofía.

CAPÍTULO III

Las otras disciplinas

En este capítulo se presenta un breve desarrollo histórico de la psicología, la neurociencia, la lingüística, la antropología y la filosofía, centrándose en las escuelas que adoptaron los distintos paradigmas de la Ciencia Cognitiva y que de algún modo contribuyeron a su desarrollo, pasando así a formar parte de ésta. Para la exposición de la psicología, lingüística y antropología he seguido básicamente el texto de Howard Gardner: *La nueva ciencia de la mente: historia de la revolución cognitiva* (Paidós, Barcelona, 1996, 1988), indicando al final de cada párrafo su referencia en dicha obra.

3.1 Psicología

El siglo XIX heredó el legado kantiano de la imposibilidad de una ciencia psicológica. Kant creía, por un lado, que la ciencia debía aplicar leyes matemáticas a los datos empíricos, y que dichos datos debían reunirse en experimentos reales, y como la psicología se ocupa de elementos que presuntamente carecen de dimensiones espaciales (los pensamientos puros), dicha experimentación no era posible. Un segundo problema consistía en que la psicología debía estudiar el instrumento del saber, la subjetividad [*self*] cognoscente; pero a la subjetividad no le es posible examinar su propio funcionamiento, y mucho menos hacerlo en forma imparcial. Por añadidura, estaba el problema del nivel de abstracción. Para llevar a cabo investigaciones científicas es preciso dejar de lado los factores accidentales y centrarse en las variables decisivas para una teoría; y es difícil, si no imposible, imponer a la compleja interacción humana esta manipulación radical del objeto de estudio.

Las objeciones mayores de Kant hacia la psicología procedían de su concepción general del ámbito del conocimiento. Según su punto de vista, la filosofía debía establecer la naturaleza del pensar, establecer las relaciones entre las diversas ciencias y formular sus fundamentos y limitaciones; dentro de este programa, la psicología era concebida como una incursión de segunda categoría, que debía contentarse con examinar los contextos sociales e históricos de ese pensar, pero sin tratar de descifrar la naturaleza del pensamiento en sí mismo.⁹⁰

Sin embargo, a mediados y del siglo XIX, el físico y fisiólogo alemán Herman von Helmholtz pretendió demostrar que gran parte de las especulaciones de Kant podían someterse al estudio empírico. Escéptico acerca de las diversas opiniones, discordantes entre sí, sobre la sorprendente velocidad del pensamiento, emprendió y consiguió la medición efectiva del tiempo que lleva transmitir impulsos a lo largo del sistema nervioso, con lo que demostró que las reacciones de la conducta humana eran medibles. Posteriormente pasó a cuestionar las convicciones kantianas de las ideas innatas del espacio, sosteniendo que el conocimiento del espacio se va creando en el individuo del mismo modo que los restantes datos vinculados al mundo material. Hizo utilizar a los sujetos experimentales un prisma que deformaba su visión de los objetos, y logró demostrar que el individuo se adapta pronto a estas distorsiones, y al poco rato vuelve a ver las cosas de un modo que, en lo fundamental, no presenta deformaciones. En una investigación paralela, estudió a ciegos que habían recuperado la visión, documentando el largo tiempo que les llevó a percibir el mundo de los objetos igual que los individuos de visión normal.⁹¹

Helmholtz hizo tres aportaciones importantes. Ante todo, señaló que los axiomas filosóficos de Kant no tenían validez absoluta,

⁹⁰ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 116.

⁹¹ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 118.

ya que era posible esclarecer ciertos aspectos del funcionamiento psíquico humano en forma empírica. En segundo lugar, estableció que había cabida para análisis moleculares (la velocidad de un impulso que recorre la fibra nerviosa), así como los molares (la forma en que ordenamientos espaciales complejos son vistos en condiciones normales y distorsionadas). Por último, al destacar la contribución del sujeto en el proceso de la percepción, fue uno de los creadores de la Ciencia Cognitiva.⁹²

A Helmholtz le interesaba demostrar que las cuestiones psicológicas podían brindar resultados cuantitativos en experiencias de laboratorio. Franz Brentano, a mediados y finales del siglo XIX, abordó las cuestiones de la psicología desde una perspectiva diferente. Rechazó que fuera posible descomponer la psicología en elementos de la experiencia o de la conciencia para examinarlos en forma aislada, así como la idea de que los procesos de pensamiento o de conciencia podían concebirse de un modo puramente mecanicista. Para Brentano, la psicología tenía como punto de partida la mente, entidad activa y creadora dotada de intenciones, pues ella implica y exige un objeto. El auténtico tema de estudio de la psicología es el acto mental intencional: juzgar, sentir, imaginar, oír, y cada uno de estos actos refleja una dirección y un propósito.⁹³

A partir de aquí, la tarea de la psicología empírica consistía en estudiar la mente del actor al ocuparse de objetos, propósitos y finalidades. Brentano subrayó los aspectos fenomenológicos de toda acción psíquica: los pensamientos y el razonamiento son inconcebibles, y mucho más inconcebible es estudiarlos, si no se toma en cuenta la experiencia fenoménica interior del sujeto. Y no se accede a ellos mediante la introspección – ya que al sujeto no le es posible observarse al mismo tiempo que vivencia – sino mediante la simple experiencia fenoménica de la propia vida mental. La obra de

⁹² Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 118.

⁹³ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 119.

Brentano no tuvo mucha repercusión en su tiempo, sin embargo, su mismo planteamiento, de lo general a lo particular, reapareció posteriormente en la escuela de Wurtzburgo, en la psicología de la *Gestalt* y posteriormente en estudios sobre la percepción de Merleau-Ponty.⁹⁴

Puede decirse que al final del siglo XIX la psicología ya estaba en marcha. Sin embargo, el florecimiento de una disciplina depende en gran medida de la fundación de instituciones y organizaciones, y fue Wilhelm Wundt, a finales del siglo XIX y principios del XX, quien llevó a cabo la tarea de establecer la psicología como una disciplina experimental separada. A él se debió que la psicología emergiera como disciplina científica, con métodos, programas e instituciones propias.⁹⁵

Según Wundt, la psicología consiste en el estudio de la experiencia consciente como tal. Dicho estudio debe emprenderse mediante la observación interna, mediante la introspección. Así como era importante separar la psicología de la física, también lo era divorciarla de la fisiología. Esta tenía como ámbito especial el estudio de los procesos mentales no conscientes – no susceptibles de examen introspectivo – a la vez que el de la psicología estaba compuesto por los procesos mentales superiores que sí podían someterse a este examen personal.

Sin embargo la teoría de Wundt fue atacada alrededor de 1900 por la “escuela de Wurtzburgo”. Estos investigadores (Külpe, Marbe, Mecer, etc.) demostraron la ausencia de imágenes en la introspección cuando los sujetos comparaban ciertos pesos, al contrario de lo que

⁹⁴ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 119.

⁹⁵ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 120.

sostenía el introspeccionismo. A pesar de la réplica de los Wundtianos, el introspeccionismo comenzó a tambalearse.⁹⁶

En Estados Unidos, el introspeccionismo tuvo gran impacto, pero no para el que quizás fuese el personaje más importante en la vida académica de mediados y finales del siglo XIX, William James. A James le interesaban las cuestiones psicológicas tal como se nos aparecen en la vida cotidiana, y buscó esclarecerlas no sólo mediante la ciencia experimental sino también mediante la literatura y la historia. Trató todos los fenómenos psicológicos, desde los procesos cerebrales hasta la voluntad, desde los órganos sensoriales hasta la atención, y desde la indagación de los hábitos hasta la conciencia de cada momento. Adoptó un enfoque pragmático, al sostener que si existen mecanismos psíquicos, ello se debe a que resultan útiles a los individuos, ayudándoles a sobrevivir y a llevar a cabo actividades importantes para su vida.⁹⁷

En 1913, también en Estados Unidos, John B. Watson comenzó el movimiento llamado conductismo, al afirmar que el tema de estudio adecuado de la psicología no era el funcionamiento de la mente sino el examen de la conducta objetiva y observable. Basándose en estudios fisiológicos sobre el arco reflejo, propuso que todas las actividades psicológicas podían explicarse comprendiendo los reflejos que se establecen en las partes superiores del sistema nervioso.

Watson rechazaba gran parte del programa de la psicología tradicional y casi todos sus métodos. No tenía en cuenta las sensaciones o las intenciones, sino que solo era pertinente la observación de la *conducta* manifiesta. La descripción y explicación de los estados y contenidos de conciencia debían ser reemplazadas por la predicción y el control de la conducta. La psicología no tenía que ocuparse de lo que ocurre en la mente de una persona, ya que todos

⁹⁶ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 123.

⁹⁷ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 125-126.

los términos mentalistas debían ser eliminados del vocabulario del psicólogo.

La espectacularidad de los resultados del conductismo (por ejemplo, al condicionar Watson en un experimento a un niño para que temiese a una rata), el rigor retórico, la eficacia de sus demostraciones, junto a las posiciones poco concluyentes del introspeccionismo o del pragmatismo anterior, hicieron que entre 1920 y 1950 la psicología en Estados Unidos fuese predominantemente conductista. La puericultura, el tratamiento de los prisioneros de guerra, la educación y muchas otras actividades sociales quedaron bajo el imperio de la retórica y las prácticas conductistas.⁹⁸

La única corriente no conductista empezaba a penetrar en Estados Unidos a través de los inmigrantes por causa de la guerra, fundamentalmente alemanes y austriacos. Esta corriente fue la llamada psicología de la *Gestalt*. En 1890, Cristoph von Ehrenfels, un discípulo de Brentano, estudiando la percepción de las melodías musicales, sostuvo que la “cualidad de la forma” perceptual de una melodía no podía considerarse simplemente la suma de las diversas notas que la componían, sino que era una cualidad global, una *Gestalt* [estructura, forma o configuración], que trascendía sus elementos constitutivos particulares. En 1912, Max Wertheimer, junto a Wolfgang Köhler y Kurt Koffka, llevó a cabo una serie de estudios sobre el movimiento aparente cuando una serie de rayos luminosos o de formas se proyectan uno tras otro, como ocurre en los letreros de neón o en las películas cinematográficas. Probó de modo convincente que la percepción del movimiento no es la suma o asociación de diferentes sensaciones elementales. A partir de estos resultados postularon la estructura cerebral que daría cuenta de estos fenómenos.⁹⁹

⁹⁸ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 127.

⁹⁹ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 129-132.

En Gran Bretaña, Barlett sostuvo que el sistema mnémico de los seres humanos incluye la formación de estructuras cognitivas abstractas o esquemas. Los esquemas surgen a partir de los encuentros previos que el sujeto ha tenido con su medio, como resultado de los cuales llega a organizar ciertas clases de información de una manera específica.¹⁰⁰

En Suiza, Jean Piaget trabajaba con el objetivo de fundar una epistemología sobre principios biológicos. Para ello se dio cuenta de que debía estudiar el desarrollo del pensamiento de los niños. Este estudio le llevaría toda la vida. Aunque muchas de sus ideas han sido posteriormente refutadas, Piaget inauguró en psicología todo un nuevo campo concerniente al desarrollo de la cognición humana, y estableció los temas que aún hoy la mantienen ocupada.¹⁰¹

El conductismo estaba presente en la teoría de sistemas de Wiener, en los que se especifica su comportamiento como orientado a un fin, y en el llamado Test de Turing donde se considera inteligente a un sistema si es capaz de engañar a un entrevistador, considerando exclusivamente la “conducta” del sistema, independientemente de su estructura y composición. El enfoque práctico del conductismo encajaba perfectamente en la euforia de la postguerra, en la que se ensalzó a la física como modelo de las otras ciencias. Según Heims:

Fue un tiempo en el que las ciencias humanas, más que los estudios humanísticos, estaban en ascenso, más solucionando problemas que reflexionando sobre contenidos. Normalmente coexisten y se superponen los modos de comprensión humanístico y científico, y los generalistas los ven como complementarios. Sin embargo, estos dos modos también compiten, y en ciertos momentos y lugares en la historia intelectual occidental se ha favorecido mucho más a uno frente al otro. En Estados Unidos, en el periodo posterior a la segunda

¹⁰⁰ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 132-134.

¹⁰¹ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 135-136.

guerra mundial, las universidades cada vez fueron haciendo más énfasis en las ciencias sociales y de la conducta a costa de las humanidades. La frase “ciencia de la conducta”, inventada en esos días para reemplazar a “psicología”, era indicativa de la tendencia de la psicología a considerarse a sí misma como una disciplina genuinamente científica.¹⁰²

Podríamos caracterizar el marco epistemológico y el estilo cognitivo de hacer ciencia como reduccionista, atomista, positivista, pragmático, conservador, mecanicista y empiricista... Las teorías, un avance fundamental de la física, eran prácticas en el sentido de que proporcionaban reglas para calcular respuestas correctas empíricamente... Otro grupo de físicos mostró su confianza en sí mismos después de la Guerra al invadir el dominio tradicional de la biología, proponiendo desvelar el código genético como se sugería en el conocido libro de Erwin Schrödinger *¿Qué es la Vida?* (1944)... Las cuestiones filosóficas más amplias respecto a la teoría darwiniana de la evolución y variación de las especies quedaron temporalmente en un segundo plano... De acuerdo con los principios conductistas, la psicología tenía que parecerse a la física en su metodología...¹⁰³

Gracias a programas de investigación como los de Piaget, Barlett, la psicología de la *Gestalt* y anteriormente von Helmholtz, Brentano y la escuela de Wurzburg, a través de sus representantes emigrados a Estados Unidos, había un pequeño grupo de investigadores que hacía frente al conductismo. Así, se comenzaron a oír voces que hablaban de estudiar los fenómenos cognitivos en sí mismos, y se introdujeron algunas ideas que hablaban de la estructura psicológica y cerebral. Como se ha mostrado en la sección 2.2, en las conferencias Macy estaba presente Kluver, psicólogo de la *Gestalt*. Por otra parte, la física y las matemáticas, como modelo a seguir por las otras ciencias, servían también a las corrientes contrarias al conductismo, al hacer incursión en la explicación de los fenómenos

¹⁰² Steve Joshua Heims, *The Cybernetics Group 1946-1953. Constructing a Social Science for Postwar America*, 4.

¹⁰³ Steve Joshua Heims. *The Cybernetics Group 1946-1953. Constructing a Social Science for Postwar America*, 8-10.

cognitivos. Esto se hacía patente en los trabajos de von Neumann, McCulloch, y otros, que comparaban la estructura de los autómatas con la del cerebro humano y comenzaban a intentar dar explicaciones de la cognición en términos físicos y matemáticos.

Haciendo uso del enfoque computacional y de la teoría de la información, el psicólogo de Harvard, George Miller, publicó en 1956 un artículo con el título: “El mágico número 7, más o menos 2: algunos límites en nuestra capacidad para procesar información”.¹⁰⁴ Este trabajo mostraba que la capacidad de los individuos para establecer distinciones absolutas entre los estímulos, para distinguir los fonemas entre sí o calcular cifras con precisión, y para recordar una serie de ítems aislados, parecía experimentar un cambio radical al acercarse a los siete elementos. Por debajo de ese número, eran capaces de realizar esas tareas con facilidad, al contrario que por encima de él. También en Harvard, en 1956, Jerome Bruner trató el problema de la categorización o agrupación de elementos en categorías. Bruner analizaba las estrategias que decían emplear los sujetos, al contrario que el conductismo que solo tenía en cuenta los resultados y no los procedimientos empleados en su obtención. En Gran Bretaña, como resultado de los trabajos durante la segunda Guerra Mundial. Cherry y Broadbent (1953-54) publicaron sus estudios sobre la capacidad de los individuos para atender y obtener información de canales “ruidosos”. Según este trabajo, los individuos tienen una capacidad limitada para la recepción y almacenamiento de la información y sólo la información que deviene consciente puede incorporarse a la memoria de largo plazo y pasar a formar parte del conocimiento activo. Broadbent introdujo etapas o “diagramas de flujo” en el procesamiento de la información, tomando ideas de la ingeniería de comunicaciones, con lo que abría la posibilidad de

¹⁰⁴ George Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information”, en *Psychological Review*, 63, 81-97. En web: <http://www.well.com/user/smalin/miller.html>

examinar las dimensiones temporales de diversos procesos psicológicos.¹⁰⁵

Se fue creando una mayor disposición a abordar en forma directa los temas relacionados con la mente, en lugar de descartarlos reemplazándolos por una larga serie de estímulos y respuestas públicamente verificables. Se consideró el modelo computacional como una representación de la información en el interior de la mente. Muchos temas no eran nuevos en psicología, pero sí era novedoso utilizar el modelo computacional y las ideas provenientes de la teoría de la información. Ulrich Neisser comenta:

Todo esto se debió a que la actividad de la computadora es en sí misma bastante afín a los procesos cognitivos. La computadora recoge información, manipula símbolos, almacena ítems en su “memoria” y luego los recupera, clasifica los datos de entrada, reconoce pautas, y así sucesivamente.¹⁰⁶

La Psicología Cognitiva nace adoptando el modelo computacional del pensamiento, con el *cognitivismo*, la idea del pensamiento como manipulación de símbolos conforme a reglas. Este enfoque puede verse en el libro de Neisser *Cognitive Psychology*, de 1967. A partir de entonces hay numerosas contribuciones dentro del enfoque computacional, como la teoría modular de la mente de Jerry Fodor.¹⁰⁷ Los temas que trata cubren, entre otros, los modelos de memoria, percepción, atención, conciencia, toma de decisiones, reconocimiento de objetos, desarrollo cognitivo, lenguaje y representación del conocimiento.

¹⁰⁵ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 106-112.

¹⁰⁶ Ulrich Neisser, *Cognition and Reality*, W.H. Freeman, San Francisco, 1976, 5-7. Citado en Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 137.

¹⁰⁷ Jerry Fodor. *The Modularity of Mind*. MIT/Bradford Press, 1983.

Pero en los años setenta abundaron las críticas al modelo computacional y a los resultados y métodos que empleaba la Psicología Cognitiva. Por un lado, se volvía a apreciar que los resultados de laboratorio muchas veces no concordaban con lo observado en la vida diaria. En este sentido, Gibson reclamó una psicología perceptual que estudiase la manera en que los seres humanos ven el mundo que los rodea al desplazarse por él e interactuar con los objetos, y Rosch sostuvo la necesidad de realizar investigaciones sobre la formación de conceptos que tuviesen en cuenta los objetos complejos del mundo. Por otro lado, los resultados en neurociencia y el resurgimiento del conexionismo hicieron que no se tuviese el modelo computacional como único referente, y que las teorías psicológicas tuvieran que ser contrastadas en mayor medida con los avances en neurociencia. Como muestra de esto, puede mencionarse lo significativo que resulta que los últimos manuales de Psicología Cognitiva incluyan capítulos de Neurociencia Cognitiva como parte de su programa.¹⁰⁸

3.2 Neurociencia

Los avances en el estudio del sistema nervioso, y más concretamente, del cerebro como órgano central de las actividades cognitivas, han tenido gran importancia en la historia de la Ciencia Cognitiva. Ya a finales del siglo diecinueve, Camilo Golgi, Santiago Ramón y Cajal y Charles Sherrington, entre otros, desarrollaron la teoría neuronal. Cajal y Golgi mantenían posiciones contrarias por lo que respecta a la naturaleza de la actividad neuronal: Cajal defendía el carácter “local” y celular de la neurona, frente al carácter “holista”, no unitario o reticular, de la red de neuronas de Golgi. La tesis de Cajal se impuso en gran medida, y aunque el debate entre localizacionistas y

¹⁰⁸ Véase Robert J. Sternberg, *Cognitive Psychology*, Wadsworth, Belmont, California, 2003, 30-64. O también Michael W. Eysenck, Mark T. Keane, *Cognitive Psychology*, Psychology Press, Hove, 2003, 2000, 16-21.

reticularistas continúa aún hoy en distintas versiones, sus investigaciones junto con las de Sherrington respecto a la naturaleza de las uniones entre neuronas o *sinapsis* pusieron las bases de la investigación cerebral y del sistema nervioso.¹⁰⁹

Sin embargo, no fue hasta mediados del siglo XX cuando se comenzaron a obtener resultados fiables sobre lo que podían ser las consecuencias cognitivas de la actividad cerebral. Algunas de estas ideas se presentaron, por ejemplo, durante la etapa cibernética en el simposio de Hixon (1948), cuando Karl Lashley, atacando al conductismo, planteó que el sistema nervioso consiste en unidades organizadas en forma jerárquica y permanentemente activas, cuyo control proviene del centro, más que de cualquier estimulación periférica.¹¹⁰ Se comenzaba así a indagar en la estructura cerebral desde el punto de vista cognitivo.

Desde los años cincuenta hasta los años setenta, en la etapa cognitivista, el paradigma dominante en cuanto a modelo del pensamiento era el computacional, con lo que se atendía fundamentalmente a las similitudes entre el ordenador y el cerebro. Aunque se buscaba contrastar las teorías con los experimentos en neurociencia, no se perseguía la total identificación de los procesos neurales con los cognitivos, ya que la hipótesis cognitivista del pensamiento como manipulación de símbolos conforme a reglas, como ejecución de un programa o secuencia de reglas, lo hacía en gran medida independiente de la máquina (estructura neural) en la que se desarrollase físicamente.

A finales de los años setenta, la insatisfacción con el modelo computacional hizo que se volviese la vista hacia modelos que reflejasen en más detalle la estructura del cerebro. Por otra parte, la

¹⁰⁹ Jennifer Mundale, “Neuroanatomical Foundations of Cognition”, en William Bechtel, Peter Mandik, Jennifer Mundale y Robert S. Stufflebeam, *Philosophy and the Neurosciences. A Reader*, 42-43.

¹¹⁰ Stanley Finger, *Origins of Neuroscience. A History of Explorations into Brain Function*, Oxford University Press, Oxford, 1994, 32-61.

neurociencia se iba sintiendo más segura sobre poder decir algo acerca de la cognición a medida que refinaba sus métodos. Los estudios anteriores sobre el cerebro estaban limitados a las lesiones producidas en las guerras mundiales, y al estudio de cerebros de animales. Pero a finales de siglo XX, se mejoraron las técnicas no invasivas, como la tomografía por emisión de positrones (PET) y la toma de imágenes por resonancia magnética funcional (fMRI).

