

¿Cómo se inventa una «ciencia del significado»?

Carlos J. Blanco Martín

Arbor CLVII, 619 (Julio 1997) 183-201 pp.

En este trabajo hacemos un apretado recorrido en la historia del cognitivismo, con especial atención a la psicología computacional y a la inteligencia artificial (I.A.). Las ideas de símbolo, representación y significado pretenden reunir a una serie de ciencias y técnicas, sobre los supuestos escombros de la tradición conductista.

La prehistoria cibernética del cognitivismo

Las nuevas tecnologías de la información y del control se consolidaron plenamente al término de la II Guerra Mundial, sobrepasando el círculo de aplicación militar y extendiéndose a ciertas áreas de la industria, la ciencia y el comercio. Hablamos de tecnologías desde el momento en que se construyeron formalismos o teoremas lógicos y matemáticos estrechamente vinculados a los diseños y aplicaciones ingenieriles que la cibernética estaba entonces promoviendo. Las contribuciones de rango gnoseológico debidas a la cibernética y a la teoría de la información tuvieron ese carácter formalista aunque muy enraizadas en el suelo de las necesidades prácticas en campos como la transmisión de mensajes o la retroacción y los automatismos. Históricamente, pueden descubrirse la líneas de conexión entre los nuevos desarrollos

de la teoría matemática de la información —como prolongación de los teoremas de la probabilidad— así como de las contribuciones lógicas y algebraicas vinculadas a la ingeniería de circuitos eléctricos y a la construcción de máquinas de cálculo contribuciones que se vieron posibilitadas por los pasos emprendidos ya por Frege, Wittgenstein, Russell y Whitehead.

La cibernética fue un movimiento un tanto heterogéneo cuyas prolongaciones llegan hasta nuestros días en las más variadas disciplinas. La gama de campos en los que la corriente, netamente interdisciplinar, realizó incursiones no pudo ser más variada: aprendizaje, neurofisiología, ingeniería, ciencias sociales... Esa amplitud de los campos se corresponde con la amalgama de métodos y conceptos que formaron parte del agregado cibernético. J. Rose (1978) señala al menos tres tipos de actividades distintas a las que el movimiento cibernético se ha consagrado: (a) Construir una teoría eficaz para simular distintos aspectos del comportamiento de sistemas vivos, teoría ésta que incluye el concepto abstracto de *máquina de Turing*, modelos matemáticos, conceptos como *algoritmo o programa*. (b) Construir modelos del comportamiento humano para que, tanto el modelo como el modelado (el hombre) lleguen al mismo resultado final en tareas psicológicas como el razonamiento y la solución de problemas, la memoria o la percepción. (c) Simulación del comportamiento de seres vivos —hombres o animales— mediante modelos estructural o anatómicamente semejantes: robótica, biónica, etc.

La simulación de conductas abarcaba desde el diseño de programas que concernían a los aspectos más abstractos de la vida, como la *inteligencia*, acudiendo a conceptos matemáticos (programa, algoritmo, máquina de Turing), pero también la cibernética se ocupaba de modelos mecánicos que emulaban la vida, pero atendiendo a los aspectos más corpóreos de un animal o de un hombre: emular un ojo, unos brazos, unas piernas (la imitación de la corporeidad mecánica total del organismo era y es el cometido principal de la robótica). Hoy en día podemos decir que la robótica y los modelos de redes neurales del conexionismo son las dos tendencias que más directamente han recibido el legado cibernético. La Inteligencia Artificial habría de desgajarse en el segundo periodo (años 60) para ocuparse de los aspectos sintácticos, semánticos y psicológicos de la inteligencia, y no tanto de los corpóreos y estructurales (aunque de hecho todo robot sofisticado necesita de programas). Es más, la propia Inteligencia Artificial se emanciparía de las inquietudes psicologistas del enfoque llamado *procesamiento de la información*. Las tres tendencias que la cibernética apadrinó, se convirtieron en

líneas divergentes: la I.A. ha seguido la línea (a), la robótica y la ingeniería han seguido la línea (c). Según este esquema tripartito, la Psicología Cognitiva desciende, al menos en parte, de la línea (b). Los enfoques del *procesamiento de la información* de la década de los sesenta tendrán que imponerse por encima de otros modelos mecánicos completamente insatisfactorios y muy apegados aún a las ideas behavioristas, traduciendo las correlaciones E-R en términos de *feed-back* o empleando las cadenas de Markov en el aprendizaje, junto a modelos de marcado carácter fisiológico como las de las redes neurales de McCulloch y Pitts o las nuevas interpretaciones del cerebro como sistema cibernético. La *teoría de la información* a veces figura como un miembro más dentro del agregado cibernético, si bien ha tenido orígenes tecnológicos propios, en especial la tecnología de la transmisión de señales. La teoría de la información sensu stricto es la teoría estadística de la información selectiva. Como teoría matemática formal es autosuficiente y no requiere de experimentos, aunque puede servir de base a distintas ciencias empíricas. Shannon y Weaver en 1948, distinguieron los elementos de una transmisión de mensajes, a saber: (i) la *fuerza de información* que produce el mensaje o secuencia de mensajes; (ii) el *transmisor* que actúa sobre el mensaje para producir la señal, (iii) el *canal o medio* por donde fluye el mensaje; (iv) el *receptor*, que realiza la operación inversa a la del transmisor, y, por último (v) el *destino*, que es a quién va dirigido el mensaje, bien se trate de una persona, bien sea una cosa (Shannon y Weaver, 1949/71).

