

UNIVERSIDAD DE CHILE
Facultad de Filosofía y Humanidades
Escuela de Postgrado

EL ROL DE LA REPRESENTACIÓN MENTAL EN LA CIENCIA COGNITIVA

Tesis para optar al grado de Magíster en Filosofía con mención en Epistemología

Autor:

Mónica Rengifo Oyarce

Profesor Patrocinante: Guido Vallejos Oportot

Santiago, Chile 2008

..	1
AGRADECIMIENTOS .	3
RESUMEN .	5
INTRODUCCIÓN: FORMULACIÓN DEL PROBLEMA . .	7
CAPÍTULO PRIMERO: LA REPRESENTACIÓN MENTAL EN EL PARADIGMA SIMBÓLICO .	11
1. SUPUESTOS SUSTANTIVOS DEL COGNITIVISMO .	11
2. LA TEORÍA REPRESENTACIONAL Y COMPUTACIONAL DE LA MENTE DE JERRY FODOR . .	13
3. PRESCRIPCIONES METODOLÓGICAS DE LA TRMC .	16
3.1. LA ESTRATEGIA DEL SOLIPSISMO METODOLÓGICO .	16
3.2. INDIVIDUALISMO METODOLÓGICO . .	18
4. FISCALISMO Y REALISMO INTENCIONAL .	20
5. NATURALIZACIÓN DE LA INTENCIONALIDAD .	22
5.1. LA TEORÍA CAUSAL INFORMACIONAL DEL CONTENIDO Y EL PROBLEMA DE LA DISYUNCIÓN . .	23
5.2. LA TEORÍA DE LA DEPENDENCIA ASIMÉTRICA .	24
6. EVALUACIÓN CRÍTICA DE LA TRMC .	25
CAPÍTULO SEGUNDO: LA REPRESENTACIÓN EN LOS SISTEMAS CONEXIONISTAS .	31
1. COMPONENTES BÁSICOS DE LOS MODELOS CONEXIONISTA SEGÚN RUMELHART .	32
2. VENTAJAS DEL CONEXIONISMO FRENTE AL MODELO CLÁSICO .	34
3. LA REPRESENTACIÓN EN LOS SISTEMAS CONEXIONISTAS .	35
4. EL DEBATE EN TORNO A LA ESTRUCTURA CONSTITUTIVA DE LAS REPRESENTACIONES MENTALES . .	37
4.1. LA CRÍTICA DE FODOR Y PYLYSHYN AL CONEXIONISMO . .	37
4.2. LA RESPUESTA DE PAUL SMOLENSKY A FODOR Y PYLYSHYN .	39
5. LA ARQUITECTURA SUBSIMBÓLICA DE SMOLENSKY . .	48
6. EVALUACIÓN CRÍTICA DEL CONEXIONISMO . .	50
CAPÍTULO TERCERO: LA REPRESENTACIÓN EN LAS TEORÍAS ALTERNATIVAS DE LA	53

CIENCIA COGNITIVA .	
1. LA COGNICIÓN COMO SISTEMAS DINÁMICOS, SEGÚN VAN-GELDER . .	54
2. LA COGNICIÓN CORPORALIZADA: ANTECEDENTES FILOSÓFICOS .	58
3. EL ENFOQUE ENACTIVO Y ANTIRREPRESENTACIONALISTA DE VARELA .	60
4. LA POSTURA ECUMÉNICA DE ANDY CLARK . .	64
4.1. REFINANDO LA NOCIÓN DE REPRESENTACIÓN . .	68
4.2. MENTE Y MUNDO: LA FRONTERA ESCURRIDIZA .	71
5. EL ROL DE LA REPRESENTACIÓN EN LAS TEORÍAS ALTERNATIVAS . .	74
5.1. REPRESENTACIÓN Y SISTEMAS DINÁMICOS .	74
5.2. COMPARACIÓN CRÍTICA DE LOS ENFOQUES DE VARELA Y CLARK .	75
5.3. EVALUACIÓN DEL ENFOQUE CORPORALIZADO DE LA COGNICIÓN . .	78
CONCLUSIONES FINALES .	81
REFERENCIAS BIBLIOGRÁFICAS .	85
ARTÍCULOS ONLINE . .	87

A Felipe, mi hijo. Por todas aquellas veces en que me he visto imposibilitada en dedicarle un diálogo atento y ameno; en pro del trabajo en esta tesis.

AGRADECIMIENTOS

A Francisco y a mamá. Por su amor, su paciencia y apoyo en los quehaceres cotidianos. Por la voz de aliento, cariñosa e incondicional que me ha acompañado en este largo tiempo de trabajo.

En especial al profesor Guido Vallejos, por su férrea defensa del paradigma clásico y la propuesta naturalista de Jerry Fodor, que me motivaron a dar cuerpo a la temática de esta investigación. Por su tiempo destinado en la corrección de este trabajo y por su acogida siempre amistosa, cada vez que he necesitado de su ayuda.

RESUMEN

La presente investigación tiene por objetivo fundamental clarificar el rol que juega la representacional mental en los variados enfoques de la Ciencia cognitiva. Teniendo presente el desarrollo de la ciencia cognitiva (década del cincuenta hasta la actualidad), podemos visualizar tres grandes marcos de investigación que dan cuenta de su desarrollo: el cognitivismo clásico, el conexionismo y las teorías alternativas. El primer marco de investigación, se expone a partir de la Teoría *representacional y computacional de la mente* de J. Fodor, con el objeto de realizar una revisión crítica del *rol formal* que asigna este enfoque a la representación mental. En su caracterización de la conducta inteligente, el cognitivismo clásico asume dos tesis fundamentales: La primera es que los procesos mentales pueden ser concebidos como procesos computacionales (lo cual significa en última instancia que pueden ser concebidos como procesos de transformación sintáctica de símbolos); la segunda, que los estados mentales son ante todo estados representacionales (lo cual significa que los símbolos procesados deben tener contenido). Uno de los problemas que surgen al interior de este enfoque es ¿cómo adquieren contenido las representaciones mentales? El paradigma cognitivista debe en consecuencia poder explicar las propiedades semánticas de los estados mentales. A la luz de este problema, se discuten las bases teóricas sobre las que descansa el enfoque clásico: el *individualismo metodológico*. Por tal razón, se expone posteriormente la crítica realizada por el conexionismo y algunas de las teorías alternativas en Ciencia Cognitiva al paradigma clásico, con el objeto de clarificar qué rol juega la *representación mental* en estos dos enfoques.

La conclusión general de esta investigación es que los supuestos metodológicos que asume la perspectiva clásica son insuficientes para explicar la cognición y dar cuenta de manera efectiva del rol que juega la representación mental en la explicación de la conducta inteligente de un agente cognitivo. Finalmente se sugiere considerar la propuesta de Andy Clark, como una posible alternativa que podría guiar la investigación actual en Ciencia Cognitiva, aún cuando dicha propuesta no constituya todavía una teoría unificada, sino más bien sólo un esquema teórico.

INTRODUCCIÓN: FORMULACIÓN DEL PROBLEMA

En la década de los cincuenta comienza el desarrollo sostenido de la nueva ciencia de la mente, denominada ciencia cognitiva. Dicha ciencia intenta obtener un conocimiento de los procesos mentales inteligentes, tales como la cognición, la percepción, los mecanismos de aprendizaje, el funcionamiento de la memoria y la toma de decisiones. La ciencia cognitiva (CC en adelante), se constituye sobre la base del aporte de diversas disciplinas, entre las cuales adquieren esencial relevancia la filosofía (en especial, la filosofía de la mente, la filosofía del lenguaje y la lógica), la psicología cognitiva, la psicolingüística, la inteligencia artificial, la neurociencia, la ciencia de la computación y, en algunos casos, la antropología.

El desarrollo de la CC, desde sus comienzos hasta la actualidad, se puede visualizar de modo esquemático, a partir de tres grandes marcos de investigación: El cognitvismo clásico, el conexionismo y las teorías alternativas. El *cognitvismo clásico*, denominado también como *paradigma simbólico* o *computacionalismo ortodoxo*, representa la perspectiva hegemónica. Su metáfora rectora, ha sido el ordenador o computador digital. A partir de esta metáfora, la mente y la cognición humana han sido modeladas siguiendo la analogía del computador *Von Neumann* (1945). La *cognición* es definida en este enfoque, como procesamiento de información, a partir de manipulación de símbolos basados en reglas. Las **representaciones mentales**, se conciben como *símbolos capaces de recibir una interpretación semántica en función de la relación particular que*

*poseen con el ambiente o con otros símbolos.*¹

En su descripción de la conducta inteligente, el cognitivismo clásico adscribe al contenido de los símbolos interpretados la propiedad de ser acerca de algo distinto de ellos mismos, esto es, la propiedad de la **intencionalidad**. Con ello da por sentado que la conducta de un agente cognitivo está determinada por estados mentales que representan los rasgos relevantes de la situación en que se encuentra y por computaciones que se realizan sobre esos símbolos. El cognitivismo considera que la única manera de explicar la inteligencia y la intencionalidad es a partir de la hipótesis de que la cognición consiste en actuar a partir de procesos mentales que se realizan sobre **representaciones** que se instancian físicamente en forma de un código simbólico dentro del cerebro o de una máquina y que son capaces de combinarse con otros símbolos para formar estructuras más complejas.