Sin embargo, siguen persistiendo las dificultades para determinar en qué grado lo que se observa corresponde a una actividad cognitiva en concreto, o en qué medida el instrumento altera la actividad cerebral, y también, cómo la actividad cerebral en el laboratorio se corresponde con una situación normal de la vida cognitiva del sujeto. Bechtel y Stufflebeam proponen tres criterios para la evaluación de los experimentos: lo concluyente de los resultados mismos, la coherencia de los resultados con los obtenidos por otros métodos, y la coherencia de los resultados con modelos teóricos plausibles.¹¹¹ Más adelante ponen un ejemplo de estas dificultades, incluyendo también los supuestos que se introducen en el experimento, para el caso de fMRI:

Debemos darnos cuenta en primer lugar, que tanto PET como fMRI introducen limitaciones considerables en las actividades cognitivas que se pueden estudiar. El sujeto tiene que desarrollar las actividades mientras está tendido sobre la espalda sin moverse dentro del scanner. Por tanto, las actividades cognitivas registradas en la imagen no serán las ordinarias de la vida, sino las especificadas para tal circunstancia. Segundo, la meta de las técnicas de tomas de imágenes es localizar procesos u operaciones cognitivas específicas identificando el incremento de flujo sanguíneo asociado con ellos (y, en la medida en que sea posible, determinar la relación entre diferentes procesos). Estos estudios dependen por tanto de forma crítica de la elaboración de tareas para las cuales los investigadores tienen de antemano una

¹¹¹ William Bechtel y Robert S. Stufflebeam, "Epistemic Issues in Procuring Evidence about the Brain: The Importance of Research Instruments and Techniques", en William Bechtel, Peter Mandik, Jennifer Mundale y Robert S. Stufflebeam, *Philosophy and the Neurosciences. A Reader*, 69.

descomposición cognitiva plausible que emplean para guiar la interpretación de los resultados. La psicología del procesamiento de la información ha propuesto descomposiciones cognitivas de cierto número de actividades en etapas sucesivas de procesamiento, y los neurocientíficos las han utilizado frecuentemente en técnicas de neuroimágenes. Pero estas mismas descomposiciones son rebatidas, principalmente por los defensores de los modelos de sistemas dinámicos, que niegan que el comportamiento sea el resultado de etapas sucesivas de procesamiento, sino el producto emergente de procesos dinámicos altamente distribuidos... Incluso si se acepta que hay una descomposición de procesos cognitivos en el cerebro, permanece la cuestión de exactamente qué descomposición utiliza el cerebro. El asunto entonces es subrayar que cualquier estudio por imágenes es sólo tan bueno como la asunción de la descomposición de los componentes de procesamiento sobre la que descansa.¹¹²

A pesar de que, como se ha indicado en esta cita, los experimentos tienen en ocasiones una notable carga teórica previa, algunos autores ven en los métodos y resultados de la neurociencia la base “científica” sobre la que edificar la Ciencia Cognitiva, pudiendo aquélla explicar todos los fenómenos cognitivos, desde la inteligencia hasta la conciencia. Por ejemplo, Clark afirma que la conciencia es “materia, debidamente dispuesta”¹¹³, y Francis Crick que “Tú, tus alegrías y tus penas, tus recuerdos y tus ambiciones, tu sentido de identidad personal y libre voluntad, son de hecho nada más que el comportamiento del vasto ensamblaje de células nerviosas y de sus moléculas”.¹¹⁴ Esto lleva a un reduccionismo de las demás disciplinas a la neurociencia, que daría la explicación última de los fenómenos, incluso de los sentimientos e identidad de la persona.

¹¹² William Bechtel y Robert S. Stufflebeam, “Epistemic Issues in Procuring Evidence about the Brain: The Importance of Research Instruments and Techniques”, en William Bechtel, Peter Mandik, Jennifer Mundale y Robert S. Stufflebeam, *Philosophy and the Neurosciences. A Reader*, 70.

¹¹³ Andy Clark, *Mindware. An Introduction to the Philosophy of Cognitive Science*, 5.

¹¹⁴ Francis Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul*, Touchstone, Nueva York, 1995, 3, citado en M. R. Bennett and P. M. S. Hacker, *Philosophical Foundations of Neuroscience*, 355.

Otros autores, en cambio, como Bechtel, Bennett y Hacker, defienden la independencia de distintas disciplinas y niveles de explicación, y de las personas y los pensamientos respecto al cerebro. Así, Bennett y Hacker opinan:

...debería ser evidente que el sujeto pensante no es el cerebro, sino el ser humano. No es el cerebro el que se concentra al hacer una operación con el debido cuidado y atención, sino el cirujano. No es el cerebro el que juega inteligentemente al tenis o realiza una gran interpretación de la sonata “Hammerklavier”, sino el jugador de tenis y el pianista... Los cerebros no pueden razonar ni bien ni mal, aunque los procesos cerebrales de la gente que es capaz de razonar bien tienen sin ninguna duda características distintivas... No sólo no es el cerebro el sujeto del pensamiento, tampoco es el lugar de los pensamientos. Al contrario de lo que a menudo afirman los neurocientíficos (que hablan, por ejemplo, de “la cabeza humana dentro de la que no tenemos ninguna duda de que ocurran los pensamientos”), el pensamiento no ocurre *en* el cerebro, sino en el despacho, en la biblioteca, o paseando por la calle. La localización *del hecho de una persona pensando* es algo que le ocurre a la persona. Se puede encontrar a los pensamientos escritos en textos, pero no en la cabeza de los seres humanos. Las personas pueden expresar pensamientos, no así los cerebros. Ya que un pensamiento es lo que es accesible mediante una pronunciación u otra representación simbólica. El hecho de que los seres humanos puedan pensar y no decir lo que piensan no implica que los pensamientos estén expresados en sus cerebros.¹¹⁵

¹¹⁵ M. R. Bennett and P. M. S. Hacker, *Philosophical Foundations of Neuroscience*, 179-180.

3.3 Lingüística

Una de las formas más directas de tener acceso al pensamiento es a través del lenguaje, por lo que es lógico pensar que un análisis de éste podrá revelar datos significativos sobre la cognición.

Saussure, en Ginebra, a finales del siglo XIX, consideró la lengua como una totalidad organizada, cuyas diversas partes son interdependientes y derivan su significación del sistema en su conjunto. Esta teoría coincide en el tiempo con la mantenida por los psicólogos de la *Gestalt* acerca de las interrelaciones entre los elementos de la percepción visual.¹¹⁶

La influencia de Saussure tuvo continuidad con la Escuela de Praga en la década de 1920, con Roman Jakobson y Nikolai Troubetskoy. Esta escuela estaba interesada fundamentalmente en la fonología, y consideraba que un fonema es un conjunto con características propias, compuesto de una serie de rasgos articulatorios diferenciados, y atribuía realidad psicológica a estas descripciones, al considerar que el sistema nervioso había evolucionado de manera que se aseguraba la producción y discriminación de los rasgos fonológicos. Esta concepción del fonema supuso un avance importante, al comprobarse que el sistema fonémico de todas las lenguas podía caracterizarse mediante un número de características, y que estas características permitían formular generalizaciones que no se podían enunciar mediante los modelos estructuralistas.¹¹⁷

En Estados Unidos, durante la misma época de principios del siglo XX, la figura principal de la recepción de las ideas de Saussure fue Leonard Bloomfield. Bloomfield desarrolló un sistema de notación para el estudio de lenguas poco conocidas, y fue un gran

¹¹⁶ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 221-222.

¹¹⁷ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 223-224.

impulsor del establecimiento institucional de la lingüística como disciplina en Estados Unidos. Adoptó una perspectiva conductista del lenguaje, concibiendo a todos los lenguajes como fenómenos puramente físicos y rechazó, para la explicación del comportamiento lingüístico, cualquier característica mental.¹¹⁸

Pero otros lingüistas y antropólogos, aunque no de tanta influencia en la época de Bloomfield, no compartían este punto de vista. Así, Edward Sapir y su discípulo Benjamín Lee Whorf propusieron la hipótesis de que los propios procesos de pensamiento del individuo están estructurados, o incluso regidos, por las propiedades de la lengua que el sujeto habla.¹¹⁹

Es ilusorio imaginar que uno se adapta a la realidad en lo fundamental sin recurrir al lenguaje, y que éste no es otra cosa que un medio accesorio para resolver problemas concretos de la comunicación o la reflexión. La verdad es que “el mundo real” se construye, en gran medida, de manera inconsciente, a partir de los hábitos lingüísticos del grupo al que pertenece... Vemos, oímos y experimentamos de la manera en que lo hacemos porque los hábitos lingüísticos de nuestra comunidad nos predisponen a ciertas elecciones en nuestras interpretaciones.¹²⁰

Disecionamos la naturaleza de acuerdo a las líneas trazadas por nuestras lenguas. Las categorías y tipos que aislamos del mundo de los fenómenos no los encontramos ahí porque nos estén mirando fijamente a la cara. Al contrario, el mundo se nos presenta como un flujo caleidoscópico de impresiones que tienen que ser organizadas por nuestras mentes – y esto significa, en gran medida, por los

¹¹⁸ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 225-226.

¹¹⁹ Para una discusión de esta afirmación, llamada *hipótesis de relatividad lingüística*, puede verse John A. Lucy, *Language Diversity and Thought. A Reformulation of the Linguistic Relativity Hypothesis*, Cambridge University Press, Cambridge, 1992.

¹²⁰ Edward Sapir, *Language*, Harcourt, Brace & World, Nueva York, 1921, 75. En Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 227.

sistemas lingüísticos de nuestras mentes. Seccionamos la naturaleza, la organizamos en conceptos, y asignamos significados como lo hacemos, en gran medida porque somos parte de un acuerdo para hacerlo de este modo – un acuerdo que se mantiene a través de nuestra comunidad lingüística, y que se codifica a través de patrones de nuestro lenguaje. El acuerdo es, por supuesto, implícito y no hecho manifiesto, *pero sus términos son absolutamente obligatorios*. No podemos hablar en absoluto sin suscribir la organización y clasificación de datos conforme a dicho acuerdo.¹²¹

A medida que decaía el conductismo y aumentaba la influencia de la teoría de la información a través del cognitvismo, se planteaban dudas sobre la validez de la lingüística de Bloomfield. Ya no se trataba de reunir datos y describirlos de forma objetiva: las expectativas del sujeto, sus conocimientos del significado de esos datos, inevitablemente los coloreaban y daban, a descripciones en apariencia objetivas, el aspecto de estar dando vuelta en círculos. Incluso algún defensor de las tesis de Bloomfield, como Hockett, declaró hacia 1955 que era imposible analizar un lenguaje objetivamente, y que había que “empatizar” con la información.¹²²

La figura más importante de la lingüística en Estados Unidos a mediados de siglo XX, Noam Chomsky, en un intento por restaurar el edificio de la lingüística estructuralista estableció un sistema de reglas que, a partir de las estructuras abstractas subyacentes, se proyectaban en las oraciones articuladas del lenguaje (gramática generativa). Chomsky combatió el conductismo, al afirmar que su punto de vista no era adecuado para dar cuenta de la habilidad de los humanos para aprender y utilizar el lenguaje. Aunque la versión de Chomsky estaba próxima al cognitvismo, en el sentido de generar la gramática a partir de unas reglas, su punto de vista era más general al considerar que las

¹²¹ Benjamín Lee Whorf, “Science and Lingüistics” *Technol. Rev.*, 42:229-231, 247-248, no. 6 (April 1940), en Benjamin Lee Whorf, *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf*, edited by John B. Carroll, The MIT Press, Cambridge, Massachussets, 1998, 1956.

¹²² Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 229.

gramáticas finitas no podían dar cuenta del lenguaje. Pero tuvo influencia en psicólogos cognitivistas como Jerry Fodor, que a mediados de los setenta defendió la existencia de un lenguaje innato del pensamiento, un conjunto de reglas para el manejo de símbolos que son parte de la estructura cognitiva humana.

A finales de los setenta y principios de los ochenta, rescatando teorías psicológicas de la *Gestalt*, las teorías de Sapir y Whorf, tomando en cuenta la formación de categorías de forma no objetiva sino conforme a la interacción con el mundo y mediadas por esta interacción (Eleanor Rosch, 1976), y considerando los experimentos sobre la percepción del color (Berlin y Kay, 1979), George Lakoff criticó la postulación de Chomsky de dos niveles de análisis: una estructura profunda, interpretada semánticamente, y una estructura de superficie, interpretada fonéticamente. También criticó la afirmación de la autonomía de la sintaxis respecto a la semántica, y de ésta respecto a la pragmática. Para Lakoff, nuestras expresiones están fuertemente arraigadas en la experiencia, están *corporeizadas*, y organizamos y expresamos, casi siempre inconscientemente, nuestro pensamiento mediante metáforas que aluden a nuestra experiencia corpórea. Se hará una exposición de esta teoría más adelante en el próximo capítulo.

La lingüística desarrollada a partir de la teoría de Lakoff se denomina Lingüística Cognitiva. Como señala Tomasello:

La Lingüística Cognitiva tuvo claramente sus orígenes como paradigma científico en 1987 con la publicación de la obra de George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* y con la de Ronald Langacker, *Foundations of Cognitive Grammar*, seguidas inmediatamente por la fundación de la Asociación Internacional de Lingüística Cognitiva y de su publicación oficial *Cognitive Linguistics*.¹²³

¹²³ Michael Tomasello, "Cognitive Linguistics", en *A Companion to Cognitive Science*, edited by William Bechtel and George Graham, 477-478.

Así, la Lingüística Cognitiva incluye entre sus temas principales la formación de categorías conforme a nuestra experiencia y su reflejo en el lenguaje y el análisis de las metáforas y las metonimias con relación a la formación de conceptos.¹²⁴

3.4 Antropología

Una aproximación al estudio del pensamiento debe tener en cuenta su carácter social. El estudio del pensamiento de los individuos en las sociedades, fundamentalmente por medio del análisis lingüístico, y adoptando supuestos próximos al paradigma más dominante en cada momento de la Ciencia Cognitiva como el cognitivismo o conexionismo, es lo que se denomina Antropología Cognitiva.

Alrededor de 1871, El inglés Edward Tylor presentó la idea según la cual tanto la cultura humana como las religiones eran producto de una evolución natural de las capacidades mentales humanas. Su trabajo sobre las sociedades primitivas contrastó las especulaciones teóricas anteriores con los datos empíricos, comenzando una era de trabajos de experimentación, de trabajos “de campo” como método de hacer antropología, que continúa hasta nuestros tiempos.¹²⁵

En Estados Unidos, la antropología tuvo su origen a principios de siglo XX de la mano de Franz Boas, quien intentó fundar una antropología científica basada en hechos experimentales. Boas combatió el evolucionismo de Tyler, dando validez tanto a la visión

¹²⁴ Puede verse, por ejemplo, William Croft y D. Alan Cruse, *Cognitive Linguistics*. Cambridge University Press, Cambridge, 2004. O también, Fredrich Ungerer and Hans-Jörg Schmid, *An Introduction to Cognitive Linguistics*, Addison Wesley Longman Limited, Londres, 1997, 1996.

¹²⁵ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 251-252.

científica del experimentador como a la subjetiva del miembro de la cultura que se estudia. Afirmó que la cultura no tenía una evolución lineal, y que las capacidades cognitivas de los pueblos indígenas y los “avanzados” eran las mismas. Subrayó la importancia del lenguaje y de la lingüística para cualquier estudio antropológico, y defendió que cada lengua debía ser considerada, al estilo de Saussure, como una totalidad orgánica. Aunque pensaba que el pensamiento influye en el lenguaje, y no al contrario, como creía su discípulo Sapir y posteriormente Whorf, indicó el carácter inconsciente del manejo de la estructura del lenguaje y la posibilidad de que ambos influyeran en el comportamiento.¹²⁶

En Francia, en los años cincuenta, Lévi-Strauss insistió en la importancia del lenguaje, aplicándolo directamente al estudio del funcionamiento de la mente, y eligió como método la lingüística estructural. Posteriormente un discípulo suyo, Dan Sperber, continuó con el enfoque lingüístico en los años sesenta, pero eligiendo la gramática generativa de Chomsky y la lingüística de Fodor. Sperber opinaba que los procesos simbólicos, en lugar de estar contruidos a partir de la experiencia, forman parte de la estructura innata de la mente.¹²⁷ Con este enfoque cognitivista, bajo la influencia de la teoría de la información y de la teoría computacional, nació lo que se llamó etnociencia, análisis componencial o Antropología Cognitiva. Sus defensores, como Ward Goodenough o Floyd Lounsbury, procuraban descubrir cómo el conocimiento de unas reglas de su cultura se refleja en la conducta, y fundamentalmente, en el lenguaje de sus miembros.¹²⁸

Pero hay problemas evidentes a la hora de identificar claramente los términos relevantes de los conceptos que se expresan, o

¹²⁶ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 255-256.

¹²⁷ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 260-264.

¹²⁸ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 269.

incluso para identificar los mismos conceptos. Así, el carácter polisémico de las palabras, o el uso de las metáforas, son demasiado elásticos en muchas ocasiones como para someterlos a un examen rígido.¹²⁹

En esta línea se manifiestan los críticos de la Antropología Cognitiva. Por ejemplo, Geertz opina que:

Nada ha contribuido más a desacreditar el análisis cultural que la elaboración de impecables descripciones de orden formal, en cuya existencia efectiva nadie cree.¹³⁰

Stephen Tyler, desengañado por la etnociencia, expresaba en la década de los setenta:

De un estilo de pensamiento que lo vacía de todo contenido, lo menos que podemos presagiar es la muerte del significado. ¿Y qué otra cosa cabe esperar de un método de análisis que pretendió demostrar que el significado emerge, misteriosamente, de la concatenación mecánica de elementos carentes de significado?... Ya sea en el arte o en la ciencia, no hay nada más claro que la pobreza intelectual del formalismo.¹³¹

Y Roger Keesing:

Durante casi quince años, los antropólogos cognitivistas han llevado la “nueva etnografía” tan lejos como ésta parecía conducirlos. Pero resulta obvio que no los llevó muy lejos, al menos en los últimos cinco años... y que las promesas mesiánicas de la controversia primitiva no se cumplirán. Los “nuevos etnógrafos” fueron incapaces de ir más allá del análisis de dominios semánticos artificialmente simplificados y

¹²⁹ Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 274.

¹³⁰ Clifford Geertz, *The Interpretation of Cultures*, 18, en Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 275.

¹³¹ Edward Tyler, *The Said and the Unsaid: Mind, Meaning and Culture*, New York. Academic Press, New Cork, 1978, 465, citado en Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 276.

circunscritos (y por lo común triviales), y esto ha desalentado a muchos de sus devotos originales.¹³²

En los años setenta, en una línea más alejada del modelo computacional, Eleanor Rosch demostró que, independientemente de los vocablos que tengan las distintas culturas para designar los colores, los individuos seleccionan, clasifican y tratan el espectro cromático de forma aproximadamente similar. En la misma década, Berlin y Kay mostraron que los términos utilizados por distintas sociedades para designar los colores siguen una pauta sistemática. Además, Rosch mostró el carácter “abierto” de los conceptos y categorías, al indicar cómo estos se agrupan en torno a prototipos centrales que maneja el ser humano, de los que no se puede dar una definición concluyente. Estas teorías llevaron a pensar que, en todo el mundo, las personas perciben y clasifican los objetos de forma aproximadamente similar, con lo que habría una “equipamiento” mental innato en todos los hombres. Estas teorías están más próximas al conexionismo, como sistemas cerebrales que se adaptan al medio, aprenden, y son capaces de seguir funcionando aun con alteraciones en su estructura.

Actualmente, al igual que ocurre en Psicología Cognitiva, los programas de Antropología Cognitiva combinan los enfoques próximos al cognitivismo y los cercanos al conexionismo, en una tarea que se lleva a cabo eliminando uno de los paradigmas, o en una labor integradora de ambos ciertamente difícil. En cualquier caso, se persigue una supuesta fundamentación científica de la disciplina al estilo de las ciencias experimentales. Desde este punto de vista la Antropología Cognitiva critica en ocasiones a sus adversarios. Como indica D’Andrade:

La alternativa era la antropología simbólica de Victor Turner, David Schneider, Clifford Geertz, y otros... Se podía leer cualquier conjunto

¹³² Roger Keesing, *Cultural Anthropology: A Contemporary Perspective*, Holt, Rinehart & Winston, New York, 1976, 307, citado en Howard Gardner, *La nueva ciencia de la mente: historia de la revolución cognitiva*, 276-277.

de prácticas culturales como un significado cultural. Sin embargo, nuevamente no existía un método de validación. Como los significados no estaban en la mente de nadie, ni siquiera en una mente inconsciente, no era posible ningún método de verificación.¹³³

3.5 Filosofía

Las teorías en filosofía que se podrían considerar parte de la Ciencia Cognitiva están dentro de lo que actualmente es la filosofía de la mente. La filosofía de la mente acompaña a la Ciencia Cognitiva desde sus orígenes en un proceso de mutua realimentación, nutriéndose de sus asunciones y proporcionándole a la vez marcos conceptuales sobre los que basar sus teorías. A la vez, también influyen sobre sus teorías las distintas corrientes filosóficas, como el racionalismo, la filosofía analítica o la fenomenología.

Esta filosofía de la mente se ocupa de problemas epistemológicos tradicionales aunque los enmarca en las discusiones y teorías más actuales en Ciencia Cognitiva. Los principales problemas son el análisis de la estructura del pensamiento, cómo este interactúa y se desenvuelve en el mundo y la estructura de la relación entre mente y cuerpo.

Los precedentes modernos de la filosofía de la mente se pueden situar en la teoría epistemológica de Descartes en el siglo XVII (racionalismo). Para Descartes, el papel que juegan los sentidos en el descubrimiento de los procesos que operan en la naturaleza es secundario respecto al que juega la razón, siendo los procesos mentales independientes de los físicos. Esta concepción considera el pensamiento como razonamiento lógico no condicionado por la

¹³³ Roy D'Andrade, *The Development of Cognitive Anthropology*, Cambridge University Press, Cambridge, 2003, 1995, 248.

experiencia, por lo que abre la puerta al uso del ordenador como máquina que modele el pensamiento, tal y como hace el cognitivismo.