No pasó mucho tiempo hasta que tales nociones fueran acogidas por la Psicología y otras ciencias sociales. Autores como George A. Miller, Broadbent y Attneave ensayaron ciertas aplicaciones del enfoque informativista en psicología. De acuerdo con el veredicto suministrado por los psicólogos cognitivos, hoy en día se admite que tales tentativas dieron en fracaso. La teoría *técnica y formalista* de la información no podía ser introducida literalmente en los modelos psicológicos que, por su parte, exigían otro tipo de unidades primitivas distintas de los arbitrarios «trozos» cuantificables (p.e. los *chunks*).

El concepto de información ha conocido muchos campos de aplicación, desde la semiótica y la sociología, hasta la mismas ciencias físico-químicas, al ponerse en correspondencia con otros conceptos como *entropía* y *neguentropía* (Mosterín, 1987) y, en categorías lingüísticas, con la noción de significado. El concepto básico para la fundación de una psicología cognitiva ha sido el concepto de la *información en cuanto que ha sido procesada* y no el sentido técnico-probabilístico de Shannon y Weaver.

Se funda la sociedad cognitiva

En ciertos textos retrospectivos se suele señalar una fecha del nacimiento institucional del enfoque cognitivo: año 1956, y el evento fue la célebre conferencia de Darmouth. En aquella reunión participaron muchas de las personalidades clave para la comprensión de los primeros pasos de la I.A. y la Psicología Cognitiva: Samuel, Selfridge, McCarthy, Newell, Minsky, Simon, etc. En su mayor parte éstos eran miembros de una nueva generación de personas jóvenes que habían recogido muchas de las herencias intelectuales de sus antepasados, los cibernéticos Wiener, McCulloch, von Neumann o Turing. En cierto modo, la elección de la conferencia de Darmouth como acto fundacional tiene todos los visos de haber sido una decisión *ad hoc*. Como el mismo Minsky ha reconocido (McCorduck, 1991, p. 117), los únicos asistentes que tenían una idea clara acerca de lo que era un programa para simular conductas inteligentes y que además podían aportar alguna muestra concreta de lo que podía hacerse en ese sentido habían sido Newell y Simon, quienes pueden ser señalados con toda justicia los *padres fundadores* de la simulación cognitiva y no meros exploradores de las posibilidades que prometían ofrecer los ordenadores. En su libro *Maquinas que Piensan*, la escritora P. McCorduck ha recogido éste y otros testimonios de ciertos asistentes a tan renombrada conferencia y, según ha concluido la autora citada, puede decirse que Darmouth tuvo más importancia social que epistemológica debido a que la gente siguió trabajando, de hecho, en las líneas y proyectos que anteriormente se habían marcado. Ahora bien, lo que sí determinó Darmouth fue marcar con sello propio ciertas pautas sociales que motivaron la imagen que, con el paso del tiempo, el grupo inicial de la I.A. y sus continuadores habrían de ganarse: hermetismo, recelo ante los extraños, etc. En el texto de McCorduck se apuntan datos de interés acerca de las fuentes básicas de financiación de que gozó aquel grupo desde el principio, de donde se extrae que el Departamento de Defensa norteamericano ha jugado un papel determinante, circunstancia que no parece haber variado en fechas recientes. Más relevancia gnoseológica cobraron ciertas publicaciones, en especial la recopilación que en 1963 Feigenbaum y Feldman sacaron a la luz, *Computers and Thought* (Feigenbaum y Feldman, 1963) donde se podían detectar la pluralidad de intereses en el nuevo campo de la I.A., no todos vinculados de manera relevante a la Psicología. Muchos programas trataban, en realidad, de lograr que la máquina resolviera problemas en áreas cognoscitivamente muy restringidas, pero fáciles o apropiadas desde el punto de vista com-

putacional, debido a la estructura formalizada y bien definida en situaciones como el ajedrez, lógica, geometría, etc. El ambiente intelectual de la psicología, al hacerse *más cognitivo*, renovó el interés por áreas un tanto descuidadas por los behavioristas (percepción, lenguaje, memoria) y motivó el resurgimiento de metodologías que, en todo caso, ya estaban disponibles en los laboratorios mucho antes de que la «revolución cognitiva» propagara una vuelta a las mismas: tiempos de reacción, tareas de recuerdo libre y con claves, análisis de errores, exposición taquistoscópica, etc. En cierto modo, aquellas eran las técnicas de experimentación con *sujetos humanos* abandonadas por el hecho de que la corriente dominante en los Estados Unidos primaba el estudio de procesos asociativos en especies no humanas. Por otro lado, los ordenadores eran artefactos que resultaban extraños para muchos psicólogos, y rara vez los empleaban de forma directa, todo lo más, como fuente de analogías y de términos más o menos útiles con los que (re)bautizarían los procesos que ellos investigaban.

Claramente, la idea de *la mente como máquina lógica* precedió al ordenador como artefacto. Herbert A. Simon declaró (McCorduck, 1991, p. 75): «Desde el punto de vista histórico, la teoría moderna de la lingüística de transformación y la teoría cognoscitiva relacionada con el proceso de información nació de una misma matriz: la matriz de ideas producida por el desarrollo de la moderna máquina de calcular digital y la comprobación de que, a pesar de que la envoltura de la computadora estaba constituida por un armazón metálico, su alma era en realidad un programa». En efecto, Alan Turing en un trabajo ya clásico de 1937, había planteado la posibilidad de una máquina abstracta capaz de manipular números mediante la aplicación de reglas formales (Turing, 1937). La idea de que esta clase de máquina pudiera ponerse en correspondencia con la mente humana ya había flotado en el ambiente incluso antes de que las máquinas computadoras fueran tecnológicamente posibles. La formulación de esta filosofía apareció en un artículo de 1950 en el que Turing había hecho su célebre pregunta (ineludible, desde entonces, en Psicología y Ciencia Cognitiva) acerca de si podría pensar una máquina (Turing, 1950). La respuesta de Turing era afirmativa y para ello se apoyaba en el célebre argumento del *juego de simulación*. Los antiguos vínculos de conexión conceptual entre las propiedades de los hombres y las propiedades de las máquinas parecen haberse sucedido a lo largo de la historia por medio de una circulación o realimentación constante entre los dos polos (hombre / máquina) de modo tal que uno de ellos se nutre o beneficia de las propiedades o cualidades del otro. Desde precedentes medievales como