Según lo anterior, vemos que el computacionalismo clásico define la cognición como *computación de representaciones simbólicas*, concibiendo a los procesos cognitivos como *procesos computacionales*. Éstos últimos son de carácter algorítmico, lo que significa que una determinada función se realiza a través de una serie de transformaciones de estructuras simbólicas. Estas transformaciones se realizan a partir de un conjunto de reglas que determinan los pasos permitidos de cada estado del sistema. Newell y Simon (1994) definieron un símbolo físico como:

Un sistema de símbolos físicos consiste en un conjunto de entidades, llamadas símbolos, (...) que pueden presentarse como componentes de otro tipo de entidad llamada expresión (o estructura del símbolo). De este modo la estructura del símbolo se compone de un número de instancias (o signos) de símbolos relacionados de alguna manera física (como cuando un signo se encuentra junto a otro).²

El cognitivismo afirma además que los estados mentales son ante todo *estados representacionales*, lo que significa que los símbolos procesados deben tener un contenido. En consecuencia podemos decir que una parte esencial de la hipótesis cognitivista postula que el dominio propio de las CC corresponde al estudio de la cognición, en cuanto **representación mental**. Lo que se intenta es representar el conocimiento que los agentes manipulan y almacenan. No obstante, para efecto de una teoría científica (psicología computacional), el cognitivismo sólo considera la dimensión formal y sintáctica de la representación, asumiendo que las representaciones están interpretadas. Justamente, por este modo de entender la cognición, este enfoque ha sido denominado también como paradigma simbólico o computacionalismo ortodoxo.

Para ejemplificar esta visión, se expondrá la concepción de la mente de un autor que al decir de Gardner es *el cognitivista cabal*³. Se trata del filósofo norteamericano J. A. Fodor. Dicho autor se ha preocupado de abordar la justificación o fundamentación de la ciencia cognitiva en casi todos sus aspectos, extrayendo consecuencias que no sólo

¹ Clark, *Estar ahí*, p. 16.

² Newell and Simon, *La ciencia de la computación como investigación empírica*, p.127.

³ Ver Gardner, *La nueva ciencia de la mente*, p. 98

involucran a la filosofía de la ciencia cognitiva sino también a la epistemología y a la metafísica. Dada la diversidad de aspectos que Fodor aborda de la CC, se hace difícil tener un cuadro unitario de sus planteamientos. Es por esto, que un primer objetivo de esta investigación consiste en intentar sistematizar las distintas dimensiones de la justificación filosófica de la CC, llevada a cabo por Fodor. Esta sistematización, permitirá señalar los problemas generales que presenta el *paradigma simbólico* dentro de la ciencia cognitiva, y en especial destacar el rol formal que le asigna a **la representación mental**.

Cabe señalar que la aceptación del nivel representacional, postulado por el enfoque cognitivista en las CC, no ha estado libre de polémica. Por un lado, se han desarrollado líneas críticas de la concepción representacional, considerando por ejemplo que ésta construcción teórica de la mente debe ser sometida al principio de economía (navaja de Occam) y que en consecuencia postular un nivel representacional separado de análisis, resulta innecesario, equívoco e incongruente, por varias razones. (Véase por ejemplo, la postura eliminativista de Paul Churchland, 1981). Además, dentro de los autores que aceptan el nivel representacional en las CC, se han dado importantes debates sobre cuál sea la forma más conveniente de conceptualizar la representación mental.

En la década del ochenta en adelante, han surgido otros enfoques en la CC que difieren notoriamente en el modo de entender la representación mental y en varios aspectos centrales de la hipótesis cognitivista. En oposición al tipo de organización que presenta el cognitivismo clásico, surgen los distintos tipos de *sistemas conexionistas*. Para el *conexionismo*, denominado también como *Paradigma subsimbólico* o PDP (*Procesamiento Distribuido en Paralelo*), una representación consiste en la correspondencia entre un estado global emergente de una red conexionista y las propiedades del mundo; por lo que no corresponde a una función de símbolos particulares, como establece el modelo clásico. Los distintos tipos de sistemas conexionistas, poseen una organización que difiere del modelo cognitivista; estos sistemas están constituidos por unidades simples de procesamiento interconectadas de modo tal que conforman una red (red neuronal artificial). La estructura de esta red está inspirada (en un sentido muy general) en las características que posee la arquitectura neuronal del sistema nervioso (de ahí el nombre *redes neuronales artificiales*). Estas unidades se comunican entre sí, gracias a que poseen un umbral de activación e inhibición que permite la propagación de esa activación a otras unidades, ya sea activándolas o inhibiéndolas (dependiendo del peso y valor de estas conexiones). En consecuencia, la conducta del sistema es el resultado de la interacción de todos estos elementos.

Posterior al enfoque del conexionismo (década del noventa en adelante), se desarrollan varias teorías alternativas en la CC como por ejemplo: La robótica situada (Brooks, 1991), la teoría de los sistemas dinámicos (Van Gelder, 1996, entre otros), la perspectiva de la *cognición corporalizada* (Varela, 1990, 1997 y Clark, 1997), que en términos generales presentan un cuestionamiento a la pertinencia de la noción de representación como supuesto fundamental de la ciencia cognitiva. En la mayoría de estas visiones, nos encontramos con un cierto escepticismo sobre los poderes causales de las propiedades mentales, y con un cuestionamiento al carácter central de la

representación mental en las CC. Por ejemplo, si tenemos presente la perspectiva de la cognición corporalizada, vemos que ésta, cuestiona el enfoque del cognitivismo clásico y del conexionismo, alegando que éstos no proporcionan una imagen real del entorno y del cuerpo como recurso activo en la cognición. En opinión de Varela y de Clark, la cognición emerge de relaciones causales más complejas entre el cerebro, el cuerpo y el ambiente, de lo que se desprende que las representaciones mentales no son las únicas propiedades causalmente relevantes para una explicación de la cognición.

A partir de lo expuesto y teniendo presente el desarrollo de la CC, podemos decir que no existe una única manera de entender la *cognición*, como tampoco una visión homogénea del rol que ocupa *la representación mental*, en la CC. En consecuencia, lo que se diga respecto de la cognición y su arquitectura en cada enfoque, determinará si se asume o no algún tipo de representación y cuál será la forma de esa representación (si es que se asume). Podemos preguntarnos entonces, *¿cuál es el rol explicativo que desempeña la representación mental en los enfoques señalados de la CC? ¿Desempeña o no la representación mental un papel relevante en las explicaciones en ciencia cognitiva?* Como también, si *¿desempeña o no la representación mental un papel causal determinante en la etiología de conductas intencionales?* La búsqueda de respuesta a estas preguntas, constituye el objetivo central de esta investigación. Esclarecer esto, nos permitirá examinar el rol o papel que desempeñaría (si es que llega a desempeñar alguno) *la dimensión representacional*, en un nuevo marco de una ciencia cognitiva. Si la respuesta frente a esta última pregunta resulta negativa, ¿significaría esto entonces, que debería desaparecer *la dimensión representacional*, para los efectos de un nuevo enfoque? Intentar esclarecer el rol que tiene la representación mental en los distintos enfoques de la CC, nos permitirá en consecuencia evaluar cuál es en definitiva la *relevancia de la representación mental* para cierta concepción global de la arquitectura cognitiva y analizar cuáles podrían ser las implicancias del rechazo de la noción de representación para una ciencia de la mente.

Para el logro de los objetivos de esta investigación, se expone en la primera parte, los supuestos centrales del *cognitivismo* a partir de la *Teoría representacional y Computacional de la mente* de Fodor, con el objeto de realizar una revisión crítica del rol que asigna este enfoque (paradigma simbólico) a la representación mental. Posteriormente se presenta el enfoque del *conexionismo* con su concepción particular de la representación mental y finalmente se exponen, algunas de las visiones presentes en las *teorías alternativas* de las CC. Cabe señalar que dentro de estas visiones se considera con mayor preponderancia la perspectiva de Clark (2001), dado que sus investigaciones nos permiten visualizar una síntesis de los problemas presentes en los paradigmas anteriores, junto con una posible integración de los variados enfoques en las CC.

CAPÍTULO PRIMERO: LA REPRESENTACIÓN MENTAL EN EL PARADIGMA SIMBÓLICO

La versión más institucionalizada y hegemónica de la ciencia cognitiva corresponde a la Teoría representacional y computacional de la mente (TRCM, en adelante), la cual se desarrolla a partir de la siguiente hipótesis: *la mente funciona como un procesador de información*. Bajo esta hipótesis, subyacen además dos supuestos sustantivos que han sido centrales en la década cognitivista: el supuesto computacional y el supuesto representacional.⁴

1. SUPUESTOS SUSTANTIVOS DEL COGNITIVISMO

(i) **El supuesto computacional** se explicita mediante la afirmación de que la mente es un mecanismo computacional. De acuerdo a esta visión, se conciben los procesos cognitivos de la mente como funcionalmente equivalentes a un computador (modelo Von Neumann). Un computador, es concebido por la Inteligencia artificial (IA, en adelante) como un

⁴ Ver B. Von Eckart (1993), *What is cognitive science*.

sistema formal automático, vale decir, como un conjunto de procedimientos de manipulación de símbolos regido por reglas (Haugeland, 1985).