Durante los siglos XVII y XVIII, en el mundo anglosajón se desarrolló un punto de vista que daba a la percepción sensorial un papel más central en la explicación del pensamiento (empirismo). Locke afirmó que el conocimiento se basa en la experiencia sensorial, y a partir de asociaciones de estas impresiones se obtienen distintos tipos de ideas.

Puede verse la obra de Kant (siglo XVIII) como un intento de síntesis entre racionalismo y empirismo. La filosofía anterior a Kant suponía que los objetos de conocimiento existen independientemente de nosotros, y a partir de ahí se preguntaba cómo podemos conocerlos. Kant mantuvo que nuestras actividades cognitivas eran parcialmente constitutivas de los objetos de los que tenemos experiencia. Sería nuestra propia participación en la construcción de los objetos de percepción lo que haría posible su conocimiento. Esta construcción se haría aplicando unas categorías a nuestros datos sensoriales. Estas categorías, como las de causa y efecto, serían innatas y se aplicarían sobre la base de unos esquemas también innatos de espacio y tiempo.¹³⁴

Pero los experimentos en psicología de von Helmholtz, la teoría de la relatividad y las geometrías no euclídeas pusieron estos esquemas innatos en duda. Por otro lado, el avance logicista hizo que se girase la atención hacia el lenguaje como instrumento para analizar el pensamiento y como fuente de problemas filosóficos. Se consideraba que los problemas filosóficos se reducían a problemas lingüísticos y que un correcto análisis científico del lenguaje los resolvería. Fue en este periodo de principios de siglo XX de positivismo lógico cuando se abordaron los intentos de formalización

¹³⁴ William Bechtel, *Filosofía de la mente. Una panorámica para la ciencia cognitiva*, Tecnos, Madrid, 1991. 31-32. Original en inglés, *Philosophy of Mind. An Overview for Cognitive Science*, Lawrence Erlbaum Associates, Inc., Hillsdale, Nueva Jersey, 1988.

del lenguaje por parte de matemáticos y filósofos, como se ha visto en el capítulo 2. Algunos de los principales filósofos de esta corriente fueron Frege, Russell y Wittgenstein en su primera época. En Ciencia Cognitiva, el cognitivismo asume explícitamente esta idea de la posibilidad de formalizar el pensamiento, y otras teorías como el conexionismo lo hacen al menos de forma implícita al utilizar el ordenador en los estudios de los modelos cerebrales que postulan.

Sin embargo, los problemas de formalización del pensamiento, como se ha visto en el capítulo 2, eran mayores de lo que en principio se pensaba. Los problemas de fijación de los significados de forma inequívoca y de saber cómo se alcanza a conocer el mundo seguían sin resolverse, impidiendo que el positivismo lógico llevase a cabo este análisis formal.

En su segunda época, Wittgenstein se apartó del intento de formalización del lenguaje, considerando que este se manejaba conforme a "juegos" y que las categorías no eran definibles mediante condiciones necesarias y suficientes sino que se formaban agrupando sus componentes por "aires de familia", lo que posteriormente recogió Rosch en su teoría de prototipos. También Quine apuntó cómo el significado depende del contexto y la imposibilidad de definir la analiticidad. Estas teorías tuvieron influencia en concepciones de la Ciencia Cognitiva en las que se tenían en cuenta el carácter abierto de la interacción del hombre con el mundo, y que consideraban a los conceptos como parte constitutiva de esta interacción, en lugar de estructuras de símbolos que se "entregaban" al sistema cerebral para que este los procesara. Algunas de estas teorías son la teoría de la lingüística cognitiva de Lakoff y Johnson, las teorías de sistemas dinámicos y el conexionismo.

La filosofía de la mente analiza la existencia y la posible relación entre estados mentales, representaciones del mundo, estados intencionales y estados neuronales. Dependiendo de la interpretación que haga de estas relaciones se encontrará más cerca del cognitivismo o del conexionismo. Recientemente las teorías conexionistas y la

neurociencia han dado cierto auge al *materialismo*, que propugna que los estados mentales no son otra cosa que estados neuronales.¹³⁵ Otras teorías recientes recuperan aspectos de la fenomenología para definir un "yo" que va más allá de los estados neuronales y de las representaciones del mundo.¹³⁶

Estas teorías recientes toman el *evolucionismo* como motor de la vida, el pensamiento y el desenvolvimiento del hombre en el mundo, de tal forma que la realidad última que define al hombre es la adaptación o acoplamiento al medio y la evolución para la supervivencia o viabilidad de la especie. Se interpreta la ética, los valores, verdades, etc., como un ecologismo de adaptación o comunión con el medio.

En una forma de abordar el fenómeno cognitivo distinta del cientismo dominante, algunos filósofos recuperan del pragmatismo americano y especialmente de Charles S. Peirce, la idea de significar, de semiosis, como proceso creador por el que fluye el pensamiento de forma dinámica, más que como mera correspondencia de éste con el mundo. Esta corriente reconoce la imposibilidad de formalizar el pensamiento como hace el segundo Wittgenstein, pero no diluye los problemas filosóficos en problemas lingüísticos. Tampoco propugna la imposibilidad de que el discurso filosófico aluda a una realidad distinta a la por él construida, ni toma el evolucionismo como motor de toda la actividad humana, sino que busca la pauta de objetividad y de verdad en la comunidad, en la discusión filosófica y científica. Esta búsqueda de la verdad sería compatible con las teorías evolucionistas

¹³⁵ Paul M. Churchland, *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*, Gedisa, Barcelona, 1999, 50-64, 75-85. Original en inglés, *Matter and Consciousness*, The MIT Press, 1984.

¹³⁶ Cf. Francisco J. Varela, Evan Thompson y Eleanor Rosch, *De cuerpo presente. Las ciencias cognitivas y la experiencia humana*, Gedisa, Barcelona, 1997. Original en inglés, *The Embodied Mind. Cognitive Science and Human Experience*, The MIT Press, 1991.

pero no estaría regido por sus reglas.¹³⁷ Uno de los máximos exponentes de este "realismo pragmatista" es Hilary Putnam, en Estados Unidos. Si bien esta corriente filosófica no se reconocería a sí misma como parte o colaboradora de la Ciencia Cognitiva o la Filosofía de la Mente, sí hay algunos intentos de aplicación de sus ideas a la explicación de problemas centrales en Ciencia Cognitiva, como la idea de *representación* del mundo en el pensamiento.¹³⁸

3.6 Resumen

En este capítulo se ha descrito brevemente siguiendo la exposición de Gardner en psicología, neurociencia, lingüística y antropología, cómo en estas disciplinas y en filosofía han ido desarrollándose algunas corrientes que adoptan y contribuyen al desarrollo de los distintos paradigmas de la Ciencia Cognitiva. La mayoría de estas corrientes sitúan las explicaciones del pensamiento y sus leyes como fenómenos que se dan exclusivamente en la "mente" de los hombres. El descubrimiento de dichas leyes, desde un punto de vista, en general, próximo a la física, matemática y biología, hace que estas corrientes estén próximas a una explicación reduccionista de la cognición.

Por otro lado, la variedad e incluso divergencia en los paradigmas de la Ciencia Cognitiva, hacen que distintas corrientes con el apellido de "cognitivas" en muchas ocasiones planteen explicaciones enfrentadas de la cognición, lo que también resulta desconcertante respecto a su denominación, y que en otros casos intenten compatibilizar los distintos enfoques, en un trabajo bastante difícil.

¹³⁷ Cf. Jaime Nubiola, *La renovación pragmatista de la filosofía analítica. Una introducción a la filosofía contemporánea del lenguaje*, EUNSA, Barañáin, Navarra, 1996, 1994.

¹³⁸ Cf. Barbara von Eckardt, *What is Cognitive Science?*

En el capítulo próximo, como ejemplos ilustrativos de cómo la Ciencia Cognitiva intenta configurar su carácter científico e interdisciplinar, y de las dificultades que encuentra para dar una explicación completa del pensamiento, se presentarán dos ejemplos de teorías en Ciencia Cognitiva: la teoría de la metáfora de Lakoff, y, de forma más breve, la teoría de espacios conceptuales de Peter Gärdenfors.

CAPÍTULO IV

Dos teorías actuales en Ciencia Cognitiva

En este capítulo se presenta la teoría cognitiva de la mente corporeizada de George Lakoff y Mark Johnson y la teoría de espacios conceptuales de Peter Gärdenfors, como ejemplos de teorías en Ciencia Cognitiva. En la teoría de Lakoff y Johnson se ofrece una explicación de los fenómenos cognitivos en base a la atención y análisis de las metáforas del lenguaje, debido a la consideración de éstas como manifestación de la mayor parte de los fenómenos cognitivos subyacentes. Se expone también la confluencia de los experimentos y teorías en distintas disciplinas, que a juicio de Lakoff proporcionan la evidencia suficiente para sostener el carácter científico de su teoría, y las implicaciones que ésta tendría para otras disciplinas, como la filosofía o las matemáticas. En la teoría de Gärdenfors se estudian los fenómenos cognitivos atendiendo a su supuesta representación topológica en el cerebro.

4.1 La teoría de la mente corporeizada de Lakoff y Johnson

4.1.1 Introducción

En el libro *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, George Lakoff establece lo que considera una explicación nueva de los fenómenos cognitivos humanos, fundamentalmente a través del estudio de la formación de categorías. En el libro *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought* presenta junto a Mark Johnson

esta misma teoría, añadiendo sus implicaciones para la filosofía y las teorías neurológicas que la apoyan.

Al contrario que las explicaciones anteriores de la cognición, que presentarían a la razón como trascendental e independiente de nuestra experiencia corpórea, Lakoff sostiene que la razón está corporeizada o encarnada de raíz, y que esta corporeización se hace patente en los aspectos imaginativos de nuestra razón, principalmente en la metáfora, aspectos que son centrales en la razón humana. En palabras de Lakoff:

En la visión tradicional, la razón es abstracta y descorporeizada. En la nueva teoría, la razón tiene bases corporales. El punto de vista tradicional ve a la razón como literal, principalmente como proposiciones que pueden ser objetivamente tanto verdaderas como falsas. El punto de vista nuevo toma los aspectos imaginativos de la razón – metonimia, metáfora e imágenes mentales – como centrales a la razón, en lugar de como periféricos y como fenómenos sin relevancia unidos a lo literal.

La explicación tradicional mantiene que la capacidad de pensar y razonar de forma adecuada es abstracta y no está necesariamente corporeizada en ningún organismo. Por tanto, la racionalidad y los conceptos son *transcendentales*, en el sentido de que trascienden, o van más allá, de las limitaciones físicas de cualquier organismo. Puede ocurrir que los conceptos y la razón abstracta estén encarnados en seres humanos, o en máquinas, o en otros organismos, pero existen de forma abstracta, independientemente de cualquier encarnación. En la explicación nueva, el significado es una cuestión de lo que es significativo para los seres que piensan y actúan. La naturaleza del organismo pensante y su funcionamiento en el entorno son de una relevancia central en el estudio de la razón.

Ambos puntos de vista consideran la categorización como la forma principal mediante la que damos sentido a la experiencia. Las categorías en el enfoque tradicional están caracterizadas únicamente por las propiedades que comparten sus miembros. Es decir, están caracterizadas: a) independientemente de la naturaleza corporal de los

seres que realicen las categorizaciones y b) literalmente, sin mecanismos imaginativos (metáfora, metonimia e imágenes mentales) formando parte de la naturaleza de las categorías. En el enfoque nuevo, nuestra experiencia corporal y la forma en la que utilizamos nuestras capacidades imaginativas son parte central de cómo construimos las categorías para dar cuenta de la experiencia.¹³⁹

Según Lakoff, su teoría no está condicionada por un punto de vista filosófico previo, como ocurre con el cognitivismo, al que no considera basado en hechos experimentales, sino condicionado por la idea filosófica de una razón trascendental y de un objetivismo respecto a la relación con el mundo. Frente a esto, se construiría sobre la base de evidencias procedentes de distintas disciplinas y con el suficiente peso científico como para afirmar que su teoría explica cómo interactuamos con el mundo y cómo se construye el pensamiento. Nuevamente en palabras de Lakoff:

Llamaremos al enfoque tradicional *objetivismo* por la siguiente razón: los intentos modernos de hacer que funcione asumen que el pensamiento racional consiste en la manipulación de símbolos abstractos y que estos símbolos obtienen su significado por medio de una correspondencia con el mundo *construida objetivamente*, es decir, independiente de la forma de comprender de los organismos. Un conjunto de símbolos situados en correspondencia con un mundo estructurado objetivamente se entiende como una *representación* de la realidad. En el enfoque objetivista, *todo* el pensamiento racional implica la manipulación de símbolos abstractos a las que únicamente se da significado a través de correspondencias con las cosas en el mundo externo...

En años recientes, se han estudiado intensamente y en detalle las categorías conceptuales en algunas de las ciencias cognitivas, especialmente en antropología, lingüística y psicología. La evidencia que se ha acumulado está en conflicto con el enfoque objetivista de la mente. Las categorías conceptuales son, en conjunto, muy distintas de

¹³⁹ George Lakoff. *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, The University of Chicago Press, Chicago, 1987, xi-xii.

cómo indica el enfoque objetivista. La evidencia sugiere un enfoque muy diferente no sólo de las categorías, sino de la razón humana en general:

- El pensamiento está *corporeizado*, *encarnado*, esto es, las estructuras que se utilizan para integrar nuestros sistemas conceptuales surgen de nuestra experiencia corporal y tienen sentido en términos de ella, más aún, el núcleo de nuestros sistemas conceptuales está directamente basado en la percepción, el movimiento corporal, y la experiencia de lo físico y lo social.
- El pensamiento es *imaginativo*, en cuanto que aquellos conceptos que no están basados directamente en la experiencia emplean la metáfora, la metonimia y las imágenes mentales, todas las cuales van más allá de la reproducción literal, o *representación*, de la realidad externa. Es esta capacidad imaginativa la que posibilita el pensamiento “abstracto” y la que lleva a la mente más allá de lo que podemos ver y sentir. La capacidad imaginativa está también *corporeizada* – indirectamente – ya que las metáforas, metonimias e imágenes están basadas en la experiencia, a menudo experiencia corporal. El pensamiento es también imaginativo en una forma menos obvia: cada vez que categorizamos algo de forma que no refleje la naturaleza, estamos utilizando las capacidades imaginativas humanas generales.
- El pensamiento tiene una *estructura ecológica*, la categorización es una consecuencia de nuestra estructura corporal. Hemos evolucionado para categorizar y si no lo hubiésemos hecho no habríamos sobrevivido.

Me referiré al nuevo enfoque como *realismo experiencial* o *experiencialismo*. El término *realismo experiencial* hace énfasis en lo que el experiencialismo comparte con el objetivismo: (a) un compromiso con la existencia de un mundo real, (b) el reconocimiento de que la realidad impone restricciones en los conceptos, (c) una concepción de la verdad que va más allá de la mera coherencia

interna, (d) un compromiso con la existencia de un conocimiento estable del mundo.¹⁴⁰

Lakoff afirma, por tanto, que su teoría es científicamente válida, y que, una vez presentadas las evidencias científicas sobre las que descansa, el estudio de las metáforas de la lengua constituye una de las mejores formas de acceder al sistema conceptual humano y su imbricación con la experiencia. Esto tiene consecuencias para todas las ramas del saber humano, incluida la filosofía, como debe ser en una teoría que intente dar una explicación “científica”, y lo más completa posible, de la cognición. Por tanto, las teorías en las distintas disciplinas deberían ser consistentes con la teoría cognitiva de Lakoff. En el caso de la filosofía, Lakoff reclama la necesidad de “una filosofía nueva, empíricamente responsable, una filosofía consistente con los descubrimientos empíricos sobre la naturaleza de la mente”¹⁴¹.

4.1.2 Fundamentos científicos e interdisciplinares

Según Lakoff y Johnson, existe un número suficiente de “evidencias convergentes” en distintas disciplinas que soportan su teoría según la cual nuestros conceptos están basados en la experiencia y nuestra forma principal de conceptualizar es a través de metáforas. Estas “evidencias convergentes”, que a continuación se presentan, proporcionarían a su teoría el necesario carácter científico:

1.- Estudios sobre el color. La teoría interaccional del color muestra que nuestra experiencia del color se produce por una combinación de cuatro factores: las longitudes de onda de la luz reflejada, las condiciones de iluminación, los tres tipos de conos en

¹⁴⁰ George Lakoff, *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, xii-xv. George Lakoff y Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, Basic Books, Nueva York, 1999, 18.

¹⁴¹ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 15.

nuestras retinas, que absorben la luz de longitud de onda larga, media y corta, y la compleja circuitería neuronal conectada a los conos.

El color no es sólo la percepción de la longitud de onda. Su constancia depende de la habilidad del cerebro para compensar las variaciones de luz. Se podría pensar que el color es una representación interna de la realidad externa consistente en las propiedades reflectantes de la superficie de un objeto. Si esto fuese verdad, las propiedades de los colores y las categorías de colores serían representaciones de los reflejos y categorías de los reflejos. Pero esto no es verdad. Los conceptos del color tienen estructura interna, con algunos colores “focales”. La categoría *rojo*, por ejemplo, contiene tanto al rojo central como a tonos no centrales como el rojo púrpura, rojo rosáceo, y rojo anaranjado. La estructura de centro-periferia de nuestras categorías es el resultado de las curvas de respuesta neuronal de nuestro cerebro para el color. Los tonos focales corresponden a frecuencias de respuesta neuronal máxima. La estructura interna de las categorías del color no está ahí fuera en los reflejos de las superficies. El color no es simplemente la representación interna de los reflejos externos. Y no es una cosa o sustancia que esté ahí fuera en el mundo.

Por tanto, nuestros conceptos sobre el color, sus estructuras internas, y la relación entre ellos, están inextricablemente unidos a nuestro cuerpo. Los colores no son objetivos ni puramente subjetivos, no son ni una quimera de nuestra imaginación ni una creación espontánea de nuestro cerebro.¹⁴²

2.- Categorías de nivel básico. Según Lakoff y Johnson, los estudios de Brent Berlin, Eleanor Rosch, y Carolyn Mervis muestran que hemos evolucionado para formar al menos una clase de categorías que se ajusta óptimamente a nuestras experiencias corporales de las cosas y a ciertas diferencias extremadamente importantes de nuestro entorno natural. Estas categorías se denominan *categorías de nivel*

¹⁴² George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 23-25.

básico. En el mundo natural, las categorías de nivel básico se corresponden con los géneros biológicos.

Las categorías de nivel básico se caracterizarían por las siguientes condiciones:

Es el nivel más alto en el que se puede representar a toda la categoría mediante una sola imagen mental. Por ejemplo, se puede tener una imagen mental de un coche, pero no de un vehículo.

Es el nivel más alto en el que los miembros de la categoría tienen formas generales que se perciben de la misma manera. Por ejemplo, se puede reconocer a un coche por su forma y clasificarlo como tal, pero no a un vehículo.

Es el nivel en el que se organiza la mayoría de nuestro conocimiento. En general, a no ser que seamos expertos, sabemos mucho más, por ejemplo, de coches que de vehículos.¹⁴³

Es el nivel que parece correlacionarse más directamente con acciones no lingüísticas en el desarrollo de la persona. Brown lo explica con el siguiente ejemplo:

Las flores se marcan con acciones olfativas, pero no hay acciones que distingan unas especies de flores de otras. Los primeros nombres que se dan a las cosas caen en el nivel de las acciones distintivas pero se prosigue nombrando para codificar el mundo en todos los niveles. Las acciones no lingüísticas no prosiguen.¹⁴⁴

Como resultado de estas características, el nivel básico tiene otras prioridades sobre los niveles superiores e inferiores: los niños lo

¹⁴³ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 26-30.

¹⁴⁴ Roger Brown, *Social Psychology*, Free Press, New York, 1965, 320. En George Lakoff, *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, 31.

nombran y entienden antes y es el primero en introducirse en el lenguaje. Se dice que este nivel es cognitivamente “básico”, en el sentido de que tiene una prioridad cognitiva en la comprensión de nuestra experiencia, mientras que los otros niveles se crearían por procesos conceptuales imaginativos que en última instancia conectarían con el nivel básico.

3.- Conceptos de relaciones espaciales. Organizamos nuestros conceptos espaciales en términos de nuestra interacción corporal con el mundo. Por ejemplo, nuestro conocimiento más básico del movimiento está basado en un esquema conceptual de fuente-recorrido-meta, directamente ligado a nuestra experiencia corporal, y los conceptos y proposiciones que utilizamos para situar los objetos y los conceptos están basados en el esquema del recipiente, donde los objetos pueden estar *en, sobre, bajo*, etc., proposiciones que expresan sensaciones ligadas a nuestra experiencia corporal que divide el espacio, estableciendo una parte interior, una frontera, y una parte exterior.¹⁴⁵

4.- Mezcla de dominios en el aprendizaje antes del uso metafórico. Los niños pequeños no diferencian sino que mezclan las experiencias subjetivas y las corporales cuando ocurren juntas. Por ejemplo, para un bebé, la experiencia subjetiva de afecto se fusiona muchas veces con la experiencia sensorial de calor al ser sostenido en brazos. Más adelante, el niño aprenderá a separar los dominios, pero el cruce de asociaciones entre ambos permanecerá en forma de, por ejemplo, la metáfora “una sonrisa cálida”. De esta forma es cómo surgen las metáforas *primarias* o directamente ligadas a la experiencia, que surgen automática e inconscientemente.

5.- Bases neurológicas en la formación de conceptos. Lakoff y Johnson reconocen que no se tiene hasta el momento ninguna evidencia neurofisiológica lo suficientemente fuerte como para

¹⁴⁵ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 30-34.

afirmar que se usan los mismos mecanismos neuronales en la percepción y el movimiento que en el razonamiento abstracto, pero sí afirman que hay buenas razones para afirmar que esto es plausible, como mostrarían las teorías de Regier, Bailey y Narayanan.