el *Ars Magna* de Llull, o las ideas mecanicistas modernas de Descartes, La Mettrie o Hobbes, la máquina de cálculo de Pascal o los primitivos diseños de Babagge o de Torres Quevedo, pueden rastrearse los recíprocos transportes de cualidades que hacían ver que, de hecho, *el hombre era una máquina* o bien que *la máquina era humana* (o *inteligente, animada, etc.*). De todas formas, la Inteligencia Artificial parece haber allanado un tipo peculiar de mecanicismo: *el mecanicismo simbólico, formalista*.

Alrededor de 1960, se dio la convergencia de varias líneas de investigación en máquinas de cálculo, psicología, lógica matemática, etc. Pero en lo que a la Psicología le concierne, fue la *simulación cognitiva* la metodología que efectivamente se incorporó junto a las tradicionales que disponían los psicólogos investigadores. Las contribuciones de Newell y Simon incluían, de manera convergente, la programación de computadoras por un lado, y los testimonios en voz alta por parte de sujetos humanos. Aquí comienza, a nuestro juicio, el *procesamiento de información en sentido psicológico* que realmente comunicó la I.A. con la psicología. En 1960, Miller, Galanter y Pribram, mediante su emblemática obra *Plans and the Structure of Behavior* (Miller et al, 1960), saludaron con fervor este nuevo marco de investigación:

«Después de estudiar la obra pionera de Newell, Shaw y Simon, resulta muy difícil mantener nuestro inicial respeto por la economía de medios como requisito de las teorías de las teorías psicológicas. Ciertamente, ya no podemos pensar que cualquiera que postule para un organismo un complicado procesamiento de información esté apelando a principios misteriosos, vitalistas, ambiguos o acientíficos. Un procesamiento de información que sea complicado y que se atenga a principios heurísticos no es tan sólo concebible, puesto que ya ha sido realizado en la práctica y se ha demostrado su existencia con ordenadores actualmente existentes».

El nuevo enfoque que estaba naciendo no había de ser considerado animista, metafísico ni mentalista. Lo que sucedió fue que hablar de *planes* (equivalente mental del concepto de *programa*) y de *imágenes* (representación psicológica) debía ser legítimo, toda vez que las *conductas superiores o complejas* podían ser descompuestas en otras partes más sencillas gracias a las nuevas metodologías. Al declararse «conductistas subjetivos», Miller, Galanter y Pribram querían situarse en un punto equidistante respecto del behaviorismo watsoniano de los reflejos condicionados, por un lado, y del introspeccionismo, por el otro. La idea de que las conductas del hombre podían organizarse jerárquicamente, según unas estructuras como los *planes* (que contienen

a su vez subplanes, recursivamente «montados» unos dentro de otros) no hacía sino renovar por completo el enfoque neobehaviorista, incluidas en él las versiones cognitivas por excelencia a saber, las que Tolman había propugnado desde los años veinte. Así, de manera análoga a un *mapa cognitivo* tolmaniano (Tolman, 1948) la acción de colocar un cuadro en la pared demanda subrutinas o subplanes como levantar el martillo para dar el golpe o, a su vez, buscar el martillo si no lo tenemos a mano, o emplear otra herramienta alternativa que sirva para los mismos propósitos, etc. *El pensamiento en voz alta* era un dato conductual (e introspeccionista) digno de reclamar su puesto entre las técnicas de descomposición de planes en subplanes jerárquicamente organizados.

Si bien la conferencia de Darmouth de 1956 ha sido elegida como acontecimiento fundacional de la I.A. y, de una forma más global, el punto de partida de una pretendida super-ciencia cognitiva, otros eventos son reclamados como hitos relevantes del cognitivismo, en especial el simposium de Hixon en 1948 donde se ha querido ver el comienzo del declive del conductismo dominante (Jeffres, 1951). En aquella reunión, Karl Lashley, antiguo discípulo de Watson, arremetió contra ciertos dogmas conductistas en una célebre conferencia titulada «El problema del orden serial de la conducta». Otro miembro asistente al simposium, el joven Noam Chomsky, contribuyó con su propia conferencia. Casi al acabar la década de los 50, Chomsky publicaría una crítica feroz al enfoque conductista más elaborado en materia de lenguaje, en una reseña al libro de B.F. Skinner, *Verbal Behavior* (Chomsky, 1959). No obstante, la crítica no se restringió al exámen crítico de una *teoría particular* sobre el lenguaje, sino que Chomsky ensayó en esa ocasión una refutación filosófica total de los postulados más fundamentales de todo género de behaviorismo: creencia en la generalidad de las leyes del aprendizaje, neutralización de los factores innatos o instintivos, consideración de los fenómenos lingüísticos como *conductas al lado de otras cualesquiera*, etc. La refutación de los postulados conductistas en materia de lenguaje, debida al auge y prestigio que la *Gramática Generativa Transformacional* comenzó a cobrar desde la década de los 50, no podía sino beneficiar a aquellos que veían en el «paradigma conductista» su común enemigo. En otras palabras: si dos fueron las fuentes principales del surgimiento del cognitivismo, a saber, la Inteligencia Artificial y la Gramática Transformacional, esas fuentes son claramente *externas* a la tradición y, por tanto, supusieron una crítica de la psicología, en concreto de la psicología conductista. Pero cuando el edificio categorial de la Psicología hubo de ser restablecido *inter-*