Pensar, consiste en consecuencia, en realizar computaciones (transformaciones) sobre símbolos. Pero el pensamiento, desde un punto de vista psicológico, en los seres con inteligencia no se realiza sobre símbolos puramente formales, sino sobre símbolos con contenido, vale decir sobre representaciones. De esta consideración emerge el supuesto representacional de la ciencia cognitiva.

(ii) El supuesto representacional, concibe a la mente como un sistema simbólico y más específicamente como un sistema representacional. Esta visión complementa el supuesto anterior (i), puesto que señala que las computaciones que caracterizan los procesos cognitivos se realizan sobre símbolos. De este modo, para los efectos de la realización de las computaciones sólo resulta necesaria la dimensión formal de las representaciones y no su dimensión semántica, esto porque las computaciones preservan el contenido de las representaciones. Sin embargo a pesar de esta consideración, vemos que las representaciones poseen además un contenido intencional, vale decir, poseen un contenido que se refiere a algo distinto de ellas mismas.

La ciencia cognitiva tradicional (cognitivismo) considera que un sistema inteligente opera a través de **representaciones**. La cognición se define, como *computaciones sobre representaciones simbólicas* (formulación computacional de la representación), presuponiendo que los símbolos (representaciones) sobre los cuales se realizan estas computaciones son símbolos interpretados. Una computación es una operación llevada a cabo sobre *símbolos* que *representan* aquello que designan, puesto que las transformaciones operan sobre la base de la forma de las representaciones, preservando el contenido de ellas. Esto significa afirmar, que las computaciones son operaciones que respetan los valores semánticos de los símbolos o están **restringidos** por ellos. En consecuencia, el nivel semántico de las expresiones simbólicas está determinado en un ordenador por la sintaxis, que refleja su proyección semántica. El cognitivista afirma que existe un **paralelismo** entre la sintaxis y la semántica similar al que ocurre en un ordenador donde la sintaxis es reflejo de la semántica. Este paralelismo nos muestra cómo son física y mecánicamente posibles la inteligencia y la intencionalidad (semántica). Los ordenadores proporcionan en consecuencia, un modelo mecánico del pensamiento, indicando que éste consiste en computaciones físicas y simbólicas.

Resumiendo, podemos decir que para caracterizar los procesos cognitivos la teoría representacional y computacional de la mente (en adelante, TRCM) se limita a una descripción de sus operaciones *sintácticas* sin considerar los aspectos *semánticos*. La TRCM considera como relevante el supuesto computacional, que abarca sólo **la dimensión formal de las representaciones mentales**. Sostiene en consecuencia, que la dimensión semántico-intencional no es un tema que forme parte de la ciencia cognitiva. En consecuencia, la CC no debería investigar más allá de lo que proyectan las relaciones sintácticas entre proposiciones. El problema de por qué cada constituyente de esas representaciones significa lo que significa y en virtud de qué tipo de relación puede determinarse ese contenido es un tema que queda en manos de la filosofía y no de la CC.

Estas restricciones impuestas por la TRCM y claramente explicitadas en los

planteamientos de J.A Fodor que revisaremos posteriormente, nos revelan un cierto **pesimismo** en cuanto a la posibilidad de comprender de qué manera nuestros sistemas computacionales versan sobre los contenidos de las representaciones mentales (dimensión semántica). Al respecto, Fodor opina que, quizás le sea imposible al cientista cognitivo hacer semántica y que aún cuando se sostenga la existencia de estados mentales (tales como creencias, deseos e intenciones) que poseen eficacia causal; finalmente la única clase de psicología que podamos llegar a tener, sea una psicología computacional, vale decir, una psicología de las operaciones mentales formales. Este enfoque sintáctico de la cognición de J.A. Fodor, ha sido denominado por él mismo con la expresión *solipsismo metodológico* (tomada de Carnap).⁵

Finalmente, podemos decir que pareciera existir cierta paradoja en la TRCM de Fodor, puesto que ésta afirma por un lado, que cualquier tentativa por comprender la cognición humana debe enmarcarse claramente en una posición mentalista, vale decir, se asume que el comportamiento cognitivo es el producto de procesos mentales internos, afirmando además que los estados mentales existen realmente, interactuando entre sí, siendo posible su estudio. La noción de **representación** (o bien de intencionalidad) como hemos visto, juega un rol clave en el cognitivismo, siendo uno de sus supuestos sustantivos. No obstante lo anterior, se afirma por otra parte, que la dimensión semántica-intencional no le concierne a una psicología computacional. Cabe señalar al respecto, que si se desea sostener que los estados intencionales tienen propiedades causales, es necesario mostrar no sólo cómo son físicamente posibles dichos estados, sino también cómo pueden causar la conducta inteligente. En consecuencia, podemos decir que la principal dificultad que presenta esta posición representacionalista, radica en lograr establecer con exactitud cómo el contenido de un determinado ítem representacional físico se determina (problema semántico). El gran desafío cognitivista, radica justamente, en lograr dar una explicación a este problema.

En lo que sigue, se expondrá en detalle la TRCM de Fodor, junto con otros aspectos de su teoría, con el objeto de lograr explicitar la problemática presentada.

2. LA TEORÍA REPRESENTACIONAL Y COMPUTACIONAL DE LA MENTE DE JERRY FODOR

Jerry Fodor postula una variante fisicalista de la mente que admite la existencia de causas mentales de la conducta inteligente. Según esta posición las actividades cognitivas se materializan, es decir, se constituyen mediante la manipulación de los símbolos o **representaciones mentales**. Fodor desarrolla una teoría representacional y computacional de la mente (TRCM) que ocupa un polémico lugar dentro del establecimiento de los fundamentos de la ciencia cognitiva. De acuerdo a esta teoría, la mente se concibe como una máquina sintácticamente manejada que opera sobre las propiedades formales de las **representaciones mentales**, preservando sus propiedades.

⁵ Expresión que explicitaremos en la sección 5.2. de esta investigación.

El punto de partida de la TRCM, está dado por una vindicación de la psicología de sentido común (PSC, en adelante), a partir de la cual es posible transitar a una psicología cognitiva computacional. Según Fodor, las generalizaciones que pueden establecerse a partir de la PSC, nos señalan que ésta puede ser entendida como una teoría. La concepción popular explica la conducta inteligente, es decir, aquél tipo de conducta guiada por metas y propósitos, apelando a **actitudes proposicionales**. Concibe al pensamiento como una relación funcional (computacional/causal) con ejemplificaciones de representaciones mentales que expresan contenido. La PSC muestra claramente la importancia de la **dimensión semántica de las representaciones mentales**, dimensión que será constreñida para efectos de una psicología computacional.

En opinión de Fodor, es posible vincular una PSC en su dimensión intencional, con una psicología computacional, mediante la postulación de la **Teoría Representacional de la Mente** (TRM en adelante) cuyo corazón está constituido además por la postulación de un **lenguaje del pensamiento**. La psicología cognitiva concibe a la mente como un dispositivo procesador de información (equivalente al modelo de la máquina de Turing), según el cual los procesos mentales son sucesiones, series o cadenas de estados mentales, vale decir, son cómputos sobre símbolos. La concepción de la mente como un dispositivo procesador de información, conduce a plantear un medio de representación en el cual los cómputos involucrados en el procesamiento son llevados a cabo. Este medio son las representaciones mentales, que suponen la existencia de un lenguaje del pensamiento (LDP en adelante), que proporciona el material y los procedimientos para computar representaciones. Fodor cree en consecuencia, en la existencia de un LDP, teoría que es explicitada en su importante ensayo, *The Language of Thought* (1975). En efecto, si los procesos mentales son efectivamente computacionales, debe haber representaciones en las cuales se ejecutan dichas computaciones. El LDP es aquel intermediario, que posibilita el pensar y por ende, las capacidades cognitivas.

La TRM por su parte, es una teoría acerca de lo que son los estados y procesos mentales. Su formulación canónica se puede resumir a partir de tres tesis fundamentales:

6

(T1) *La explicación psicológica es típicamente nómica y es completamente intencional*. Esto significa que, las leyes que invoca la explicación psicológica expresan (de manera típica) relaciones causales entre estados mentales que se escogen por referencia a sus contenidos.

(T2) *Las representaciones mentales son las portadoras primitivas del contenido intencional*. Esto significa que, tanto ontológica como explicativamente, la intencionalidad de las representaciones es previa a la intencionalidad de los lenguajes naturales y es previa del mismo modo, a la intencionalidad de las actitudes proposicionales. Las representaciones mentales son las portadoras originales de la intencionalidad y a partir de ellas reciben intencionalidad derivada, primero las actitudes proposicionales, y posteriormente gracias a estas, las expresiones lingüísticas. En consecuencia, la TRM afirma que las representaciones mentales se relacionan con las actitudes proposicionales del siguiente modo: *Para cualquier organismo O, y para cualquier actitud A hacia la*

⁶ Ver Fodor, *Conceptos. Donde la Ciencia Cognitiva se equivocó*, pp. 24 - 30

*proposición P, hay una relación (computacional/ funcional) R y una representación mental tal que RM significa P, y O tiene A si y sólo si mantiene R con RM.*⁷

(T3) *El pensamiento es computación.* Lo que significa, dicho en pocas palabras, que las instancias de las representaciones mentales son símbolos; y que las instancias de los símbolos son objetos físicos con propiedades semánticas. A partir de esto, se establece además que: *los procesos mentales son secuencias causales de muestras de representaciones mentales.* Una cadena de pensamiento es equivalente a una secuencia causal de *muestras* de representaciones mentales que expresan las proposiciones. Por ejemplo si *X cree que va a llover* y por lo tanto *X va a un sitio cubierto*, ello es equivalente a tener una *muestra* de una representación mental que significa '*iré a un sitio cubierto*', que está causada por una *muestra* de representación mental que significa '*va a llover*'.

