

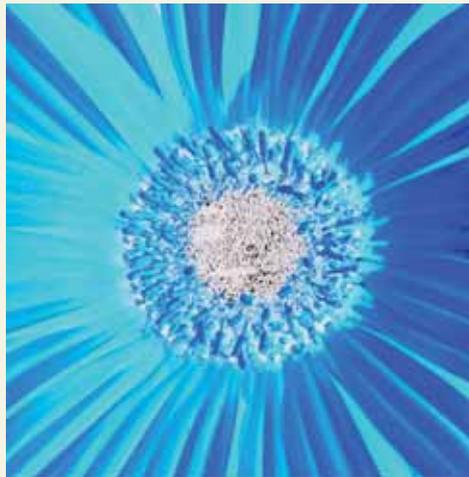
Arquitecturas de la Cognición

Mente = Cerebro + Cuerpo + Medio

Conferencia pronunciada con motivo de la concesión a

Francisco José Calvo Garzón

*del Premio Jóvenes Investigadores de la Región de Murcia 2005
convocado por la Fundación Séneca - Agencia Regional de Ciencia y Tecnología*



Francisco José Calvo Garzón

ARQUITECTURAS DE LA COGNICIÓN

Mente = Cerebro + Cuerpo + Medio

ARQUITECTURAS DE LA COGNICIÓN

Mente = Cerebro + Cuerpo + Medio

© Francisco José Calvo Garzón

© Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia

1ª ed.: Murcia 2007

D.L.: MU-855-2007

Edición realizada para la Fundación Séneca

por *QUADERNA EDITORIAL*

Telf. 968 343 050 - quaderna@quaderna.es

Impreso en España. Todos los derechos reservados.

Prohibida la reproducción total o parcial sin permiso expreso
y por escrito de los titulares del Copyright.

A ti

De bella, mi imaginación se encoge.

Dibujar tu cuerpo, parir un poema ya escrito.

Un sueño: bañarme en tus ojos,
secarme dentro de tu vientre...,
no conocer el olvido.

Beber en tus labios: un vino aún no concebido.

Y reír, reír sobre la mar en mi regazo...
en tu azotea de seda, abierta y serena,
mi pluma recoge...
recoge y siembra.

ÍNDICE

Cómo leer esta charla	9
Nota preliminar	11
1. Introducción.....	15
2. La arquitectura de la cognición: Clasicismo <i>versus</i> conexionismo	21
3. La eliminación de lo mental.....	31
4. Ciencia cognitiva corporeizada.....	39
5. Apéndice: Discurso de agradecimiento	51
6. Referencias bibliográficas	55

¿CÓMO “LEER” ESTA CHARLA?

Tras pasar por unos cuantos borradores, he decidido escribir dos charlas por el (buen) precio de una. De esta manera, el librito que tienes entre tus manos lo puedes leer de dos formas no excluyentes, y ¡no tan distintas después de todo!:

- a) Obviando las notas a pie de página (muy aconsejable para una primera lectura), o
- b) Siguiendo las notas a pie de página conforme avances en el texto principal.

En cualquier caso, pido disculpas de antemano por las simplificaciones que incluso la presentación *en modo tipo (b)* conlleva por razones obvias.

NOTA PRELIMINAR

La preparación de este material me ha dado la oportunidad de mirar atrás y reflexionar con calma sobre los orígenes de mi interés por la ciencia y la filosofía. El objetivo de la ciencia cognitiva es... explicar la cognición (¡prometo decir algo más al respecto!). El punto de partida es la adopción de un marco materialista, en el mejor sentido de la palabra: la mente es parte del mundo físico, material. Pero, ¿cómo es posible que de un sustrato físico, el cerebro, emerja *lo mental*? Ésta no es cuestión baladí. Hay competencias que parecen ser exclusivas de la especie humana o, en el peor de los casos, sólo tienen algunas especies animales. Digo “en el peor de los casos”, ya que tendemos a aplicar nuestros sesgos antropocéntricos al estudio de la cognición. ¿Quién osaría ver en la conducta de una hormiga o de una abeja, o incluso en las respuestas heliotrópicas de una planta, atisbos de mentalidad?¹

El subtítulo de esta charla, “Mente = Cerebro + Cuerpo + Medio”, abre una serie de interrogantes, algunos de los cuales pretendo abordar aquí. Poner “mente” en una ecuación en la que el resto de variables (“cerebro”, “cuerpo” y “medio”) son objeto directo de estudio de las ciencias naturales, invita a una reflexión sobre el modo en que, recurriendo a la manida expresión, la mente humana trata de comprender la propia mente.

1 La pregunta es retórica. Como puedes apreciar, no estoy pensando en experimentos clásicos en etología cognitiva o psicología comparada con respecto al abanico de competencias que exhiben los *grandes simios*. Peter Carruthers (2004) defiende la tesis de que algunos insectos tienen mente en el sentido cognitivo que aquí nos interesa. En Calvo Garzón (en revisión) elaboro un *reductio ad absurdum* de la tesis defendida por Carruthers. Para un estudio empírico de habilidades cognitivas particulares (atención, percepción, aprendizaje, memoria, etc.) en la literatura especializada sobre cognición animal, ver Bekoff y col. (2002).

Por otro lado, esta ocasión también me ha permitido recuperar a aquellos filósofos que más me marcaron durante los primeros estadios de mi carrera investigadora, y ver cómo, también a veces sin darme cuenta, me he ido alejando de ellos. Si tuviera que citar aquellas obras que han dejado un poso especial en mi quehacer académico, éstas han sido: *Word and Object*, de Willard V.O. Quine, *Neurophilosophy*, de Patricia Churchland, *A Neurocomputational Perspective*, de Paul Churchland, y *Being There*, de Andy Clark. El orden de interés en estas obras es cronológico, y el orden en que me he ido despegando de todas ellas también. De Quine me quedo con el estrecho vínculo entre conocimiento y experiencia. De los Churchland, con el materialismo eliminativista. Por último, de Clark me quedo con su visión abierta de los muchos retos a los que todavía tiene que enfrentarse la ciencia cognitiva. Sin embargo, ni soy conductista Quineano (en lo que me imagino coincidiré con la mayoría), ni comparto el estrecho materialismo eliminativista de los Churchland (aquí seguro que discreparé con casi todo el mundo)², ni mi posición se puede calificar de conciliadora, más allá de una siempre sana explotación heurística de los postulados de la ortodoxia dominante. Espero ir desenredando esta madeja de aparentes inconsistencias en mi biografía intelectual a lo largo de esta charla.

Si estoy en lo cierto al proponer el crítico subtítulo de esta charla, no es posible comprender la mente sin entender a su vez las interacciones del cerebro con el propio cuerpo y con el medio en que habita. La mente *es* el cerebro, el cual está radicalmente *corporeizado* y *situado*. El resultado es la descentralización, en un marco anti-representacionista y anti-computacionalista, de la cognición³.

2 Esto se debe a que el tipo de eliminativismo que defiende deja a los Churchland, si me apuran, en el lado de Jerry Fodor. Ver los encendidos intercambios entre Fodor y los Churchland en McCauley, R. (ed.) (1996), así como Fodor y Lepore (1999) y Churchland (1998). Como las cosas no son siempre blancas o negras, para una revisión crítica de la semántica a la Churchland, ver Calvo Garzón (2000; 2003a).

3 Por romper el hielo con respecto a estos conceptos fundamentales, y como tributo a la obra de Quine, mi primera fuente de inspiración, la primera moraleja que se desprende de un marco Quineano es el escepticismo en torno a la existencia de significados. Lo que está en tela de juicio es el objetivismo aplicado al campo de la semántica, de tal manera que no hay actos mentales en los que apoyarse a la hora de poner en correspondencia *las palabras* y *las cosas*. Obviamente, una afirmación de esta índole dependerá en buena medida del tipo de evidencia que consideremos relevante. En este sentido, el acercamiento de Quine a la semántica estuvo críticamente moldeado por su naturalismo y el conductismo dominante en la filosofía y en las ciencias experimentales de

El lector crítico ya estará en alerta al sospechar que esta nota preliminar no es sino el preámbulo a una hipótesis *reactiva* (es decir, no cognitiva) de la cognición (!), con todas las paradojas que nos saltan a la cara. Sin embargo, estoy convencido de que las implicaciones de la visión que trataré de esbozar aquí son tan polémicas como esperanzadoras a la hora de enterrar una serie de mitos en filosofía y ciencia cognitiva, y de tratar de comprender aquello que, tal y como nos empeñamos en seguir creyendo, nos hace únicos.

Rincón de Seca, 30 de marzo de 2007

los años 50. Si bien toda actividad humana, y esto incluye a la actividad lingüística, debe ser estudiada bajo el prisma de las ciencias naturales, el resultado en Quine es la constatación de que toda evidencia objetiva se reduce a la evidencia conductual. La pregunta es: ¿qué clase de hechos podemos admitir más allá de la conducta observable del sujeto? Medio siglo después, con el avance en las técnicas de neuroimagen cerebral y con el desarrollo de sofisticados diseños experimentales, no hay una razón clara por la que un acercamiento naturalizado a la semántica no pueda abrir la *caja negra*, además de observar las disposiciones conductuales de los hablantes. Sin embargo, y todavía siguiendo a Quine, podemos distinguir tres posibles niveles de explicación a la hora de enfrentarnos a cuestiones semánticas: el nivel mentalista, el conductista y el de la neurociencia. En el caso de Quine, las razones para trabajar al nivel intermedio tienen que ver en parte con su esfuerzo por naturalizar la epistemología (Quine, 1969), y con consideraciones concernientes al proceso mediante el cual los hablantes ordinarios adquieren su lengua materna (Quine, 1970). Ésta parece ser de hecho la postura de Quine, después de todo. En ‘Facts of the Matter’ Quine clarifica su comprensión fisicista de lo fáctico: “Mental states and events do not reduce to behavior, nor are they explained by behavior. They are explained by neurology, when they are explained. But their behavioral adjuncts serve to specify them objectively. When we talk of mental states or events subject to behavioral criteria, we can rest assured that we are not bandying words; there is a physical fact of the matter, a fact ultimately of elementary physical states” (Quine, 1979, p. 167). De la misma manera, Quine cierra ‘Mind and Verbal Dispositions’ afirmando: “Until we can aspire to actual physiological explanation of linguistic activity in physiological terms, the level at which to work is the middle one; that of dispositions to overt behavior. Its virtue is not that it affords causal explanations but that it is less likely than the mentalistic level to engender an illusion of being more explanatory than it is” (Quine, 1975, p. 95).

INTRODUCCIÓN

Todo manual comienza invariablemente presentando la Ciencia Cognitiva como el estudio científico multidisciplinar de la mente-cerebro a través de las numerosas disciplinas que la integran. Entre éstas destacan la psicología, la neurociencia, la inteligencia artificial, la informática, la lingüística, la antropología y la filosofía (ver figura 1). Dadas las peculiaridades de cada una de estas disciplinas, una primera observación es que es preciso trabajar desde el plano del análisis conceptual, tal y como hace la filosofía, así como desde los planos experimentales y de modelización, como es el caso de la neurociencia y la inteligencia artificial, respectivamente. Sin embargo, los progresos en distintos frentes científicos hacen que la geometría del “hexágono cognitivo” se desdibuje, resultando una tarea ardua su delimitación. Dentro de la neurociencia podemos distinguir, entre otras especialidades, la neurociencia celular y molecular, la neurociencia cognitiva, la neurociencia computacional, etc., y lo mismo se sigue para el resto de vértices en la figura (lingüística cognitiva, lingüística computacional, etc.). Por otro lado, el tipo de competencias que son específicamente objeto de estudio de la ciencia cognitiva no está claramente delimitado, y en última instancia la labor científica se ramifica en función de si nos interesa por ejemplo el estudio del lenguaje, de la memoria, del aprendizaje, y así un largo etcétera. Si a esto añadimos el creciente interés en el estudio de las emociones, del contexto sociocultural, o de la cognición animal (no humana, se entiende), por poner unos pocos ejemplos, se hace patente la complejidad y dificultad de la empresa⁴.

4 Si estás interesado en el enfoque multidisciplinar que ofrece la Ciencia Cognitiva, Wilson y Keil (1999) es una muy buena obra de carácter enciclopédico que reúne los distintos frentes.

Asimismo, lo que en un principio se consideraba conexiones “en obras” o “en proyecto” (ver líneas discontinuas en figura 1) se ha convertido, especialmente en la última década, en vínculos estables. De este modo encontramos Departamentos e Institutos de Filosofía, especialmente en Estados Unidos y Reino Unido, en los que se ofrecen programas en Filosofía de la Ciencia Cognitiva, Filosofía de la Psicología, Filosofía de la Neurociencia, o Filosofía de la Inteligencia Artificial, entre otros⁵.

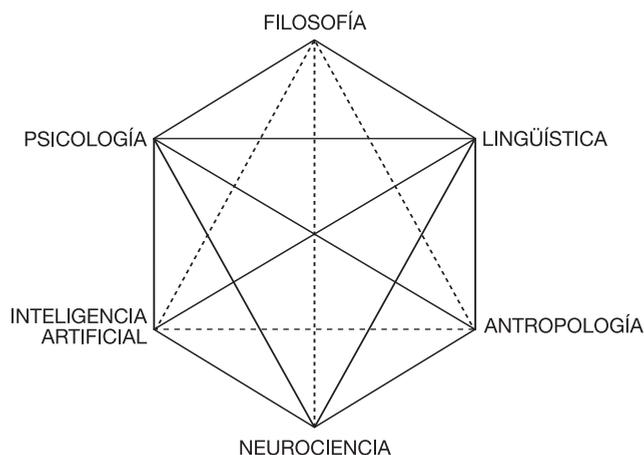


Figura 1: El hexágono cognitivo de Gardner (1985).