namente, sin la ayuda de los lingüistas ni de los expertos en computadoras, la psicología practicada *como tal* no fue tan distinta de la que se practicara en décadas anteriores, ya por conductistas, ya por estructuralistas, gestaltistas o funcionalistas. Así pues, en contra de esa versión estándar, rechazamos la idea de que *una ciencia pueda refutar a otra in toto*. En el caso presente, en vez de abogar por una pretendida naturaleza *multiparadigmática* de la Psicología (que habría de constatarse hoy mediante la coexistencia misma de distintas escuelas y la falta de acuerdo entre ellas), nos inclinamos más bien a subrayar el carácter *abierto* de las ciencias psicológicas y conductuales. Sustenta nuestra tesis la comprobación de que las operaciones metodológicas de la Psicología no conocieron graves distorsiones tras una supuesta *revolución* y el consiguiente *cambio paradigmático*. Por ejemplo, Chomsky y sus seguidores hicieron cuanto les fue posible por sustituir unas *ideas filosóficas* acerca de la naturaleza del lenguaje y el origen del conocimiento, por otras más de su gusto. Las discusiones acerca del innatismo, la representación (¿simbólica?) del conocimiento, y la interacción entre el cuerpo y la mente, se mantuvieron en el *plano meta-teórico*.

El conductismo cae en desgracia

En la vida «interna» de la propia psicología behaviorista, tuvo lugar un proceso de «liberalización» (también se podría decir, de manera más literaria, una *traición*) de los postulados y dogmas de la escuela, gracias a la cual aumentó (o, más bien, se *recuperó*) un conjunto de áreas de investigación que habían quedado arrinconadas en los años precedentes: conducta instintiva, percepción y pensamiento). Koch (1974) describe agudamente el modo en que los términos «estímulo» y «respuesta» empujaron su escala y se introdujeron en el seno de la mente o del cerebro, practicando así una suerte de psicología E-R pero hacia dentro: así, los psicólogos podían hablar alegremente acerca de un órgano o parte del cerebro/mente que envía señales («estímulos») a otro, que se despierta y «reacciona» otras partes internas. Como ya advirtió Koch, muchas de las explicaciones pretendidamente novedosas sobre la conducta inteligente ya habían sido facilitadas por gestaltistas, piagetianos o fisiólogos en las décadas anteriores. Según el autor, el nuevo periodo neo-neo-behaviorista (y que hoy consideraríamos *cognitivo*) consistió en la interpolación de un segundo sistema E-R (en sentido watsoniano primitivo), entre los estímulos y las res-

puestas *reales*, las del organismo. La Psicología experimental no sufrió tanto el impacto del generativismo o del computacionalismo (que más bien tuvo lugar en el plano filosófico o metateórico) como el declive interno del propio programa conductista. Así, Rychlack sostiene la tesis de que toda psicología cognitiva implica una teoría mediacional y, por tanto, pese a los nuevos tecnicismos, no han tenido lugar grandes cambios revolucionarios desde Tolman:

«Basically, all cognitive theories are mediational formulations; they focus on the presumed factors which 'come between' what has always been thought of as the stimulus-input and the eventual response-output of a behaving organism. Mediation theory is therefore an effort to frame processes 'within in' the flow of efficient causation. Because cognitive theories use concepts of encoding and feedback the impression is given that there is something more active going on in the organism than earlier mediational models such as those of Tolman (1967) or Hull (1943) were able to capture. When we look at these modern cognitive theories in terms of causal conceptions, it is clear that there is no fundamental difference in the style of explanation compared to the earlier mediational concepts».(Rychlack1984, p. 128).

En cuanto al revisionismo epistemológico del conductismo todo comenzó con los conceptos de «variable interviniente» y «constructo hipotético» (MacCorquodale y Meehl, 1953). La diferencia entre estos dos conceptos es bastante «técnica», y está incrustada en la discusión especializada de la epistemología de aquellos tiempos. Hoy nos basta con señalar que los *constructos hipotéticos* constituyen una especie de artificio, «una ficción con la que se pueden expresar hechos de la experiencia». Como dicen MacCorquodale y Meehl, pueden ser sustancias, conceptos o procesos, que ensaya el investigador para expresar de un modo cuantitativo o legaliforme sus hallazgos. Son necesarios para la verdad de los enunciados, mas no *suficientes*. He aquí la diferencia formal con respecto a las *variables intervinientes*. Estos son conceptos cuya verdad está constituida por condiciones necesarias y suficientes. Los enunciados contenidos en las *variables intervinientes* son los que Carnap denominaba *conceptos disposicionales*. En la ciencia física, estos conceptos disposicionales son los que se corresponden con fenómenos tales como «solubilidad», «resistencia», etc. Las variables intervinientes quedarían conectadas por «leyes empíricas» o funciones, las cuales no presentan un «excedente de significado», una substancialización. Mario Bunge (1969), como epistemólogo, pone el acento en la diferencia entre lo observable y lo inobservable. Las variables intermedias son conceptos

que median entre conceptos observacionales. En sus palabras: son «simples auxiliares calculísticos». Por contra, las construcciones hipotéticas son entidades no-observables, pero *inferibles*. En física, el átomo es un concepto no observable, dice Bunge, y no por ello es un concepto fantasmal. Otro ejemplo, en economía, es el concepto de «coste de producción». Pues bien: las ciencias, aunque partan de la observación, según la epistemología bungiana (y según la mayor parte de la filosofía anglosajona) están plagadas de contenidos no observables, "abstractos", pero completamente racionales:

«La diferencia entre variables intermedias y construcciones hipotéticas no es absoluta, sino relativa a la teoría en que se presenten, así como a la posición filosófica adoptada. Un mismo concepto puede manejarse como construcción hipotética en determinado planteamiento, y como intermedia en otro planteamiento. Así, por ejemplo, 'hábito', 'impulso' y 'estado mental' son a lo sumo variables intermedias para la psicología conductista (por ejemplo la de C. Hull) mientras que en otros planteamientos (el de Tolman, por ejemplo) pueden ser construcciones hipotéticas referentes a propiedades reales o sistemas de propiedades del sistema nervioso. En general, la actitud externalista o fenomenista no atribuirá realidad más que a los correlatos de conceptos observacionales, y considerará a todos los conceptos no-observacionales como expedientes auxiliares —y sospechosos— que median e intervienen entre los conceptos observacionales» (Bunge, 1969, p. 114).

Nos parece que en el fondo de esta distinción, late, de manera larvada, la distinción entre «realismo» y «ficcionalismo», acaso no muy distante de aquella que enfrentó a Galileo contra Belarmino a propósito del sistema copernicano. La discusión acerca de si un determinado modelo matemático o legaliforme es una descripción realista de la naturaleza tal cual ella es, o bien «un mero auxilio de nuestra mente». Pero la distinción, tan metafísica como se nos muestra a este respecto, se nos complica y enmaraña terriblemente al asumir, con Bunge, que no puede ser sino relativa. Y es que los constructos hipotéticos y las variables intervinientes no sólo dependen del «problema» como tal (tomado objetivamente), sino de lo que los científicos creen al tratar dicho problema (un punto de vista *emic*, autoconceptual).

La ciencia cognitiva y la búsqueda del significado

La búsqueda de una irreductibilidad esencial de la Psicología se ha apoyado en la noción de *significado* —o contenido—, tal y como

ha venido preocupando tradicionalmente a los filósofos del lenguaje y la semántica filosófica, y se convierte en el ejemplo pre-eminentemente de una *relación funcional* (quizá ajena o desprendida de la causalidad). La noción de contenido pretende ser salvada respecto de dos supuestos peligros de reduccionismo: a) el *fisicalismo*, que persigue reducir todo «hecho mental» a la neurociencia (Churchland, Crick) o incluso a la microfísica (J. Kim); y b) el *sintacticalismo*, dominante en la ciencia cognitiva en relación con las técnicas formalistas de la Inteligencia Artificial (S. Stich). Sin embargo, ese *contenido* mental, irreductible, quiere ser incardinado en «explicaciones naturalistas», esto es, acordes con las que han venido ofreciendo las ciencias naturales. Asumen implícitamente muchos filósofos preocupados del *contenido* mental, que ni la teoría de la evolución ni la neurociencia han de entrar en contradicción con el hecho de que el cerebro humano es capaz de funcionar con alguna intencionalidad, esto es, de ser un sistema que se *refiere* o *apunta* hacia objetos que desbordan al propio sujeto portador de un cerebro. La noción de significado se presenta como un concepto metafísicamente *relacional*.

La psicolingüística cognitiva, muy apegada inicialmente a las teorías chomskianas, a modo de prolongación acerca de la actuación que completase la Gran Teoría de la Competencia del Lenguaje, hubo de emanciparse progresivamente a lo largo de estos años como especialidad propiamente psicológica, sin ir a la zaga servilmente de la Lingüística.

En los 70, Marvin Minsky desarrolla la noción de *descripción simbólica*: los rasgos de una situación son representados por símbolos simples o primitivos. Las relaciones entre símbolos son representadas por otros símbolos. El armazón o marco (*frame*), fue una estructura que Minsky desarrolló para representar situaciones estereotipadas (Minsky, 1975). Los *marcos* eran estructuras de nodos y relaciones que podían reunirse en sistemas de orden superior (*frame-systems*). Un número elevado de conceptos parecidos se desarrollaron en la misma época, tanto en Psicología como en Inteligencia Artificial. Schank y Abelson (1977) elaboraron los guiones (*scripts*) y otras estructuras destinadas a manejar computacionalmente los conceptos que hacen referencia a situaciones sociales o ceremoniales. Ulric Neisser (Neisser, 1976) abogaba decididamente por los *esquemas* en el sentido de Bartlett, muchos de cuyos rasgos se asemejaban a las estructuras antes mencionadas.