En opinión de Fodor la TRM brinda un mecanismo plausible para explicar la conexión entre las propiedades causales y las propiedades semánticas, de los estados mentales:

El truco consiste en combinar la postulación de representaciones mentales con la metáfora del ordenador. De este modo los ordenadores nos muestran cómo conectar las propiedades semánticas de los símbolos con las propiedades causales.⁸

Para Fodor, las propiedades causales de un símbolo se conectan con las propiedades semánticas por medio de sus **sintaxis**. La sintaxis de un símbolo se reduce a su forma y ella es la que permite graficar cómo puede un símbolo interactuar causalmente en virtud de sus estructuras sintácticas. Según la metáfora del ordenador, los procesos mentales se consideran como cómputos, vale decir, transformaciones realizadas en virtud de reglas sintácticas que operan sobre los símbolos mentales. Las representaciones mentales son símbolos que expresan proposiciones. De esta manera, la sintaxis juega el papel de mediar entre las propiedades causales (formales) de las representaciones y las propiedades semánticas de las proposiciones expresadas por esos símbolos, permitiendo de este modo que los procesos sintácticos que operan sobre las representaciones preserven la verdad de las proposiciones que estas expresan:

Lo que la TRM toma en préstamo de los ordenadores es, en primer lugar, la receta para mecanizar la racionalidad: usar una máquina guiada sintácticamente para aprovechar los paralelismos entre las propiedades sintácticas y las propiedades semánticas de los símbolos.⁹

Según Fodor, una psicología computacional no puede dar cuenta del problema semántico de las actitudes proposicionales, puesto que para él una psicología computacional trata con propiedades de las actitudes que no son semánticas (sino, sólo sintácticas, como hemos visto).

Resumiendo, podemos decir que la TRM resulta comprometida con lo siguiente:

⁷ Fodor, *Psicosemántica*, p. 40

⁸ *Ibid.*, p. 40.

⁹ *Ibid.*, p. 47.

1. Afirma que los procesos mentales son transformaciones de representaciones mentales, cuyos contenidos deben estar representados explícitamente.
2. Afirma que los procesos mentales son secuencias causales de instanciaciones.
3. Justifica la presencia del constructo de representación mental en las explicaciones de la psicología.

En opinión de Fodor, la TRM combinada con la *Metáfora del computador* proporciona realmente un avance para la ciencia cognitiva contemporánea. Justamente, por este motivo, su teoría ha sido denominada como *Teoría Representacional y Computacional de la Mente*¹⁰ (*TRCM en adelante*). Dicha teoría concibe la imagen de la mente como una máquina que maneja la sintaxis, afirmando con ello la existencia de símbolos mentales con propiedades semánticas y sintácticas.

3. PRESCRIPCIONES METODOLÓGICAS DE LA TRMC

Con el objeto de que se pueda asumir el carácter científico de la TRCM, Fodor y los partidarios de esta teoría, agregan ciertas prescripciones de carácter metodológico para impedir que se introduzcan factores que no sean explicativamente relevantes, en las hipótesis de esta teoría. Una de estas prescripciones corresponde al individualismo metodológico y otra al solipsismo metodológico. Cabe señalar al respecto, que en *Psicosemántica* (1987) Fodor insiste en que es muy importante distinguir entre ambas doctrinas, dado que en general se confunden.

3.1. LA ESTRATEGIA DEL SOLIPSISMO METODOLÓGICO

En su escrito *Solipsismo metodológico como estrategia de investigación en ciencia cognitiva* Fodor (1980) distingue entre la versión de la TRM (teoría representacional de la mente) y la versión de la TCM (teoría computacional de la mente), considerando que la segunda corresponde a una teoría más fuerte de la ciencia cognitiva, que establece que los estados y procesos mentales son computacionales, o al menos aquellos a los que la psicología cognitiva debe atender. Según Fodor, los procesos computacionales son, a la vez, simbólicos y formales. Simbólicos porque son definidos sobre representaciones y formales porque se aplican a las representaciones en virtud de sus propiedades físicas de nivel superior. Las operaciones formales son aquellas que son especificadas sin hacer referencia a propiedades semánticas de las representaciones tales como verdad, referencia y significado.

El hecho de considerar estos procesos como siendo de índole puramente formal, significa establecer una restricción o condición para la psicología cognitiva (computacional), la cual es denominada por Fodor *condición de formalidad*. En consecuencia, la aceptación que Fodor hace de la teoría computacional lo conduce a la

¹⁰ Ver Lower y Rey, *Meaning in Mind: Fodor and his Critics*, Editors Introduction, p. XV.

imposición del solipsismo metodológico:

El solipsismo metodológico es, de hecho, una teoría empírica que dice que los procesos mentales son computacionales y, por consiguiente, sintácticos.¹¹

La doctrina del solipsismo metodológico nos señala un constreñimiento para la individuación de los estados mentales por sus contenidos. Según este constreñimiento se sostiene que los estados psicológicos se individualizan sin relación a su evaluación semántica. Lo anterior significa establecer que la psicología debe ocuparse exclusivamente de los estados psicológicos en el *sentido estrecho* (*the narrow sense*), es decir, aquellos estados psicológicos que no presuponen la existencia de ningún individuo, salvo la del sujeto al que el estado es adscrito. En consecuencia, se afirma que para la perspectiva de la psicología mentalista *lo que está en el mundo* es algo relevante, puesto que no ocasiona diferencias en los contenidos intencionales de los sujetos. Lo anterior significa afirmar que *el contenido amplio* (*the wide sense*) no es determinante en los contenidos intencionales. Lo anterior resulta ser una consecuencia problemática para Fodor, dado que no es posible explicitar con esto el rol del contenido semántico en la producción del comportamiento inteligente de un agente cognitivo.

Las distinciones entre *contenido amplio* y *contenido estrecho*, fueron tomadas por Fodor, de Putnam (1975) quien en su artículo *El significado del significado* distinguió entre *los estados psicológicos en sentido amplio* (*the wide sense*) y *los estados psicológicos en el sentido estrecho* (*the narrow sense*). Un estado psicológico en sentido estrecho es aquel que no presupone la existencia de cualquier otro individuo o cosa, para su atribución (es decir, estados psicológicos que no hacen referencia al mundo). Todos los otros estados psicológicos serían estados en sentido amplio. En este escrito, Putnam analiza la siguiente pregunta: *¿están los significados en la cabeza?* Para dar respuesta a ella, introduce el conocido ejemplo de ficción sobre *La Tierra Gemela*. En palabras simples, su argumentación puede resumirse en lo que sigue: Imaginemos que existe un planeta similar en casi todo a la tierra, *La Tierra Gemela*. Las peculiares diferencias existentes en dicho planeta, estarían dadas en que el líquido llamado agua en la Tierra Gemela no es H₂O sino un líquido diferente cuya forma química puede ser abreviada como XYZ. Se supone además que XYZ no es distinguible informalmente de H₂O, salvo mediante una prueba química. En dicho planeta existe además un individuo idéntico a mí (incluso a mi microestructura neurofisiológica), vale decir, un gemelo. Putnam sostiene que a pesar de la igualdad de los hechos químicos y de los neurofisiológicos, cuando cada gemelo emite la forma de palabras *el agua moja*, esta significa algo diferente en cada caso (en el de la tierra gemela podría significar algo así como *el gagua moja*). De ello se sigue también que el contenido del pensamiento que tienen ambos gemelos cuando piensan *que el agua moja* es diferente. Con ello se concluye que ambos gemelos tienen una diferencia en el contenido intencional de sus pensamientos sin que varíen sus contenidos amplios. Putnam introduce, además, el tema de la superveniencia, afirmando que los pensamientos no supervienen sobre la microestructura neurológica de los individuos, puesto que en el caso descrito acerca de la diferencia intencional entre ambos gemelos, no se sigue una diferencia en la microestructuras de sus cerebros, sino que ambos siguen siendo siempre molecularmente idénticos. En síntesis, Putnam concluye en

¹¹ Fodor, *Psicosemántica*, p. 75.

este escrito, que los significados **no están en la cabeza**, vale decir, considera que lo que determina el significado (o referente) del término agua, depende de la conexión que tenga dicho término con el mundo (significado amplio).