Por otro lado, ni siquiera sabemos bien qué significa enfrentarse a un problema de manera multidisciplinar. En el mundo académico insertar el término “multidisciplinar” en la solicitud de un proyecto es un valor añadido. Pero, ¿en qué sentido contribuye al progreso científico? Cualquier investigador que haya asistido a un congreso multidisciplinar habrá comprobado cómo los distintos *corrillos* hablan de *sus cosas*. ¿Alguien ha visto a un neurocientífico preguntarse por la normatividad, o a un filósofo por la elección de variables en un diseño experimental?⁶

5 Monografías en cada una de estas áreas que puedes consultar son: Bechtel y Graham (1998); Symons y Calvo Garzón (2008); Bechtel y col. (2001); y Boden (1990).

6 Recuerdo una reunión científica antes de fin de año en la Universidad de las Islas Baleares en la que, bajo el título de “Embodimenting cognition: Towards an integrated approach?”, se dieron

Por último, a la hora de investigar la cognición es preciso decidir qué nivel de descripción/explicación es el adecuado. Basándonos en la clásica distinción tripartita de David Marr (1982) podemos distinguir entre el nivel computacional, el algorítmico y el de la implementación. El nivel computacional nos proporciona una descripción abstracta de la tarea que el sistema cognitivo realiza; el algorítmico facilita un nivel de descripción funcional de dicha actividad, especificando *cómo* el sistema procesa la información; y finalmente, el tercero describe el sustrato físico en el que se implementa el procesamiento de la información que realiza el sistema. Una influyente tradición filosófica privilegia el nivel algorítmico sobre el de la implementación. Si asumimos que distintos sustratos físicos pueden implementar la misma función (piensa por ejemplo en una partida de ajedrez sobre un tablero convencional o en una pantalla de ordenador), el nivel adecuado de análisis no puede ser otro que el algorítmico. Al fin y al cabo, ¿de qué me sirve conocer pormenorizadamente los detalles de implementación (neuronal, pongamos por caso), si otro sistema cognitivo con *otra hardware* puede sentir o pensar lo mismo? En mi opinión, no hay mayor error que se pueda cometer en ciencia cognitiva que pensar que esto efectivamente es así⁷.

Para tu tranquilidad, o tal vez para todo lo contrario, no pretendo dar una respuesta clara a estas cuestiones. No hay chistera ni conejo a los que preten-

cita destacados investigadores en inteligencia artificial, ingeniería electrónica, robótica y neurociencia básica, junto a un grupo de filósofos. Pues bien, a pesar de la supuestamente clara diferencia entre los aspectos descriptivos que tienen que ver con el mundo, y los aspectos normativos que tienen que ver con las razones, no fue posible que (literalmente) ningún no-filósofo comprendiera *qué es eso de la normatividad*. ¡Aquí me temo que estoy en el bando de los no-filósofos!

7 Sin entrar en detalles, el funcionalismo es la teoría filosófica de la mente que está detrás de la supuesta irrelevancia del nivel de implementación. La idea central es que nuestras creencias y deseos (lo que en la sección tercera introduciremos como nuestras actitudes proposicionales) vienen constituidas por su papel funcional; esto es, por sus relaciones causales con otros estados de creencia o de deseo, y con respuestas conductuales. Es la identificación de dichos estados a través de su papel funcional lo que nos permite entenderlos como realizables de múltiples maneras (diferentes implementaciones pueden manifestar el mismo estado de creencia, por ejemplo). Si bien esta visión funcionalista fue el motor en los años 60 del valor explicativo de la causación mental desde un marco materialista que evitaba los problemas a los que debía enfrentarse tanto el dualismo como el reduccionismo (ver especialmente Putnam, 1975, y Fodor, 1975; 1998a; 1998b), hay un creciente malestar con la visión funcionalista dominante en la actualidad. Aparte de argumentos clásicos como la *Habitación China* de Searle (1980), hay muy buenas razones, que desgraciadamente no puedo entrar a analizar aquí ahora, para no ignorar la relación de la descripción algorítmica con los detalles de implementación (ver especialmente Bickle, 2003, y Churchland, 2005).

da recurrir en esta introducción. Al fin y al cabo, ¿se trata de poner retos asumibles sobre la mesa! El objetivo último es la comprensión de la mente-cerebro, tarea para la cual es preciso remangarse y trabajar codo con codo con otros investigadores en psicología, neurociencia, lingüística, informática, antropología, etc., desde muchos frentes en lo que concierne a metodologías, paradigmas experimentales, niveles de descripción y modelos de explicación científica.

Ahora bien, a pesar de la heterogeneidad con que nos enfrentamos al reto de comprender la cognición, es preciso señalar que la ciencia cognitiva es en la actualidad primordialmente *cognitivista*, valga la redundancia, en el sentido de que se aborda el estudio de la cognición tomando como punto de partida los conceptos de computación y de representación mental⁸. Sucintamente, por “cognitivismo” vengo a referirme al paradigma según el cual los procesos cognitivos consisten en la manipulación computacional de representaciones mentales. Sin embargo, no está claro qué queremos decir al afirmar que manipulamos representaciones. De hecho, ni siquiera hay consenso sobre qué propiedades poseen, o han de poseer, dichas representaciones mentales, más allá de estipular que se trata de entidades físicas interpretables semánticamente⁹.

En esta breve introducción he señalado dos movimientos o tendencias mayoritarias que se dan en el estudio de la cognición, que creo están haciendo un flaco favor a la disciplina. Por un lado, la tendencia a ignorar los detalles de implementación, y por el otro, el tomar los conceptos de computación y de repre-

8 Tras el abandono del conductismo hace ya más de medio siglo, el cognitivismo se ha convertido en el paradigma dominante en ciencia cognitiva. Incluyo bajo la etiqueta “conductismo” tanto al conductismo psicológico (Watson, 1924; Skinner, 1957) como al de corte analítico en filosofía (Ryle, 1949). En cuanto al cognitivismo, visiones ortodoxas bien conocidas se encuentran en Marr (1982) y Pylyshyn (1984). Lingüistas e informáticos pueden tener en mente para leer el resto de esta sección a Chomsky (1980) y su teoría de reglas y representaciones, y a Newell y Simon (1976) y su hipótesis de sistemas de símbolos físicos, respectivamente. Una incursión histórica desde la perspectiva del origen multidisciplinar del cognitivismo se puede consultar en Garnham (2008).

9 Como primera aproximación podemos decir que los estados internos de un sistema desempeñan un papel mediador en virtud del contenido informacional asociado a ellos (en Calvo Garzón, en prensa, llamo a esta precondition el “principio de conectividad”). Para que este papel sea crítico, debemos ser realistas. Al decir que los estados representacionales tienen poderes causales queremos decir que un sistema contiene estados físicos que están por otros estados y que pueden desempeñar un papel causal en la conducta del sistema debido al contenido del estado en cuestión. Los estados representacionales son estados causalmente eficaces, de tal manera que el sistema se puede comportar adaptativamente mediante la explotación de sus estados informacionales. Es decir, las representaciones son *para el sistema* (ver nota 28, abajo).

sentación como los pilares fundamentales de la ciencia cognitiva. En esta charla abordaré cuestiones relacionadas con la cognición desde mi experiencia personal, en la medida en que me he ido interesando como filósofo por los conceptos de representación y computación, o mejor dicho, por las alternativas a esos conceptos. Las interrelaciones con las vertientes de experimentación y de modelización son obvias, y emergerán en las siguientes secciones, así que no os aburriré aquí ahora explicitándolas.

LA ARQUITECTURA DE LA COGNICIÓN: CLASICISMO VERSUS CONEXIONISMO

La primera pregunta que se nos plantea apunta a la *arquitectura* de la cognición: ¿qué propiedades estructurales son fundamentales en relación con la conducta observable de un sistema cognitivo? Una manera de enfrentarse a esta cuestión es rastreando exhaustivamente todas aquellas arquitecturas propuestas hasta la fecha, comparando entonces sus propiedades estructurales y evaluándolas respecto a un listado consensuado de criterios, como por ejemplo, flexibilidad conductual, capacidad de aprendizaje o plausibilidad biológica. Esta estrategia ya ha dado sus frutos, y se han propuesto criterios de satisfacción con el fin de preservar la adecuación computacional de las hipótesis arquitectónicas en cuestión¹⁰. Sin embargo, si bien es cierto que se ha avanzado mucho en estas líneas, para nuestros propósitos no es preciso realizar dicho rastreo, ya que el listado de matices y diferencias de grado sería interminable. Por ello, en el resto de esta sección introduciré dos sistemas que podemos considerar paradigmáticos en la filosofía de la ciencia cognitiva, y que explotan nociones de computación diferentes: el clasicismo y el conexionismo.

La pregunta del millón de euros es: ¿qué estructura tienen las representaciones en el cerebro? O visto desde la perspectiva de la manipulación de esas representaciones, ¿cómo codifica información un sistema cognitivo? De acuer-

10 Ver Newell (1990). Formas de explicar la cognición van desde teorías clásicas “duras” (Newell y Simon, 1972) y modelos híbridos (Anderson, 1993) hasta redes neuronales conexionistas no clásicas (O’Reilly y Munakata, 2000). Hay incluso quien propone una competición, un *decathlon cognitivo* (Anderson y Lebiere, 2003), que permita evaluar hipótesis arquitectónicas rivales. En Calvo Garzón (2003b) exploró el conexionismo no-clásico en el contexto del decathlon cognitivo propuesto por Anderson y Lebiere (2003). En cualquier caso, como veremos más adelante, sea cual sea la clase de restricciones estructurales que prefiramos, todas ellas incurren en el mismo tipo de problemas, por tratarse de modelos elaborados desde el prisma cognitivista.

do con la conocida hipótesis del *Lenguaje del Pensamiento*, la mejor metáfora del modo en que la información se codifica es el lenguaje humano. Al igual que las oraciones de un lenguaje natural tienen una estructura, que es lo que nos permite combinar y recombinar palabras para formar nuevas oraciones, las representaciones mentales necesitan una estructura combinatoria que permita manipular las representaciones mentales en función de sus propiedades sintácticas¹¹. De acuerdo con esta metáfora lingüiforme, la *independencia del contexto* es el rasgo fundamental que permite explotar la sintaxis combinatoria. Las representaciones mentales están formadas por un inventario de constituyentes atómicos, de tal manera que los constituyentes aparecen en distintos pensamientos como ejemplares sintácticamente idénticos, portando el mismo contenido conceptual. Su combinación y recombinación es lo que permite al agente cognitivo producir y comprender nuevos pensamientos.

De esta manera, según la hipótesis del *Lenguaje del Pensamiento*, el pensamiento PACO QUIERE A LA CHICA está relacionado de manera sistemática con LA CHICA QUIERE A PACO, en el sentido de que nuestra capacidad de producir o comprender el primer pensamiento está directamente conectada con nuestra capacidad de producir o comprender el segundo. Esto se puede explicar si ambos pensamientos están formados por los mismos constituyentes (PACO, QUIERE, A, LA, CHICA); constituyentes que desempeñan el mismo papel semántico en los dos pensamientos (esto es, con independencia del contexto). Así, existen estados físicos en mi cerebro que se corresponden con estos pensamientos y que podemos analizar en términos de otros estados físicos que se corresponden a su vez con sus componentes atómicos (PACO, CHICA, etc.)¹².

11 De acuerdo con la hipótesis del *Lenguaje del Pensamiento* (LOT, en inglés), el estudio de las habilidades cognitivas superiores, y en particular, la explicación de la sistematicidad, productividad y coherencia inferencial del pensamiento, precisa del tratamiento de representaciones mentales de carácter simbólico. Decimos que nuestros pensamientos son sistemáticos en la medida en que nuestra capacidad de producir o comprender el pensamiento AB (en lo sucesivo utilizaré mayúsculas para expresar conceptos) está directamente conectada con nuestra capacidad de producir o comprender BA. Por otro lado, los pensamientos son productivos en el sentido de que tenemos la habilidad de producir o comprender un número indefinido de pensamientos de creciente complejidad: A, AB, ABC, ...—si bien restricciones neurofisiológicas respecto a la capacidad del hardware biológico aparecerán irremisiblemente. Por último, el pensamiento manifiesta coherencia inferencial ya que la producción, o por lo menos la comprensión, del pensamiento A&B ocasiona la comprensión del pensamiento A y del pensamiento B (ver Fodor, 1975; 1987; y Fodor y Pylyshyn, 1988).

12 En el marco Fodoriano, los conceptos son categorías de LOT, y nuestro empleo de conceptos

En definitiva, el concepto clásico de constituyente viene a decir que las partes atómicas son transportables. Se trata de un modelo en el que los símbolos, entendidos como aquellos objetos físicos discretos que se encuentran en una relación de representación con respecto a un estado de cosas del mundo, son almacenados en una memoria, pudiendo ser recuperados y transformados mediante algoritmos que especifican cómo componer dichos símbolos sintácticamente y cómo transformarlos. Para el clasicismo, la manifiesta coherencia inferencial que exhibe la cognición superior y que se manifiesta por ejemplo cuando la producción, o por lo menos la comprensión, del pensamiento A&B ocasiona la comprensión del pensamiento A y del pensamiento B, sólo puede explicarse desde este prisma. La concepción que subyace a esta postura es que el pensamiento puede entenderse como un tipo de procesamiento inferencial de tipo lógico¹³. En resumidas cuentas, podríamos sintetizar la visión clasicista en el siguiente eslogan:

Clasicismo:
La cognición es manipulación de símbolos
de acuerdo con reglas algebraicas

no es otra cosa que la ocurrencia de ejemplares de LOT que pertenecen a dichas categorías. Dicho en jerga filosófica, nuestros conceptos son palabras-TIPO de LOT, y nuestro empleo de conceptos consiste en la ocurrencia de palabras-EJEMPLARES de LOT. La hipótesis de trabajo de LOT es que debe existir algún tipo de relación causal entre las cadenas de LOT y las cadenas de castellano, en nuestro caso, que refleja una similitud sintáctica entre cadenas de LOT y cadenas de castellano. Existen distintas versiones de LOT (el lector interesado puede consultar Field (1978) y Harman (1973), aunque el tratamiento explícito más temprano se encuentra en Sellars, 1968). No es necesario discriminar entre versiones. El rasgo que es compartido por todas ellas y que nos interesa aquí es el de “constituyente clásico” (ver abajo). Para una rápida revisión de algunos problemas a los que se enfrenta el seguidor de LOT, ver Clark (1994).