En la década de los 70 es la época de los críticos. Podemos dividir dos bloques de críticas según el frente abierto en la guerra contra la metáfora del ordenador. Un bloque es el constituido por aquellos que van a deslegitimar o rebajar las pretensiones de la Inteligencia Artificial

tras una revisión de los fundamentos metafísicos, epistemológicos o ideológicos que subyacen a los programas. Aquí podemos reseñar las críticas de Dreyfus (1972) y las de Weizenbaum (1976). En el ámbito de la Psicología, se va extendiendo una cierta actitud consistente en «detenerse y mirar hacia atrás». Neisser, en su manual *Cognitive Psychology* (Neisser, 1968) ya había expresado algunas de sus reservas hacia la trayectoria que la Psicología estaba recorriendo. Gran parte de cuanto se hacía en experimentación con humanos no era sino una acumulación de trivialidades ajenas a los asuntos realmente importantes de la vida cotidiana. Los sujetos de experimentación eran sumergidos en contextos muy artificiales y con una «libertad operatoria» muy restringida, con el fin de que pudieran ajustarse los datos conductuales a la mecánica y el funcionamiento de los ordenadores:

«El estudio del procesamiento de información tiene actualidad y prestigio, pero todavía no se ha comprometido abiertamente con una concepción de la naturaleza humana susceptible de aplicación más allá de los límites del laboratorio. Y dentro del laboratorio, sus supuestos básicos no van más allá del modelo de computador al cual debe su existencia. Todavía no disponemos de una explicación acerca del modo de actuar o interactuar de los individuos en el mundo ordinario».

Desde los 70 comienza a hablarse, cada vez más, de una *Ciencia Cognitiva* única, cuya misión será aglutinar todo un conjunto de disciplinas que tendrían en común el estudio del *conocimiento*, en el más amplio sentido (no sólo el conocimiento humano, sino el animal y el artificial).

El enfoque de las categorías o conceptos naturales, en realidad, quiere puentear ese hiato que siempre se ha abierto entre percepción y significado, y es una alternativa seria («empirista») al excesivo intelectualismo de la psicología computacional. En realidad, un concepto natural es más bien un *patrón de reconocimiento* de objetos-estímulos, que le sirve al sujeto para *decidir* si es miembro de una categoría o no lo es. Entre los miembros de una determinada categoría hay una serie de *rasgos comunes* o atributos, y que recuerdan mucho a las *invariantes* de que hablaba Gibson en su psicología perceptual. Se trata de desentrañar los mecanismos cognitivos que pueden subyacer a una operatoria «inductiva», esto es, que el sujeto enfrentado ante casos particulares que percibe, llega a «construir» categorías naturales («universales») que están estrechamente enraizados con los preceptos de que ha partido. Esta operatoria inductiva difiere esencialmente del

modelo que diríamos «hipotético-deductivo», esto es, de los modelos cognitivos que proceden de arriba (de la teoría o la hipótesis) hacia abajo (deducción), o si se quiere: de inclusión de un caso particular en el concepto-clase del que se había partido.

El enfoque de los prototipos es constructivista y circular: hay unas restricciones *a priori* en la sensorialidad (actuación de los módulos periféricos), pero también la mente impone sus esquemas o prototipos innatos, que aguardan ser completados con los datos de actualización y detalle que provienen del exterior. El enfoque naturalista se funda en las correlaciones entre los rasgos perceptivos y la extensión de ese concepto, para lo cual se hace un amplio uso de la teoría de la probabilidad y de la lógica borrosa. Por el contrario, el enfoque clásico del tipo «arriba-abajo» (*top-down*) se basa en una caracterización de la *intensión* del concepto (como conjunto de rasgos esenciales o definitorios de un concepto), y mediante una serie de condiciones satisfechas, aplicarse a una serie de objetos (la *extensión*) que cumplen con los requisitos previamente establecidos, luego que los hemos inspeccionado perceptivamente (Gardner, 1985, García-Albea, 1993, p. 129).

Las categorías constan de un núcleo central, formado por los casos más claros u obvios. En torno a ese núcleo, a diferentes distancias (representables como «grados de probabilidad») se sitúan otros miembros de la categoría más o menos similares al caso prototípico. Esto quiere decir que las categorías tienen un alto grado de estructuración interna, según la psicóloga E. Rosch. Ahora bien, el modelo de los conceptos naturales no tiene por qué verse comprometido con las tesis formalistas y rígidamente innatistas del Léxico del Pensamiento: categorías naturales no tiene por qué implicar categorías innatas. Ello puede deberse a un aprendizaje ontogenético mediante el cual los niños aprenden antes los casos paradigmáticos o «centrales» que los casos periféricos o «raros». Por ejemplo, se aprende primero el término «coche», que otros términos emparentados, más difíciles y menos probables en cuanto a su ocurrencia en el medio: «automóvil». Esto quiere decir que el niño no basa sus aprendizajes tanto en estructuras sintácticas o formales de carácter universal, abstracto, desligado de atributos perceptivos, como lo hace por medio de ejemplos «nominales», sencillos y claros, donde el representante «coche» representa con distinción a la mayoría de los objetos que tienen ruedas.

El enfoque de Chomsky y Fodor acerca del aprendizaje de nuevos conceptos, *niega* de hecho, que tal aprendizaje exista en absoluto. Realmente no puede haber conceptos nuevos (en un sentido absoluto en cuanto a su «novedad»), sino, a lo sumo, meros complejos de los conceptos