En opinión de Fodor (1987), las intuiciones de Putnam sugieren que las actitudes se individualizan atendiendo a sus propiedades relacionales mostrando además que las actitudes proposicionales (como creer, desear) no supervienen sobre los estados cerebrales. Contrario a esta visión, Fodor considera que la teoría de la superveniencia mente/cerebro (y/o identidad mente/cerebro) es la mejor teoría que permite explicar cómo es posible la causación mental. Para él, la posición de Putnam atenta contra la noción de contenido y suscita problemas para las teorías de la mente basadas en las actitudes proposicionales. Fodor afirma que las actitudes se individualizan no relacionadamente, asumiendo con ello una noción de *contenido restringido*¹² (afirmación, que corresponde en sentido estricto a la doctrina del solipsismo metodológico).

Cabe señalar para precisar este concepto, que cuando se alude a los propósitos de una taxonomía de estados mentales que se supone que es implementado computacionalmente se considera la noción de **contenido estrecho**. En cambio, para los propósitos de una teoría que especifique las condiciones bajo las cuales una representación mental es instanciada, se considera la noción de **contenido amplio**.

Resumiendo podemos concluir, que el *solipsismo metodológico* resulta ser una tesis que se sustenta en la condición de formalidad. Dada esta condición, si los estados mentales requieren ser individualizados por su contenido, entonces deben serlo por su contenido estrecho. En consecuencia, las propiedades semánticas amplias de los estados mentales no desempeñen ningún rol en la especificación de las generalizaciones psicológicas.

3.2. INDIVIDUALISMO METODOLÓGICO

Podemos caracterizar al individualismo en psicología como una visión respecto qué son los estados mentales y cómo han de individuarse, clasificarse o taxonomizarse. Esto significa que el individualismo propone una modalidad de individuación de las clases naturales para los propósitos de la explicación causal en psicología, por lo que resulta ser también un constreñimiento para la explicación psicológica. Fodor (1987) considera que el individualismo es un principio metodológico de carácter general para la ciencia, que se sigue del propio fin científico de la explicación causal, y que establece que todas las taxonomías científicas deben obedecerle. El individualismo metodológico es una doctrina que postula básicamente que los estados psicológicos se individualizan en relación a sus *poderes causales*, considerando que aquello que es causal y explicativamente relevante para la psicología son aquellos procesos y estados mentales que supervienen sobre la estructura del cerebro.

En psicosemántica Fodor argumenta en defensa del individualismo metodológico, llevando a cabo para ello, una refutación de las historias de Putnam (de la *tierra gemela*) y de Burge (en relación al concepto *BRISKET*).¹³ A modo de resumen, nos centraremos

¹² Posteriormente Fodor (1994) modifica esta visión, dejando de lado el contenido estrecho.

sólo en la discusión entre Fodor y Putnam sobre la relevancia del contenido amplio, en la determinación de la individuación de los estados mentales.

En *El argumento modal a favor del contenido estrecho*, Fodor analiza dos argumentos antinómicos para introducir su justificación a favor del contenido estrecho. Los reproduciré para poder continuar con su análisis.

Por un lado existe el argumento A (que es el que Fodor sostiene):

1. Mi gemelo y yo somos duplicados moleculares.
2. Por lo tanto, nuestras conductas (reales y contrafácticas) son idénticas en los aspectos relevantes.
3. Por lo tanto, los poderes causales de nuestros estados mentales son idénticos en los aspectos pertinentes.
4. Por lo tanto, mi gemelo y yo pertenecemos a la misma clase natural a los efectos de la explicación psicológica, y el *individualismo* es verdadero.

Por otro, el argumento B:

1. Mi gemelo y yo somos duplicados moleculares.
2. Sin embargo, nuestras conductas (reales y contrafácticas) son diferentes en los aspectos pertinentes.
3. Por lo tanto, los poderes causales de nuestros estados mentales son diferentes en los aspectos pertinentes.
4. Por lo tanto, mi gemelo y yo pertenecemos a clases naturales diferentes, y el *individualismo* es falso.

Fodor pretende mostrar que el argumento B es falso, puesto que para él las diferencias en el contenido intencional amplio no determinan diferencias en las clases naturales para los efectos de la explicación psicológica. Esto significa que los estados mentales que difieren en las *propiedades intencionales amplias* no difieren en los poderes causales. Lo que el sujeto tiene en su mente sería lo que causaría, para Fodor, su conducta y no aquello a lo que los estados mentales se refieren (contenido amplio). Según él, existir en un mundo de H₂O o de XYZ (siguiendo la historia de Putnam) no afecta la conducta de los individuos a menos que afecte sus estructuras internas. Lo anterior apoyaría la individuación estrecha.

Los argumentos que da Fodor para apoyar sus ideas están centrados en el establecimiento de una condición necesaria que permita distinguir, cuándo una diferencia en las causas es una diferencia en los poderes causales. Fodor llama a ésta *condición C*. Para entender en qué consiste esta condición considérese el siguiente ejemplo: *C1 (causa uno) tiene PC1 (propiedad causa 1), C2 tiene PC2, E1 (efecto 1) tiene PE1, E2 tiene PE2, y la diferencia entre las causas es responsable de la diferencia entre los efectos, en el sentido de que E1 no habría tenido PE1 (en vez de PE2) y que C1 tuvo PC1 (en vez de PC2). Lo que se requiere saber es: ¿cuándo el hecho de que esa diferencia en las causas es responsable de esa diferencia en los efectos, hace PC1 y PC2 poderes*

¹³ Ver Fodor, capítulo segundo de *Psicosemántica y Un argumento modal a favor del contenido estrecho*.

*causales?*¹⁴ La respuesta que da Fodor es la que la diferencia en las causas no debe estar **conceptualmente conectada** con las diferencias de los efectos, para que la diferencia en las causas cuente como tal. Es decir, para que la diferencia entre ser PC1 y ser PC2 sea una diferencia de poderes causales, tiene que darse, al menos, que los efectos de ser PC1 difieran de los efectos de ser PC2. Fodor agrega que se requiere además que esta diferencia entre los efectos esté relacionada de manera **no conceptual** con la diferencia entre las causas. Sin embargo, las diferencias en el contenido amplio no satisfacen *per se* esta condición. Así, pues, la diferencia entre los estados mentales de los gemelos no cuenta como una diferencia en el poder causal en virtud de su responsabilidad por las diferencias intencionales entre las conductas de los gemelos. En consecuencia, Según Fodor lo que está mal en el argumento B es que la inferencia 2' a 3' es errónea.

Sintetizando, vemos que Fodor está en contra del argumento B y del contenido amplio. Su conclusión frente al análisis presentado, consiste en sostener que el *individualismo* es una modalidad apropiada de individuación de los estados mentales (por poderes causales), porque las propiedades de contenido amplio no satisfacen la condición C. Esta afirmación arranca del hecho de no encontrar una condición necesaria que permita distinguir cuándo una diferencia en las causas produce una diferencia en los *poderes causales*. En consecuencia, podemos decir que lo que el individualismo postula es que la diferencia entre los pensamientos que un sujeto tiene sobre *agua* en vez de sobre *gagua* (u otro) no resulta ser un *poder causal* que sea responsable de las propiedades de la conducta de ambos sujetos referente a *agua* o *gagua*, respectivamente.

4. FISCALISMO Y REALISMO INTENCIONAL

Hay dos compromisos fundamentales que asume la teoría de Fodor. Uno corresponde a su adhesión **al fiscalismo**, por una parte, y por otra al **realismo intencional**. Los fiscalistas sostienen que las leyes introducidas en cualquier ciencia especial están determinadas por la existencia de procesos físicos más básicos que implementan tales leyes. Por lo tanto, los procesos mentales suponen la existencia de procesos físicos capaces de llevar a cabo dichos fenómenos. Fodor señala una disposición jerárquica de las ciencias, según la cual existe la física como ciencia básica, pero existen también las ciencias especiales, como por ejemplo la psicología. De este modo es posible formular generalizaciones y leyes psicológicas con independencia del ámbito físico. Fodor sostiene que las leyes introducidas en cualquier ciencia especial están determinadas por la existencia de procesos físicos más básicos que implementan tales leyes (fiscalismo). En consecuencia, los procesos mentales suponen la existencia de procesos físicos capaces de llevar a cabo dichos fenómenos. Fodor postula que las ciencias especiales están en una relación de **supervenencia**¹⁵ sobre dominios más básicos, como por ejemplo con el nivel neurobiológico, que a su vez superviene sobre el nivel físico.

¹⁴ *Ibíd.*, p. 384.

Existirían así, varias ciencias especiales, pero sólo una ciencia básica que sería la física. La psicología intencional equivale entonces a una ciencia especial, estando posibilitada para formular enunciados de leyes. Sus leyes, similares a las de cualquier ciencia especial, contemplan la cláusula *ceteris paribus* (siendo todo lo demás igual) y son autónomas con respecto a las leyes más básicas.

Aceptando las consideraciones metodológicas precedentes, Fodor postula que las propiedades del ámbito de lo mental son parte del orden del mundo: **realismo intencional**. A partir de éste, se afirma que las actitudes proposicionales (estados mentales con un contenido proposicional) cumplen un rol esencial en las explicaciones psicológicas de sentido común y que poseen una existencia real, siendo semánticamente evaluables y causalmente eficaces. Fodor postula además un realismo científico, que sostiene que existen entidades, propiedades y leyes independientemente de nosotros. Para él, una psicología científicamente adecuada debe contener leyes que cuantifican sobre los fenómenos intencionales. En su opinión, existen estados mentales que corresponden a actitudes proposicionales, que poseen propiedades intencionales genuinas que se implican causalmente en la producción de la conducta.