13 La siguiente cita muestra la literalidad del manifiesto Fodoriano: “[The] symbol structures in a Classical model are assumed to correspond to real physical structures in the brain and the combinatorial structure of a representation is supposed to have a counterpart in structural relations among physical properties of the brain... [The] relation ‘part of’, which holds between a relatively simple symbol and a more complex one, is assumed to correspond to some physical relation among brain states” (Fodor y Pylyshyn, 1988, p. 13). Con esta declaración de intereses, el clasicismo es sinónimo de computación programada, en tanto en cuanto el concepto clásico de constituyente explota la metáfora del ordenador digital. El objetivo de la computación programada es derivar un

Sin embargo, en los años 80 se produce una revolución en la ciencia cognitiva que desemboca en la consolidación del conexionismo como modelo computacional de cálculo alternativo al clasicismo. El objetivo es explicar los complejos patrones de la conducta humana sin recurrir a símbolos ni algoritmos explícitos. Los fundamentos de este nuevo paradigma derivan de observaciones sobre la organización del procesamiento de información en el cerebro, entendiendo la cognición como el resultado emergente de la interconectividad de unidades básicas de procesamiento dispuestas en una red neuronal conectada en paralelo¹⁴. Con el fin de presentar el concepto de constituyente no-clásico que explota el paradigma conexionista, en contraposición al de constituyente clásico, es preciso presentar antes los componentes y la dinámica básica de una red neuronal.

Las unidades en una red neuronal conexionista reciben señales de otras unidades o del exterior mediante conexiones de distinta polaridad y “peso” (el peso indica la fuerza de la conexión entre unidades). Las unidades pueden ser excitadas o inhibidas en función de las conexiones existentes y de sus valores, adquiriendo nuevos valores de activación que pueden tener como resultado la emisión de una señal (figura 2).

algoritmo en un lenguaje de programación que permita resolver un problema en cuestión. Debido al mayor grado de familiaridad del lector medio para con el paradigma clásico, dedicaré la mayor parte de esta sección a introducir la alternativa conexionista en mayor detalle. Por otro lado, aunque son muchos los problemas a los que ha de enfrentarse el clasicismo, no voy a detenerme en este punto. El lector interesado puede consultar Winograd y Flores (1986), Deyfrus (1992, edición revisada), o más recientemente Bickhard y Terveen (1995).

14 Podemos pensar en las unidades como si fueran “neuronas artificiales” y en las conexiones entre ellas como “sinapsis” entre neuronas artificiales. Hablaré indistintamente de redes conexionistas o de redes neuronales artificiales para referirme al procesamiento distribuido en paralelo tal y como se define en la llamada *Biblia del Conexionismo* (Rumelhart, McClelland y col., 1986; McClelland, Rumelhart y col. 1986). Desde una perspectiva histórica sería más preciso decir que ambas hipótesis, la clásica y la conexionista, surgieron en la misma época. Los fundamentos de la neurocomputación ya se encuentran en McCulloch y Pitts (1943) y Hebb (1949). Otro hito histórico de la disciplina es Rosenblatt (1959). Sin embargo con la devastadora crítica de Minsky y Papert (1969) se produjo un estancamiento en el paradigma conexionista del que no se salió hasta los 80 con la *Biblia del Conexionismo*. Para una introducción al conexionismo orientada a la comunidad filosófica, ver Bechtel y Abrahamsen (1991). Otra introducción para psicólogos se encuentra en McLeod y col. (1998). Un recorrido crítico más reciente que abarca desde los orígenes en McCulloch y Pitts (1943), y Hebb (1949), hasta las disputas de los últimos años se puede consultar en Sharkey y Sharkey (2008). Asimismo, el lector interesado en la metodología de las neurosimulaciones puede consultar Plunkett y Elman (1997), y McLeod y col. (1998).

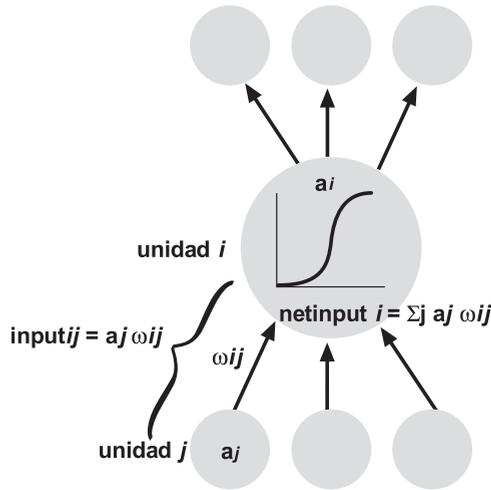


Figura 2: Operación computacional básica en una red neuronal. La unidad i (a) integra los inputs de las unidades de la capa anterior, creando un netinput; (b) transforma el netinput en el valor de activación a_i ; y por último (c) envía la señal de salida a_i a las unidades siguientes.

En función de cómo se conectan las unidades entre sí, existen muchos tipos de arquitecturas conexionistas¹⁵. En una red *feedforward* (ver figura 3, a continuación) las unidades se organizan en una capa input, una capa output, y una o más capas ocultas, de tal manera que las unidades de cada capa se conectan exclusivamente con unidades de la siguiente capa (en el caso de la capa output, el resultado representará la respuesta del sistema). El patrón de activación de cualquier capa se corresponde con la secuencia de valores de activación de las unidades que configuran la capa, pudiendo tratar de este modo los patrones de activación como vectores en un espacio n -dimensional, donde n corresponde al número de unidades en la capa.

Cada unidad suma la información de las unidades de la capa anterior, realiza una computación a partir de esta suma (ej., decidir si excede un umbral determinado), y pasa el resultado a las unidades de la capa siguiente (figura 2), o

¹⁵ Podemos distinguir grosso modo entre redes *feedforward* y redes interactivas. Mientras que en una red estándar acíclica *feedforward* la activación fluye en un único sentido (de las capas inferiores a las superiores), en las redes interactivas la activación fluye cíclicamente. Bechtel y Abrahamsen (1991) exponen algunas de las más relevantes (para un resumen taxonómico, ver Hanson y Burr, 1990).

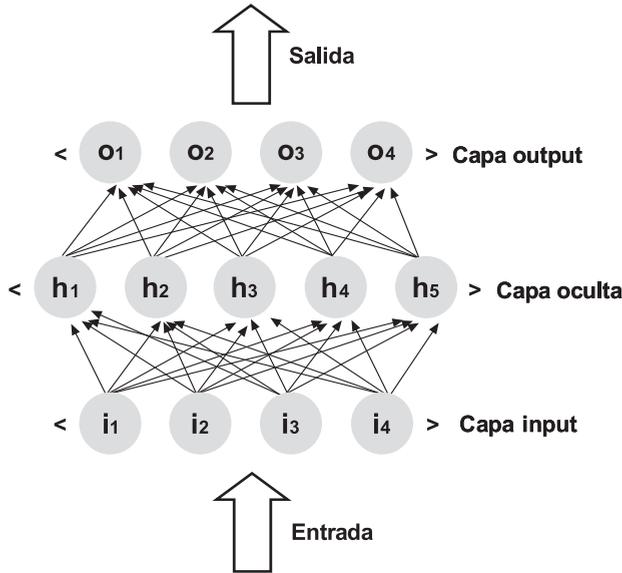


Figura 3: Arquitectura de una red *feedforward* con patrones de activación vectorial input, oculto y output, $\langle i_1, i_2, i_3, i_4 \rangle$, $\langle h_1, h_2, h_3, h_4, h_5 \rangle$, y $\langle o_1, o_2, o_3, o_4 \rangle$, respectivamente.

constituye parte de la respuesta del sistema, en el caso de que se trate de unidades output. Así, inicialmente los valores de activación de la capa input se propagan a las unidades de la siguiente capa. Cada unidad oculta sufrirá de este modo una variación en su nivel de activación. Los resultantes valores de activación ocultos pueden verse como un nuevo patrón de activación a tratar como un vector oculto en un espacio n-dimensional. Nuevamente, las unidades ocultas propagarán su activación a las unidades de la siguiente capa (oculta o output). Similarmente, se formará un nuevo vector n-dimensional como resultado de los nuevos niveles de excitación o de inhibición de las n unidades ocultas o output. De este modo la activación se propaga desde la capa input a la output, pasando por la/s capa/s oculta/s.

De esta manera, la resolución de un problema pasa por la transformación de la información input en información output. El patrón de activación de la primera capa representa el estímulo presentado al modelo, el cual se transforma de manera gradual hasta producir el patrón de activación de la última capa, que constituiría la respuesta del modelo ante el patrón presentado al sistema. Las redes neuronales operan de esta manera ejecutando transformaciones vec-

toriales. Las computaciones se pueden así entender como la transformación de un vector de activación en otro¹⁶.

De acuerdo con el conexionismo no existen ni símbolos almacenados ni reglas explícitas que gobiernen su manipulación. Las categorías cognitivas clásicas (proposiciones, conceptos, etc.) son producto de multitud de procesos microcognitivos que ocurren en forma de intrincadas redes de unidades elementales que funcionan en paralelo (simultáneamente). La memoria, de este modo, no consiste en piezas de información independientes unas de otras que se almacenan por separado. En una red neuronal, la extracción de información conlleva el flujo de excitaciones e inhibiciones a través de todo sistema, afectando a todo lo que tenga relación con esa información. El resultado es la activación de muchas unidades diferentes¹⁷.

16 Para comprender la diferencia entre modelos clásicos y conexionistas es preciso prestar atención al modo en que se almacena la información en el sistema y a la forma en que “aprende” una red neuronal a detectar la estructura de los estímulos con que ésta se activa inicialmente. Si el valor de activación de una unidad depende de la activación que recibe de las unidades con las que se conecta, y de la fuerza con que esa señal es recibida a través de la conexión sináptica (figura 2), modificando el valor de esas conexiones, podemos modificar la manera en que la red procesa la información. Teniendo en cuenta que la influencia de una unidad sobre otra depende de la fuerza de la conexión, los pesos de las conexiones entre unidades es lo que permite a la red recrear todos los patrones de activación presentados. De esta manera, el aprendizaje consiste en el ajuste de los pesos de las conexiones con el fin de que la red produzca el patrón de activación correcto en cada caso. Rumelhart, Hinton y Williams (1986) articularon inicialmente el algoritmo de aprendizaje de *retropropagación* como una generalización de la regla delta empleada en el perceptrón de Rosenblatt (para una ilustración, ver Bechtel y Abrahamsen, 1991, y Elman y col., 1996). Es preciso tener en cuenta que las redes neuronales artificiales conllevan una gran simplificación con respecto a las “redes biológicas”. Algoritmos de ajuste de pesos, como retropropagación, dependen de alguna forma de supervisión externa. El punto crítico es que los pesos de las conexiones sólo pueden ajustarse mediante la derivación de una medida de error a partir de la respuesta “deseada”. Este rasgo no tiene una implementación biológica directa conocida. Por otro lado, es preciso recordar que restricciones impuestas por la investigación neurobiológica están permitiendo refinar más y más los modelos conexionistas. Hay análogos conexionistas de la sinaptogénesis, o redes que permiten explicar aspectos del cortex visual, de las afasias, etc. (ver Elman y col., 1996).

17 La información que la red procesa no se encuentra explícitamente almacenada en ninguna parte del sistema. Toda la información presente es aquella que se representa *activamente* como un patrón de activación en cada momento dado de procesamiento. Cuando la red no emplea dicha información, ésta no está en ningún sitio. En última instancia, podemos decir que el conocimiento que la red tiene reside en los pesos que han sido generados durante la fase de aprendizaje. Decimos así que la memoria en un modelo conexionista se encuentra implementada en la matriz de

Así, los modelos clásicos y los conexionistas difieren en el modo de representar los constituyentes. Mientras que en el modelo simbólico los constituyentes son independientes del contexto, los constituyentes conexionistas son subsimbólicos, esto es, dependientes del contexto¹⁸. Las redes conexionistas que aquí nos interesan representan estados como conjuntos de valores *distribuidos*, de tal manera que una unidad oculta tomada por separado carece de interpretación semántica¹⁹. Por otro lado, la información se almacena de manera superpuesta. Las redes conexionistas aprenden mediante el ajuste de la matriz de pesos con el fin de producir la correlación input/output apropiada para todo el conjunto de entrenamiento. De esta manera, todo el conocimiento que tiene la red se almacena en una única matriz de conexiones, no exhibiendo el procesamiento el carácter composicional de las representaciones clásicas. Las unida-

conexiones. De esta manera, memoria y procesamiento en teoría conexionista son dos caras de la misma moneda. La memoria consiste en los pesos de las conexiones entre unidades (los patrones de conectividad con que diseñamos la red). Por otro lado, el procesamiento viene determinado por el mismo conjunto de pesos.