Narayanan propuso un modelo conexionista en el que descubrió que la misma estructura neuronal que puede ejercer el control motor es la que caracteriza la estructura conceptual lingüística, y que los mismos mecanismos neuronales que pueden controlar los movimientos corporales pueden realizar inferencias lógicas sobre la estructura de las acciones en general. También propuso la teoría de que las “asociaciones” que se crean durante el periodo del desarrollo del niño previo al surgimiento de las metáforas, se realizan neuronalmente como activaciones que producen conexiones permanentes que definen dominios conceptuales. Estas conexiones formarían las bases anatómicas de las activaciones de una fuente a un destino que constituyen las metáforas. Regier propuso un modelo conexionista que reconociese y clasificase objetos espaciales, lo que daría una idea de cómo las estructuras neuronales del cerebro que realizan trabajos de percepción también pueden trabajar en la formación de conceptos. Por último, Bailey desarrolló un modelo neuronal que reconocía perceptualmente los movimientos de la mano a la vez que categorizaba dichos movimientos con distintos nombres.¹⁴⁶

¹⁴⁶ Cf. George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 38-42. Las obras de los autores citados en las que se exponen estas teorías son: S. Narayanan, *Embodiment in Language Understanding: Sensory-Motor Representations for Metaphoric Reasoning About Event Descriptions*, Ph.D. Dissertation, Department of Computer Science, University of California, Berkeley, 1997 y “Talking the Talk Is Like Walking the Walk: A Computational Model of Verbal Aspect”, en M.G. Shafto y P. Langley, eds., *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, Erlbaum, Mahwah, N.J, 1997. T. Regier, *The Human Semantic Potential: Spatial Language and Constrained Connectionism*, The MIT Press, Cambridge, Mass., 1996. D. Bailey, *A Computational Model of Embodiment in the Acquisition of Action Verbs*, Ph.D. Dissertation, Computer Science Division, EECS Department, University of California, Berkeley, 1997.

4.1.3 La teoría de la mente corporeizada

Sobre las bases expuestas en la sección anterior, Lakoff y Johnson desarrollan su teoría cognitiva de la mente corporeizada, fundamentalmente a través de la teoría cognitiva de la metáfora. Según esta teoría “la esencia de la metáfora es comprender y experimentar una clase de cosas en términos de otra”¹⁴⁷, y éste es el procedimiento cognitivo principal en los seres humanos. La metáfora consiste en la proyección de una clase de cosas o experiencias, desde un dominio “fuente”, a un dominio “destino” o clase diferente de experiencias, para entender estas últimas en términos de las primeras.

Dentro de las metáforas hay dos tipos: una metáfora “primaria” o “atómica”, que es la formada directa e inconscientemente en el curso del desarrollo del niño por contacto con el mundo. Esta metáfora surgiría espontáneamente en el sistema neuronal como asociación de una experiencia, principalmente sensoriomotora, con una experiencia subjetiva a través de la interconexión de áreas del cerebro dedicadas a ambos tipos de experiencia. Ejemplos de esta metáfora podrían ser: “mañana es un *gran* día”, “los precios *suben*”, “¿están los tomates *en* la categoría de frutas o de vegetales? , “¿Cómo *encajan* las *piezas* de esta teoría?, etc. El segundo tipo de metáfora sería una metáfora “compleja”, construida a partir de las metáforas primarias y formas de conocimiento común, como modelos culturales o conocimientos y creencias ampliamente aceptados en una cultura. Por ejemplo, señalan Lakoff y Johnson, en nuestra cultura hay un sentimiento y creencia en que se debe tener un propósito en la vida, una “meta que alcanzar”, y que si no lo tienes estás “perdido” sin “rumbo” y no sabes “qué camino seguir”. Todas estas frases son parte de la metáfora *una vida*

¹⁴⁷ George Lakoff and Mark Johnson, *Metáforas de la vida cotidiana*, Ediciones Cátedra, 1986, 5. Original en inglés, *Metaphors we Live by*, University of Chicago Press, Chicago, 1980.

con un objetivo es un viaje, que es una metáfora compleja formada a partir de las metáforas primarias *los propósitos son destinos* y *las acciones son movimientos*, junto con la creencia en la necesidad de tener un objetivo en la vida y el hecho de que un viaje es una serie de movimientos para llegar a un destino.¹⁴⁸ De esta forma se construirían los conceptos, creando metáforas complejas a partir de metáforas básicas. El mecanismo utilizado en esta construcción sería el de la *fusión conceptual* de dominios según la teoría de Fauconnier y Turner. La teoría de la fusión conceptual establece que la capacidad imaginativa de las personas hace que se puedan activar a la vez distintos dominios conceptuales o “espacios mentales”¹⁴⁹. Esta activación simultánea haría que en ocasiones se formaran conexiones nuevas entre ellos, dando lugar a inferencias y dominios nuevos.

Lakoff y Johnson aceptan que no todo el pensamiento es metafórico, por ejemplo, la frase “estos colores son similares” es literal, mientras que “estos colores están bastante próximos” utiliza la metáfora *similitud es proximidad* . Este pensamiento “literal” podría ser adecuado en algunas circunstancias, pero en general carecería de la riqueza del pensamiento metafórico. La metáfora sitúa el significado en la experiencia, más allá del lenguaje, al evocar un mundo de sensaciones no expresable literalmente, de ahí su contenido cognitivo y su riqueza. Como señala Bustos, citando a Aristóteles:

“Lo mejor del mundo, con mucho, es ser un maestro de la metáfora” (Poética, 1.459a, 5-7). La base de su consideración era doble: por una parte, el dominio de la metáfora “es lo único que no se puede aprender de los demás, y es también la impronta del genio” (1.459a), porque no se basa en mecanismos que se puedan aplicar de forma regular y precisa. En este sentido, la metáfora no es “lógica”, sino que implica el ejercicio de la imaginación y de la sensibilidad. Por otro lado, “una buena metáfora encierra una percepción intuitiva de las semejanzas en

¹⁴⁸ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 60-61.

¹⁴⁹ Cf. Gilles Fauconnier y Mark Turner, *The Way we Think. Conceptual Blending and the Mind's Hidden Complexities*, Basic Books, Nueva York, 2002, 40-50.

las cosas que no son similares” (1.459a): la imaginación y la sensibilidad no se aplican en el vacío, no elaboran fantasmas arbitrarios; constituyen un instrumento necesario en la adquisición de conocimiento en la medida en que ayudan a penetrar la estructura de la realidad. La metáfora se encuentra en la base de nuestras abstracciones, posibilitando la elaboración de nuevos conceptos, que abarcan nuevas realidades... En la expansión del conocimiento, la metáfora desempeña un papel fundamental. Permite captar la estructura de lo desconocido en virtud de lo ya conocido, manifestando la homogeneidad oculta de la realidad... La metáfora es una integración, entre algo con lo que ya se está familiarizado y algo que hasta entonces es extraño, pero que la metáfora nos lleva a percibir como propio.¹⁵⁰

Además, las metáforas primarias, y otras metáforas cuyo uso durante largo tiempo ha hecho que pasaran a ser norma, forman parte inevitablemente de nuestro sistema conceptual, por lo que las utilizamos inconscientemente, aun cuando pensemos que estamos razonando de una forma literal y objetiva. Por tanto, utilizamos las metáforas en la vida corriente y en todas las disciplinas, desde las ciencias hasta la filosofía, la política o la literatura, de forma más o menos consciente. Por tanto, para comprender una teoría o un concepto habría que estudiar inevitablemente las metáforas que utiliza, fundamentalmente en el sentido de ver cómo están ligadas con metáforas primarias y otros fenómenos cognitivos básicos, privilegiados, que surgen de nuestra interacción corporal con el mundo.

Lakoff ha presentado, junto a otros autores, numerosos estudios de la aplicabilidad de la teoría de la metáfora en distintos campos. Como ejemplo en matemáticas, ofrece una explicación de la aritmética de los números naturales a través de sus metáforas primarias constitutivas como la teoría más cercana a experiencias

¹⁵⁰ Eduardo de Bustos, *La metáfora. Ensayos transdisciplinares*, Fondo de Cultura Económica de España, Universidad Nacional de Educación a Distancia, Madrid, 2000, 48-49.

corporales¹⁵¹. Esto podría explicar por qué Hilbert, en su intento de formalización de sistemas lógicos, eligiera la aritmética de los números naturales como un sistema que surge de la experiencia inmediata y que no es reducible a ningún otro. Como aplicación en política, Lakoff explica cómo, según él, ciertos valores culturales arraigados y metaforizados de forma convencional en la sociedad norteamericana se encuentran incluidos en el mensaje del partido republicano pero no entre los liberales.¹⁵² En literatura, Lakoff y Turner dan una explicación de la poesía como la combinación imaginativa de metáforas básicas para crear otras más complejas.¹⁵³ En filosofía, analizan las metáforas subyacentes a los conceptos de espacio, tiempo, y los tópicos más comunes y corrientes en filosofía a lo largo de la historia.¹⁵⁴

Una fuente de objeciones a la validez de la teoría cognitiva de la metáfora como explicación de los principales mecanismos de pensamiento está relacionada con la insuficiencia de las teorías sobre las que se apoya a la hora de explicar todo el pensamiento, y con su más que dudoso carácter científico. Respecto a la primera, Bustos presenta cómo algunos psicólogos han reclamado la necesidad de complementar la teoría de separación de dominios abstractos a partir de experiencias sensoriales básicas en el bebé, con una teoría de formación de sistemas de reglas:

Aunque escépticos respecto al papel central de la metáfora y la analogía en la constitución del conocimiento abstracto, la principal cuestión que algunos psicólogos han planteado es la necesidad de dotar a la semántica cognitiva de una dimensión evolutiva real. De

¹⁵¹ George Lakoff y Rafael Núñez. *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics into Being*, Basic Books, Nueva York, 2001, 50-103.

¹⁵² George Lakoff, *Moral Politics: What Conservatives Know that Liberals Don't*, University of Chicago Press, Chicago, 1996.

¹⁵³ George Lakoff y Mark Turner, *More than Cool Reason: A Field Guide to Poetic Metaphor*, University of Chicago Press, Chicago, 1989.

¹⁵⁴ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*.

acuerdo con su conjetura, y la naturaleza de ciertos datos iniciales, entre los que se encuentran los aportados por las propias investigaciones de Piaget, es preciso matizar ciertas tesis radicales de la semántica cognitiva. Entre ellas, la de que la constitución de sistemas formales para gestionar el conocimiento es una derivación de sistemas más básicos de representación del significado, de su *encarnación (embodiment)* en imágenes esquemáticas. Según sugieren, es más probable que los sistemas formales de reglas se desarrollen y evolucionen al mismo tiempo y de forma independiente de los esquemas imaginativos y que desempeñen un papel importante en el almacenamiento, producción y conservación del conocimiento, tal como postula la teoría computatoria del significado.¹⁵⁵

Cintia Rodríguez y Cristiane Moro han señalado la necesidad de introducir el signo como portador de significado, y, a partir de la interacción del bebé con los objetos del medio, de sus experiencias comunicativas, y de su desarrollo cognitivo, han situado ontogénicamente el significado en medio del mundo y de la comunidad:

Hemos tratado de despertar al objeto del “impasse” en que se halla, gracias a su inclusión dentro de la pragmática: los objetos también se usan, sirven para hacer cosas en el mundo. Como consecuencia, no pueden quedarse al margen ni de la comunicación ni del análisis semiótico. Como consecuencia, y como la comunicación circula antes de que el niño posea el lenguaje, básicamente por el territorio de los signos no lingüísticos, hemos tenido que revitalizarlos y unirlos a los usos de los objetos, que es la manera en que en muchas ocasiones cobran vida. Por fin, que el signo pueda concebirse como inferencia abre la puerta hacia el pensamiento de manera unitaria con lo que ocurre en el terreno de la comunicación con los otros y con el objeto. El resultado de todo ello es que ni el objeto puede quedarse al margen del análisis semiótico, ni lo cognitivo puede seguir encapsulado en el sujeto solipsista, al margen de los procesos comunicativos, que es el lugar donde los signos viven.¹⁵⁶

¹⁵⁵ Eduardo de Bustos. *La metáfora. Ensayos transdisciplinares*, 200.

¹⁵⁶ Cintia Rodríguez y Christiane Moro, *El mágico número tres. Cuando los niños aún no hablan*, Paidós Ibérica, Barcelona, 1999, 262-263.

El punto de vista semiótico, además de integrar experiencia con desarrollo simbólico, apunta a la difícil cuestión de cómo se podría integrar la perspectiva individual y la social en la determinación del significado. Estas perspectivas, como señala Bustos, se hallan en tensión a lo largo de las distintas teorías en Ciencia Cognitiva:

Esta tensión se refleja, por ejemplo, en la diferente orientación de las obras de Piaget y Vygotsky, que no se puede reducir a una discrepancia en el *énfasis* de sus teorías, en lo individual la de Piaget y en lo social la de Vygotsky. En definitiva, se trata de la cuestión de la *fuerza* del significado, el proceso mediante el cual se constituye y el *locus* donde se debe analizar. Buena parte de la filosofía contemporánea de la mente o de la ciencia cognitiva es, en este aspecto, de tradición individualista, esto es, concibe el significado como, o en relación con, un fenómeno mental individual. La comunicación a su vez se conceptualiza como el traslado de representaciones mentales entre individuos... En cambio, un enfoque wittgensteniano, en sentido general, abogaría por el abandono de esta perspectiva individualista, afirmando la naturaleza *social* del significado y la comunicación, de acuerdo con sus acostumbrados pilares conceptuales: el carácter no intrínseco del significado, el papel constitutivo de la coordinación de las acciones, la función reguladora del auditorio, la naturaleza fluida del significado, etc. Pero la utilidad de tal enfoque, más que extraer conclusiones o formular sugerencias en esa línea, sería la de mantener un equilibrio entre lo individual/social, esto es, el punto de lo *relacional*, que es donde se aloja la fuente de la estructuración de lo psicológico individual y lo social-lingüístico.¹⁵⁷

Respecto al carácter científico de los fundamentos de la teoría cognitiva de la metáfora, conviene señalar cómo, por ejemplo, las bases neurológicas sobre formación de conceptos en las que se apoya están a su vez tomadas de un modelo conexionista construido artificialmente. Ya se ha mostrado anteriormente (sección 3.4) que en

¹⁵⁷ Eduardo de Bustos, *La metáfora. Ensayos transdisciplinarios*, 191-193.

estos modelos se introducen comportamientos que en bastantes aspectos no parecen corresponderse con el funcionamiento del cerebro, pero que en los modelos resultan útiles para conseguir su estabilidad y que “funcionen”, por lo que al menos en este punto queda entre paréntesis su validez neurológica. Además, también se ha indicado (sección 2.4) cómo en opinión de algunos autores, las mismas técnicas de imágenes neuronales están sujetas a crítica a la hora de fijar resultados generales de la cognición, por su carácter artificial y la introducción de carga teórica en ellos. Otros puntos de sus fundamentos, como la hipótesis de la existencia de un nivel básico de comprensión, están basados en experimentos psicológicos que hacen referencia en su definición a nociones un tanto vagas, como la de imagen mental.

La noción misma de metáfora parece escapar en ocasiones a la interpretación cognitiva de Lakoff y Johnson. Según Bustos, no siempre se utiliza la metáfora con un valor cognitivo, en el sentido de conceptualizar una realidad en términos de otra, sino que su uso es más amplio, incluyendo una forma de percibir la realidad y de compartirla:

La poesía, tanto la popular como la culta, no sólo desarrolla, sino que también hace una profusa utilización de imágenes metafóricas, de representaciones sensitivas que entrañan nuevas formas de percibir una realidad, no de conceptualizarla. Para decirlo con Octavio Paz: “El poeta nombra las cosas: éstas son plumas, aquéllas son piedras. Y de pronto afirma: las piedras son plumas, esto es aquello. Los elementos de la imagen no pierden su carácter concreto y singular: las piedras siguen siendo piedras, ásperas, duras, impenetrables, amarillas de sol o verdes de musgo: piedras pesadas. Y las plumas, plumas: ligeras”.¹⁵⁸

Cooper recuerda que la metáfora ha sido utilizada, en la literatura y el arte modernos, con propósitos estrictamente opuestos a los cognitivos, esto es, con la intención de diluir el sentido de realidad,

¹⁵⁸ Eduardo de Bustos, *La metáfora. Ensayos transdisciplinarios*, 282.

para difuminar la dualidad texto/mundo, como un medio para exorcizar la opacidad del propio lenguaje.¹⁵⁹

Dentro del ámbito de las relaciones humanas y personales, la función “retórica” de la metáfora que con más detenimiento se ha analizado es la del cultivo de la intimidad. Intimidad en el sentido de experiencia exclusivamente vivida, pero compartida; intimidad pues que no afecta sólo al individuo, sino a su relación con otro u otros con los que puede “formar sociedad” para propósitos socio-comunicativos... La metáfora puede verse así como una especie de invitación por parte de quien la produce, una sugerencia tácita a participar en un juego de agudeza y penetración que, como una danza, requiere la armonía de los “movimientos” de los participantes.¹⁶⁰

Tomando las bases de la teoría cognitiva de la metáfora como plausibles, como ciertamente pueden ser, junto con los numerosos ejemplos de metáforas que presentan Lakoff y Johnson, se puede aceptar que hay una “evidencia convergente” del carácter corpóreo del pensamiento. Pero dado el carácter hipotético en muchos casos de esas bases, junto con las pocas explicaciones respecto a cómo se produce la formación de conceptos nuevos en comparación con los numerosos ejemplos que dan los autores de cómo estos se expresan en términos metafóricos “corpóreos”, también podríamos obtener una “evidencia convergente” de que estamos ante una teoría científicamente tentativa que no explica toda la cognición, si es que toda se puede explicar desde la ciencia. La teoría cognitiva de la mente corpórea privilegia el carácter corpóreo del pensamiento frente al carácter abstracto, al menos en su dimensión *explicativa*. En esta teoría se dan numerosos ejemplos de cómo la razón está corporeizada, ligada a nuestro sistema neuronal y a nuestros sistemas psicológicos básicos, y cómo todo esto se refleja en el lenguaje. Por así decirlo, es una teoría que *trae* los conceptos a nuestras experiencias sensoriomotoras, y que sitúa en éstas el fenómeno de la comprensión. Cabe preguntarse cuál es el factor que hace que construyamos conceptos abstractos sobre esta

¹⁵⁹ Eduardo de Bustos, *La metáfora. Ensayos transdisciplinarios*, 297.

¹⁶⁰ Eduardo de Bustos, *La metáfora. Ensayos transdisciplinarios*, 298-299.

base corpórea, falta decir algo sobre cómo nuestros conceptos *escapan, van más allá* de nuestras experiencias, aunque acaben expresándose en términos de éstas.

4.1.4 Implicaciones filosóficas

En su interacción corporal con el mundo, se produce en los hombres una categorización primaria muy ligada a la experiencia, que es la que se correspondería según Lakoff y Johnson con el llamado “nivel básico”. Estas categorizaciones estarían determinadas únicamente por nuestra estructura neuronal y nuestra interacción con el mundo. La mayoría de las categorías se formarían inconscientemente de este modo. Las categorías más abstractas se formarían por proyecciones metafóricas a partir de las experiencias del nivel básico. A esta forma de categorizar o conceptualizar hay que añadir una explicación de por qué formamos unas categorías y no otras. Lakoff y Johnson dan una explicación evolucionista:

La categorización es por tanto una consecuencia de nuestra encarnación. Hemos evolucionado para categorizar; si no lo hubiésemos hecho, no habríamos sobrevivido. La categorización no es, en gran parte, un producto de nuestro razonamiento consciente. Categorizamos en la forma que lo hacemos porque tenemos el cerebro y el cuerpo que tenemos y porque interactuamos con el mundo de una determinada manera... Lo importante no es simplemente que tenemos cuerpo y que el pensamiento está de alguna forma corporeizado. Lo importante es que la peculiar naturaleza de nuestro cuerpo moldea las mismas posibilidades de conceptualización y categorización... Un concepto corporeizado es una estructura neural que es parte de, o que usa, el sistema sensoriomotor de nuestro cerebro.¹⁶¹

¹⁶¹ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 18-20.

Por tanto, lo únicamente “real” sería nuestra interacción con el mundo, guiada por un motivo evolucionista de supervivencia. Sobre la base de este motor evolutivo atribuiríamos realidad a nuestros conceptos o experiencias:

Lo que queremos decir con “real” es lo que necesitamos establecer conceptualmente para ser realistas, es decir, para funcionar con éxito para sobrevivir, para conseguir metas, y llegar a una comprensión manejable de las situaciones en las que nos encontramos.¹⁶²

A partir de aquí, lo que es “verdad” se hace depender de esta interacción con el mundo y de nuestra forma corporeizada y evolutiva de comprender. También se rechaza la dualidad objetivo/subjetivo y se sustituye a ambas por un “sistema” interactivo-adaptativo hombre/mundo.

Como se ha visto antes, Lakoff y Johnson denominan a su postura “realismo corporeizado” o “experiencialismo”, y lo distinguen del relativismo:

Como el realismo corporeizado niega, sobre bases empíricas, que exista una y solo una forma correcta de describir el mundo, puede parecer a algunos que es una forma de relativismo. Sin embargo, aunque trata el conocimiento como relativo – relativo a la naturaleza de nuestros cuerpos, cerebros e interacciones con nuestro entorno – no es una forma de relativismo extremo, porque explica cómo el conocimiento estable y real es posible tanto en ciencia como en nuestras interacciones diarias en el mundo. Esta explicación tiene dos aspectos. Primero, están los conceptos corporeizados directamente, como los conceptos de nivel básico, los conceptos de relaciones espaciales, y los conceptos que estructuran hechos espacialmente. Estos conceptos tienen un origen evolutivo y nos permiten funcionar de forma extremadamente exitosa en nuestras interacciones diarias con el mundo. También forman la base de nuestro conocimiento científico estable.

¹⁶² George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 109.

Segundo, las metáforas primarias hacen posible la extensión de estos conceptos corporeizados a dominios teóricos abstractos. Las metáforas primarias son cualquier cosa menos construcciones sociales arbitrarias, ya que están altamente determinadas tanto por la naturaleza de nuestro cuerpo y cerebro como por la realidad de nuestras interacciones diarias.