La psicología popular en consecuencia, es una teoría que puede ser evaluada de acuerdo a los patrones que se aplican a cualquier teoría. Posee un vocabulario (términos teóricos) que hace referencia a procesos mentales y supone, que tales estados y procesos son parte del mundo natural. La atribución de estos estados y procesos permite interpretar, explicar y predecir las acciones de los demás, lo que señala claramente que las explicaciones y predicciones, proporcionada por la psicología popular, suponen la existencia de regularidades y leyes que soportan contrafácticos (y que están metafísicamente validadas).

Cabe señalar por último, que una ciencia especial como la psicología presenta problemas frente a los estreñimientos indicados, cuando asume propiedades tales como la intencionalidad. Fodor considera que la **intencionalidad** superviene sobre propiedades fisicalistas, lo que implica postular que hay condiciones fisicalistas bajo las cuales los símbolos mentales son instanciados. Fodor se compromete con la idea de que las generalizaciones psicológicas tienen un carácter eminentemente intencional y, más aún, que las propiedades mentadas por dichas generalizaciones supervienen en y son realizadas por propiedades fisicalistas. Fodor tiene que mostrar entonces que las propiedades intencionales mentadas por los términos teóricos de las leyes intencionales supervienen en última instancia sobre propiedades físicas. Sólo si consigue esto, puede entonces afirmar que las propiedades intencionales son causalmente eficaces.

Lo anterior resulta equivalente a **naturalizar el contenido de las representaciones mentales**. Para Fodor una teoría de la mente satisfactoria debe dar cuenta del carácter naturalista de las propiedades intencionales mentadas en las generalizaciones de la psicología. Por ello, una parte importante de su teoría está destinada a proporcionar al menos condiciones fisicalistas de la **intencionalidad**, estableciendo condiciones suficientes de carácter fisicalistas bajo las cuales los símbolos del *mentalese* podrían ser instanciados.¹⁶

¹⁵ Ver Kim, J. (1984), *Concepts of Supervenience*.

5. NATURALIZACIÓN DE LA INTENCIONALIDAD

Como vimos en la sección precedente, Fodor debe ser capaz de mostrar que las propiedades intencionales son causalmente eficaces en la explicación de la conducta. Como fisicalista, Fodor debe poder dar esta explicación en términos de la ciencia natural, es decir, debe mostrar que las propiedades intencionales (mentadas por los términos teóricos de las leyes intencionales), supervienen en última instancia sobre propiedades físicas. La TRM de Fodor, estaría coja si la **representación mental** fuese algo inexplicable, vale decir, si se la asume como un concepto primitivo sin explicación. Recordemos que anteriormente, según la tesis (2) de la formulación canónica de la TRM, establecimos que *las representaciones mentales son las portadoras primitivas del contenido intencional* (ver sección 2, pág 8). No obstante, aún no hemos dado una explicación acerca de este carácter primitivo de la intencionalidad.

En consecuencia, Fodor debe poder brindar una solución al problema de la intencionalidad (o problema de Brentano), lo que significa responder: ¿cómo se puede explicar en términos de la ciencia natural el hecho de que una representación mental es *sobre algo*? (carácter intencional de la representación mental).

Considerando los planteamientos de Fodor ya expuestos, es necesario tener presente, que el problema *de la intencionalidad (problema de Brentano)* no puede ser resuelto dentro de los márgenes de una psicología computacional, dado que el problema mismo no surge al interior de una concepción computacional de la mente. El problema de la naturalización de la intencionalidad, surge cuando se trata de establecer cómo los constituyentes de representaciones complejas (conceptos o representaciones simples y atómicas adquieren propiedades semánticas). El problema de la naturalización surge en la teoría de Fodor: *Cuando se hace necesario especificar las condiciones bajo las cuales el contenido de un símbolo no-lógico y primitivo de Mentalés (como 'caballo', para usar uno de los ejemplos favoritos de Fodor) es determinado. Tales condiciones deben especificarse en términos no-semánticos y no intencionales*.¹⁷

Si las actitudes proposicionales son relaciones con oraciones de mentalés, Fodor debe en consecuencia, como fisicalista, encontrar un modo de *naturalizar la semántica de mentalés*, vale decir, debe poder ofrecer una teoría naturalista del contenido para mentalés. Esto significa que debe ser, *capaz de decir, en un idioma que no sea ni intencional ni semántico, en qué consiste el que un símbolo de mentalés tenga una cierta interpretación en un cierto contexto*.¹⁸

En *Psicosemántica* (1987) y más recientemente en *A Theory of Content* (1990),

¹⁶ Ver Fodor (1990), *A Theory of Content and other essays*.

¹⁷ Vallejos, *Realismo intencional y naturalización del contenido*, p. 14

¹⁸ García Suárez, *Modos de significar*, p. 337.

como también en *el Olmo y el experto* (1994) Fodor ofrece solución a este problema, presentando una **teoría causal e informacional del contenido**.

5.1. LA TEORÍA CAUSAL INFORMACIONAL DEL CONTENIDO Y EL PROBLEMA DE LA DISYUNCIÓN

De manera simple y resumida podemos decir que la *Teoría Causal Informacional* de Fodor postula, que el contenido de un símbolo porta información acerca de los eventos que han producido sus instanciaciones, existiendo una relación nomológica entre el contenido del símbolo y los contextos causales que produjeron sus instanciaciones. Así, por ejemplo, el símbolo primitivo y no-lógico de Mentalés CABALLO expresa la propiedad *caballo*, dado que existe una relación de covarianza nómica entre las instancias de la propiedad *caballo* y las instanciaciones del símbolo CABALLO. La idea central presente en las teorías causales- informacionales es que CABALLO se refiere a los caballos, cuando efectivamente pensar 'caballo' está correlacionado fiablemente con los *caballos*. De este modo (y tal como Dretske lo expresa)¹⁹, un pensamiento tiene el contenido *eso es un caballo*, sólo si porta información de que efectivamente *eso es un caballo*. Portar información requiere necesariamente de una correlación fiable entre pensamientos de ese tipo y caballos. Por esta razón estas teorías han sido denominadas informacionales.

Fodor postula además de la teoría causal e informacional del contenido una teoría de la dependencia asimétrica para poder solucionar el problema de la disyunción o de la representación fallida. El problema de la disyunción consiste en que no se puede establecer cuándo las muestras de un símbolo (como por ejemplo CABALLO) son causadas efectivamente por un caballo, o en caso contrario, por una hipótesis falsa (como por ejemplo una vaca en una noche oscura). Lo anterior significa que no es posible dar cuenta de la representación fallida: lo que ocurre es que la correlación entre CABALLO y los caballos no es perfecta, dado que alguna persona puede tomar por CABALLO a una vaca en una noche oscura, lo que implica que no se puede determinar con precisión, cuándo CABALLO corresponde correlativamente con su instanciación en el mundo real.

Varios filósofos han intentado resolver este problema (Dretske, 1981, Millikan, 1984, Papineau, 1984) brindando distintas soluciones al respecto. Millikan y Papineau han introducido consideraciones teleológicas en la semántica como intento de hacer frente a este problema. Las *teorías teleológicas de la representación* postulan que las creencias son estados que responden a propósitos biológicos y que en consecuencia su significado debe ser especificado en términos de las condiciones con las que se supone biológicamente covarían. Si lleváramos estas teorías al ejemplo que hemos estado analizando, esto significaría que nuestras creencias sobre 'caballos' están por *caballos* y no por *vacas en una noche oscura* porque su propósito biológico (en términos de selección natural) es tener estas creencias cuando están presentes *caballos* y no *vacas en una noche oscura*. Los defensores de estas teorías generalmente se adhieren a una explicación del propósito biológico en términos de la selección natural, así por ejemplo el

¹⁹ Drestke (1981), *Knowledge and The Flow of Information*.

rasgo *R* tiene el propósito *P*, dado que se da ahora como resultado de una selección pasada a través de un proceso que favoreció a cosas con *P*.

Fodor objeta las teorías teleológicas porque considera que no logran en definitiva resolver *el problema de la disyunción*. El argumento principal que presenta Fodor es que apelar a mecanismos de selección natural no decide entre adscripciones de contenido confiablemente equivalentes y que en consecuencia nos deja frente a una indeterminación intencional. Tomemos como ejemplo a una rana que responde frente a la presencia de una mosca en su campo visual. En opinión de Fodor, el estado de su cerebro es el mismo que en el que está cuando responde frente a una manchita negra. Las teorías teleológicas en cambio dirían que ese estado representa moscas y no manchitas negras porque su propósito biológico es covariar con moscas. Fodor objeta el argumento anterior preguntando ¿por qué no decir que su propósito biológico es covariar con moscas –o- manchitas-negras? El sistema cognitivo de la rana no presenta un mal funcionamiento cuando responde a manchitas negras. En consecuencia, Fodor señala que las teorías teleológicas no logran realmente resolver el problema de la disyunción. Para solucionarlo, Fodor propone su *teoría de la dependencia asimétrica*, que explicaremos a continuación.