18 Un ejemplo clásico de sensibilidad al contexto se encuentra en la historia de la ‘taza-de-café’ de Smolensky (1991) (ver también McClelland y Kawamoto, 1986). La dependencia contextual queda codificada por la localización espacial precisa de un patrón de activación en el espacio oculto representacional.

19 Las activaciones de una red neuronal pueden interpretarse de manera localista o distribuida. En un modelo localista, cada unidad sirve para representar un concepto (ver, por ejemplo, el modelo de reconocimiento de palabras de Rumelhart y McClelland, 1982). Estas unidades localistas son atómicas, no estando sujetas a posterior descomposición. La semántica a asignar en una interpretación localista es función de los patrones de conectividad entre estas unidades atómicas. Como las unidades localistas son atómicas, y la semántica de la red en estos modelos se desarrolla como resultado de la combinación de estas unidades no descomponibles, las representaciones localistas son equivalentes funcionalmente a las representaciones simbólicas del clasicismo. Por otro lado, en un modelo distribuido las unidades, tomadas individualmente, no pueden interpretarse semánticamente. Las representaciones se procesan simultáneamente a través de varias (¡muchas!) unidades. Así, como un único agregado de unidades puede representar muchos conceptos, es preciso atender al patrón de activación en su conjunto a la hora de saber qué concepto se está representando en cada momento. En las redes totalmente distribuidas, el patrón de activación más pequeño semánticamente interpretable corresponde al conjunto completo de unidades ocultas. No es posible identificar partes discretas en los patrones de activación ocultos con componentes particulares de las oraciones procesadas. Sin embargo, hay redes recurrentes totalmente distribuidas que permiten capturar la estructura gramatical de las oraciones a las que son expuestas (Elman, 1990; 1998). La estructura gramatical queda reflejada codificando variaciones gramaticales como pequeñas variaciones en los patrones de activación relevantes en el espacio representacional. La contribución sintáctica que cada palabra hace a la oración se mide por el nivel de activación de la palabra, tal y como queda codificado en el espacio oculto.

des individuales y los pesos portan información de grano más fino que la que se procesa clásicamente. Los constituyentes pueden mantenerse dependientes del contexto mediante la codificación de la localización precisa de cada patrón de activación. De acuerdo con esta visión alternativa, no hay una representación canónica de cada constituyente conceptual que sea común a todas las estructuras moleculares en las que el constituyente haga acto de presencia. En su lugar, encontraremos una representación distinta que codifica *cada* contexto molecular distinto, en función de las regularidades estadísticas presentadas a la red y las subsiguientes modificaciones realizadas en la matriz de conexiones²⁰. El eslogan ahora podría ser:

Conexionismo:
La cognición es manipulación de subsímbolos
de acuerdo con reglas estadísticas

En definitiva, la cognición es computación de entidades físicas interpretables semánticamente (esta vez, subsímbolos), y consiste en la manipulación de dichos subsímbolos mediante reglas estadísticas implícitas que se implementan en la matriz de conectividad de la red neuronal. De esta manera, el conexionismo ofrece una visión de la cognición que difiere sustancialmente de la visión expuesta por Fodor. Sin embargo, más allá de las diferencias arquitectónicas entre modelos clásicos y conexionistas, existe un tipo de implicaciones filosóficas que tiene que ver de manera más general con la psicología popular, y que nos servirá para evaluar el pulso entre clasicistas y conexionistas.

20 De esta manera, las representaciones idiosincráticas no son palabras-EJEMPLAR de la misma palabra-TIPO, tal y como mantiene LOT y su forma clásica de constituyente (ver fn. 12). Mientras que en LOT, ejercitar el concepto expresado por “chica” es la instanciación de su correspondiente expresión en LOT (CHICA), la hipótesis de trabajo conexionista es que el repertorio conceptual que expresa “chica” en una preferencia de “Paco quiere a la chica” es el estado interno real con el que la red conexionista equipara “chica”. Por supuesto, son numerosos los intercambios entre defensores del modelo clásico y el conexionista. Fodor y Pylyshyn (1988), Smolensky (1988), Fodor y McLaughlin (1990), y Smolensky (1991) son algunas de las referencias clásicas del debate. Ver también Pinker y Prince (1988) para una poderosa crítica del paradigma conexionista que se centra, no en argumentos de corte Fodoriano, sino en la adquisición del lenguaje en niños.

LA ELIMINACIÓN DE LO MENTAL

La psicología popular trata de explicar la conducta humana en términos de nuestras creencias, deseos, temores, anhelos, etc. Para situarnos, podemos pensar en la psicología popular de la misma manera que hablamos habitualmente de los remedios tradicionales (lo que podríamos denominar “medicina popular”) en contraposición al conjunto de conocimientos científicos y técnicos que recibe un estudiante de medicina y cirugía en cualquier universidad. Un ejemplo servirá: a la hora de tratar de explicar por qué he guiñado el ojo dirigiéndome a la primera fila, podría argumentar que *creía que* los niños ahí sentados eran mis hijos y que *esperaba que* mi gesto despertara una sonrisa en ellos, ya que no se *esperarían que* a mitad de charla fuera a hacer eso (todo ello teniendo presente que de no haber habido familiares míos sentados en primera fila, no habría osado hacer un gesto tan informal porque *pensaría que* no procedía, etc.). El rasgo crítico es que la explicación de un patrón de conducta se basa en la actitud que el sujeto mantiene respecto a una determinada proposición: mi creencia de *que p*, mi deseo de *que q*, etc. (estados como mi creencia de *que la proposición p* reciben en filosofía el nombre de “actitudes proposicionales”).

Ahora bien, como hemos visto en la sección anterior, las representaciones almacenadas en una red neuronal no son representaciones de *proposiciones abstractas*. Los modelos conexionistas no suscriben la teoría computacional de la mente en su formato clásico. La actividad cognitiva no consiste en operaciones formales sobre representaciones internas de acuerdo con reglas sintácticas. Las representaciones que emergen en una arquitectura conexionista pueden estar distribuidas entre todas las unidades de procesamiento, de tal manera que para identificar el contenido semántico de un patrón de activación se requiere medir el nivel de activación de todas las unidades que conforman la capa. De-

cimos en estos casos que la representación está “totalmente distribuida”. Por otro lado, decimos que una red es “superpuesta” cuando la información que procesa se almacena en la matriz de conectividad constituida por todas las conexiones existentes entre todas las unidades. De este modo, cuando las representaciones son totalmente distribuidas y se almacenan de manera superpuesta, podemos hablar de redes *totalmente superpuestas*. El resultado es que las redes neuronales totalmente superpuestas no confirman la teoría computacional de la mente, al menos en su formato clásico.

Algunos filósofos arguyen que si las redes conexionistas *totalmente superpuestas* reflejan adecuadamente la dinámica de la cognición, la psicología popular está equivocada, y debemos eliminar las actitudes proposicionales de nuestras explicaciones científicas²¹. En particular, desde el momento en que la psicología popular individualiza los estados mentales en términos de su contenido proposicional, la psicología popular se compromete con la tesis de la modularidad proposicional. Según esta tesis, las actitudes proposicionales son: (I) semánticamente interpretables; (II) funcionalmente discretas (podemos adquirir o perder determinadas actitudes proposicionales, sin alterar otras); y (III) causalmente discretas, de tal manera que pueden desempeñar un poder causal en la producción de otras actitudes proposicionales y finalmente en la producción de la conducta observable.

Hay numerosos modelos en la literatura clásica que asumen la tesis de la modularidad proposicional. Podemos pensar en una red semántica de representación de la memoria del tipo de Collins y Quillian (1972). En el modelo de Collins y Quillian (figura 4), las proposiciones se representan mediante nodos junto con etiquetas que los enlazan con varios conceptos.

De esta manera, las proposiciones almacenadas en la memoria forman un entramado de estados funcionalmente discretos (cada nodo por separado) y semánticamente interpretables (etiquetas enlazadas). Una característica de estos modelos es que podemos añadir/suprimir proposiciones particulares sin necesidad de hacer reajustes en el resto de la red²².

21 Éste es el tipo de argumentación condicional expuesto por Ramsey y col. (1991). El antecedente analítico a esta línea de argumentación conexionista se encuentra en Feyerabend (1963), Rorty (1970), y Churchland (1981).

22 Su contenido es evaluable y proyectable en el sentido de poder figurar en leyes causales y soportar generalizaciones nomológicas (Goodman, 1965).

Red semántica de Collins y Quillian

Proposiciones

- 1 Los perros tienen pelo
- 2 Los perros tienen patas
- 3 Los gatos tienen pelo
- 4 Los gatos tienen patas

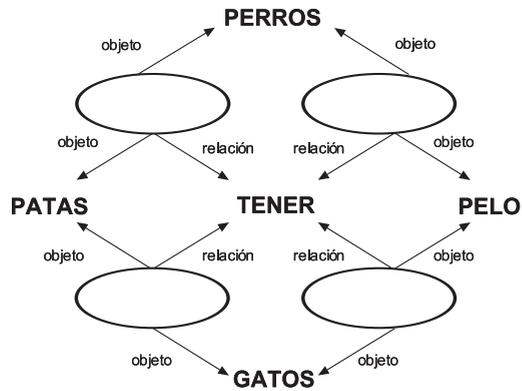


Figura 4: Representación de la memoria según la red semántica de Collins y Quillian (1972) (adaptado de Ramsey y col., 1995).

Red semántica de Collins y Quillian

Proposiciones

- 1 Los perros tienen pelo
- 2 Los perros tienen patas
- 3 Los gatos tienen pelo

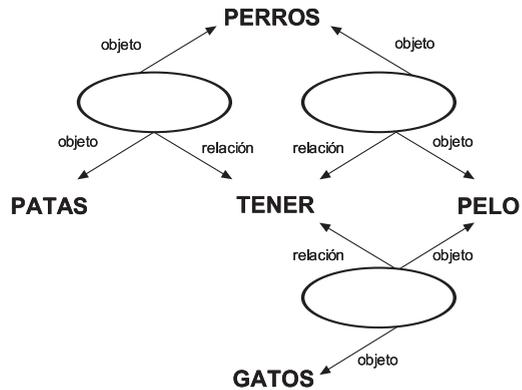


Figura 5: Representación de la memoria según la red semántica de Collins y Quillian (1972) con una proposición eliminada (adaptado de Ramsey y col., 1995).

En la sección anterior vimos que una red neuronal manipula subsímbolos cuyo contenido *representacional* está totalmente distribuido. De esta manera, es claro que las redes neuronales totalmente superpuestas son compatibles con (I). Esto es, son interpretables semánticamente²³. ¿En qué consiste pues la supuesta incompatibilidad entre el conexionismo y las actitudes proposicionales? La línea argumental eliminativista explota que las redes neuronales totalmente superpuestas son incompatibles con los rasgos (II) y (III) de la tesis de la modularidad proposicional. Es decir, si bien es cierto que podemos interpretar semánticamente los patrones de activación que se generan como fruto del procesamiento y de la modificabilidad sináptica en la red, no encontramos en la red patrones de activación a identificar con actitudes proposicionales que sean funcionalmente discretos y que tengan eficacia causal.

En la dinámica de una red neuronal no encontramos actitudes proposicionales lógicamente independientes con distintos poderes causales. Al estar la información almacenada de manera totalmente superpuesta, no tiene sentido hablar de la eficacia causal que desempeña la representación de una proposición en particular. El contenido de cada proposición viene determinado por la superposición de todos los recursos representacionales; recursos que son función de todo el rango de patrones input/output vistos. En definitiva, el empleo de representaciones totalmente distribuidas y técnicas de almacenaje de super-

23 Paul Churchland, sin ir más lejos, un ávido defensor del materialismo eliminativista, ha articulado al mismo tiempo una teoría de la representación mental conocida como “semántica de espacios vectoriales”, siendo la idea básica que un sistema de procesamiento de información representa determinados aspectos de la realidad mediante posiciones en un espacio caracterizado geoméricamente: “[the] brain represents various aspects of reality by a position in a suitable state space, and the brain performs computations on such representations by means of general coordinate transformations from one state space to another” (Churchland, 1986, p. 280). Churchland propone que interpretemos los conceptos como puntos en el espacio que determina la dimensionalidad de la capa oculta. Estos puntos se corresponderían con las puntas de los vectores determinados por los niveles de activación de las unidades que componen la capa oculta. Las características semánticas de un concepto pueden de este modo verse en función del lugar que el concepto (o sea, el punto) ocupa en el hiperespacio oculto caracterizado geoméricamente. Churchland propone así que podemos hablar de similitud entre conceptos en términos de la proximidad de sus respectivas posiciones absolutas en el espacio representacional, identificado en relación a un número de dimensiones semánticamente relevantes. Churchland ha desarrollado su modelo con el fin de responder a una serie de retos. En particular, Churchland (1996; 1998) ahora define la similitud semántica, no en término de posiciones absolutas, sino relativas (para una crítica a la extensión del modelo de Churchland, ver Calvo Garzón, 2000b; 2003).

posición es incompatible con la tesis de la modularidad proposicional. La conclusión eliminativista es, por un lado, que si el conexionismo está en lo cierto, la tesis de la modularidad proposicional no se cumple (no encontramos eficacia causal discreta entre los estados del sistema), por lo que la psicología popular está equivocada; y segundo, que si la psicología popular está equivocada, deberíamos eliminar las actitudes proposicionales de nuestra ontología científica.