El realismo corporeizado, sin embargo, reconoce un aspecto central en el pensamiento relativista, que es que en muchos casos importantes, los conceptos *sí* que cambian con el tiempo, varían entre culturas, tienen múltiples estructuras inconsistentes, y reflejan condiciones sociales. El realismo corporeizado también proporciona mecanismos para caracterizar estos cambios, variaciones, multiplicidades, y ejemplos de “construcciones sociales”. La formación de metáforas complejas y otras fusiones conceptuales parece ser el principal mecanismo.¹⁶³

Lakoff y Johnson intentan aclarar su “evolucionismo”. Explican que no se trata del concepto equivocado que nos hemos formado comúnmente del evolucionismo darwinista como la metáfora “la evolución es la supervivencia del mejor competidor”, junto con la metáfora “la evolución es el cambio en la naturaleza” y la creencia popular de que “la evolución produce el mejor resultado”. La combinación de las tres ha producido la idea de que el cambio natural es la supervivencia del mejor competidor, lo que produce el mejor resultado. En general afirman que no se puede definir la evolución en términos sociales o morales ya que es un proceso natural. Queda, por tanto, la evolución como adaptación al medio, como supervivencia biológica de la especie. Pero ya que la teoría de Lakoff y Johnson debe ser “científica”, deben existir las suficientes “evidencias convergentes” como para asignar a la evolución, a la adaptación al medio, el estatus ontológico que ellos le asignan. Seguramente existan estas evidencias convergentes, como probablemente existan otras que converjan en que el evolucionismo como adaptación al mundo no es el

¹⁶³ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 96-97.

único motor de nuestro pensamiento. En cualquier caso, Lakoff y Johnson no presentan estas evidencias más o menos científicas, al contrario de cómo hacen al presentar las bases de su teoría corporeizada de la cognición.

Cabe preguntarse cómo se hacen compatibles en su teoría el compromiso con la existencia de un conocimiento estable del mundo y con un concepto de verdad que va más allá de la mera coherencia interna. Parece que al reconocer a la evolución, a la adaptación al medio, como lo “objetivo”, entonces las teorías son más o menos objetivas y realistas cuanto mejor nos permitan desenvolvernos en el mundo, sirviendo esto tanto para las teorías científicas como para el uso más común de los conceptos. Pero esta pauta de “adaptación” no hay que buscarla en el pensamiento de un solo individuo, sino que es la comunidad la garante de que se tienda hacia esas “verdades”. Podría decirse que la cognición es un fenómeno individual con un “ajuste evolutivo” social. Esto, que no sin dificultades, nos podríamos imaginar cómo funcionaría en las ciencias “duras”, al contrastar las teorías con los resultados experimentales, resulta todo un embrollo en ciencias sociales o en moral, donde lo que se considere de acuerdo a una “adaptación” al medio está por definir. En la teoría de Lakoff y Johnson surge una moral dependiente únicamente de nuestras experiencias corporales. Lakoff reconoce que no todo es metafórico en moral, pero afirma que la mayoría de nuestras posiciones en moral conectan con experiencias sobre bienestar corporal y algunos *derechos* fundamentales, como el de la propiedad privada.

Igualmente refieren la espiritualidad de la persona a su interacción corporal con el mundo, aunque no explican cómo se produce:

Si no existe la mente descorporeizada, entonces ¿cuál es el lugar en el que se da la experiencia espiritual real que la gente tiene en distintas culturas del mundo? Esta experiencia es corporeizada. Debe ser una consecuencia de lo que esté ocurriendo en nuestros cuerpos y nuestros cerebros. Cómo exactamente el cuerpo y el cerebro hacen que surja la

experiencia espiritual es una cuestión empírica para la Ciencia Cognitiva que está mucho más allá del ámbito de este libro.¹⁶⁴

Lakoff y Johnson exponen una espiritualidad que a su juicio tenga en cuenta esta interacción del hombre con el mundo:

Una espiritualidad corporeizada requiere una actitud estética ante el mundo que es central al cuidado de sí mismo, al cuidado de los demás, y al cuidado del mundo. La espiritualidad corporeizada requiere comprender que la naturaleza no es inanimada e inferior a lo humano, sino animada y más que humana. Requiere placer y diversión en la conexión corporal con el aire y la tierra, con el mar y el cielo, con las plantas y los animales – y el reconocimiento de que todos son más que humanas, más que lo que cualquier humano podría conseguir jamás. La espiritualidad corporeizada es más que una *experiencia* espiritual. Es una relación ética con el mundo físico.¹⁶⁵

En cuanto a su idea de persona, su concepción viene dada por la tensión en su teoría entre la parte bien explicada del carácter corpóreo de la cognición, y la parte sólo mencionada de su carácter no corpóreo. Para explicar más su posición, sobre todo, en mi opinión, para salvar el paso de lo cognitivo a lo neuronal, se alinean con una explicación de la cognición en distintos “niveles”: Nivel cognitivo, nivel neurocomputacional y nivel neurobiológico. La validez de cada nivel estaría justificada porque contribuiría en algo a la explicación de la cognición y su validez “ecológica o evolutiva” vendría dada porque habría “verdades” en cada nivel que no se podrían dar en los otros, por lo que los tres se hacen necesarios para “adaptarnos y manejarnos en este mundo en lo que refiere a la cognición”, con lo que se justificaría su validez científica. Así, se autodefinen como fisicalistas y no eliminacionistas:

¹⁶⁴ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 564.

¹⁶⁵ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 566.

En una teoría fisicalista eliminacionista, la explicación fluiría sólo en una dirección, de abajo arriba, tomando solo a la neurobiología como real y a los otros niveles como epifenómenos... Nuestro paradigma es fisicalista en el sentido de que no afirma que existan entidades místicas no físicas como el alma, el espíritu o la mente cartesiana descorporeizada. En último término, el cerebro es todo neuroquímica y neurofisiología. Pero es no eliminacionista de dos formas. Primero, se toma cada nivel como real, como teniendo una ontología teórica necesaria para explicar los fenómenos. Segundo, las explicaciones y las motivaciones fluyen en ambas direcciones.¹⁶⁶

Lo que es poco claro en su teoría es cómo se daría la interacción entre los distintos niveles. Además, parece que hay aspectos que escapan a esta explicación y que Lakoff y Johnson prefieren no abordar:

Las personas no son simplemente cerebros ni circuitos neuronales. Ni experiencias cualitativas y patrones de interacción corporal. Ni estructuras y operaciones del inconsciente cognitivo. Los tres (niveles) están presentes y mucho más que no discutimos aquí. Las explicaciones en los tres niveles son necesarias (aunque ciertamente insuficientes) para dar cuenta de forma adecuada de la mente humana.¹⁶⁷

Pero, en mi opinión, para que lo “cognitivo” influya en lo “neuronal” hace falta tener en cuenta a la persona como creador de conceptos, inducciones, creencias, etc. Sin una autonomía, una libertad que permita salirnos de la concepción de la persona como la “interacción de un sistema neuronal con un mundo” de donde saldrían todas nuestras actividades cognitivas, no puede haber la variabilidad que hay en las experiencias vitales y en la cognición humana.

¹⁶⁶ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 112-113.

¹⁶⁷ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 104.

Respecto a los conceptos más abstractos, Lakoff se limita a decir que el hombre tiene la capacidad de “conceptualizar”, por medio de su “imaginación”. La metáfora nos lleva a percibir algo que nos es extraño como algo propio, pero, ¿porqué algo nos es extraño? ¿cuándo y cómo se sorprende el hombre? ¿tiene esto una base neurológica? ¿quién conceptualiza, las personas o los sistemas neurológicos? ¿qué es la imaginación? Si para atar y entender lo que pensamos y sentimos lo traemos a la experiencia, igualmente y de forma natural nos dirigimos hacia lo desconocido. Tan fuerte y tan humana es la realidad de estar anclado en la experiencia, como la de soltar amarras hacia dominios nuevos, muchas veces con origen en la experiencia, por lo que es éste un viaje en ambos sentidos.

Una concepción completa de la persona debe integrar ambos aspectos y tratarlos como un todo, pero distinguiendo a la vez uno de otro. Tan radical y desatinado es el “pienso, luego existo” como sería el “existo (como organismo biológico), luego pienso (como persona)”. No hay contradicción, como quiere hacernos creer Lakoff al presentarlo como un dualismo, en sostener que la persona consta de alma y cuerpo como entes distintos aunque dependientes, y que el fenómeno cognitivo se encuentra influenciado por ambos. Al contrario, sólo de esta forma, y situando al Amor como principio motor que sustituya al evolucionismo, se puede entender una moral que considere a cada persona por lo que ella misma es, cuerpo y alma, más allá de consideraciones naturalistas y evolucionistas, aunque compatible con éstas. Lakoff presenta como incompatible con su teoría lo que hacen según él ciertas formas de cristianismo, aunque no se detiene a analizar otras (más acertadas y reconocidas) que serían compatibles y completarían su teoría:

Se supone que los cristianos deben vivir una vida santa enfocada en trascender todas las cosas de este mundo – deseos corporales, posesiones materiales, fama, éxito mundanal y una vida larga. Esta concepción ajena al mundo y descorporeizada de la espiritualidad y la transcendencia minusvalora la relación de uno con el mundo, con el entorno natural, y con todos los otros aspectos de la existencia

corporeizada. Se ordena a los cristianos que actúen moralmente hacia los demás y a ser buenos cuidadores de la tierra porque esto es lo que Dios quiere de ellos para que se salven, para que así puedan ir al cielo y estar unidos a Dios en una realidad más allá de este mundo. Por tanto, esta forma de cristianismo ata la moral a la desencarnación de mente y alma.¹⁶⁸

4.2 Los “espacios conceptuales” de Peter Gärdenfors

Peter Gärdenfors expone su teoría cognitiva en el libro *Conceptual Spaces. The Geometry of Thought*. En una línea situada entre el cognitivismo y el conexionismo, esta teoría adopta el punto de vista representacional según el cual conocer algo es tener una “representación” mental de ese algo que hay en el mundo, de forma que a lo que tenemos acceso es a la “representación” y no directamente al mundo.

Gärdenfors opina que la mayoría de las representaciones tienen una estructura geométrica, siendo esto así por la misma estructura topológica del cerebro, su modo de funcionamiento, su adaptación evolutiva y el carácter básico, en el sentido de nivel de comprensión directa de Lakoff y Johnson, de las percepciones espaciales.

Desde este punto de vista, presenta una modelización geométrica de algunos tipos de representaciones. A esta modelización geométrica la denomina “espacios conceptuales”, por ser, según él, especialmente apta para representar conceptos. Los espacios conceptuales tendrían la ventaja de dar cuenta de ciertos fenómenos cognitivos, como el razonamiento inductivo, que hasta ahora no han sido explicados satisfactoriamente por la otras dos corrientes (niveles, según Gardenfors) que tratan de explicar los fenómenos cognitivos: el simbolismo y el conexionismo.

¹⁶⁸ George Lakoff and Mark Johnson, *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, 564.

Gärdenfors entiende el simbolismo, los espacios conceptuales y el conexionismo como los tres niveles necesarios para abordar el fenómeno cognitivo. Estos niveles se distinguen, entre otras cosas, en su grado de granularidad de explicación. Sin embargo, a lo largo de todo el trabajo reclama que la verdadera comprensión, la inducción y la producción científica se producen en el nivel conceptual. Este nivel constituye la novedad principal en su teoría, al ser el integrador de los otros dos niveles y el lugar donde se dan los fenómenos cognitivos más relevantes.

Ya que los espacios conceptuales son una modelización geométrica de las representaciones mentales, es necesario explicar lo que es una estructura geométrica. Toda estructura geométrica viene dada por unas dimensiones y una métrica. Estos dos son los componentes básicos de los espacios conceptuales, y en la medida en que acertemos en determinarlos podremos hablar de éxito en la modelización de representaciones.

Se puede tomar como ejemplo el caso de los colores. A la hora de describir cómo percibimos los colores en nuestro desenvolvimiento ordinario, podemos hacerlo mediante tres dimensiones: tono, cromaticidad y brillo. Con estas tres variables (o dimensiones), y siendo el tono una variable polar y las otras dos variables lineales se genera una peonza que es el espacio conceptual de los colores. Con estas dimensiones se genera un espacio geométrico en el que cualquier color que veamos se puede situar en un punto del espacio. Son éstas tres, por tanto, las dimensiones relevantes.

Queda por explicar cómo se han elegido estas dimensiones, por qué una es polar y las otras lineales, y ver cómo se ha determinado la forma de la peonza resultante. Las respuestas a estas preguntas se dan mediante la realización de experimentos con personas basados en los juicios de similaridad que emiten éstas respecto, en este caso, distintos colores que se les presentan. Hay distintas técnicas para detectar las

dimensiones relevantes¹⁶⁹, y se basan fundamentalmente en la elección del número mínimo de dimensiones que hace que se puedan explicar las similitudes entre los objetos presentados, fijando el número de variables o dimensiones cuando al añadir una nueva dimensión no varía significativamente la exactitud en los grados de similitud emitidos por los sujetos. Respecto a por qué una dimensión es polar y las otras son lineales, también se detecta en el experimento, al situar gráficamente la variabilidad del tono y ver que se organiza en un círculo, mientras que la cromaticidad y el brillo lo hacen sobre una línea recta. La forma de la peonza (la métrica del espacio) queda determinada por la proximidad entre tono, brillo y cromaticidad, y se corresponde en el experimento con el grado de similitud entre colores detectado por el sujeto.

Como ya se ha expuesto, Gärdenfors considera que a lo que tenemos acceso “objetivo” es a nuestras representaciones, y no al mundo. Sin embargo, las dimensiones constitutivas de las representaciones se determinan mediante experimentos basados en la interacción con el mundo. ¿Cómo se resuelve esta aparente contradicción? Según Gärdenfors, y esta es la explicación subyacente en toda su teoría, la geometría de espacios conceptuales resultante es una de las muchas soluciones evolutivas que se puede dar al fenómeno en cuestión. Los humanos nos hemos ido adaptando al entorno y nuestras representaciones son las mejores “soluciones” que hemos encontrado para actuar en el entorno. Pero esto no significa que la estructura conceptual *represente* el mundo, en el sentido que sea una imagen fiel de él, sino que lo *presenta* adaptativamente. Esto estaría en concordancia con que las representaciones sean culturalmente dependientes. Por ejemplo, en nuestra cultura occidental el tiempo tiene forma lineal pero en otras culturas su forma es circular.

¹⁶⁹ Ver Peter Gärdenfors, *Conceptual Spaces. The Geometry of Thought*, The MIT Press, Cambridge, Massachusetts, 2000, 21-31, 142-150, 221-230, y Mikael Johansson: *Obtaining Psychologically Motivated Spaces with MDS*. En web: http://www.lucs.lu.se/ftp/pub/LUCS_Studies/Lucs87.pdf

Las dimensiones se pueden clasificar en integrales y separables. Una dimensión es integral cuando no se puede asignar a un objeto un valor en esa dimensión sin asignarle un valor en otra dimensión. Como ejemplo se puede poner el tono y el brillo en el espacio de los colores. Una dimensión es separable cuando se le puede dar un valor sin tener que hacerlo a otra dimensión. Un ejemplo es la talla y el color. La detección de dimensiones separables e integrales se hace mediante experimentos psicológicos. A su vez, la clasificación de las dimensiones no es absoluta. Por ejemplo, en mecánica newtoniana espacio y tiempo son separables pero en mecánica relativista no.

Una noción importante es la de dominio. Un dominio es un conjunto de dimensiones integrales que son separables de todas las demás (por ejemplo, tono, brillo y cromaticidad). Se puede entonces definir un espacio conceptual como una colección de uno o más dominios. Estos dominios pueden o no estar correlacionados. Un ejemplo de dominios correlacionados podría ser el de la madurez y el del color en el espacio de las frutas.

Gärdenfors presenta la forma simbólica, conceptual y subconceptual de representación como los tres niveles de granularidad de representación y argumenta que dependiendo del hecho cognitivo en cuestión, será necesaria una mayor o menor explicación en cada uno de los niveles, pero sólo el conjunto de los tres dará una explicación completa del hecho cognitivo. El nivel de mayor granularidad es el subconceptual o conexionista, que es el nivel más cercano a la interacción con el mundo por medio de los sentidos y donde se maneja el mayor número de variables, entendiendo variables como entradas del mundo al sistema. El nivel con menor granularidad es el simbólico, que “sintetiza” en leyes matemáticas o proposiciones lingüísticas lo que se representa en los otros dos niveles inferiores. Este nivel se construye sobre los otros dos y, por tanto, depende de ellos. En medio de estos dos niveles y como nexo de unión se sitúa el nivel conceptual.

Gärdenfors presenta las limitaciones de los niveles simbólico y subconceptual. Los problemas que no sería capaz de solucionar el nivel simbólico, entre otros, serían los siguientes: Las representaciones proposicionales no serían adecuadas para representar conexiones causales o interacciones dinámicas ya que no proporcionan ninguna forma natural de separar diferentes dominios de información. Esto está unido al hecho de que los portadores centrales de las representaciones simbólicas en los lenguajes de primer orden son los predicados, y se supone que los predicados se *dan* al sistema, no cuestionándose nunca cómo se forman ni su valor cognitivo. Otro problema es la imposibilidad de estudiar la emergencia evolutiva de las representaciones simbólicas, ni explicar la inducción creativa ni el conocimiento nuevo, puesto que no hay categorías conceptuales disponibles para especificar la situación antes de que surjan los símbolos. Sin embargo, un sistema con éxito debe ser capaz de aprender propiedades radicalmente nuevas de sus interacciones con el mundo y no solo formar combinaciones de los predicados dados. El problema es ver de dónde surgen los predicados y cómo evolucionan, pero para conseguir esto es necesario ir por debajo del nivel simbólico.

El nivel subconceptual viene dado por el paradigma conexionista cuya implementación son las redes neuronales artificiales. De acuerdo con el conexionismo, los procesos cognitivos no se deberían representar por manipulación de símbolos, sino por la dinámica de los patrones de actividades en las redes neuronales.

Las limitaciones de la interpretación conexionista son, según Gärdenfors, que a diferencia de los espacios conceptuales de bajas dimensiones en estos sistemas no se pueden interpretar las dimensiones. Además, estos sistemas necesitan un gran número de entradas (un entrenamiento muy largo) para aprender estructuras relevantes, e incluso si sabemos que la red ha aprendido a categorizar la entrada de forma correcta, puede que no seamos capaces de describir lo que representa la red emergente. Otro problema es que al

asumirse implícitamente que hay un dominio dado para los receptores de la red, la restricción lleva a confinar las capacidades representacionales de los sistemas. En particular, una red no puede generalizar fácilmente lo que ha aprendido de un dominio a otro, lo que quiere decir que estos sistemas no son aptos para tratar las inferencias.

Situado entre ambos niveles estaría el nivel conceptual. Este nivel sería independiente del simbólico ya que las dimensiones relevantes también serían independientes de las representaciones simbólicas (las dimensiones relevantes se forman con base a experimentos psicológicos con objetos). Según Gärdenfors, los espacios conceptuales resolverán el problema de la fundamentación de los símbolos, al menos de los símbolos que se refieran a dominios perceptivos. Se puede dar significado a los símbolos conectándolos con varias construcciones en los espacios conceptuales. Por tanto, para llegar a tener significado, el nivel simbólico depende del conceptual. Gärdenfors define el estatus ontológico de los espacios conceptuales: “son entidades teóricas que se pueden usar para explicar y predecir varios fenómenos empíricos respecto a la formación de conceptos”.¹⁷⁰

Gärdenfors busca una fundamentación empírica de los espacios conceptuales en la neurociencia: “si la naturaleza utiliza representaciones geométricas, esto supondría un apoyo a la teoría de espacios conceptuales. Por tanto, considero relevantes a los datos de la neurociencia y la psicología a la hora de evaluar diferentes formas de representar la información en sistemas artificiales.”¹⁷¹ Otra razón sería que hay indicios para pensar que el procesado de información de varios sistemas sensoriomotores es mucho más fundamental para el funcionamiento cognitivo del cerebro que los procesos implicados en la manipulación de símbolos. Parece ser que hay una distribución geométrica de funciones cerebrales, unos mapas topográficos (por ejemplo, vista y oído). Según Gärdenfors, los mapas topográficos

¹⁷⁰ Peter Gärdenfors, *Conceptual Spaces. The Geometry of Thought*, 31.

¹⁷¹ Peter Gärdenfors, *Conceptual Spaces. The Geometry of Thought*, 48.

generan distintos dominios de representación ya que la mayoría preservan la modularidad de los sentidos. La razón es que los distintos tipos de neuronas receptoras son sensibles a distintas características de nuestro entorno y estas características se mantienen distinguibles en los mapas superiores en el sistema de proyecciones. Según Gallistel: “El cerebro usa un esquema espacial para representar combinaciones de propiedades de estímulos, incluso cuando esas propiedades no son espaciales.”¹⁷² Respecto a por qué utilizar los espacios conceptuales en formas de razonamiento abstractas, Gärdenfors opina que la hipótesis más simple para el científico cognitivo es que el modo de representación geométrico también se usa para estas formas. Para dominios más altos que el básico no se puede dar cuenta de la similaridad de forma perceptiva (por ejemplo, en la clasificación de las plantas), siendo la tendencia general en el desarrollo de categorizaciones de un dominio hacia una menor dependencia de la similaridad perceptiva. Lo que sí se hace es transferir la similaridad con base perceptual a dominios abstractos por medio de mapeos metafóricos.