5.2. LA TEORÍA DE LA DEPENDENCIA ASIMÉTRICA

Fodor propone *la teoría de la dependencia asimétrica* (TDA, en adelante) para resolver *el problema de la disyunción*, enunciado anteriormente. Fodor considera que para resolver el problema de la disyunción se necesita romper la simetría, vale decir marcar una diferencia entre *caballo / vaca*. Esto se lograría según Fodor, al seleccionar una relación causal semánticamente relevante, y evidenciar que el resto de las relaciones causales heterogéneas son **asimétricamente dependiente** de la primera relación causal seleccionada.

La TDA establece entonces que: CABALLO, significa *caballo* y no *vaca en una noche oscura*, porque el que haya ejemplares de 'caballo' causados por *vaca en una noche oscura* depende de que haya ejemplares de 'caballo' causados por *caballos*, pero no a la inversa. Lo anterior, resulta equivalente a afirmar que *las instancias falsas son metafísicamente dependientes de las verdaderas*, de manera tal, *que las verdades no son ontológicamente dependientes de las falsedades*. De acuerdo a esto, *los mecanismos que rinden falsedades son, de alguna manera, parásitos de los mecanismos que rinden verdades*²⁰ y no al revés. La TDA afirma que las instancias de 'caballo' causadas por vacas son **asimétricamente dependientes** de instancias de 'caballo' causadas por instancias de la propiedad caballo. Esta teoría entiende las muestras en términos de una dependencia empírica entre las relaciones causales, siendo una **teoría naturalista, atomista y fisicalista**.

En su versión de *A Theory of Content*, la teoría de Fodor dice que 'X' significa 'X' si:²¹

²⁰ Fodor, *Psicosemántica*, p. 157

²¹ Extraído del libro de García Suárez, *Modos de significar*, p. 344

(1) 'Los Xes causan "X" es' es una ley.

(2) Algunos 'X'es son efectivamente causados por Xes.

(3) Para todo Y no igual a X, el que los Yes causen efectivamente 'X' es asimétricamente dependiente de que los Xes causen 'X' es.

(4) Hay algunos 'X'es no causados por Xes.

La condición (1) es la condición de información, que nos asegura que efectivamente los 'X' porten información sobre los Xes. Esta condición afirma además, que existe una ley que conecta la propiedad de ser un X con la propiedad de ser un ejemplar de 'X'. La condición (2) es una concesión a las teorías causales "fácticas" y sirve para acomodar los casos de la *tierra gemela*. La condición (3) es la condición de dependencia asimétrica que acabamos de explicar y que permite resolver el problema de la disyunción. La condición (4) permite dar cuenta de la robustez del significado. La robustez es una característica del significado que permite establecer que los ejemplares de 'caballo' puedan ser causados *de cualquier manera*, sin necesidad de que haya caballos en la cercanía. Esto significa que 'caballo' puede emitirse en distintas circunstancias (a la luz del día en presencia de un caballo, en una noche oscura en presencia de un caballo o alegremente en presencia de cualquier otra cosa. En todas estas circunstancias, significa que esto es un caballo.

Como conclusión general de esta sección, podemos agregar la siguiente cita con respecto al problema de la naturalización del contenido:

El problema de la naturalización del contenido se reduce, entonces, al problema de articular una teoría que establezca las condiciones suficientes bajo las cuales un conjunto de propiedades es acerca de otro conjunto de propiedades. Más concretamente, es establecer bajo qué condiciones, por ejemplo, el símbolo 'caballo' exprese la propiedad 'caballo', perteneciente a Mentalés, puede ser acerca de, expresar o significar la propiedad caballo. .²²

El proyecto de naturalización de la intencionalidad consiste en poder explicar cómo es posible que una cosa tenga la propiedad de representar, significar o referirse a otra sin presuponer en esta explicación la noción misma de representación, significado o referencia u otra que las implique. Establecer condiciones suficientes del contenido intencional implica establecer una teoría que no presuponga a su vez nociones semánticas o intencionales y que sea compatible con una ontología materialista inspirada en las ciencias naturales.

6. EVALUACIÓN CRÍTICA DE LA TRMC

Se ha intentado exponer sistemáticamente algunos aspectos centrales de la concepción de la mente de Fodor, con el objeto de clarificar el rol que asigna este filósofo a la representación mental en la ciencia cognitiva. Acorde a nuestro objetivo de la investigación, podemos preguntarnos: *¿cuál es el rol explicativo que desempeña la*

²² Vallejos. *Realismo intencional y naturalización del contenido*, p. 16.

representación mental en este enfoque de la CC? o bien, *¿cuál es la relevancia cognitiva de la representación mental en el paradigma clásico?*

Según lo expuesto, podemos resumir que en el paradigma clásico (computacionalismo) es la sintaxis lo que orienta la causalidad de la conducta intencional y no la semántica. La teoría de Fodor (1994) es formalista, dado que considera que la semántica no forma parte de la investigación de la ciencia cognitiva y que en consecuencia el contenido de los estados mentales no desempeña ningún rol explicativo en la explicación de la cognición y de la conducta humana. De esto se sigue que la TRCM destaque sólo el rol formal de la representación mental, sin lograr establecer con claridad *el rol que le cabe al contenido* (de las representaciones mentales) en la causación y explicación del comportamiento. El hecho de reducir la semántica a la sintaxis hace que un ordenador manipule símbolos sin considerar lo que se representa realmente en esos símbolos (visión solipsista).

La TRCM de Fodor, pretende conciliar dos postulados altamente incompatibles: el postulado computacional con el postulado representacional. El problema radica justamente en pretender conciliar un postulado que enfatiza *las propiedades formales de las representaciones mentales*, con uno que enfatiza *las propiedades semánticas de los símbolos en cuestión*. Para intentar dar solución a este problema, Fodor desarrolla la estrategia del *solipsismo metodológico*, imponiendo una condición de formalidad en el tratamiento de los estados y procesos mentales. Como consecuencia de esto, se establece que para los propósitos de una psicología computacional, sólo se consideran las propiedades sintácticas de los símbolos involucrados y no las propiedades semánticas, vale decir, aquellas que hacen referencia a la conexión de las representaciones con aquello que representan.

Como vimos en la sección 3.1. de este trabajo, Fodor distingue al interior de la noción de contenido, entre la *dimensión amplia* y la *dimensión estrecha*. La dimensión restringida del contenido permite honrar la condición de formalidad impuesta por la práctica de la psicología cognitiva, en tanto deja de lado los rasgos semánticos de las representaciones mentales. De acuerdo a esto, vemos que las generalizaciones de la psicología cognitiva cuantifican sobre estados mentales que son individuados en virtud de la *dimensión restringida* del contenido. No obstante, la dimensión intencional del contenido de las representaciones, se refiere a la *dimensión amplia* del contenido. Lo que decide el carácter intencional de los estados mentales, corresponde al contenido amplio (dimensión semántica de la representación) que es asumida por la filosofía y más específicamente por la metafísica y no por la psicología computacional. La semántica de las representaciones se aborda en el nivel de los fundamentos filosóficos de la ciencia cognitiva, y debe ser posible dar cuenta de ella en términos naturalistas.

En consecuencia, podemos decir que la TRCM es un intento de conciliar el paradigma simbólico con una teoría que destaque *el rol que le cabe al contenido en la causación y explicación del comportamiento*, en su dimensión estrecha. No obstante, vemos que la propuesta fodoriana no logra enfatizar con suficiente plausibilidad el rol en cuestión. El carácter ontológico de su *teoría del contenido amplio* y la asignación predominante en la causación del comportamiento al *contenido restringido*, son factores que atentan contra el objetivo de mostrar claramente *el rol de las estructuras semánticas*

en la producción y explicación del comportamiento.

Por último, como evaluación final, podemos desprender los siguientes aspectos críticos a considerar de la TRCM propuesta por Fodor.

1. Concepción de la psicología como ciencia especial autónoma. Una de las ventajas de asumir la doctrina del individualismo metodológico, postulada por los partidarios de la TRCM, radica en la posibilidad de concebir a la psicología como una ciencia especial autónoma. Fodor se presenta como un defensor de la psicología concebida como una ciencia especial. Su proyecto, consigue fundamentar la psicología, en un sustrato ontológico que es común a todas las ciencias naturales. No obstante, la autonomía de la psicología es relativa, en tanto sus propiedades mentales dependen de otras propiedades de dominios más básicos, de ciencias como: la neurobiología, la química hasta llegar a la física. En consecuencia, Fodor necesita mostrar que las leyes de la psicología pueden ser implementadas por mecanismos regidos por las leyes de una ciencia básica como la física, lo que significa resolver en alguna medida el problema de la *naturalización de la intencionalidad*, explicitado anteriormente.

Sin embargo, aún cuando se desprenden consecuencias positivas para la psicología, podemos decir que la teoría de Fodor deja de lado un aspecto importante en la fundamentación de esta disciplina como ciencia autónoma, porque no consigue en definitiva explicar, de qué forma las propiedades semánticas están imbricadas en la producción de la conducta. La teoría de Fodor, no logra mostrar claramente *cómo el contenido está incorporado a los factores que participan en la producción del comportamiento*. Por esta razón, podemos decir, que aún cuando su teoría del *contenido restringido* cumpla un rol en la explicación causal del comportamiento, resulta ineficiente para efectos de proporcionar una solución al problema de la intencionalidad, vale decir, explicar cómo adquieren su contenido las representaciones mentales.