Hasta ahora hemos visto que, como hipótesis empíricas sobre la arquitectura de la cognición, el clasicismo y el conexionismo difieren. Según el clasicismo, la actividad cognitiva se reduce a la manipulación de símbolos de acuerdo con reglas explícitas; según el conexionismo, a la manipulación de subsímbolos de acuerdo con reglas estadísticas implícitas. Y sin embargo hay razones para pensar que no son tan distintos, o que sus diferencias no tienen por qué llevarnos necesariamente a mantener tesis tan radicales como la de la eliminación de lo mental.

Recordando “los niveles de Marr” de la introducción, se podría argumentar que mientras que el clasicismo es una hipótesis centrada en el nivel algorítmico, al conexionismo le interesa el nivel de implementación del sistema. Después de todo, introdujimos el conexionismo diciendo que sus fundamentos derivaban de observaciones sobre la organización del procesamiento de información en el cerebro. Parece pues que clasicismo y conexionismo pueden ser dos caras de la misma moneda, constituyendo el conexionismo una implementación del clasicismo. El clasicismo, a su vez, especifica, a un nivel superior de descripción, qué restricciones deben operar “desde arriba”. Por ejemplo, se podría argumentar que para que un sistema mantenga coherencia inferencial debe manipular constituyentes que sean independientes del contexto. No obstante, no todos aceptan el debate planteado en estos términos. El conexionismo pretende ser una alternativa *algorítmica* al clasicismo.

Desafortunadamente, visto desde esta perspectiva, el debate entre cognitivistas clásicos y conexionistas está lleno de ejemplos que ilustran la situación de *impasse* que vive la ciencia cognitiva en la actualidad: la ortodoxia clasicista sigue explotando argumentos tipo *pobreza-del-estímulo* con el fin de identificar habilidades cognitivas que se resistan a una explicación puramente estadística: “¿Cómo puede una red neuronal inducir las regularidades gramaticales que subyacen al input lingüístico que recibe si los patrones presentados al sistema, en base a los cuales modifica su matriz de conectividad, son una mínima parte del total de combinaciones gramaticales legales?”. El conexionista, por otro

lado, típicamente contesta con datos ecológicos que pueden ser explotados estadísticamente y que permite a las redes conexionistas retener su adecuación computacional. Dicho claro y pronto, la estrategia conexionista ha consistido en mostrar que el estímulo no es tan pobre después de todo, y que el procesamiento distribuido en paralelo puede dar cuenta de fenómenos de generalización que tradicionalmente han apuntado a una explicación clasicista²⁴. El clasicista, tal y como ha señalado Gary Marcus (2001) en repetidas ocasiones, contesta a su vez que el conexionista es sólo capaz de mantener su adecuación computacional bien a base de alimentar a su red neuronal con símbolos explícitos en las representaciones de la capa input, bien mediante formas de aprendizaje supervisado a través de las cuales el sistema implementa un algoritmo clásico.

Con la re-emergencia del conexionismo en los 80, se llegó a creer que nos encontrábamos ante un nuevo paradigma a la hora de entender la cognición. Sin embargo los numerosos intercambios entre cognitivistas clásicos y conexionistas hacen presagiar que, después de todo, clasicismo y conexionismo son el mismo perro (cognitivista) con distintos collares. La arquitectura de la cognición está en disputa, pero los supuestos computacionales no son tan distintos.

24 La estrategia de defensores del clasicismo como Fodor y Pylyshyn (1988) o Marcus y col. (1999), por ejemplo, es la de identificar habilidades cognitivas que desafíen una interpretación estadística. Por su parte, los conexionistas (Seidenberg y Elman, 1999) contestan identificando datos ecológicos susceptibles de ser explotados estadísticamente con el fin de mantener la adecuación computacional de sus modelos respecto de los clásicos. Debates que ilustran a la perfección esta dialéctica son el “gran debate” sobre el tiempo pasado de los verbos en inglés (Pinker y Ullman, 2002; Ramscar, 2002), el debate de la sistematicidad (Fodor y Pylyshyn, 1988; Elman, 1998), el debate álgebra-versus-estadística (Marcus y col., 1999; Seidenberg y Elman, 1999; Calvo y Colunga, 2003), y en los últimos 4 años, el debate sobre la segmentación del habla (Peña y col., 2002; Calvo Garzón, 2005). Que esta dialéctica de confrontación es hasta cierto punto espuria se puede advertir si recapitamos brevemente sobre qué quiere decir que un sistema sea computacionalmente adecuado. En un sentido no trivial, se trata de que las alternativas arquitectónicas sean computacionalmente equivalentes respecto a sus competencias futuras. Por definición, una competencia debe tener un elemento de generalización que vaya más allá de la base observacional estipulada de una manera no trivial. Entendemos, por ejemplo, que hay signos de una competencia gramatical cuando el sistema es capaz de generalizar más allá del corpus gramatical de habituación. Esto es, cuando distinguimos como gramaticales o no gramaticales estructuras no procesadas con antelación. Sin embargo, en este sentido, trivialmente se sigue la pobreza del estímulo. Por definición la base observacional es limitada y sesgada con respecto al dominio de aplicación de la competencia cognitiva. De no ser así, tendríamos una asimetría entre explicación y predicción en los modelos en liza.

En resumidas cuentas, el debate se reduce a saber si la cognición acaba siendo manipulación de símbolos de acuerdo con reglas algebraicas explícitas, tal y como piensa el clasicismo, o manipulación de subsímbolos de acuerdo con reglas estadísticas implícitas. Mientras que el computacionalismo sintáctico clásico precisa de constituyentes libres de contexto, el computacionalismo conexionista se basa en la explotación de transformaciones vectoriales. Sin embargo, la diferencia estriba meramente en que el contenido, en el segundo caso, es subsimbólico, y los constituyentes, dependientes del contexto. La disputa, como puede observarse, es entre defensores del paradigma del procesamiento de información de corte representacional. A pesar de ello, y aunque las cosas no son o blancas o negras, el debate entre clasicistas y conexionistas se ha movido en esta dirección desde la re-emergencia del conexionismo a mitad de los 80. Es por ello que el debate en los términos actuales se encuentra en un callejón sin salida. Clasicistas y conexionistas cuestionan la arquitectura de la cognición, pero los pilares de la disciplina permanecen inalterados. O por lo menos así lo parecía hasta la aparición de una tercera aproximación al estudio de la cognición: la ciencia cognitiva corporeizada.

CIENCIA COGNITIVA CORPOREIZADA

El estancamiento al que aludíamos al final de la sección anterior ha venido acompañado del desarrollo de nuevas hipótesis arquitectónicas. El concepto de arquitectura cognitiva edificado sobre los pilares de las nociones de computación y representación presupone una separación entre *agente* cognitivo y medio. La cognición es “manipulación de ...”; actividad ésta que requiere de un agente que opere internamente sobre la información que recoge a través de sus módulos periféricos. Las dudas en torno a este modelo computacional han hecho que muchos investigadores encuentren en la ciencia cognitiva corporeizada una forma de escapar del callejón sin salida en el que vimos se encontraban clasicistas y conexionistas. La ciencia cognitiva corporeizada rechaza la metáfora de la cognición como mecanismo de procesamiento de información, evitando recurrir así a la teoría computacional de la mente en sus formatos simbólico o subsimbólico²⁵. La cognición se entiende como fenó-

25 A diferencia de los modelos basados en el paradigma de procesamiento de información, la ciencia cognitiva corporeizada trata de modelar y explicar la conducta de sistemas concretos identificándolos con conjuntos de variables que cambian de manera continua en el tiempo. Los modelos dinamicistas (Thelen y Smith, 1994) son los que más eco han tenido en la literatura filosófica, y tienen un papel destacado en áreas tales como control motor, percepción, desarrollo, y robótica situada. Port y Van Gelder (1995) es una buena introducción a la literatura específicamente dinamicista. Para una revisión de los principios del dinamicismo aplicados a distintas áreas, como son la teoría del desarrollo, la interacción social, la emergencia, el desarrollo motor, o la temporalidad, ver Calvo y col. (2008a; 2008b). Sin embargo, en lo sucesivo hablaré de “ciencia cognitiva corporeizada” (Clark, 1997) en un sentido más general. Este sentido incluye el “no-cognitivismo” de la percepción ecológica (Gibson, 1979; Turvey y Carello, 1995), la cognición corporeizada (Varela, Rosch, Thompson, 1991; Clancey, 1997), la cognición distribuida (Hutchins, 1995), los sistemas simbólicos perceptivos (Barsalou, 1999), algunas clases de conexionismo (Rolls y Treves, 1999), el interactivismo (Bickhard y Terveen, 1995) y los sistemas dinámicos (Port y Van Gelder, 1995), entre otros. Por

meno auto-organizado, emergente y extendido que ha de estudiarse teniendo en cuenta, en su dimensión temporal, las interacciones continuas entre factores neuronales, corporales y medioambientales. Según este modelo, no tiene sentido distinguir entre el sistema cognitivo como tal y el medio en el que éste se desenvuelve. El resultado es la explicación de la conducta desde una perspectiva distribuida en la que cerebro, cuerpo y medio constituyen un sistema acoplado (ver fig. 6).

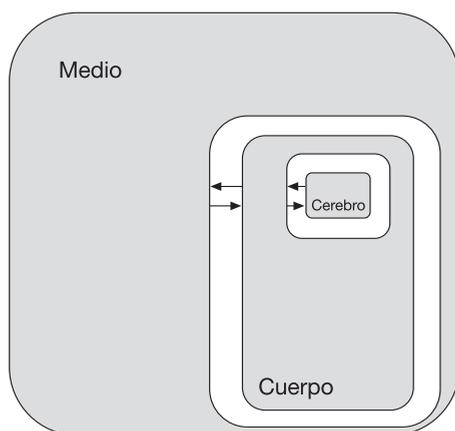


Figura 6: Cerebro, cuerpo y medio en constante y mutua interacción (figura adaptada de Beer, 1995).

Bajo este prisma, en el que tanto factores corporales como medioambientales son parte esencial de la ecuación cognitiva, podemos introducir una hipótesis de continuidad (cognitiva) atendiendo a dos puntos. Por un lado, en tanto que hipótesis empírica sobre la arquitectura de la cognición, la ciencia cognitiva corporeizada incorpora una restricción de homogeneidad según la cual:

- 1) La cognición superior está al mismo nivel que los procesos nerviosos y corporales, pudiendo explicarse, de esta manera, en términos de continuidad y temporalidad.

supuesto, esta variedad de visiones tiene su reflejo en distintos programas de investigación, si bien el objetivo último, el desarrollo de una alternativa al cognitivismo, bien de corte clasicista o conexionista, es compartido (ver Gomila y Calvo, 2008).

Por otro lado, la hipótesis incorpora un requisito de interactividad según el cual:

- ii) interacciones de niveles múltiples tienen lugar de modo continuo desde la actividad cognitiva superior hasta la percepción, el control motor, y los procesos no cognitivos en general²⁶.

De esta manera, tanto la homogeneidad como la interactividad hacen de la continuidad en un contexto de descentralización uno de los conceptos fundamentales a la hora de explicar la cognición. El interés reside en la interacción continua de todos los componentes de la ecuación cognitiva. Hablar de continuidad es simplemente el resultado de entender la actividad cognitiva como la actividad de un sistema físico compuesto (cerebro + cuerpo + medio) acoplado en el espacio-tiempo²⁷.

Una de las principales implicaciones de este proceso de descentralización es la minimización de la carga representacional del sistema²⁸. De hecho, la des-

26 En Calvo Garzón (en prensa) desarrollo esta hipótesis de continuidad (cognitiva).

27 La continuidad no es sino el resultado de entender literalmente la cognición como el resultado emergente de las interacciones que aparecen en la figura 6. Por otro lado, si la actividad cognitiva superior crece a partir de principios sensorio-motrices, la hipótesis fundamental es que no hay un momento en el desarrollo ontogénico en que esto deja de ser así. Es decir, no hay un momento de maduración en que la cognición deja de estar corporeizada. De esta manera, un modelo de mentalidad debe respetar los mismos principios de no-linealidad, dependencia temporal y continuidad generalmente invocados en las explicaciones de las interacciones corporales y la actividad neuronal (ver Freeman, 2000). El trabajo pionero en teoría del desarrollo motor de Esther Thelen ilustra a la perfección la naturaleza corporeizada a la que venimos a referirnos en esta sección. Como botón de muestra, en una serie de trabajos Thelen estudió la desaparición de los reflejos de avance neonatales (el característico pataleo de los bebés), que tiene lugar alrededor de los primeros dos meses de vida, y cuya explicación “descentralizada” tiene que ver con el depósito de grasa en las piernas del bebé conforme crece. Como se puede apreciar, la explicación resulta ser poco *cognitiva*: el mayor peso de las extremidades inferiores debido a la acumulación de grasa es lo que impide al bebé levantarlas, desapareciendo el reflejo característico de los dos primeros meses. De hecho, cuando el bebé es sumergido en agua, el reflejo reaparece (ver Thelen y col, 1994, para los detalles). En palabras de Thelen y col.: “We are not building representations at all! Mind is activity in... the real time of real physical causes”.

28 La economía representacional puede deberse al hecho de que una considerable cantidad de información se encuentra accesible en el entorno del sistema. De este modo, no es preciso almacenar un gran número de elementos representacionales manipulables que “estén por” la realidad externa con la que interaccionamos. Esta visión ha dado pie a una forma de “realismo directo”, en el sentido de que no hace falta manipular elementos internos que medien entre el sistema y la

centralización de la cognición suele muchas veces venir acompañada de una línea de argumentación de marcado carácter antirrepresentacionista. Un ejemplo ya clásico en la literatura filosófica lo proporciona el *gobernador centrífugo de Watt*, así que vamos a dar un pequeño rodeo por la revolución industrial para llegar a nuestro destino²⁹.