Según la teoría de espacios conceptuales, una propiedad es un conjunto de regiones de un dominio, y un concepto (un espacio conceptual) es un conjunto de dominios. Por tanto, una propiedad es un caso especial de concepto. Una propiedad es natural cuando se representa por una región convexa del espacio conceptual al que pertenece. La convexidad depende de la métrica del espacio. Una región A es convexa si para dos puntos x e y que estén en A , cualquier punto entre x e y también está en A . A partir de aquí Gärdenfors hace una interpretación geométrica de distintos tipos de conceptos. Por ejemplo, las propiedades se definen sobre la base de los criterios topológicos de conectividad y convexidad. Una propiedad natural, en el sentido de categoría de nivel básico en la teoría de Lakoff y Johnson, sería una región convexa de un dominio en un sistema

¹⁷² Gallistel, C.R. *The Organization of Learning*, The MIT Press, Cambridge, Massachusetts, 1990, 501. Citado en Peter Gärdenfors, *Conceptual Spaces. The Geometry of Thought*, 52.

conceptual. Al haber una “geometría conceptual”, el manejo de propiedades y conceptos admite un tratamiento computacional. En este tratamiento las estructuras de representación son espacios conceptuales consistentes en un número de dominios con ciertas estructuras geométricas o topológicas, y las unidades de computación son los puntos del espacio tomados como vectores.

Como aplicación de su teoría, Gärdenfors desarrolla una explicación de la inducción sobre la base de criterios de proximidad en los espacios conceptuales, dando relevancia a las propiedades naturales al afirmar que son las que se utilizan en la inducción. La generalización viene dada por la reducción de dimensiones relevantes y la creación de una nueva idea o concepto se interpreta como correlaciones nuevas entre dominios.

Gärdenfors se sitúa en la línea de gran parte de los científicos cognitivos al hacer un planteamiento de la cognición como estructurada en niveles. Esta estructura revela - en mi opinión - el reconocimiento del problema mente-cuerpo como el fundamental en Ciencia Cognitiva. Sabemos que hay unos mecanismos físicos y biológicos en la cognición y sabemos que creamos unas estructuras de símbolos con las que comunicamos referentes del mundo, pero el misterio está en el paso de los unos a las otras. La variabilidad y la sorpresa en la creatividad humana hacen que las explicaciones de los conceptos como entidades que surgen de la interacción física con el mundo resulten insuficientes. A su vez, no entendemos cómo estos conceptos tienen un funcionamiento neuronal que pueda estar motivado en parte por algo situado más allá del funcionamiento del propio “sistema”. Se intenta dar una explicación evolucionista en términos de adaptación al medio y de supervivencia de la especie, explicación que parece insuficiente, al menos en el caso de la cognición humana.

Analizando el carácter científico de la teoría de espacios conceptuales, hay que señalar que pueden surgir muchas dudas respecto a su validez. En primer lugar, no existe un acuerdo sobre la

validez de los experimentos a la hora de elegir las dimensiones relevantes ni la métrica del espacio conceptual en cuestión, a lo que se puede añadir que la distinción del espacio vectorial puede depender del contexto, lo cual lo empieza a hacer difuso como para extraer conclusiones relevantes en su posterior tratamiento, sea éste computacional o no. Por otra parte, excepto para campos muy restringidos, tampoco es definitiva la organización espacial del pensamiento, ni que ésta se rija por unas leyes matemáticas con una estructura geométrica y una métrica. En este sentido, Gärdenfors en ningún momento habla de que su teoría sea aplicable a todas las representaciones mentales, aunque el libro está escrito como si así fuese. Pero tiene un carácter fundamental saber si *todo* el pensamiento se puede interpretar geoméricamente o no, y en su caso, qué pensamiento no es representable geoméricamente. Por ejemplo, resulta difícil representarse geoméricamente el *amor* o la *justicia*. Incluso dentro del pensamiento representado geoméricamente, falta una explicación geométrica de cómo se correlacionan los dominios o se reducen las dimensiones, teniendo que volver a los experimentos psicológicos.

Respecto a las implicaciones filosóficas, Gärdenfors no aplica su teoría a gran parte de los problemas filosóficos, al contrario que Lakoff y Johnson. Sí define su postura como de acuerdo a una “semántica cognitiva”, próxima al experiencialismo y evolucionismo de Lakoff y Johnson, pero más representacional, ya que aunque toma como base el evolucionismo y relación psicológica con el medio, una vez se han producido éstos, mantiene que sólo tenemos acceso a las representaciones que hay en nuestras mentes, aunque admite que hay un “ajuste social” mediante una “mutua coordinación de significados”. Este aspecto más “simbólico” y “computacional” del pensamiento hace que sus espacios conceptuales admitan, en principio, un tratamiento computacional.

4.3 Resumen

En este capítulo se han presentado dos teorías sobre la cognición. A mi juicio, la teoría de Lakoff y Johnson descubre y muestra aspectos del vínculo corpóreo del pensamiento, y la teoría de Gärdenfors intenta explicar la cognición desde la interacción neurológica con el mundo hasta el pensamiento simbólico, postulando un nivel geométrico-conceptual. El problema es que en las dos teorías se intenta construir *toda* una teoría de la cognición sobre los aspectos que resaltan, y es en el desarrollo de la teoría completa donde se aprecian sus deficiencias.

Por otra parte, se ha intentado mostrar el carácter científico e interdisciplinar de las teorías, llegando a la conclusión de que ambos son en gran parte tentativos más que contrastados firmemente y aceptados por la comunidad científica.

CAPÍTULO V

Claves para el debate filosófico de la Ciencia Cognitiva

En este capítulo se expone cómo el problema central mente-cuerpo planteado en la Ciencia Cognitiva lleva a la necesidad de ampliar la noción de experiencia y de ciencia para poder encontrar un marco explicativo para la cognición, y cómo esto, a su vez, nos encamina hacia una concepción de la cognición y del hombre abierta y en permanente desarrollo. Esta concepción, junto con la asunción del misterio que lleva implícito, hace posible abrir nuevas vías explicativas respecto a las posibilidades de la razón humana y del hombre que van más allá de la consideración de éste como mero “sistema”.

Permítanme decirle al estudiante que, si quiere que su pensamiento abstracto llegue a algún sitio, en lugar de que simplemente gire alrededor de un eje fijo, debe dejar que vuele su imaginación, a la vez que ase firmemente la cuerda con la que la controla¹⁷³.

¹⁷³ Charles S. Peirce, en Richard Robin, *Annotated Catalogue of the Papers of Charles S. Peirce*. Amherst: University of Massachusetts Press, 1967, 413:264. Citado en Cornelis de Waal. *On Peirce*. Wadsworth Philosophers Series, 2001, Prefacio.

5.1 Introducción: filosofía y Ciencia Cognitiva

En los capítulos anteriores se ha expuesto cómo la Ciencia Cognitiva intenta dar cuenta del pensamiento desde sus distintas perspectivas. Estas perspectivas tienen en común el tratamiento del fenómeno de la cognición desde el punto de vista de las ciencias naturales, como la física y más recientemente la biología, y de las matemáticas. Incluso en el caso de las disciplinas de la Ciencia Cognitiva más “sociales” como la antropología, la psicología o la lingüística, permanece la explicación de los hechos cognitivos con base a los mecanismos físico-matemáticos o biológico-matemáticos del cerebro. Debido a su carácter formal, el éxito de este tipo de explicaciones permitiría hacer predicciones muy exactas y crear modelos de los fenómenos cognitivos, principalmente a través del uso del ordenador.

Estas aproximaciones al fenómeno del pensamiento, como se vio en el capítulo 1, tienen su origen en la creencia en la viabilidad de realizar un estudio “científico” de la cognición, debido fundamentalmente al avance de las matemáticas en el estudio de funciones recursivas, su aplicación en los ordenadores, y a los avances en neurociencia. A esto se sumaría el descontento respecto a las explicaciones puramente filosóficas del pensamiento que se habían producido anteriormente.

Durante los años treinta, algunos matemáticos y neurólogos emprendieron los estudios cognitivos en todo su alcance. También participaron en el proyecto psicólogos, lingüistas, antropólogos y filósofos, aplicando en sus respectivos campos los métodos y explicaciones de las matemáticas y las ciencias naturales. Las actitudes de los científicos cognitivos respecto al papel de la filosofía eran, y son, variadas y dispares: unos abogan por una paulatina desaparición de ésta a medida que avanzan las explicaciones

científicas¹⁷⁴, con lo que desaparecería lo que consideran que hay de “superstición”, “charlatanería”, etc., respecto a la explicación de los fenómenos cognitivos. Otros consideran que nunca podrá desaparecer la reflexión filosófica sobre el pensamiento, y en su mayoría no suscribe el proyecto de la Ciencia Cognitiva en toda su extensión¹⁷⁵. En un difícil equilibrio entre estas dos posturas están los que aceptan un tipo de reflexión filosófica en línea con los modelos del pensamiento propuestos por las disciplinas centrales de la Ciencia Cognitiva, pudiéndose considerar a esta filosofía como una “filosofía cognitiva” o “filosofía de la mente”¹⁷⁶. En este trabajo no se discutirá el papel de la filosofía como disciplina que forma parte de la Ciencia Cognitiva, ya que insistimos en que se considera a la Ciencia Cognitiva una empresa que utiliza los métodos de las matemáticas y las ciencias naturales de manera central, y por tanto se hablará de la filosofía sin condicionarla a este planteamiento positivista. Se mostrarán las dificultades con las que se encuentra la Ciencia Cognitiva y se analizará si tienen alguna raíz cuya determinación esté fuera del alcance de sus métodos y que sin embargo se pueda delimitar y mostrar mediante la reflexión filosófica.

No hace falta comenzar por una reflexión filosófica sobre la Ciencia Cognitiva para darse cuenta de cuáles son estas dificultades. Basta con observar su incapacidad para establecer resultados generales respecto a la cognición y para predecir comportamientos humanos. Como se indicaba en el capítulo 1, las principales cuestiones que se resistían al tipo de análisis propuesto por la Ciencia Cognitiva eran el carácter social del pensamiento, las emociones, la creatividad, el aparente carácter imprevisible de la cognición, etc. En definitiva, la Ciencia Cognitiva ha avanzado en la creación de modelos y simulación de fenómenos en un ámbito muy restringido de la

¹⁷⁴ John Haugeland, "Farewell to GOFAI?", en *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, 103.

¹⁷⁵ John Searle, "Ontology is the question", en *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, 203.

¹⁷⁶ Cf. Jerry Fodor, "The Folly of simulation", en *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, 85-100.

actividad cognitiva. Fuera de este ámbito sus resultados son muy limitados, y como la Ciencia Cognitiva aspira a dar una explicación completa y global de *todo* el fenómeno cognitivo, esto hace que se dude de la adecuación de sus métodos al campo de estudio de la cognición y, por tanto, de su pretendido carácter científico garante del proyecto.

En tareas cognitivas restringidas la Ciencia Cognitiva va produciendo resultados que nos asombran, sobre todo al observar los artilugios que parecen razonar y comportarse como las personas. No tenemos sin embargo una idea clara de lo que será capaz de explicar y producir, y nos preguntamos si el hecho de que aún no obtenga resultados en campos más amplios es tan sólo una cuestión de tiempo o si hay algo en la mente humana que hace imposible su estudio y simulación o recreación mediante las ciencias naturales. Esto no deja de tener un cierto grado de tarea imposible, al ser en cierto modo un intento de adelantarnos al futuro, ya que puede ser que lo que ahora nos parezca inalcanzable llegue algún día a explicarse mediante las ciencias naturales. Las fronteras entre lo que es susceptible de descubrirse y explicarse mediante el estudio científico experimental y lo que cae dentro de la reflexión filosófica no son nítidas. Un discernimiento total entre ambos requeriría tanto una idea clara de cuál es el ámbito de cada uno, como de la delimitación del campo que se pretende estudiar, y ya se vio en el capítulo 1 que esta diferenciación no es nada fácil. En este sentido, tiene interés observar las disputas y equivocaciones en los juicios respecto a los avances en inteligencia artificial por parte de algunos filósofos y científicos. Por ejemplo, ante la predicción de Herbert Simon en 1957 de que una máquina sería el campeón del mundo de ajedrez en pocos años, el filósofo Hubert Dreyfuss aseguró que un niño de diez años ganaría de forma aplastante al programa de ajedrez de Simon. Poco después, Seymour Papert retaría a Dreyfuss a una partida de ajedrez contra "MacHack", una máquina construida en el MIT. Dreyfuss aceptó y

perdió la partida¹⁷⁷. Por otro lado, Simon también predijo que en diez años un ordenador descubriría un importante teorema matemático y que la mayoría de las teorías psicológicas importantes se expresarían en forma de programas de ordenador, y ninguna de estas previsiones se han cumplido¹⁷⁸.

Cabe preguntarse, ateniéndonos a los resultados, si la Ciencia Cognitiva no lo explica todo porque aún no está lo suficientemente madura (como argumentaría von Eckardt o Thagard) o no lo explica todo porque su planteamiento no es válido de raíz. En este punto la pregunta puede ser filosófica: si nos referimos a si lo que no conocemos es el mecanismo tal y cual por el que se da una determinada actividad cognitiva, de la misma manera que aún no sabemos cómo acabar con el cáncer, parece claro que la filosofía debe retirarse para dejar paso a la Ciencia Cognitiva, como lo ha hecho de las ciencias naturales. Pero si a lo que nos referimos es a si la actividad cognitiva *es únicamente* un mecanismo formulable matemáticamente en último término, o un sistema, como asume la Ciencia Cognitiva, o si hay algo más que esto, parece lógico pensar que la respuesta se encuentre fuera del alcance de la Ciencia Cognitiva. Puede incluso pensarse que el pretendido carácter interdisciplinar de la Ciencia Cognitiva pueda solucionar este problema, pero como hemos comentado anteriormente, para lograr una mínima vertebración entre las disciplinas que se consideran a sí mismas parte de la Ciencia Cognitiva, todas deben asumir ese corte científico-positivista que les proporciona el apellido de “cognitivas”.

¹⁷⁷ Cf. Daniel Crevier, *AI: the Tumultuous History of the Search for Artificial Intelligence*, 122-124.

¹⁷⁸ Cf. Daniel Crevier, *AI: the Tumultuous History of the Search for Artificial Intelligence*, 108-109.

5.2 El problema mente-cuerpo

El problema mente-cuerpo se encuentra en el centro de esta discusión. Parece claro que nuestra actividad mental tiene un sustrato biológico que consiste principalmente en actividad del cerebro, lo que induce a pensar que la actividad mental no es otra cosa que actividad cerebral y corpórea. Este es el punto de vista de la mayor parte de las teorías en Ciencia Cognitiva. En contra de esta opinión están los que consideran lo mental como un fenómeno que no es reducible a la actividad físico-biológica. En los dos casos se presentan dificultades en la explicación del fenómeno cognitivo: para los primeros las dificultades consisten en explicar lo que en cierto modo se podría considerar más distintivamente humano: sentimientos, carácter creativo y social del hombre, etc. Los que mantienen esta posición sostienen que si no se dan estas explicaciones en términos matemático-físico-biológicos es porque la Ciencia Cognitiva aún no es lo suficientemente madura. Basan su optimismo en los resultados obtenidos en inteligencia artificial y neurociencia, y rechazan las críticas tachándolas de filosóficas, religiosas y poco científicas. La libertad, el amor, la creatividad, etc., serían epifenómenos que se reducirían en última instancia a estados neuronales, de los que si no se puede dar una explicación satisfactoria a día de hoy es porque el número de variables a tratar en cualquier fenómeno cognitivo relevante es tan alto que no se puede manejar matemáticamente, o porque existe una incertidumbre inherente en los fenómenos físicos por su propia naturaleza, o, en versiones más sutiles y sofisticadas matemáticamente, porque el pensamiento consista en un tipo de algoritmo del que no se puede saber el resultado hasta que se ejecuta, con lo que es imprevisible en cierto modo¹⁷⁹. Sin embargo, sí sería previsible que esta incertidumbre o complejidad se pudiese manejar matemáticamente en algún momento, con lo que sería un concepto totalmente “científico”. Si esto que propone la Ciencia Cognitiva es verdad, las repercusiones sobre la filosofía serían enormes, ya que

¹⁷⁹ Cf. Stuart Kauffman, *At Home in the Universe. The Search for the Laws of Self-Organization and Complexity*, Oxford University Press, Nueva York, 1995, 21-22.

ésta, como actividad mental, estaría totalmente condicionada y explicada por los resultados en Ciencia Cognitiva.

En el proceso de formulación matemática de la actividad cognitiva la Ciencia Cognitiva ha ido pasando por varias fases, y en todas estas fases se han creado modelos formales del pensamiento. Estos modelos implican una antropomorfización de la máquina que utilizaba estos modelos, al considerar que se había conseguido “trasladar” el pensamiento humano a la máquina, y a la vez una mecanización de lo humano, ya que los procesos cognitivos se podían realizar en una máquina. Los modelos matemáticos han ido cambiando para adecuarse a los descubrimientos del funcionamiento cerebral, pasando del modelo computacional en serie al modelo de procesamiento distribuido en paralelo, han pasado a tener en cuenta los procesos biológicos, genéticos y químicos además de los físicos, y se ha buscado integrar la interacción con el medio, mediante teorías dinámicas y adaptativas, de una forma más rica que como meras “entradas” a un sistema aislado. Como uno de los fenómenos más relevantes y complejos de la cognición se encuentra el tema de la *autoconsciencia*, que se explicaría como una propiedad del sistema de reconocerse y observarse a sí mismo.

Hasta aquí se tendría el enfoque para abordar el *cómo* de la actividad cognitiva. Pero aún es necesario añadir una explicación sobre el *motor* que guía los actos, las decisiones, las intenciones y la vida del hombre, en definitiva una explicación científica del *porqué*. La tendencia actual más extendida es la de buscar este “motor” de la vida y del pensamiento humano en el evolucionismo. La adaptación al medio y la supervivencia serían las “funciones” fundamentales “grabadas” en los genes que explicarían en última instancia el flujo de la vida y del pensamiento. El carácter social del hombre se habría producido como un “ajuste” adaptativo-evolutivo de la especie al medio. Este ajuste sería más sofisticado que el de los animales, por tener el hombre un desarrollo cerebral mayor, pero en último término estaría gobernado únicamente por los mismos principios. Por supuesto, de estas funciones evolutivas también se busca un modelo

matemático. De esta forma se completaría una explicación “científica” de la cognición humana.

Frente a estas teorías, otros mantienen que la persona y su actividad mental no son reducibles a su actividad orgánica. Si la posición reduccionista tiene dificultades para explicar lo que parece más distintivo del hombre, en este otro caso la dificultad parece estar en explicar “científicamente” la actividad cognitiva. La dificultad estaría en explicar qué es lo mental que, aun necesitando un sustrato físico, no puede reducirse a éste. Cómo ya señalaba, por ejemplo, Peirce:

No hay duda de que toda la fisiología del sistema nervioso muestra la dependencia de la mente del cuerpo... La cuestión es si los fenómenos mentales están controlados exclusivamente por la ley mecánica ciega...: pero hay “objeciones obvias” a esta idea¹⁸⁰.

Una de las objeciones a la posición materialista es la opinión fundada que podemos tener de que un tratamiento científico reduccionista que solo tuviese en cuenta los aspectos biológicos de la cognición sería semejante a aquél que trataría de explicar “lo que es una tarjeta de crédito mirándola por el microscopio”¹⁸¹, en donde se percibe cómo se deja de lado el inevitable carácter social del pensamiento. El paso que da la Ciencia Cognitiva al situar los aspectos físico-biológicos como únicos determinantes de los aspectos sociales está rebatido por la interpretación que sostiene también la influencia contraria. Por ejemplo, para Peirce, como señala De Waal, el niño va adquiriendo la autoconsciencia y el sentido de sí mismo con el tiempo, en interacción con las personas y el mundo. A la vez va

¹⁸⁰ Charles S. Peirce, *Collected Papers*, editado por Charles Hartshorne, Paul Weiss, y Arthur Burks. Cambridge, Massachusetts, Harvard University Press, 1931-1958, volumen 7, párrafo 274. Citado en Susan Haack: “Not Cynism, but Synechism: Lessons from Classical Pragmatism”, en *Transactions of the Charles S. Peirce Society*, primavera 2005, vol. XII, no. 2, 252.

¹⁸¹ Jaime Nubiola, “Perspectivas actuales en la filosofía de lo mental”, en *Espíritu* (Barcelona), XLIX, 2000, 13-24. <http://www.unav.es/users/Articulo45.html>

desarrollando conceptos como verdad o falsedad, bondad y maldad, etc. En este mismo desarrollo se configura como esencial en el hombre el carácter falible y comunitario de su conocimiento:

La explicación de Peirce sobre el origen del yo muestra que, para Peirce, algunos conceptos metafísicos clave emergen a la vez y están conectados íntimamente con el yo (self). Más aún, la explicación de Peirce subraya la naturaleza comunitaria del yo, ya que se encuentra íntimamente ligado al testimonio de los otros. En ausencia de este testimonio, no puede emerger el yo. La explicación de Peirce muestra también que ser falibles no es un obstáculo temporal que debemos superar, sino que constituye nuestra misma esencia. El conocimiento perfecto disolvería el yo y destruiría nuestra individualidad¹⁸².

Geertz, desde un análisis cultural del origen del hombre, subraya también la influencia del carácter social del hombre sobre su desarrollo biológico, y la imposibilidad de separar ambos:

Entre las estructuras culturales, el cuerpo y el cerebro, se creó un sistema de realimentación positiva en el cual cada parte modelaba el progreso de la otra; un sistema en el cual la interacción entre el creciente uso de herramientas, la cambiante anatomía de la mano y el crecimiento paralelo del pulgar y de la corteza cerebral es sólo uno de los ejemplos más gráficos. Al someterse al gobierno de programas simbólicamente mediados para producir artefactos, organizar la vida social, o expresar emociones, el hombre determinó sin darse cuenta de ello los estadios culminantes de su propio destino biológico. De manera literal, aunque absolutamente inadvertida, el hombre se creó a sí mismo... Como nuestro sistema nervioso central se desarrolló en gran parte en interacción con la cultura, es incapaz de dirigir nuestra conducta u organizar nuestra experiencia sin la guía suministrada por sistemas de símbolos significativos... De manera que esos símbolos son no meras expresiones o instrumentos o elementos correlativos a nuestra existencia biológica, psicológica y social, sino que son requisitos previos de ella. Sin hombres, no hay cultura, pero igualmente, sin cultura no hay hombres¹⁸³.

¹⁸² Cornelis de Waal, *On Peirce*, 80-81.

¹⁸³ Clifford Geertz, *La interpretación de las culturas*, 54-55.