innatos. ¿Será cierta esta teoría al *modo geométrico*, según la cual construimos figuras —del pensamiento— a partir de unos primitivos y de unos trámites de derivación (construcción sin aprendizaje? ¿No será más bien, y por el contrario, que partimos de las propias figuras del ambiente y de las percepciones, tal como nos vienen dadas en sus aspectos, y que por medio de contrastes, comparaciones, análisis y otras operaciones varias, llegamos al «concepto»? Así pues, los objetos «abstractos», no existen en un sentido propiamente perceptual. Una psicofísica de la percepción no se ve comprometida necesariamente por toda una serie de problemas (filosóficos, en buena medida) relativos a la significación. Escribió Gibson (Reed y Jones, 1982, p. 417) que los objetos abstractos (en un sentido perceptual) no existen: «*There are [abstract objects], to be sure, realities that cannot be perceived directly but only known indirectly by means of instruments, measuring operations, pictures, and language, but it seems to me mistaken to call such realities 'abstract' as against 'concrete'*». En efecto, Gibson considera todos esos objetos complejos y abstractos de nuestra experiencia, ligados estrechamente a las *operaciones constructivas*, que implican percepciones concretas. Los «objetos», para Gibson, en un sentido meramente perceptual, sólo son superficies de sustancias, topológicamente cerradas, (objetos segregados) o bien, superficies casi cerradas (objetos segregados). Pero los objetos perceptuales que sobresalen de un fondo de experiencia no constituyen los únicos elementos de la percepción: concavidades y convexidades, esquinas, cierres, etc...La variedad de cosas que percibimos es enorme, y, por ende, hay una necesidad de *categorizar* imperiosa ante todos esos elementos propios de un conocimiento *implícito*, en orden a hacer de todo ello, un conocimiento *explícito*. De esa operatividad intelectual a partir de las percepciones brota el mismo significado, según la doctrina de Gibson (Reed y Jones, 1982, p. 417). Lo abstracto no sirve como criterio de «significatividad», reñido con lo concreto (perceptual). El funcionalismo computacional admite que los símbolos, con todas sus propiedades, pueden formar parte del sistema cerebral, y, de manera isomórfica, en las computadoras o sistemas cualesquiera capaces de manejar símbolos. Así, los símbolos son una clase de entidades que, básicamente se intercambian entre sujetos, diríamos de forma interpersonal o social. Hay una reordenación continua de los símbolos intracerebrales, en virtud de un intercambio con las configuraciones simbólicas de los otros, esto es, un intercambio de representaciones (Sperber, 1985). Escribe D. Hofstadter (1989, p. 412): «(...) nos planteamos la posibilidad de un isomorfismo entre cerebros en el nivel simbólico: una correspondencia que no sólo vincule entre

*sí los símbolos de cada uno de los cerebros, sino también de los respectivos patrones de desencadenamiento. Esto significa que los símbolos correspondientes están conectados dentro de cada cerebro, en formas que también son correspondientes entre cerebro y cerebro». Así las cosas, programar una computadora, introducirle unos ciertos *inputs* será, de una manera u otra, *dialogar* con ella... Estamos ante la hipótesis de los *sistemas de símbolos físicos*, explicitada por Newell (1980) no hace mucho tiempo, pero asumida de forma muy frecuente y más o menos encubierta desde tiempo atrás. El cerebro, tanto el animal como el electrónico, es un sistema que maneja símbolos, los procesa (combina, reordena, compara, clasifica...), en un ámbito *sintáctico*, predominantemente, sobre el cual se aplican interpretaciones de significado. Esos símbolos, no son entidades fantasmales, sino que se ven instanciados en un soporte físico o *hardware*, que no necesita estar organizado con materia biológica (Simon, 1987), sino de cualquier otra sustancia física. En este ámbito es en el que se pretende mover la Ciencia Cognitiva, en el plano *abstracto* en el sentido de que se extraen o se segregan las peculiaridades sustanciales de las máquinas, tanto humanas como artificiales, así como los propios sistemas formales de la lógica y de las matemáticas.*

El *conexionismo*, partiendo del legado de la Inteligencia Artificial y, más remotamente, de la vieja idea cibernética de las *redes neurales*, pretende instaurar un nivel *subsimbólico* que, a pesar de ciertas declaraciones armonistas supone una verdadera crítica y reorganización del computacionalismo simbólico en Psicología y, a la vez, una acentuada tendencia a abandonar la idea misma de lo *psicológico*, reduciéndolo a una mera *emergencia de origen neurobiológico* mediante el cómputo numérico que subyace al procesamiento de los símbolos .

Conclusiones

A modo de resumen y tabla de conclusiones extraídas a la luz del desarrollo del cognitivismo, podemos decir que (a) el éxito gradual del cognitivismo en Psicología respondió a razones internas dadas en el propio contexto tecnológico y metodológico en que la tradición psicológica ha actuado; (b) por mucho que se haya querido enfatizar el impacto de una nueva clase de tecnologías de la información y la computación, con excepción de la *simulación cognitiva*, las técnicas y métodos propios de los psicólogos no han sufrido grandes variaciones respecto de períodos anteriores; (c) la *metáfora del ordenador* y la teoría de la información

significaron una fuente de términos con los que los psicólogos cognitivos *se representaron su trabajo*, si bien el ejercicio de sus operaciones permaneció enmarcado en *contextos* behavioristas, si bien hubo un regreso a ciertos conceptos de la *Gestalt* o de otras escuelas psicológicas más antiguas; (d) a pesar de las discusiones epistemológicas, promovidas a raíz de la I.A., en torno a la *inteligencia de las máquinas*, lo que ocurrió efectivamente en Psicología fue un regreso al estudio de la *cognición humana*, en detrimento de la preferencia conductista por otras especies como sujetos de investigación; (e) disciplinas externas a la Psicología, como la Gramática Generativa, la Filosofía de la Mente o la Tecnología Computacional, sólo pudieron afectar a los propios *postulados filosóficos imperantes en la Psicología* sustituyéndolos por otros, pero no así en las metodologías y tecnologías psicológicas, salvo de forma indirecta, por medio de una compleja *dialéctica entre categorías e ideas*.