En una máquina de vapor el movimiento del pistón se convierte en el movimiento giratorio del motor que mueve la máquina. El problema es que la velocidad del motor fluctúa en función de una serie de factores, como son la temperatura del horno, la carga, etc., haciendo falta un mecanismo para que

realidad a la que el sistema se adapta, o ante la que el sistema se muestra competente. ¿Pero qué significa que un estado interno “está por” otro estado externo al sistema? ¿Qué quiere decir “mediar”? En definitiva, ¿qué es realmente una representación? Son muchas las respuestas que se han dado a este problema fundamental de la filosofía, y no voy a desarrollarlas aquí hoy. Un recorrido bastante exhaustivo se encuentra en Ryder (2008). Para un análisis pormenorizado de los muchos problemas del representacionismo, ver también Bickhard y Terveen (1995). Además, una condición para que un estado sea representacional sería la satisfacción de un principio de disociación según el cual para que un estado sea representacional, el estado debe a veces poder estar por cosas o sucesos que simplemente no existan o que no sean accesibles durante un intervalo de tiempo. Asimismo, debemos ser capaces de identificar estados físicos específicos con los papeles (semánticos) que desempeñan. De ahí se sigue un principio de reificación. Un estado puede contar como representacional sólo si se puede detectar y permitiendo el establecimiento de un paralelismo entre el estado en cuestión y el papel que juega en el establecimiento de una conexión entre los estados input y output. Estos dos principios se pueden elaborar, junto con un principio de mínima duración (‘principle of minimal lastingness’, Calvo Garzón, en prensa), constituyendo una definición tripartita de representación. Según este principio adicional de mínima duración, para que un estado físico que satisface el principio de conectividad sea representacional debe exhibir un grado de persistencia equivalente como mínimo al grado mostrado por el input a ser representado. Si bien este principio está relacionado con las posturas de Haugeland (1991) y Clark (1997), hay una serie de diferencias que no consideraré aquí (ver Calvo Garzón, en preparación). Podemos entonces sintetizar estos principios diciendo que un estado físico tiene contenido semántico si puede identificarse espacio-temporalmente como causalmente eficaz en la conexión de los estados input y output del sistema de tal manera que el estado en cuestión se mantiene latente mientras que el estado input que sintoniza decae o deja de estar presente.

29 Ver Van Gelder (1995; 1998). Otros ejemplos de la literatura antirrepresentacionista incluyen el trabajo pionero de Rodney Brooks con “mobile robots” (1991) o el estudio de la fonotaxis con “grillos robot” de Barbara Webb (1994). De acuerdo con Brooks, la clave está en atender a la inteligencia de manera incremental, tal y como hizo en el diseño de sus criaturas inteligentes de creciente complejidad (*mobile robots*). De esta manera desaparece la necesidad de postular estados representacionales. Tal y como Brooks apunta en su estudio: “When we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model... Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems”.

funcione a velocidad constante. Esto se puede lograr ajustando la válvula que controla la cantidad de vapor que entra en el pistón. Curiosamente, el problema al que se enfrenta la máquina de vapor nos sirve para ilustrar la diferencia entre sistemas cognitivos representacionales y sistemas extendidos no-representacionales.

Una solución al problema de las fluctuaciones en el motor pasaría, desde la perspectiva computacional, por implementar un algoritmo que: (a) midiera la velocidad del motor, (b) la comparara con la velocidad deseada, y en caso de que éstas difirieran, (c) midiese la presión del vapor, (d) calculara la alteración deseada en la presión y el ajuste necesario de la válvula reguladora, y por último (e) realizara el ajuste requerido; y así sucesivamente. De esta manera el sistema tomaría como input una serie de representaciones simbólicas de diversos aspectos del motor que manipularía con el fin de convertir los resultados en ajustes de la válvula reguladora. Es decir, explotamos un mecanismo cíclico que es computacional y representacional. Se realizan transformaciones discretas secuenciales de una estructura de datos en otra, o transformaciones vectoriales de un vector input en un vector output. Estableciendo un paralelismo con un sistema cognitivo, volvemos inexorablemente al punto de partida de esta charla: la cognición es manipulación de símbolos/subsímbolos de acuerdo con reglas algebraicas/estadísticas.

Sin embargo la solución a la que llegó James Watt fue bien distinta. Ésta consistió en conectar un eje vertical al motor (figura 7), de tal manera que la velocidad de rotación del eje depende directamente de la velocidad del motor. Al eje van cogidos dos brazos con una bola de metal en sus extremos. De esta manera, al girar el eje, la fuerza centrífuga hace que las bolas se vayan hacia fuera y hacia arriba. Como el movimiento de los brazos está conectado directamente a la válvula reguladora, el resultado es que conforme la velocidad del motor aumenta, los brazos se elevan, cerrando la válvula y restringiendo así el flujo de vapor.

Por el contrario, conforme la velocidad disminuye, los brazos caen, abriéndose la válvula y permitiendo de esta manera un mayor paso de vapor.

Es tentador pensar que el ángulo de los brazos en el gobernador de Watt en cierto modo representa la velocidad del motor, y que precisamente gracias a ello es por lo que el gobernador es capaz de controlar la velocidad. Sin embargo, esta idea sólo puede ser correcta si presuponemos que todo sistema en que se identifiquen correlaciones es un sistema representacional, lo cual es muy

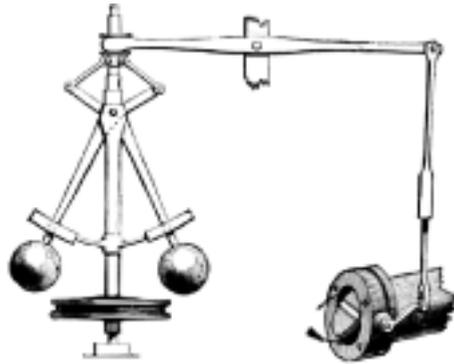


Figura 7: Gobernador centrífugo de Watt.

discutible³⁰. En cualquier caso, aunque ello fuera posible, no ayudaría a comprender la relación recíproca existente entre la velocidad del motor y el ángulo de los brazos, donde en todo momento uno determina y es determinado por la conducta del otro. El punto crítico es que, en lugar de ciclos de inputs, representaciones simbólicas/subsimbólicas, computaciones atemporales gobernadas por reglas explícitas/implícitas, y outputs, en el gobernador centrífugo de Watt tenemos la influencia mutua continua de dos cantidades en tiempo real, donde el estado de una cantidad determina continuamente cómo la otra va acelerando y viceversa.

Desde esta perspectiva, volviendo al paralelismo anterior, la cognición sería el resultado emergente de las interacciones continuas de conjuntos de variables emparejadas. La idea central que subyace a este nuevo paradigma es el rechazo, o la transformación radical, del dogma de que la mente debe describirse como computando y representando. El clasicismo y el conexionismo ignoraban la dimensión temporal de la cognición³¹. Una primera conclusión es

30 De lo poco en lo que se está de acuerdo en filosofía de la ciencia cognitiva, una cosa es que no vale con la existencia de una correlación para entender que un estado representa a otro (ver Haugeland, 1991; Clark, 1997; y Calvo Garzón, en prensa).

31 En Calvo Garzón (2004b) articulo una noción de constituyente en el que la conducta combinatoria del sistema se basa en la actividad sensoriomotriz y el parámetro del tiempo. Una cita visionaria respecto a la dimension temporal de la actividad cognitiva y a los problemas que ignorarla conlleva se encuentra en Bickhard y Richie (1983): "... interactive competence in an environment requires skill, and skill requires not just formal 'information processing' or 'symbol manipulation' capabilities, but also timing and temporal coordination capabilities, and Turing machine theory has no natural way of introducing such considerations. There is no natural timing unit in Turing machine theory, nor even any sense in which the processing steps take the same or determinate

que no tiene sentido hablar de representaciones estáticas que son manipuladas por parte del sistema, ya sean simbólicas y subsimbólicas. Sólo tiene sentido hablar de estados del cerebro + cuerpo + medio en tanto en cuanto éstos interactúan en tiempo real. En definitiva, la idea de “acoplamiento” sustituye a las representaciones atemporales de clasicistas y conexionistas.

De acuerdo con esta visión alternativa, es obvio que no puede haber una representación canónica de cada constituyente conceptual que sea común a todas las estructuras moleculares en las que el constituyente haga acto de presencia. Pero del mismo modo, tampoco encontramos una representación distinta que codifique *cada* contexto molecular, como en el caso del conexionismo. La cognición ha de estudiarse teniendo en cuenta las interacciones continuas entre factores neuronales, corporales y medioambientales, así que no nos sirve de nada atender a representaciones discretas, ya sean simbólicas o subsimbólicas. El nuevo eslogan que podríamos usar en el caso de la ciencia cognitiva corporeizada sería:

Ciencia Cognitiva Corporeizada:

La cognición es un fenómeno auto-organizado, emergente y extendido que ha de estudiarse en tiempo real teniendo en cuenta las interacciones continuas entre factores neuronales, corporales y medioambientales.

Sin embargo, quizás hemos ido demasiado lejos en el establecimiento del paralelismo entre el gobernador centrífugo de Watt y un sistema cognitivo genuino. Al fin y al cabo, ¿qué tiene que ver la regulación del flujo de vapor con las competencias y capacidades de un *agente*? En particular, se puede ver en el concepto de continuidad un problema para la perspectiva corporeizada cuando cuestiones de representación surgen en el contexto de la actividad cognitiva *superior*, ya que no parece claro que determinados estados del sistema tengan que evolucionar de manera continua. Pensemos por ejemplo en la explicación

multiple amounts of time as each other: the steps are simply serially ordered with no metric time considerations at all. There is, for example, nothing equivalent to an oscillator in Turing machine theory, and no way to construct one without adding to the fundamental assumptions of the formalism. ...” (p. 90).

de mi conducta cuando cojo la botella de agua perfectamente visible delante de mí, en contraposición al movimiento que realizo cuando se encuentra detrás de un obstáculo que no me permite verla. En el segundo caso, he cogido la botella no en base a información disponible directamente que me permite hacer ajustes *online* conforme estiro el brazo para alcanzarla, sino a partir de una memoria más duradera de la posición de la botella³². La hipótesis de trabajo representacionista diría simplemente que cualquier modelo que consiga explicar la cognición en términos de corporeización tendrá que incorporar el carácter inherentemente discontinuo de los estados representacionales. De este modo, conforme nos alejamos de los patrones de activación *online* de bajo nivel, la supuesta continuidad comienza a disolverse. Otra manera de decir lo mismo es que hay una separación entre conducta de bajo nivel que cumple con los requisitos de homogeneidad y de interactividad arriba expuestos, y conducta de alto nivel que no los cumple.

Por desgracia, esta discusión nos lleva otra vez al riesgo de colapso que vimos en la sección anterior del conexionismo con respecto del clasicismo, donde el nivel algorítmico parecía imponer restricciones difíciles de cumplir a no ser que supusiera la trivialización del nuevo paradigma. En este caso estamos nuevamente ante el riesgo de asimilación de la ciencia cognitiva corporeizada por parte del conexionismo, lo que a su vez implica el colapso en el clasicismo³³.

32 La existencia de un estado interno que cumpla los requisitos apuntados en la nota 28 es lo que supuestamente explicaría mi capacidad para alcanzar la botella sin estar a la vista. Por otro lado, Clark y Grush (1999) defienden un principio de disociación a partir de la teoría de la emulación de Grush (2004). En Calvo Garzón (2004a) ofrezco una visión antirrepresentacionista del marco teórico de Grush. Desde la propia perspectiva corporeizada se viene cuestionando si la ciencia cognitiva puede dar cuenta de este tipo de casos; lo que se ha venido en llamar “problemas ávidos de representación” (Clark y Toribio, 1994; Clark, 1997). Ver Eliasmith (1996) para una crítica del concepto de continuidad dinamicista. Asimismo, como señalan Spencer y Schöner (2003), los estados del sistema sujetos a un tratamiento representacional no tienen por qué evolucionar de manera continua. Spencer y Schöner introducen el *Dynamic Field Approach* (DFA), una extensión de los modelos corporeizados, con el fin de incorporar la dinámica de estados representacionales (Erlhagen y Schöner, 2002).

33 Como simple ilustración de esta dinámica, en un comentario de la línea de investigación antirrepresentacionista de Brooks, Vera y Simon (1993a) afirman que “[sensory] information is converted to symbols which are then processed and evaluated in order to determine the appropriate motor symbols that lead to behavior” (p. 34). Del mismo modo, Vera y Simon (1993b) arguyen que las affordances Gibsonianas, ejemplo prototípico del realismo directo, “far from removing the need

Recordando el movimiento anterior, el clasicista pregunta al conexionista: “¿Cómo es posible que expliques la coherencia inferencial del pensamiento humano si no es mediante la manipulación de símbolos?”. Sea cual sea la implementación del sistema, la restricción viene dada desde el momento en que el problema se formula en esos términos. Ahora, se le pregunta al defensor de la ciencia cognitiva corporeizada: “¿Cómo es posible defender la continuidad en términos de homogeneidad e interactividad de un sistema cognitivo extendido si hay situaciones cuya resolución requiere de la explotación de discontinuidades?”.