Pero la tendencia al materialismo perdura en gran parte de las teorías cognitivas, al explicar todos los fenómenos en términos de lo que percibimos por los sentidos, asignándoles las mismas propiedades que a las cosas materiales. Entendemos inmediatamente que no habría cultura sin cerebros, pero nos cuesta más imaginarnos por qué no habría cerebros sin cultura. Ya Wittgenstein advirtió sobre los errores terminológicos que pueden dar lugar a errores conceptuales, y Peirce destacó la permanencia del modo de pensar cartesiano en esta concepción de la cognición:

El error que estamos expuestos a cometer podría expresarse así: estamos buscando el uso de un signo, pero lo buscamos como si fuese un objeto que coexistiese con el signo... Como las frases están en alguna parte, buscamos un lugar para el pensamiento¹⁸⁴.

La filosofía moderna nunca ha sido capaz de desprenderse del todo de la idea cartesiana de la mente como algo que “reside”, -este es el término- en la glándula pineal. Todo el mundo se ríe de esto hoy en día, y sin embargo todo el mundo continúa pensando sobre la mente de esta misma manera, como algo que está en esta o en aquella persona, que le pertenece y que se corresponde con el mundo real¹⁸⁵.

Por otra parte, la experiencia parece ser una realidad irreducible a posibles componentes, una realidad que no se puede separar del pensamiento y sus procesos biológicos, pero que tiene al menos la misma autonomía e influencia sobre el pensamiento que la que tienen los procesos biológicos. Citando de nuevo a Wittgenstein:

“Una máquina piensa, desea” parece en cierto modo carente de sentido. Es como si hubiésemos preguntado “¿tiene color el número 3?”. Pues en un aspecto de la cuestión, la experiencia personal, lejos

¹⁸⁴ Ludwig Wittgenstein, *Los cuadernos azul y marrón*, Tecnos, Madrid, 1968, 1998, 31-34.

¹⁸⁵ Nathan Houser et al., *The Essential Peirce. Selected Philosophical Writings*, Bloomington: Indiana University Press, 1992-98, vol. 2, 199. Citado en Cornelis de Waal, *On Peirce*, 78.

de ser el producto de procesos físicos, químicos y fisiológicos, parece ser la base misma de todo lo que decimos con algún sentido sobre tales procesos. Considerándolo de este modo, nos inclinamos a utilizar nuestra idea de un material de construcción de otro modo equivocado y a decir que el mundo entero, mental y físico, está hecho de un único material¹⁸⁶.

Peirce aborda el problema de la génesis y estructura del pensamiento analizando si tenemos la capacidad de distinguir intuiciones puras, que no hagan referencia a otras, y lo aplica al análisis de la autoconciencia como intuición pura. Acaba afirmando que no tenemos tal capacidad, sino que un pensamiento nos remite a otro. No hay intuiciones puras, sino que pensamos a través de signos que hacen referencia a otros signos, y así sucesivamente. Como origen del pensamiento se sitúa la experiencia, por lo que sólo somos capaces de pensar a través de ésta, y no tenemos acceso a la mente si no es a través del razonamiento hipotético que parte de los hechos externos. Esto tiene su origen en una semiosis del pensamiento que comienza durante el desarrollo del niño, se produce por mediación de las personas y en interacción con el medio que le rodea, y se desarrolla mediante la experimentación y el error. Como se ha expuesto, el niño no nace autoconsciente, sino que se va generando esta capacidad en su interacción con el mundo. Todos los conceptos, aún los más abstractos, nos remiten en última instancia a una experiencia exterior. Mediante la introducción de una referencia continua de un signo a otro y la consideración del tiempo, Peirce llega a la conclusión de que la cognoscibilidad y el ser son lo mismo y que el hombre es un signo que hace referencia a algo exterior, que el hombre es un signo externo.

El significado se sitúa así en la comunidad, aunque esté soportado por cada hombre. Es la “red” de significaciones tejida por los hombres la que da sentido a los significados, y no los procesos biológicos y químicos o los estados mentales del cerebro, aunque sin éstos el hombre no tuviese acceso a aquellos. A su vez, como dice Wittgenstein, es la experiencia personal la que nos hace poder hablar

¹⁸⁶ Ludwig Wittgenstein, *Los cuadernos azul y marrón*, 79.

de cualquier cosa, incluso de los procesos físicos, químicos y biológicos que la hacen posible.

Pero si los significados se encuentran en los signos y éstos a su vez en la comunidad, es necesario dotarlos de un nivel de realidad equivalente al de los fenómenos físicos. Y, aunque unos dependan de otros y no puedan existir por separado, se debe establecer claramente la entidad propia de cada uno de ellos. El problema se sitúa entonces en identificar lo que es “real” para el hombre. Debido a nuestra corporeización, inevitablemente consideramos más “real” aquello a lo que podemos acceder a través de los sentidos. La teoría cognitiva de la metáfora nos ha hecho ver esto más claramente al mostrar cómo expresamos los conceptos en términos sensoriales. También Peirce se sitúa (en cierto modo ya que su concepción de experiencia es más amplia), en línea con la teoría cognitiva de la corporeización al afirmar que tenemos una “consciencia menos viva” de los conceptos a medida que se hacen más abstractos y se alejan de la experiencia, por lo que en última instancia nuestros pensamientos nos remiten a ésta como origen del significado¹⁸⁷. Y afirma también que “la sensación y el poder de abstracción o atención pueden considerarse, en un sentido, como los únicos constituyentes de todo pensamiento”¹⁸⁸. Pero a la vez es verdad que estas referencias últimas a lo sensitivo parecen resultar insuficientes, al considerar la validez, utilidad, necesidad, y por tanto el carácter “real” de los razonamientos abstractos al contrastarlos con la experiencia. Aunque hablemos en términos sensoriales, muchas veces hablamos de algo que no “vemos”, ni “sentimos”, a través de nuestro cuerpo.

¹⁸⁷ Cf. Charles S. Peirce. "Cuestiones acerca de ciertas facultades atribuidas al hombre", en Nathan Houser et al., *The Essential Peirce. Selected Philosophical Writings*, vol. 1, 11-27. Hay versión electrónica en inglés en: <http://members.door.net/arisbe/menu/library/bycsp/question/qu-frame.htm> y en español en: <http://www.unav.es/gep/QuestionsConcerning.html>.

¹⁸⁸ Charles S. Peirce. "Algunas consecuencias de cuatro incapacidades", en *Collected papers*, volumen 5, párrafo 295. Hay versión electrónica en inglés en: <http://members.door.net/arisbe/menu/library/bycsp/conseq/cn-frame.htm> y en español en: <http://www.unav.es/gep/AlgunasConsecuencias.html>.

Este “algo”, que no sentimos a través de nuestro cuerpo, y que sin embargo se impone como una “realidad” tan fuerte al menos como la que percibimos directamente por la experiencia sensorial se hace especialmente patente en la necesidad que tiene el hombre de contestar a las preguntas que se hace sobre su existencia y su sentido.

5.3 La búsqueda del sentido

Es necesario tener en cuenta tanto la experiencia personal, en el sentido más amplio, del hombre, o mejor, de *cada* hombre, como su dimensión social. La sociedad, aunque tenga sus fuerzas propias, es fruto de la experiencia personal. La cultura refleja las preguntas que el hombre se hace para orientar su existencia y las respuestas que da a éstas. Cada hombre necesita explicarse y “solucionarse” a sí mismo además de integrarse en unas normas sociales, por lo que no es suficiente para la comprensión del hombre y de la cognición humana complementar el análisis físico-experimental del hombre con el simbólico-social. Si nos parásemos en formular teorías físico-matemáticas del funcionamiento cerebral junto con descripciones del comportamiento social, cultural o psicológico del hombre, seguiríamos dejando a un lado la pregunta por el sentido del hombre. No se trata de completar el “sistema” físico de neuronas con un “sistema” social de símbolos.

Aunque esto se pudiese considerar como una propiedad neurológica emergente de la conciencia, el hombre necesita explicarse a sí mismo, necesita responder a las preguntas que se plantea. Pero, ante esta situación, el hombre se queda perplejo, no se explica a sí mismo. Puede analizar, descubrir que tal sentimiento se corresponde con tal actividad neuronal, pero no se da sentido a sí mismo, no es capaz de explicarse la vida, el sufrimiento y la muerte. Para el hombre las respuestas a estas preguntas son un misterio. No hay una “autoconciencia neurológica final” que encuentre en su interior, que

le tranquilice y dé sentido a su vida. Éstas son preguntas fundamentales que puede intentar responder buscando respuestas “científicas”, como el evolucionismo, que responde en términos de adaptación y supervivencia de la especie. Sin embargo, si bien el evolucionismo explica el flujo de la vida en sus aspectos naturales, no constituye una solución al problema de la existencia para cada persona. No creo que a nadie deje tranquilo ni orientado en la vida una explicación del tipo: “haces esto y lo otro porque el objetivo es la supervivencia del individuo y de la especie humana, y esta información la tienes codificada en tus genes”. El no-responder de la consciencia a esta pregunta junto con su persistencia hacen que sea una necesidad profunda del hombre que no puede contestar por sí solo. A partir de aquí se podría considerar esta necesidad humana como un fenómeno biológico que se podría controlar o suprimir. Sería el hombre un sistema con una “función” que le hace necesitar explicarse su vida y la cual se podría “corregir”. Anulando esta necesidad el hombre simplemente seguiría el curso evolutivo de la especie sin cuestionarse el sentido de la vida. Pero anular esta supuesta función no supondría eliminar el problema, sino que estaríamos en el mismo punto, al haber dado el control a otra “función”: aquella que juzga según criterios “prácticos” cómo debe vivir el hombre, modificando para ello los procesos biológicos y psicológicos conforme a estos criterios. Independientemente del carácter especulativo de estas supuestas modificaciones, es relevante la pregunta de cuáles serían las características más propiamente humanas que no deberíamos alterar.

En esta búsqueda de lo más profundamente humano, Levinas, siguiendo a Husserl en el análisis fenomenológico de las capas pasivas de la conciencia, intentó profundizar más allá de la característica intencional de ésta. Habría, según Levinas, una conciencia no intencional pasiva, que no es acto ni consiste en conceptualizar, representar, objetivar o reconocer cognitivamente, sino que es experiencia preconceptualizante, pura pasividad que consistiría en la vivencia de ser apelada desde fuera. Esta conciencia no intencional sería lo más profundamente humano, antes que

cualquier acto cognitivo, lo que daría sentido al hombre a través de su característica fundamental, que es ser apelada desde fuera. Sería, en palabras de Levinas, donde se revelaría el sentido último de mi “yoidad”, sería lo más ciego y sordo de nosotros mismos y a la vez lo que nos lleva más lejos y nos hace seres trascendentes y espirituales:

La trascendencia no se reduce a una experiencia de la trascendencia por cuanto se trata de un estupor previo a toda *posición* del sujeto y a todo contenido percibido o asimilado. Trascendencia o despertar que es la vida misma de lo humano, inquietada de antemano por lo Infinito¹⁸⁹.

Si lo más íntimo, auténtico, humano, en nosotros es esta conciencia interpelada, entonces es en el exterior donde se construye lo interior. Pero no en un exterior representado y ordenado por lo interior, no construido como un orden de un sí-mismo, sino en lo exterior como realidad fundamental donde cobra sentido lo interior. Se trata, como expresa de forma tan radical y con tanta fuerza Levinas, al buscar la realidad fuera de sí mismo y rechazar la posición egoísta del punto de vista interior que pone en el centro al sujeto, de construir la realidad no como otro modo *de ser*, sino de otro modo *que ser*¹⁹⁰. No es que el hombre integre en su “mismidad” lo exterior, no es que asimile la realidad como un orden de sí-mismo, es que el hombre es exterioridad, vida en común, y en esa exterioridad se encuentra a sí mismo. Esta apelación desde el exterior viene dada según Levinas por el prójimo, especialmente por el humilde, el que sufre, que introduce “una perturbación absoluta que no es de este mundo”¹⁹¹, y así se parece a Dios:

El hombre en tanto que Otro nos viene de fuera, separado –o santorostro. Su exterioridad, es decir, su apelación por mí, es su verdad. Mi

¹⁸⁹ Emmanuel Levinas, “La filosofía y el despertar”, en *Entre nosotros. Ensayos para pensar en otro*, Pre-textos, 1993, 2001, 112.

¹⁹⁰ Cf. Emmanuel Levinas, *De otro modo que ser o más allá de la esencia*, Ediciones Sígueme, Salamanca, 1995.

¹⁹¹ Emmanuel Levinas, “¿Un Dios hombre?”, en *Entre nosotros. Ensayos para pensar en otro*, 73.

respuesta no se agrega a un "núcleo" de su objetividad como un accidente, sino que *produce* así su verdad (que su "punto de vista" sobre mí no podría abolir). Esta excedencia de la verdad sobre el ser y sobre su idea que sugerimos por la metáfora de "curvatura del espacio intersubjetivo", significa la intención divina de toda verdad. Esta "curvatura del espacio" es, tal vez, la presencia misma de Dios¹⁹².

Para Levinas, la ética, la respuesta del hombre a la llamada del prójimo es lo esencial, lo que precede a la metafísica y a la ontología. La conciencia intencional puede acoger la llamada ética del prójimo o puede intentar acallarla, ya que también una inclinación no-intencional a no atender esta llamada sino a ser para sí mismo, que el hombre no se explica ni representa, late en la profundidad de su ser. Una inclinación que no responde a la pregunta del sentido pero que no soporta el exceso de misterio que supone aceptar que el hombre se realice dándose, que se encuentre perdiéndose. Entonces se podría decir que la conciencia intencional, como acto objetivante de la conciencia no-intencional, puede ayudar a la voluntad a "cruzar el umbral de la esperanza"¹⁹³ en un acto libre, a "no tener miedo"¹⁹⁴, sino a "dejarse conducir"¹⁹⁵. En este acto, el misterio, sin ser conocimiento, aflora en lo cognitivo:

Gracias a la capacidad del pensamiento, el hombre puede encontrar y reconocer esta verdad. En cuanto vital y esencial para su existencia, esta verdad se logra no sólo por vía racional, sino también mediante el abandono confiado en otras personas, que pueden garantizar la certeza y la autenticidad de la verdad misma. La capacidad y la opción de confiarse uno mismo y la propia vida a otra persona constituyen

¹⁹² Emmanuel Levinas, *Totalidad e infinito*, Ediciones Sígueme, Salamanca, 2002, 295.

¹⁹³ Juan Pablo II, *Cruzando el umbral de la esperanza*, Plaza & Janés, Barcelona, 1994, 216.

¹⁹⁴ Juan Pablo II, *Cruzando el umbral de la esperanza*, 218.

¹⁹⁵ Juan Pablo II, *Cruzando el umbral de la esperanza*, 218.

ciertamente uno de los actos antropológicamente más significativos y expresivos¹⁹⁶.

5.4 Líneas futuras de investigación

Quiero proponer dos líneas de investigación:

1.- La teoría del carácter corporal de los conceptos en Lakoff y Johnson y su relación con la teoría del sentido común en Peirce. Relación con el conocimiento científico.

Tomando como punto de partida la existencia de una referencia corporal sobre la que comprendemos los conceptos, y de una necesidad humana de formar conceptos abstractos a partir de la experiencia, cabe investigar cómo son las relaciones e influencias entre la experiencia misma, o los conceptos más "próximos" a la experiencia, y los más abstractos.

En este sentido, existe una relación entre lo que Lakoff y Johnson llaman "categorías de nivel básico" como más ligadas a la experiencia y las primeras que desarrolla el hombre, y las creencias básicas en Peirce. Ambas parecen tener un valor adaptativo-evolutivo en el desarrollo humano.

A partir de aquí es necesario esclarecer cómo se produce el conocimiento científico y la diferencia en sus fines y métodos respecto al conocimiento del sentido común. El conocimiento científico es más abstracto y persigue la búsqueda de la verdad, en lugar de estar regido por principios evolucionistas, según Peirce. Sin embargo, el trasfondo del lenguaje y la experiencia ordinaria permanece en todo el

¹⁹⁶ Juan Pablo II, *Fides et Ratio*, 1998, número 33. Por ej, en: http://www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jp-ii_enc_14091998_fides-et-ratio_sp.html

proceso¹⁹⁷. Lakoff y Johnson no tienen en cuenta el valor de la búsqueda de la verdad en ciencia, y refieren todo conocimiento en última instancia a su valor adaptativo al medio.

Cabe también preguntarse si existen estructuras conceptuales innatas en el hombre que nos hagan categorizar como lo hacemos, y si limitan o no nuestra razón. El tiempo y el espacio parecen formarse desde muy temprano, según Lakoff y Johnson, en conjunción con la experiencia y una cierta estructura cerebral. Whewell concede que partimos de unas ideas innatas: las ideas de espacio y tiempo, al menos. Estas ideas no son creables o derivables de la experiencia. Sin embargo esto no implica que sean anteriores o separables en el origen de las sensaciones, ya que desde el principio hemos percibido las cosas en el espacio y en el tiempo.

Respecto a la relación entre las creencias básicas, las estructuras conceptuales innatas, y la ciencia, es interesante analizar la influencia de unas en otras. La ciencia tiene un trasfondo de sentido común, de mundo "corporeizado", y la estructura cognitiva humana hace que los conceptos abstractos se refieran y se terminen de comprender en la experiencia. Pero el pensamiento abstracto tiene a la vez su propia autonomía, y parece escapar a sus posibles estructuras cognitivas. Tal podría ser el motivo de, por ejemplo, la diferencia entre la noción de espacio en ciencia y la creencia básica de espacio.

2.- *Un análisis de la conciencia y semiosis en Peirce y Levinas.*

Para Levinas lo más radicalmente humano es la apelación exterior a la conciencia no intencional, no representada. Esta apelación no se produce de la misma forma en el niño que en el adulto. Peirce concibe cualquier pensamiento como un proceso de

¹⁹⁷ Cf. David Carnicer, "¿Puede el sentido común ser crítico? Investigando las creencias básicas en Peirce.", en *Información Filosófica* 2/1, 2005, 45-66. En web: <http://www.philosophica.org/if/art/carnicer.pdf>

semiosis a través de algo externo. Cabe preguntarse y analizar si hay una semiosis de la conciencia no intencional, o si no hay tal cosa como una conciencia no intencional y una autoconciencia intencional, y si existen ambas conciencias, cuál es la relación entre ambas.

CONCLUSIONES

A continuación se enumeran las principales conclusiones de este trabajo:

1.- La dificultad en encontrar un modelo matemático que dé cuenta de toda la actividad cognitiva humana, y la viabilidad de que éstos modelos puedan explicar aspectos de la cognición.

La cognición humana es inseparable del ser humano en su totalidad como ser trascendente y abierto al misterio, y este misterio aflora en lo cognitivo. Una explicación completa del fenómeno cognitivo implicaría una explicación completa del ser humano, lo que no está al alcance de ninguna ciencia.

Sin embargo, tampoco tenemos un conocimiento apriorístico de cuáles son exactamente los fenómenos cognitivos explicables científicamente y cuáles no, por lo que es lógico dejar la puerta abierta a la ciencia en su exploración y acotación de las bases científicas de la actividad cognitiva. Las ideas científicas irán ganando o no la aceptación de la comunidad científica según se consoliden o desaparezcan con el tiempo, y así podrán ir explicando aspectos de la cognición. Como explica Geertz:

Todos se abalanzan a esa idea como si fuera una fórmula mágica de alguna ciencia positiva como si fuera el centro conceptual alrededor del cual es posible construir un sistema general de análisis... pero una vez que nos hemos familiarizado con la nueva idea, una vez que ésta forma parte de nuestra provisión general de conceptos teóricos, nuestras expectativas se hacen más equilibradas en lo tocante a los usos reales de dicha idea, de suerte que así termina su excesiva popularidad. Sólo unos fanáticos persisten en su intento de aplicarla

universalmente; pero pensadores menos impetuosos al cabo de un tiempo se ponen a considerar los problemas que la idea ha generado. Tratan de aplicarla y hacerla extensiva a aquellos campos donde resulta aplicable y donde es posible hacerla extensible y desisten de hacerlo en aquellos en que la idea no es aplicable ni puede extenderse. Si era válida se convierte entonces verdaderamente en una idea seminal, en una parte perdurable de nuestro arsenal intelectual. Pero ya no tiene aquel promisorio, grandioso alcance de su aparente aplicación universal que antes tenía¹⁹⁸.

2.- El dudoso carácter científico de la Ciencia Cognitiva al intentar dar una explicación de todo el fenómeno cognitivo.

En relación con las dificultades expuestas anteriormente, es lógico dudar del carácter científico de la Ciencia Cognitiva como garante de la validez del estudio y explicación de todo el fenómeno cognitivo.

Se puede pensar en salvar esta situación ampliando la concepción de ciencia, buscando una actividad científica que incluya metodologías distintas a las de las ciencias naturales y las matemáticas. En este caso la forma de medir su carácter científico no tendría por qué ajustarse exactamente al caso de una ciencia natural. Por ejemplo, en el caso de las ciencias sociales, un análisis del fenómeno social que no estuviese guiado por la metodología de las ciencias naturales requeriría un análisis simbólico diferente, y éste:

... en lugar de comenzar con una serie de observaciones e intentar incluirlas bajo el dominio de una ley, comienza con una serie de significantes e intenta situarlos dentro de un marco inteligible. Esta teoría no es predictiva, pero el marco teórico dentro del cual se hacen dichas interpretaciones debe ser capaz de seguir dando interpretaciones defendibles a medida que aparecen nuevos fenómenos¹⁹⁹.

¹⁹⁸ Clifford Geertz, *La interpretación de las culturas*, 19.

¹⁹⁹ Clifford Geertz, *La interpretación de las culturas*, 36-37.

Aun en este caso, otras dificultades, como la interrelación efectiva entre las distintas disciplinas que estudian la cognición, y fundamentalmente el carácter trascendente de la persona impedirían la validación “científica” completa de esta metodología. Otras concepciones más abiertas, como la pragmática de Peirce, pueden ofrecer un buen punto de partida en el establecimiento de una ciencia de la cognición:

No llamo ciencia a los estudios solitarios de un hombre aislado. Sólo cuando un grupo de hombres, más o menos en intercomunicación, se ayudan y se estimulan unos a otros al comprender un conjunto particular de estudios como ningún extraño podría comprenderlos, [sólo entonces] llamo a su vida ciencia²⁰⁰.