Nota

¹. Los comentarios acerca de la conferencia de Darmouth pueden encontrarse entre otros lugares, en el epílogo histórico de Newell, A. y Simon, H.A. en su *Human Problem Solver*, Plenum Press, Englewood Cliffs, N.Jersey, 1972; Delclaux, I. «Introducción al procesamiento de información en psicología» en Delclaux, I. Y Seoane, J. *op. cit.* pps. 21-38; Gardner, H. *La nueva ciencia de la mente. Historia de la Revolución Cognitiva*. Paidós, Buenos Aires, 1987; Minsky, M. *The Society of Mind*. Heinemann, Londres, 1987; McCorduck, P. *Máquinas que piensan. Una incursión personal en la historia y las perspectivas de la inteligencia artificial*. Tecnos, Madrid, 1991.

Bibliografía

- BUNGE, M. (1969): *La investigación científica*. Ariel, Barcelona.
- CHOMSKY, N. (1959): «Review of Verbal Behavior by B.F. Skinner», *Language*, 35, 1, 26-78.
- DREYFUS, H. (1972): *What computers can't do. A critique of artificial reason*. Harper & Row, N. York.
- FEIGENBAUM, E. A. y FELDMAN, J. (1963): *Computers and thought*. McGraw-Hill, N. York.
- GARCIA-ALBEA, J. (1993): *Mente y conducta*. Trotta, Madrid.
- GARDNER, H. (1987): *La nueva ciencia de la mente. Historia de la revolución cognitiva*. Paidós, Buenos Aires.
- HOFSTADTER, D. R. (1989): *Gödel, Escher, Bach: Un eterno y grácil bucle*. Tusquets, Barcelona.
- HULL, C. L. (1943): *Principles of behavior*. Appleton-Century-Crofts, N. York,.

- JEFFRESS, L. A. (1951): *Cerebral mechanisms in behavior, The Hixon Symposium*. Wiley, N.York.
- KOCH, S. (197): «Psychology and emerging conceptions of knowledge as unitary», pps. 1-45, en Wann, T.W. (ed.), *Behaviorism and phenomenology*.
- LASHLEY, K. (1951): «The problem of serial order in behavior», pps. 112-35 en Jeffress, L.A. (ed.) *Cerebral mechanisms in behavior*.
- McCORDUCK (1991): *Máquinas que piensan. Una incursión personal en la historia y las perspectivas de la inteligencia artificial*. Tecnos, Madrid.
- McCULLOCH, W. S.(1949): «The brain as a computing machine», *Electrical Engineering* LVIII, 492-97.
- MILLER, G. A., GALANTER, E. y PRIBRAM, K. (1960): *Plans and the structure of behavior*. Holt, Nueva York.
- MINSKY, M.: «A framework for representing knowledge», en Winston, P. (ed.), *The psychology of computer vision*.
- MOSTERIN, J.(1987): «Conceptos de información», pps. 115-129 en *Actas II Congreso de Lenguajes Naturales y Lenguajes Formales, Barcelona*.
- NEISSER, U. (1967): *Cognitive psychology*, Appleton Century-Crofts, N. York, Trad. española: *Psicología cognoscitiva*. Trillas, México, 1979.
- (1974): *Cognition and reality. Principles and implications of cognitive psychology*, Freeman, S. Francisco,. Trad. española: *Procesos cognitivos y realidad. Principios e implicaciones de la psicología cognitiva*. Marova, Madrid, 1981.
- NEWELL, A. (1980): «Physical symbol systems», *Cognitive Science*, 4, 135-83. Versión española en pps. 51-71 de Norman, D. (ed.), *Perspectivas de la ciencia cognitiva*.
- REED, E. y JONES, R. (1982): *Reasons for realism. Selected essays of James J. Gibson*. LEA, Hillsdale.
- ROSCH, E. H. «Cognitive representations of semantic categories». *Journal of Experimental Psychology: General*, (1975), 104, 192-233.
- ROSE, J.(1978): *La revolución cibernética*. Fondo de Cultura Económica, México, 1978.
- ROSE, J. R. y MOS, L. P. (1984): *Annals of theoretical psychology* (2 vols.). Plenum Press, N. York.
- RYCHLACK, (1984): «Teleological psychological theory», pps. 115-50 en Royce, J. y Mos, L. (ed.), *Annals of theoretical psychology*. Plenum Press, Londres-Nueva York
- SCHANK, R. y ABELSON, R. (1977): «Scripts, plans, goals, and understanding», LEA, Hillsdale,. Trad. española: *Guiones, planes, metas y comprensión*. Paidós, Barcelona, 1984.
- SHANNON, C. y WEAVER, W. (1949): *The mathematical theory of communication*. Univ. of Illinois, Urbana, 1949 (reimp. en 1971).
- SIMON, H. A., (1987): *Ciencia cognitiva: la más nueva ciencia de lo artificial*, pps. 25-39 en Norman, D. (ed.), *Perspectivas de la ciencia cognitiva*. Paidós, Barcelona.
- SPERBER, D. (1985): *Ciencias cognitivas, ciencias sociales y materialismo*, 85, 39-62
- TOLMAN, E.Ch. (1948): «Cognitive maps in rats and men», *Psychological Review*, 55, 189-208.
- TURING, A. M. (1937): «On computable numbers, with an application to the *Entscheidungsprobleme*», *Proc. Lond. Math. Soc., ser. 2*, 42, 230-65, y en 43, 544-46.
- (1950): «Computing machinery and intelligence». *Mind*, LIX , 236.
- WANN, T. W. (1974): *Behaviorism and phenomenology*. Univ. of Chicago Press, Chicago.

WEIZENBAUM, J. (1972): *Computer power and human reason*. Freeman, S. Francisco.
Trad. española: *La frontera entre el ordenador y la mente*. Pirámide, Madrid, 1976.

WINSTON, P. (1975): *The psychology of computer vision*. McGraw-Hill, N. York.