Curiosamente, este reto viene de defensores de la perspectiva corporeizada que se resisten a renunciar al representacionalismo. Conforme el concepto de computación se descentraliza, el representacionista tiene la tentación de quedarse con lo mejor de cada casa, y optar por nociones “blandas” de computación y representación. De acuerdo con una forma blanda de representación, no se pueden postular entidades abstractas que vayan más allá de los patrones de conducta contextualizados del sistema³⁴. Sin embargo, eso equivale a hacer de los estados de cognición descentralizados de la perspectiva corporeizada una réplica de los constituyentes no-clásicos del conexionismo, donde los estados eran representacionales, pero subsimbólicos en lugar de simbólicos. De esta manera, afirmar que un sistema explota discontinuidades conlleva el riesgo de convertirlo nuevamente en una arquitectura de procesamiento de información, donde la información discreta almacenada se manipula, esta vez, estadísticamente.

for internal representations, are carefully and simply encoded internal representations of complex configurations of external objects, the encodings capturing the functional significance of the objects” (p. 41).

34 Clark y Grush (1999) y Wheeler y Clark (1999) juegan con la idea de que los estados físicos de un sistema cuentan como representacionales sólo si se puede establecer un paralelismo entre los estados en cuestión y el papel que desempeñan en el establecimiento de una conexión entre los estados input y output del sistema. Esto es, respaldan la identificación de estados físicos específicos con el papel semántico que cumplen. Sin embargo, el reconocimiento de la posibilidad de que el sistema no puede descomponerse se debe en su opinión al hecho de que los estados internos del sistema se encuentran conectados causalmente de manera masiva y recíproca. Pero esto es solamente un reto que la neurobiología (representacional) está encantada de aceptar: a saber, el reto de identificar y asignar papeles representacionales a los sub-estados y procesos que la neurobiología descubra en el futuro. O dicho de otra forma, esto sólo va en contra de los logros alcanzados a fecha de hoy.

La cuestión de fondo (y nueva pregunta del millón de euros) es: ¿tiene sentido desde una perspectiva genuinamente corporeizada distinguir entre conducta reactiva (*online*) y cognitiva (*offline*)? ¿No es posible que los problemas formulados en términos de discontinuidades no sean sino un subproducto de una manera de plantear la cuestión en términos del paradigma de procesamiento de información? En mi opinión, si pensamos desde esta perspectiva podemos reconsiderar todo el paisaje. Las preguntas formuladas desde el paradigma de procesamiento de información ya vienen viciadas. El clasicismo y el conexionismo preguntan por aquellos estados internos del sistema que en virtud de su contenido representacional permiten al sujeto mostrarse competente en una tarea dada. Esta forma de acercarse al estudio de la cognición presupone una respuesta en la que explotamos estados representacionales más o menos estables (piénsese en las explicaciones que daba la psicología popular en la sección anterior). Por el contrario, desde la perspectiva que defiendo aquí, deberíamos preguntarnos, sin preconcepciones impuestas “desde arriba”, por aquellas transformaciones continuas sensorio-neuronales que permiten al sujeto mostrarse competente en una tarea dada.

Nótese que el primer modo de hacer la pregunta implica la existencia de discontinuidades. Sin embargo, esto no es más que volver a la estrategia del clasicismo contra el conexionismo. Así, nuevamente, rechazo la carga de la prueba: dándole la vuelta a la tortilla, si las representaciones requieren discontinuidades, entonces no hay representaciones ya que las interacciones de un sistema físico compuesto (cerebro + cuerpo + medio) acoplado en el espacio-tiempo son continuas³⁵.

Dicho esto, es preciso enfatizar que una arquitectura cognitiva corporeizada es una hipótesis de trabajo empírica que debe hacerse explícita. Es solamente a partir de atender a los detalles de lo que la ciencia cognitiva corporeizada nos ofrece que podemos evaluar si realmente estamos en algo distinto de la ciencia cognitiva clásica y conexionista. Desde el rechazo del conductismo, la visión cognitivista dominante se caracteriza por la condición *sine qua non* de que se asigne un papel explicativo a procesos computacionales y a estados internos representacionales. Sin embargo, es a fecha de hoy una hipótesis empírica abier-

35 De hecho, a partir de la idea de que los sistemas cognitivos no son discretos puede seguirse que no sean computacionales (ver Globus, 1992, y Van Gelder, 1995).

ta si la neurociencia necesita retener los postulados de la psicología o no. En cualquier caso es algo que solamente lo resolverá el desarrollo de programas de investigación ambiciosos con nuevas hipótesis de investigación³⁶. Para ello la propuesta es mirar a los fenómenos cognitivos y no cognitivos de la misma manera, y ver hasta dónde podemos llegar en la generación de predicciones e interpretaciones no clásicas de los datos recogidos. En definitiva, vamos a evaluar la significación de una ciencia cognitiva corporeizada integrada.

Con el fin de conectar problemas teóricos fundamentales sobre la cognición y la representación con los recientes desarrollos experimentales es preciso mejorar las vías de comunicación entre la filosofía de la ciencia cognitiva y las ciencias cognitivas propiamente dichas. Sólo de esta manera se podrá dar respuesta a cuestiones como qué cuenta como un sistema (mínimamente) cognitivo³⁷. Uno de los problemas principales de la falta de progreso es que filósofos y científicos se encargan de cuestiones conceptuales y empíricas, respectivamente, por separado. No obstante, si la ciencia cognitiva pretende ser algo más que la suma de las partes, que la suma de las disciplinas que la integran, es preciso romper los *corrillos* y abordar cuestiones fundamentales en el plano conceptual, experimental y de modelización de manera conjunta. En definitiva, es preciso compartir agendas de investigación con todas aquellas disciplinas que tienen algo que aportar al estudio científico de la cognición³⁸.

36 Ver por ejemplo Colunga y Smith (2008) y Johnson y col. (2008).

37 Otras cuestiones están relacionadas con el modelo de explicación científica preferido. ¿Consiste la explicación en una descripción matemática de los cambios de un sistema extendido en el tiempo? ¿Qué relación de eficacia causal hay entre los postulados de la ciencia cognitiva corporeizada y los componentes del sistema?

38 Agradezco muy especialmente a mis colegas Lilian Bermejo y Javier Valenzuela las correcciones realizadas en un tiempo récord al último borrador de esta charla.

APÉNDICE: DISCURSO DE AGRADECIMIENTO

Quisiera empezar por agradecer a la *Fundación Séneca* el haberme hecho sentir joven. Luego, he envejecido, todo hay que decirlo, tratando de escribir el librito que tenéis entre vuestras manos, pero no hay mayor placer que recibir el “Premio Jóvenes Investigadores de la Región de Murcia”, insisto en lo de “jóvenes”, a una edad en la que empieza uno a inquietarse con demasiada frecuencia por cosas como las patas de gallo, los michelines que ya preocupan de cara al verano (¡todavía quedan unos meses para los buenos propósitos!), etc. También quería agradecer muy especialmente a la Facultad de Filosofía y al Departamento de Filosofía de la Universidad de Murcia el apoyo recibido durante estos años. Obviamente no hay investigación que hacer sin apoyo institucional, y aunque suene raro en una disciplina en la que la investigación se hace sin laboratorio, no he encontrado en la Facultad y en el Departamento más que facilidades para desarrollar mi carrera investigadora.

En esta charla he presentado distintas hipótesis arquitectónicas sobre la cognición. Siendo coherente con el modelo descentralizado y distribuido que privilegio, ¡me veo en la obligación de dar las gracias literalmente a *todo*, a todas las fuentes de perturbación que han impactado sobre mí! Así que, hablando de perturbaciones (especialmente de madrugada) me gustaría empezar por mis hijos Hortensia y Paquillo. Si un premio que lleva el nombre de “Jóvenes Investigadores” te hace sentir joven (e investigador), el ver cómo crecen tus hijos, así de rápido, da mucho que pensar. Ya conocéis la historia. Esos arbolitos que hay que ir regando y enderezando desde que son pequeños, que son esponjas, que te maravillan con el brillo y limpieza de su mirada; y como decían *Les Luthiers* en sus *Consejos Para Padres*: “Porque los hijos son la alegría de la vida, querido, con sus risas, con sus juegos, con sus preguuuuntas” (no me sale en argentino). En fin, que son la razón de todo.

Por supuesto, de los hijos a los padres sólo hay un paso. A pesar de no poder recordar esos primeros pasos en primera persona, es un regalo verlos en los hijos de uno, e imaginar por dónde tuvieron que pasar tus propios padres, día tras día, noche tras noche. Sacándote de la cama cada mañana para ir al cole, preguntándote cómo te ha ido todos los días, qué has aprendido, etc., y como sabéis de lo que hablo, no voy a seguir... De hecho, volviendo al motivo de mi charla, un debate candente en ciencia cognitiva es el debate naturaleza-entorno. Las espadas están en alto. Los modelos clasicistas tienden a identificarse con visiones nativistas (ya sabéis, el estímulo es muy pobre, así que de alguna manera la capacidad ya viene codificada en el paquete genético). Los conexionistas, por otro lado, contestan diciendo que el estímulo es bastante más rico de lo pensado, y que el número de restricciones impuestas desde la perspectiva nativista es mínimo. Obviamente, no pretendo aquí decir nada al respecto, aunque por la línea de argumentación de esta charla, podréis imaginar por dónde van los tiros. En cualquier caso, lo que quería decirles a mis padres es que tenga quien tenga razón en el debate naturaleza-entorno, *todo se lo debo a ellos* (naturaleza y entorno). En esta ocasión tan especial, me gustaría decirles que sé cómo se tienen que sentir. Me muero de pensar en estar en su pellejo y ver a Paquillo o a Hortensia aquí arriba recibiendo un premio. Me gustaría darles las gracias por todo infinitamente, porque realmente este premio es para ellos.

Siguiendo en esta línea de cognición distribuida, y otra vez viendo en mis hijos lo que significa crecer como hermanos, me gustaría darles también un beso gordísimo a mis hermanas Maena y Pingo, que han sufrido todos los pavos de hermano pequeño por los que he ido pasando, y que me han llevado de la mano como hermanas mayores, y nos hemos disfrutado como colegas. Lo que se comparte como hermanos es algo que no tiene precio y que no se puede experimentar en ningún otro tipo de relación, así que, otra vez, gracias infinitas.

Y hablando de colegas, también tengo muchas cosas que agradecer a mis colegas. Son un eslabón esencial de esta cadena distribuida (y como la mayoría de las cosas que les agradezco son inconfesables, si no, no serían *cosas de colegas*, pues no digo nada más).

Me he dejado este último párrafo *pá mi mujé*. ¿Qué puedo decir de mi media naranja, la persona con la que he compartido todo hasta ahora? Ella pensará: pues no compartas tanto. Te podías haber ahorrado las noches predoctorales en vela, cuando te venía una idea y tenías que escribirla antes de que se

te fuera la inspiración, las “vacaciones” con portátil en las épocas *apretás* de plazos de entrega de trabajos, y seguro que muchas más cosas que recordará ella mejor que yo. En cualquier caso, ¿qué puedo decir de la persona que aparece en la foto de mi Primera Comunión que no sepa ya? Y por si todavía no lo sabe, yo se lo diré. Eso sí, en privado, que hay niños delante.

Y como me he quedado sin palabras, me callo ya. Gracias a todos por todo.

REFERENCIAS BIBLIOGRÁFICAS

- Anderson, J.R. (1993) *Rules of the mind*, Erlbaum.
- Anderson, J.R. y Lebiere, C. (2003) "The Newell Test for a theory of cognition", *Behavioral and Brain Sciences* 26: 587-637.
- Barsalou, L.W. (1999) "Perceptual symbol systems", *Behavioral and Brain Sciences* 22: 577-609.
- Bechtel, W. y Abrahamsen, A. (1991) *Connectionism and the Mind: And Introduction to Parallel Processing in Networks*, Blackwell.
- Bechtel, W. y Graham, G. (eds.) (1998) *A Companion to Cognitive Science*, Blackwell.
- Bechtel, W., Mandik, P., Mundale, J. y Stufflebeam, R.S. (eds.) (2001) *Philosophy and the Neurosciences: A reader*, Blackwell.
- Beer, R.D. (1995) "A Dynamical Systems Perspective on Agent-Environment Interaction", *Artificial Intelligence* 72: 173-215.
- Bekoff, M., Allen, C. y Burghardt, G.M. (eds.) (2002) *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, MIT Press.
- Bickhard, M.H., y Richie, D.M. (1983) *On the Nature of Representation: A Case Study of James Gibson's Theory of Perception*, Praeger Publishers.
- Bickhard, M.H. y Terveen, L. (1995) *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*, Elsevier.
- Bickle, J. (2003) *Philosophy and Neuroscience: A Ruthlessly Reductive Approach*, Kluwer.
- Boden, M.A. (1990) *The Philosophy of Artificial Intelligence*, Oxford Univ. Press.
- Brooks, R.A. (1991) "Intelligence without representation", *Artificial Intelligence* 47: 139-59.
- Calvo Garzón, F. (2000) "A Connectionist Defence of the Inscrutability Thesis", *Mind and Language* 15: 465-80.