3.- La dificultad en integrar las distintas disciplinas dentro del modelo que se adopte, de tal forma que resulte relevante dicha integración.

Ya se vio en los capítulos 1 y 2 cómo las aproximaciones interdisciplinarias al estudio de la cognición siempre han resultado controvertidas. Estas controversias se originan tanto por factores de poder académico, al intentar hacer prevalecer la visión de una determinada disciplina, como por la dificultad de crear teorías que por un lado sean lo suficientemente concretas para resultar relevantes y contrastables con la experiencia, y por otro sean lo suficientemente amplias para poder resultar significativas y dar cabida en ellas a los puntos de vista sobre la cognición que las distintas disciplinas ofrecen.

4.- El reconocimiento del mismo grado de realidad, irreducibilidad e interdependencia de los hechos o fenómenos físicos y de los conceptuales o relacionales, y el carácter unitario e irreducible de la experiencia humana como lugar donde cobran sentido.

²⁰⁰ Charles S. Peirce, "The Nature of Science", en *The Charles S. Peirce Papers*, 1966, Cambridge, MA: Harvard University Library, MS 1334, Adirondack Summer School Lectures, 1905.

De esta manera es posible decir tanto que el ser, la consciencia “emerge” de nuestro sistema neuronal, como que nuestro sistema neuronal y físico “soporta” al ser y a la consciencia.

Peirce encuentra una metáfora muy ilustrativa de esto:

De la misma manera que decimos que un cuerpo está en movimiento, y no que el movimiento está en un cuerpo, deberíamos decir que nosotros estamos en el pensamiento y no que los pensamientos están en nosotros²⁰¹.

Y Whewell remarca que hechos y teorías, concepciones y cosas, son anverso y reverso de la misma moneda:

Ninguna aprehensión de las cosas es puramente ideal, ninguna experiencia de las cosas externas es puramente sensitiva. Si se conciben como *cosas*, la mente debe haber llegado a esta convicción a través de los sentidos, si se *conciben* como cosas, las sensaciones han tenido que ser unidas por concepciones. Si *pensamos* en alguna *cosa*, debemos reconocer la existencia tanto de pensamientos como de cosas. *La antítesis fundamental de la filosofía es una antítesis de elementos inseparables*. No solo no se pueden mostrar estos elementos separados, sino que tampoco se pueden concebir ni describir separados²⁰².

Que hechos y teoría, cosas y conceptos sean diferentes pero inseparables, es debido a que se producen dentro de la unidad del ser humano, y más concretamente, dentro de su experiencia, entendiendo ésta en el sentido más amplio, que engloba la producción mental completa.

²⁰¹ Nathan Houser et al., *The Essential Peirce. Selected Philosophical Writings*, vol. 1, 42.

²⁰² William Whewell, *The Philosophy of the Inductive Sciences (founded upon their history)*, 1840, Londres, en *The Historical and Philosophical Works of William Whewell*, vol. V, Editado por G. Buchdal y L.Laudan. Frank Cass & Co Ltd., 1967, 38-46.

5.- *La autoconsciencia como fenómeno cognitivo causante de que el hombre se interese por la verdad de lo que se le presenta, incluido el mismo hombre. La plausibilidad, a la vez que se presenta como misterio, de la Verdad como don de Amor ofrecido inevitablemente al hombre desde el exterior y que le da sentido. La influencia en la cognición, y la respuesta a esta Verdad como lugar en el que se dan la mano pensamiento, experiencia y trascendencia.*

Desde el análisis de la autoconsciencia es plausible llegar a reconocer en el hombre la necesidad de la apertura al misterio, que no a lo misterioso y oculto. Un misterio que se puede sentir próximo y del que se puede hablar, que debe ser compatible y reinterpretable con los nuevos avances científicos y que puede iluminar la razón, un misterio que está en la raíz de que “la razón vaya más allá de lo que puede formalizar”²⁰³, al reconocer esa influencia exterior.

6.- *El rechazo de la concepción del hombre como un “sistema” que interactúa con el mundo.*

En Ciencia Cognitiva se buscan modelos del pensamiento que “funcionen”. Es necesario incorporar la amplitud y el carácter ilimitado y sorprendente de la razón y de la persona, que, como indica Gabriel Marcel, debido a su “imperfección”, es imposible reducir al funcionamiento de un “sistema”:

Es sin duda necesario reaccionar con fuerza contra la idea clásica de *autarkia*, de la suficiencia de sí mismo para sí mismo. Lo perfecto no es lo que se basta a sí mismo o, en todo caso, su perfección sería la de un sistema, no la de un ser. ¿En qué condiciones puede presentar un valor espiritual la relación que liga a un ser con aquello de lo que tiene necesidad? Parece que debe darse ahí una reciprocidad, un despertar. Sólo puede ser espiritual una relación de un ser con otro ser... Lo que

²⁰³ Hilary Putnam, *Representation and Reality*, The MIT Press, Cambridge, Massachussets, 1988, 1989, 118.

cuenta es el comercio espiritual entre dos seres, y aquí no se trata de respeto sino de amor²⁰⁴.

7.- *La filosofía como actividad donde se muestra más claramente que la razón va más allá de sí misma.*

La Ciencia Cognitiva, o las distintas disciplinas que la forman, pueden explicar ciertos aspectos de la cognición. La actividad filosófica también está sujeta a estas explicaciones, que pueden revisar, corregir y replantear las respuestas que se da los problemas filosóficos²⁰⁵.

Pero la filosofía no parte de supuestos previos como las ciencias, y por eso se enfrenta continuamente a la revisión de estos supuestos. Se enfrenta a todo lo que se le presenta al hombre, incluido él mismo, y en esta actividad está presente el Misterio. Como indica Levinas:

Tal es la filosofía: lenguaje de la trascendencia y no mero relato de una experiencia; lenguaje en el que el locutor pertenece al relato, lenguaje necesariamente personal, también, que ha de entenderse más allá de lo que dice, que ha de interpretarse. La filosofía no es otra cosa que la "intriga" intersubjetiva de los filósofos, que nadie puede resolver y en la que no se permite relajamiento de la atención ni falta de rigor²⁰⁶.

La filosofía consiste en mostrar claramente que las cosas no están nada claras, y hace esto presentando una visión sinóptica de los problemas. En este proceso escruta todos los aspectos de la realidad hasta llegar al Misterio. Con terminología de Wittgenstein, Marcel y

²⁰⁴ Gabriel Marcel, *Diario metafísico*, 207. Citado en Emmanuel Levinas. "Una nueva racionalidad. Sobre Gabriel Marcel", en *Entre nosotros. Ensayos para pensar en otro*, 112.

²⁰⁵ Cf. Mariano Artigas, *La mente del universo*, En <http://www.unav.es/cryf/lamentedeluniverso.html>

²⁰⁶ Emmanuel Levinas, "La filosofía y el despertar", en *Entre nosotros. Ensayos para pensar en otro*, 112.

Levinas, respectivamente, podemos decir que a filosofía ofrece una visión sinóptica de los problemas, horada la realidad y acaricia el Misterio.

BIBLIOGRAFÍA

La bibliografía abordada en las cuestiones de este trabajo es enorme. Por ejemplo, en <http://cogweb.ucla.edu/Bibliography.html> se recogen millares de referencias. En esta sección me he limitado a reunir por orden alfabético de autores todas las obras consultadas y expresamente citadas en el cuerpo del texto y en las notas a pie de página.

American Society for Cybernetics, <http://www.asc-cybernetics.org/>

Artigas, M., *La mente del universo*,
<http://www.unav.es/cryf/lamentedeluniverso.html>

Bechtel, W., *Filosofía de la mente. Una panorámica para la ciencia cognitiva*, Tecnos , Madrid, 1991.

Bechtel, W. y Abrahamsen, A., *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, Blackwell Publishers Ltd, Malden, Massachusetts, 2002, 1991.

Bechtel, W., Mandik, P. y Mundale, J., “Philosophy Meets the Neurosciences”, en Bechtel, W., Mandik, P., Mundale, J. y Stufflebeam, R.S., *Philosophy and the Neurosciences. A Reader*, Blackwell Publishers Ltd, Malden, Massachusetts, 2001, 4-22.

- Bechtel, W. y Stufflebeam, R. S., "Epistemic Issues in Procuring Evidence about the Brain: The Importance of Research Instruments and Techniques", en Bechtel, W., Mandik, P., Mundale, P., y Stufflebeam, R.S., *Philosophy and the Neurosciences. A Reader*, Blackwell Publishers Ltd, Malden, Massachusetts, 2001, 55-81.
- Bennett, M.R., and Hacker, P.M.S., *Philosophical Foundations of Neuroscience*, Blackwell, Malden, Massachusetts, 2004, 2003.
- Brown, R., *Social Psychology*, Free Press, Nueva York, 1965.
- Bustos, E., *La metáfora. Ensayos transdisciplinarios*, Fondo de Cultura Económica de España, Universidad Nacional de Educación a Distancia, Madrid, 2000.
- Carnicer, D., "¿Puede el sentido común ser crítico? Investigando las creencias básicas en Peirce", en *Información Filosófica* 2/1, 2005.
- Casson, R. W., "Cognitive Anthropology", en Robert A. Wilson, R.A. y Kiel, F.C. (eds.), *The MIT Encyclopaedia of the Cognitive Sciences (MITECS)*, The MIT Press, Cambridge, Massachusetts, 2001, 1999, 120-121.
- Churchland, Paul M., *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*, Gedisa, Barcelona, 1999.
- Cicourel, A., "Cognition and Cultural Belief", en Baumgartner, P., y Payr, S. (eds.), *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, 1995. 47-58.
- Clark, A., *Mindware. An Introduction to the Philosophy of Cognitive Science*, Oxford University Press, Oxford, 2001.

Cognitive Linguistics Journal. Mouton de Gruyter,
www.degruyter.de/rs/384_386_ENU_h.htm

Cognitive Psychology Journal, Elsevier,
http://www.elsevier.com/wps/find/journaldescription.cws_home/622807/description#description

Collins, A., "Why Cognitive Science?", en *Cognitive Science*, vol. 1, 1977.

Crevier, D., *AI: the Tumultuous History of the Search for Artificial Intelligence*, Basic Books, Nueva York, 1993.

Croft, W. y Cruse, D. A., *Cognitive Linguistics*, Cambridge University Press, Cambridge, 2004.

D'Andrade, R., *The Development of Cognitive Anthropology*, Cambridge University Press, Cambridge, 2003, 1995.

De Waal, C., *On Peirce*. Wadsworth, Belmont, California, 2001.

Dupuy, J-P., *The Mecanization of the Mind. On the Origins of Cognitive Science*, Princeton University Press, Princeton, Nueva Jersey, 2000.

Echeverría, J., *Filosofía de la Ciencia*, Akal, Madrid, 1995.

Eysenck, M. W. y Keane, M. T., *Cognitive Psychology*, Psychology Press, Hove, 2003, 2000.

Fauconnier, G. y Turner, M., *The Way we Think. Conceptual Blending and the Mind's hidden Complexities*, Basic Books, Nueva York, 2002.

- Finger, S., *Origins of Neuroscience. A History of Explorations into Brain Function*, Oxford University Press, Oxford, 1994.
- Fodor, J., *The Modularity of Mind*, MIT/Bradford Press, Cambridge, Massachusetts, 1983.
- "The Folly of simulation", en Baumgartner, P., y Payr, S. (eds.), *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, 1995, 85-100.
- Gallistel, C.R., *The Organization of Learning*, The MIT Press, Ciudad, 1990.
- Garbarino, M. S., *Sociocultural Theory in Anthropology. A Short History*. Waveland Press, Inc., Prospect Heights, Illinois, 1983.
- Gardner, H., *La nueva ciencia de la mente: historia de la revolución cognitiva*, Paidós, Barcelona, 1996, 1988.
- Garrido, M., "El *ego cogito* del siglo XX", en Ernest Nagel, E. y Newman, J. R., *El Teorema de Gödel*, Tecnos, Madrid, 2000, 1999.
- Gärdenfors, P., *Conceptual Spaces. The Geometry of Thought*, The MIT Press, Cambridge, Massachusetts, 2000.
- Geertz, C., *La interpretación de las culturas*, Gedisa, Barcelona, 1990, 1973.
- Georges, M. T., *The Problem of Storing Common Sense in Artificial Intelligence. Context in CYC*, Universidad de Navarra, Pamplona, 2000.
- <http://www.unav.es/gep/TesisDoctorales/MarkGTesis.pdf>

- Gödel, Kurt, "On Formally Undecidable Propositions of Principia Mathematica and Related Systems 1", <http://home.ddc.net/ygg/etext/godel/godel3.htm>
- Haack, S., "Not Cynicism, but Synechism: Lessons from Classical Pragmatism", en *Transactions of the Charles S. Peirce Society*. primavera 2005, vol. XII, no. 2.
- Haugeland, J., "Farewell to GOFAI?", en Baumgartner, P., y Payr, S. (eds.), *Speaking Minds. Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, 1995. 101-114.
- Heims, S. J., *The Cybernetics Group. 1946-1953. Constructing a Social Science for Postwar America*. The MIT Press, Cambridge, Massachusetts, 1993, 1991.
- Johannesson, M., *The Problem of Combining Integral and Separable Dimensions*, http://www.lucs.lu.se/ftp/pub/LUCS_Studies/Lucs87.pdf
- Juan Pablo II, *Cruzando el umbral de la esperanza*, Plaza & Janés, Barcelona, 1994.
- *Fides et Ratio*, 1998, http://www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jp-ii_enc_14091998_fides-et-ratio_sp.html
- Kauffman, S., *At Home in the Universe. The Search for the Laws of Self-Organization and Complexity*, Oxford University Press, Nueva York, 1995.

- Lakoff, G., *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, The University of Chicago Press, Chicago, 1987.
- *Moral Politics: What Conservatives Know that Liberals Don't*, University of Chicago Press, Chicago, 1996.
- Lakoff, G. y Johnson, M., *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*, Basic Books, Nueva York, 1999.
- Lakoff, G., y Núñez, R.E., *Where Mathematics comes from. How the Embodied Mind brings Mathematics into Being*, Basic Books, Nueva York, 2001.
- Lakoff, G. y Turner, M., *More than Cool Reason: A Field Guide to Poetic Metaphor*, University of Chicago Press, Chicago, 1989.
- Levinas, E., "La filosofía y el despertar", en *Entre nosotros. Ensayos para pensar en otro*, Pre-textos, Valencia, 2001.
- *De otro modo que ser o más allá de la esencia*, Ediciones Sígueme, Salamanca, 1995.
- *Totalidad e infinito*, Ediciones Sígueme, Salamanca, 2002.
- Lucy, J. A., *Language Diversity and Thought. A Reformulation of the Linguistic Relativity Hypothesis*, Cambridge University Press, Cambridge, 1992.
- McCarthy, J., Minsky, M. L., Rochester, N. y Shannon, C. E., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. August 31, 1955*, <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

- McCulloch, W. S., “What is a Number, that a Man may know It, and a Man, that he May know a Number?”, en McCulloch, W. S.: *Embodiments of Mind*, The MIT Press, Cambridge, Massachusetts, 1970, 1965. 1-18.
- McCulloch, W. S., y Pitts, W. H., “A Logical Calculus of the Ideas Inmanent in Nervous Activity”, en *Bulletin of Mathematical Biophysics*, University of Chicago Press, Chicago, 1945, 89-93.
- Miller, G., “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information”, en *Psychological Review*, 63, 81-97.
<http://www.well.com/user/smalin/miller.html>
- Mitchell, M., *An Introduction to Genetic Algorithms*, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1998, 1996.
- Mundale, J., “Neuroanatomical Foundations of Cognition”, en Bechtel, W., Mandik, P., Mundale, J. y Stufflebeam, R.S., *Philosophy and the Neurosciences. A Reader*, Blackwell Publishers Ltd, Malden, Massachusetts, 2001, 37-54.
- Newell, A. y Simon, H. A., *The Logic Theory Machine, a Complex Information Processing System*, The Rand Corporation. P-868, 1956.
<http://shelf1.library.cmu.edu/IMLS/MindModels/logictheorymachine.pdf>
- Norman, D.A., “¿Qué es la ciencia cognitiva?”, en *Perspectivas en Ciencia Cognitiva*, Paidós, Barcelona, 1987.

Nubiola, J., "Perspectivas actuales en la filosofía de lo mental", en *Espíritu* (Barcelona), XLIX, 2000, 13-24.

<http://www.unav.es/users/Articulo45.html>

——— *La renovación pragmatista de la filosofía analítica. Una introducción a la filosofía contemporánea del lenguaje*, EUNSA, Barañáin, Navarra, 1996, 1994.

Peirce, C. S., "The Nature of Science", en *The Charles S. Peirce Papers*, 1966, Cambridge, MA: Harvard University Library, MS 1334, Adirondack Summer School Lectures, 1905.

——— "Questions Concerning Certain Faculties Claimed for Man", en Nathan Houser et al., *The Essential Peirce. Selected Philosophical Writings*, Bloomington: Indiana University Press, 1992-98, vol. 1, 11-27.

<http://members.door.net/arisbe/menu/library/bycsp/question/qu-frame.htm>

——— "Some Consequences of Four Incapacities", en *Collected Papers*, editado por Charles Hartshorne, Paul Weiss, y Arthur Burks. Cambridge, Massachusetts, Harvard University Press, 1931-1958, volumen 5, párrafo 295.

<http://members.door.net/arisbe/menu/library/bycsp/conseq/cn-frame.htm>.

Priest, G. y Tanaka, K., "Paraconsistent Logic", en *Stanford Encyclopedia of Philosophy*,

<http://plato.stanford.edu/entries/logic-paraconsistent/>

Putnam, H., "Against the new associationism", en Baumgartner, P., y Payr, S. (eds.), *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, New Jersey, 1995. 177-188.

——— *Representation and Reality*, The MIT Press, Cambridge, Massachusetts, 1989, 1988.

Rodríguez, C. y Moro, C., *El mágico número tres. Cuando los niños aún no hablan*, Paidós Ibérica, Barcelona, 1999.

Rosenblatt, F., “The Perceptron: a Probabilistic Model for Information Storage and Organization in the Brain”, en *Psychological Review*, 65. 1958, 368-408.

Rosenblueth, A., Wiener, N. y Bigelow, J., “Behavior, Purpose and Teleology”, en *Philosophy of Science*, Vol. 10, No. 1 (Jan., 1943).

Rumelhart, D. E., “From Searching to Seeing”, en Peter Baumgartner and Sabine Payr (eds.), *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, Nueva Jersey, 1995. 189-202.

Rumelhart, D. E., McClelland, J. L. y el PDP Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, The MIT Press, Cambridge, Massachusetts, 1986.

——— *Introducción al procesamiento distribuido en paralelo*, Alianza Editorial, Madrid, 1992.

Schunn, C. D., Crowley, K. y Okada, T., “The Growth of Multidisciplinarity in the Cognitive Science Society”, en *Cognitive Science* 22(1), 1998, 107-130.

<http://www.lrdc.pitt.edu/Schunn/research/papers/multidisc.htm>

[1](#)

Searle, J. R., “Ontology is the Question”, en Baumgartner, P., y Payr, S. (eds.), *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*, Princeton University Press, Princeton, New Jersey, 1995, 203-214.

Simon, H. A., *The Sciences of the Artificial*, The MIT Press. Cambridge, Massachusetts, 1996, 1969.

Solomon, S. N., “A Brief History of Cognitive Anthropology”, 2000, <http://www.geocities.com/xerexes/coganth.html>

Sperber, D., “Why Rethink Interdisciplinarity?” Abril 2003, <http://www.interdisciplines.org/interdisciplinarity/papers/1>

Sternberg, R. J., “Intelligence”, en Robert A. Wilson and Frank C. Kiel (eds.) *The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, The MIT Press, Cambridge, Massachusetts, 2001, 1999, 409.

——— *Cognitive Psychology*, Wadsworth, Belmont, California, 2003.

Thagard, P., *Mind: Introduction to Cognitive Science*, The MIT Press, Cambridge, Massachusetts, 1998, 1996.

——— “Cognitive Science”, en *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/cognitive-science/>

Tomasello, M., “Cognitive Linguistics”. En Bechtel, W. y Graham, G., *A Companion to Cognitive Science*, Blackwell, Malden, Massachusetts, 1999, 1998, 477-487.

Turing, A., “Computing Machinery and Intelligence”, en *Mind*, Vol. 59, octubre 1950, 433-60. <http://www.abelard.org/turpap/turpap.htm>

——— “On Computable Numbers, with an Application to the Entscheidungsproblem”, <http://www.abelard.org/turpap2/tp2-ie.asp>

- Ungerer, F. y Schmid, H-J., *An Introduction to Cognitive Linguistics*, Addison Wesley Longman Limited, London, 1997, 1996.
- Varela, F. J., *Conocer: las ciencias cognitivas, tendencias y perspectivas*. Gedisa, Barcelona, 1998, 1990.
- Varela, F. J., Thompson, Evan, Rosch, Eleanor, *De cuerpo presente. Las ciencias cognitivas y la experiencia humana*, Gedisa, Barcelona, 1997.
- Von Eckardt, B., *What is Cognitive Science?*, The MIT Press, Cambridge, Massachusetts, 1996, 1993.
- Whewell, W., *The Philosophy of the Inductive Sciences (founded upon their history)*, 1840, Londres, en Buchdal, G. y Laudan, L. (eds.), *The Historical and Philosophical Works of William Whewell*, vol. V, Frank Cass & Co Ltd., Londres, 1967.
- Whorf, B., *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf*, Carroll, J. B. (ed.), The MIT Press, Cambridge, Massachusetts, 1998, 1956.
- Wiener, N. *Cibernética o el control y comunicación en animales y máquinas*, Tusquets Editores, Barcelona, 1998, 1985.
- Wittgenstein, L., *Los cuadernos azul y marrón*, Tecnos, Madrid, 1998, 1968.
- Zach, R., "Hilbert's Program", en *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/hilbert-program/>