- Calvo Garzón, F. (2003a) "Connectionist Semantics and the Collateral Information Challenge", *Mind and Language* 18: 77-94.
- Calvo Garzón, F. (2003b) "Non-classical connectionism should enter the decathlon", *Behavioral and Brain Sciences* 26: 603-604.
- Calvo Garzón, F. (2004a) "Issues of Implementation Matter for Representation", *Behavioral and Brain Sciences* 27: 397-98.
- Calvo Garzón, F. (2004b) "Context-Free versus Context-Dependent Constituency Relations: A False Dichotomy", *Proceedings of the American Association for Artificial Intelligence*, Tech. Report FS-04-03: 12-6.
- Calvo Garzón, F. (2005) Non-adjacent transitional probabilities and the induction of grammatical regularities, *Proceedings of the American Twenty-Seventh Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum.
- Calvo Garzón, F. (en prensa) "Towards a general theory of antirepresentationalism", *British Journal for the Philosophy of Science*.
- Calvo Garzón, F. (en revisión) "On the Cognitive Architecture of Insects and Other Information-Processing Systems: A Reply to Carruthers".
- Calvo Garzón, F. (en preparación) "The Collapse of Scientific Intencional Realism".
- Calvo Garzón, F. y Colunga, E. (2003) "The Statistical Brain: Reply to Marcus' *The Algebraic Mind*", *Proceedings of the American Twenty-Fifth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum.
- Calvo Garzón, F., Laakso, A. y Gomila, T. (2008a) *Special Issue on Dynamics and Psychology*, vol. 1, *New Ideas in Psychology*.
- Calvo Garzón, F., Laakso, A. y Gomila, T. (2008b) *Special Issue on Dynamics and Psychology*, vol. 2, *New Ideas in Psychology*.
- Carruthers, P. (2004) "On being simple minded", *American Philosophical Quarterly* 41: 205-20.
- Chomsky, N. (1980) *Rules and Representations*, Blackwell.
- Churchland, P.M. (1981) "Eliminative Materialism and the Propositional Attitudes", *Journal of Philosophy* 78: 67-90.
- Churchland, P.M. (1989) *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, MIT Press.
- Churchland, P.M. (1996) "Learning and Conceptual Change: The View from the Neurons", en Clark, A. y Millican, P.J.R. (1996) *Connectionism, Concepts and Folk Psychology: The Legacy of Alan Turing*, Vol. 2, Clarendon Press.

- Churchland, P.M. (1998) "Conceptual Similarity Across Sensory and Neural Diversity: the Fodor/Lepore Challenge Answered", *The Journal of Philosophy* 95: 5-32.
- Churchland, P.M. (2005) "Functionalism at forty: A critical retrospective", *The Journal of Philosophy* 102: 33-50.
- Churchland, P.S. (1986) *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, MIT Press.
- Clancey, W.J. (1997) *Situated Cognition*, Cambridge University Press.
- Clark, A. (1994) "Language of Thought (2)", en Guttenplan, S. (ed.) (1994) *A Companion to the Philosophy of Mind*, Blackwell.
- Clark, A. (1997) *Being There: Putting Brain, Body, and World Together Again*, MIT Press.
- Clark, A. y Grush, R. (1999) "Towards a Cognitive Robotics", *Adaptive Behavior* 7: 5-16.
- Clark, A. y Toribio, J. (1994) "Doing Without Representing?", *Synthese* 101: 401-31.
- Collins, A. y Quillian, M. (1972) "Experiments on Semantic Memory and Language Comprehension", en Gregg, L. (ed.) (1972) *Cognition in Learning and Memory*, Wiley.
- Colunga, E. y Smith, L. (2008) "Flexibility and variability: Essential to human cognition and the study of human cognition", en Calvo y col. (eds.) *Special Issue on Dynamics and Psychology*, vol. 1, *New Ideas in Psychology*.
- Dreyfus, H.L. (1992) *What Computers Still Can't Do: A Critique of Artificial Reason*, MIT Press.
- Eliasmith, C. (1996) "The Third Contender: A Critical Examination of the Dynamicist Theory of Cognition", *Philosophical Psychology* 9: 441-463.
- Elman, J. (1990) "Finding Structure in Time", *Cognitive Science* 14: 179-211.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. y Plunkett, K. (1996) *Rethinking Innateness: A connectionist perspective on development*, MIT Press.
- Elman, J. (1998) "Generalization, Simple Recurrent Networks, and the Emergence of Structure", en Gernsbacher, M.A. y Derry, S. (eds.) *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum.
- Field, H. (1978) "Mental Representations", *Erkenntnis* 13: 9-61.
- Erlhagen, W. y Schöner, G. (2002) "Dynamic Field Theory of Movement Preparation", *Psychological Review* 109: 545-72.

- Feyerabend, P. (1963) "Materialism and the Mind-Body Problem", *Review of Metaphysics* 17: 49-66.
- Fodor, J. (1975) *The Language of Thought*, Harvester Press.
- Fodor, J. (1987) *Psychosemantics*, MIT Press.
- Fodor, J. (1998a) *Concepts: Where Cognitive Science Went Wrong*, Clarendon Press.
- Fodor, J. (1998b) *In Critical Condition*, MIT Press.
- Fodor, J. y Lepore, E. (1999) "All at Sea in Semantic Space: Churchland on Meaning Similarity", *The Journal of Philosophy* 96: 381-403.
- Fodor, J. y McLaughlin, B. (1990) "Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work", *Cognition* 35: 183-204.
- Fodor, J. y Pylyshyn, Z. (1988) "Connectionism and Cognitive Architecture", *Cognition* 28: 3-71.
- Freeman, W.J. (2000) *Neurodynamics. An Exploration in Mesoscopic Brain Dynamics*, Springer-Verlag.
- Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution*, Basic Books.
- Garnham, A. (2008) "Cognitivism", en Symons, J. y Calvo Garzón, F. (eds.) *The Routledge Companion to Philosophy of Psychology*, Routledge Philosophy Companions.
- Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Lawrence Erlbaum.
- Globus, G.G. (1992) "Towards a Noncomputational Cognitive Neuroscience", *Journal of Cognitive Neuroscience* 4: 299-310.
- Gomila, T. y Calvo, F. (2008) "Embodying Cognition: Are There Grounds for Consensus?", en *Handbook of Embodied Cognitive Science*, Elsevier.
- Goodman, N. (1965) *Fact, Fiction and Forecast*, Bobbs-Merrill.
- Grush, R. (2004) "The Emulation Theory of Representation: Motor Control, Imagery, and Perception", *Behavioral and Brain Sciences* 27: 377-442.
- Hanson, S.J. y Burr, D.J. (1990) "What Connectionist Models Learn: Learning and Representation in Connectionist Networks", *Behavioral and Brain Sciences* 13: 471-518.
- Haugeland, J. (1991) "Representational Genera", en Ramsey, W., Stich, S. y Rumelhart, D.E. (eds.) (1991) *Philosophy and Connectionist Theory*, Lawrence Erlbaum.
- Harman, G. (1973) *Thought*, Princeton University Press.

- Hebb, D.O. (1949) *The Organization of Behavior*, John Wiley.
- Hutchins, E. (1995) *Cognition in the Wild*, MIT Press.
- Johnson, J.S., Spencer, J.P. y Schöner, G. (2008) "Moving to higher ground: The dynamic field theory and the dynamics of visual cognition", en Calvo y col. (eds.) *Special Issue on Dynamics and Psychology*, vol. 1, *New Ideas in Psychology*.
- Marcus, G.F. (2001) *The algebraic mind: Integrating connectionism and cognitive science*, MIT Press.
- Marcus, G.F., Vijayan, S., Bandi Rao, S. y Vishton, P.M. (1999) "Rule learning in seven-month-old infants", *Science* 283: 77-80.
- Marr, D. (1982) *Vision*, W.H. Freeman.
- McCauley, R. (ed.) (1996) *The Churchlands and Their Critics*, Blackwell.
- McClelland, J.L., Rumelhart, D.E. and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2, MIT Press.
- McClelland, J.L. y Kawamoto, A.H. (1986) "Mechanisms of Sentence Processing: Assigning Roles to Constituents of Sentences", en McClelland, J., Rumelhart, D. y col., pp. 272-325.
- McCulloch, W. y Pitts, W. (1943) "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics* 7: 115-133.
- McLeod, P., Plunkett, K. y Rolls, E. (1998) *Introduction to Connectionist Modeling of Cognitive Processes*, Oxford University Press.
- Minsky, M. y Papert, S. (1988) *Perceptrons: An Introduction to Computational Geometry*, MIT Press.
- Newell, A. (1990) *Unified Theories of Cognition*, Harvard University Press.
- Newell, A. y Simon, H. (1972) *Human Problem Solving*, Prentice-Hall.
- O'Reilly, R. y Munakata, Y. (2000) *Computational Explorations in Cognitive Neuroscience*, MIT Press.
- Peña, M., Bonatti, L., Nespor, M., y Mehler, J. (2002) "Signal-Driven computations in Speech Processing", *Science* 298: 604-607.
- Pinker, S. y Prince, A. (1988) "On language and connectionism: analysis of a parallel distributed processing model of language acquisition", *Cognition* 28: 73-193.
- Pinker, S. y Ullman, M. (2002) "The past and future of the past tense", *Trends in Cognitive Sciences* 6: 456-463.
- Plunkett, K. y Elman, J. (1997) *Exercises in Rethinking Innateness*, MIT Press.

- Port, R. y Van Gelder, T. (1995) *Mind as Motion*, MIT Press.
- Putnam, H. (1975) *Mind, Language and Reality*, Cambridge University Press.
- Pylyshyn, Z. (1984) *Computation and Cognition*, MIT Press.
- Quine, W.V. (1960) *Word and Object*, MIT Press.
- Quine, W.V. (1969) *Ontological Relativity and Other Essays*, Columbia University Press.
- Quine, W.V. (1970) "Methodological Reflections in Current Linguistic Theory", *Synthese* 21: 386-98.
- Quine, W.V. (1975) "Mind and Verbal Dispositions", en Guttenplan, S. (ed.) (1975) *Mind and Language*, Oxford University Press.
- Quine, W.V. (1979) "Facts of the Matter", en Shahan, R.W. y Swoyer, C.V. (eds.) (1979) *Essays on the Philosophy of W.V. Quine*, Harvester.
- Ramscar, M. (2002) "The role of meaning in inflection: why the past tense doesn't require a rule", *Cognitive Psychology*, 45: 45-94.
- Ramsey, W., Stich, S. y Rumelhart, D.E. (eds.) (1991) *Philosophy and Connectionist Theory*, Lawrence Erlbaum.
- Rolls, E.T. y Treves, A. (1998) *Neural networks and brain function*, Oxford University Press.
- Rorty, R. (1970) "In Defense of Eliminative Materialism", *Review of Metaphysics* 24: 112-21.
- Rosenblatt, F. (1959) *Principles of Neurodynamics*, Spartan Books.
- Rumelhart, D.E., Hinton, G.E. y Williams, R.J. (1986) "Learning Internal Representations by Error Propagation", en Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press.
- Rumelhart, D.E. y McClelland, J.L. (1982) "An Interactive Activation Model of Context Effects in Letter Perception: Part II. The Contextual Enhancement Effect and Some Tests and Extensions of the Model", *Psychological Review* 89: 60-94.
- Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press.
- Ryder, D. (2008) "The Many Problems of Representation", en Symons, J. y Calvo Garzón, F. (eds.) *The Routledge Companion to Philosophy of Psychology*, Routledge Philosophy Companions.
- Ryle, Gilbert (1949). *The concept of mind*, Hutchinson and Company, Ltd.

- Searle, J.R. (1980) "Minds, brains, and programs", *Behavioral and Brain Sciences* 3: 417-457.
- Seidenberg, M. y Elman, J.L. (1999) "Networks are not hidden rules", *Trends in Cognitive Sciences* 3: 288-289.
- Sellars, W. (1968) *Science and Metaphysics: Variations on Kantian Themes*, Humanities Press.
- Sharkey, A. y Sharkey, N. (2008) "Connectionism", en Symons, J. y Calvo Garzón, F. (eds.) *The Routledge Companion to Philosophy of Psychology*, Routledge Philosophy Companions.
- Skinner, B.F. (1957) *Verbal behaviour*, Prentice-Hall.
- Smolensky, P. (1988) "On the Proper Treatment of Connectionism", *Behavioral and Brain Sciences* 11: 1-74.
- Smolensky, P. (1991) "Connectionism, Constituency and the Language of Thought", en Loewer, B. y Rey, G. (1991) *Meaning in Mind: Fodor and His Critics*, Blackwell.
- Spencer, J.P. y Schöner, G. (2003) "Bridging the Representational Gap in the Dynamic Systems Approach to Development", *Developmental Science* 6: 392-412.
- Symons, J. y Calvo, F. (eds.) *The Routledge Companion to Philosophy of Psychology*, Routledge Philosophy Companions.
- Thelen, E. y Smith, L. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press.
- Turvey, M. y Carello, C. (1995) "Some dynamical themes in perception and action", en Port, R. y Van Gelder, T. (eds) *Mind as Motion*, MIT Press.
- Van Gelder, T. (1995) "What Might Cognition Be, if not Computation?", *Journal of Philosophy* 92: 345-81.
- Van Gelder, T. (1998) "The Dynamical Hypothesis in Cognitive Science", *Behavioral and Brain Sciences* 21: 615-665.
- Varela F.J., Thompson E., y Rosch E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press.
- Vera, A.H. y Simon, H.A. (1993a) "Situated Action: A Symbolic Interpretation", *Cognitive Science* 17: 7-48.
- Vera, A.H. y Simon, H.A. (1993b) "Situated Action: Reply to William Clancey", *Cognitive Science* 17: 117-33.
- Watson, J. (1924) *Behaviorism*, Norton.
- Webb, B. (1994) "Robotic Experiments in Cricket Phonotaxis", en Cliff, D.,

- Husbands, P., Meyer, J.A., y Wilson, S.W. (eds.) *From Animals to Animats 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behaviour*, MIT Press.
- Wheeler, M. y Clark, A. (1999) "Genic Representation: Reconciling Content and Causal Complexity", *British Journal for the Philosophy of Science* 50: 103-35.
- Wilson, R.A. y Keil, F.C. (eds.) (1999) *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press.
- Winograd, T. y Flores, C.F. (1986) *Understanding Computers and Cognition. A New Foundation for Design*, Ablex.