2. Carácter ontológico de la naturalización de la intencionalidad. Según lo expuesto, podríamos decir que el proyecto de naturalización de la intencionalidad o semántica llevada a cabo por Fodor, radica finalmente en resolver un problema atinente a la ontología de la ciencia en general, como también de la ontología de la psicología en particular (como vimos en el punto anterior). Para lograr naturalizar la intencionalidad, Fodor proporciona condiciones suficientes del contenido intencional compatibles con una ontología materialista inspirada en las ciencias naturales. En consecuencia, la solución de Fodor no consiste en identificar propiedades intencionales con propiedades fisicalistas, sino en mostrar cómo éstas primeras pueden ser implementadas por las segundas. Las propiedades intencionales son propiedades naturales en la medida que pueden ser implementadas por propiedades fisicalistas, lo que implica que dichas propiedades no deben desaparecer del ámbito de una ontología de las ciencias:

Por el contrario, su presencia dentro de dicha ontología es validada si es posible mostrar la posibilidad de su implementación física. El problema de la naturalización del contenido adquiere así un propósito ontológico, de lo que se trata es mostrar que las propiedades intencionales forman parte del ámbito natural y que, por lo tanto, pueden ser objeto de una ciencia natural.²³

²³ *Ibid.*

En consecuencia, la naturalización del contenido de Fodor, sólo muestra coherencia con los compromisos fundamentales asumidos por este filósofo, es decir, con el fisicalismo y el realismo intencional (explicados en la sección 4). La naturalización sólo es coherente con la tesis de la superveniencia, según la cual las propiedades mentadas en las generalizaciones de las ciencias especiales están en una relación de superveniencia sobre dominios de propiedades más básicos (hasta llegar al nivel físico), resolviendo con ello un problema de ontología general de las ciencias.

3. Distanciamiento de las bases biológicas-cerebrales y ambientales de la cognición. Si consideramos otros factores que debieran ser relevantes en la explicación de la cognición humana, vemos que la TRCM deja fuera las bases biológicas-cerebrales y ambientales de la cognición, mostrando con ello un fuerte alejamiento de la realidad biológica. El modelo cognitivista hace una abstracción de las realizaciones biológicas de las estructuras cognitivas, mostrando con ello un profundo *prejuicio anti-biológico*. Por otra parte, en su intento de dar una explicación científica de la conducta inteligente y de la cognición, los cognitivistas introducen prescripciones metodológicas (individualismo metodológico y solipsismo metodológico), que dejan fuera los aspectos sociales y ambientales de la cognición.

Estas bases (biológicas-cerebrales y ambientales) de la cognición son consideradas como centrales en la visión de la teorías alternativas (que expondremos en el capítulo tercero de esta investigación), siendo en definitiva el punto de crítica a este paradigma simbólico de la ciencia cognitiva. Así, por ejemplo en *Estar ahí*, Clark (1991) nos propone que la interacción *cerebro-cuerpo-mundo* es constitutiva de la mente, señalando con ello la indisolubilidad entre la mente y el cuerpo.

4. Limitaciones de la arquitectura de la mente bosquejada por la TRCM. La arquitectura de la mente bosquejada por la TRCM, hace referencia a la arquitectura del computador Von Neumann clásico (IA clásica), la cual presenta importantes limitaciones. La primera de ellas, se debe a que el procesamiento de la información bosquejado por esta arquitectura, no permite dar cuenta de una serie de tareas más complejas de la cognición humana. El procesamiento de la información simbólica de este modelo, se basa en reglas secuenciales que son aplicadas una por vez. Este modo de procesar la información, implica una gran limitación, cuando la tarea requiere gran cantidad de operaciones secuenciales como por ejemplo en el análisis de imágenes y finalmente nos conduce a una especie de *cuello de Botella Von Neumann*.²⁴ Otra limitación importante que presenta el procesamiento simbólico se debe a que éste no logra dar cuenta de la *degradación grácil* (deterioro biológico). Por ello, una pérdida o disfunción de cualquier parte de sus símbolos o reglas del sistema deriva en un grave daño total de éste. Estas limitaciones que presenta el modelo cognitivista, junto con otros aspectos, serán cuestionadas principalmente por los sistemas conexionistas, que presentaremos en el capítulo segundo de esta investigación.

Por las razones presentadas, junto a otras críticas que expondremos en los capítulos

²⁴ Se forma un *cuello de botella Von Neumann* cuando los contenidos de la memoria sólo pueden ser examinados uno a la vez. Los computadores Von Neumann manejan la información unidad por unidad y aunque haya varias unidades trabajando en forma paralela, ellos trabajan en forma secuencial y lineal.

siguientes, la ciencia cognitiva ha continuado su desarrollo presentado un conjunto de teorías variadas, que han intentado suplir las deficiencias del paradigma simbólico expuestas en los puntos anteriores, en especial: el carácter formalista de la teoría de Fodor, el distanciamiento de las bases biológicas –cerebrales y ambientales de la cognición, como también las limitaciones de la arquitectura de la mente del modelo clásico (Von Neumann). En lo que sigue, revisaremos algunas de estas teorías, siguiendo con nuestro objetivo propuesto.

CAPÍTULO SEGUNDO: LA REPRESENTACIÓN EN LOS SISTEMAS CONEXIONISTAS

En la década de los ochenta irrumpen en la CC los modelos conexionistas, cuestionando el punto de vista clásico de la mente, propio del paradigma simbólico. Como vimos en el capítulo anterior, el enfoque clásico de la mente ha sido denominado *enfoque de reglas y representaciones*²⁵, dado que postula la existencia de *representaciones estructuradas sintácticamente* y de *reglas formales precisas*, que refieren a la estructura sintáctica de esas representaciones. Este modelo de la mente estaba inspirado en la arquitectura del computador clásico, modelo Von Neumann (Inteligencia artificial clásica) y postulaba que la cognición es equivalente al modo en que el computador procesa la información (manipulación de símbolos según reglas específicas).

Los modelos conexionistas postulan en cambio, que la base arquitectónica de la cognición humana, como también sus características para el procesamiento de la información, son radicalmente distintas a la del *hardware* del computador. El modelo conexionista está *neuralmente inspirado*, y toma como sus unidades de procesos elementales, a las neuronas (de forma abstracta), considerando que la computación es llevada a cabo por una interacción a través de estas unidades de proceso. Las unidades

²⁵ Tieson, *Una introducción al conexionismo*, p. 360. En Rabossi.

básicas del sistema ya no corresponden a la estructura proposicional (como plantea Fodor), sino a una microunidad (las neuronas). Estos modelos presentan componentes de tipo neural que al conectarse presentan propiedades globales, correspondientes al ejercicio de aptitudes cognitivas; dejando con ello atrás, las descripciones abstractas, propias del paradigma simbólico. En consecuencia, los símbolos (o representaciones estructuradas sintácticamente) no son ítemes importantes a considerar en este enfoque, sino que lo importante corresponde a las conexiones que surgen entre estas numerosas unidades que conforman la red.

El surgimiento de los modelos conexionistas, se debió en gran medida a un intento de respuesta frente a las promesas incumplidas y resultados decepcionantes de la BAIA (Buena y anticuada inteligencia artificial)²⁶. Estos resultados decepcionantes, condujeron a pensar que los cerebros humanos no actúan de la manera convencional que los computadores lo hacen y que la inteligencia humana es en definitiva, diferente a la *inteligencia computacional*.

Las ideas anteriores se apoyaban además, en otras consideraciones establecidas por los partidarios del conexionismo, que marcaban importantes diferencias con el modelo clásico. Una diferencia sustancial entre ambos enfoques, radica en el tiempo que cada uno asigna al procesamiento de la información. El conexionismo señala que el tiempo de procesamiento neural es más lento que la velocidad de procesamiento de los computadores. Las neuronas operan a una escala de tiempo de milisegundos (o incluso centésimas de segundo), mientras que la velocidad de procesamiento de los computadores se mide en nanosegundos. En consecuencia, si medimos el tiempo de procesamiento en milisegundos, obtenemos que las neuronas son *'diez elevado a seis'* veces más lentas que los computadores. Además, las neuronas procesan de modo paralelo / distribuido y no de forma lineal / secuencial, como postula la arquitectura clásica.

1. COMPONENTES BÁSICOS DE LOS MODELOS CONEXIONISTA SEGÚN RUMELHART

Para explicar los componentes generales que presentan los sistemas conexionistas, consideraremos la aproximación realizada por Rumelhart (1989) en *The Architecture of Mind: A Connectionist Approach*. En este texto, Rumelhart señala que la metáfora del computador (modelo Von Newman) ha sido una herramienta adecuada para explicar los procesos de la mente, porque nos ha permitido **implementar** la simulación de estos procesos, vale decir, de la cognición mental. Sin embargo, en su opinión, esta metáfora no resulta adecuada para realizar un **modelamiento** de la mente humana, puesto que no permite graficar su arquitectura fundamental. Un modelamiento adecuado de la mente, estaría representado por la arquitectura del cerebro (modelo neural), mediante la cual, se muestra que la información es procesada de una manera muy distinta a como lo hacen

²⁶ John Haugelang denominó así: BAIA a la inteligencia artificial clásica.

los computadores. Su propuesta implica entonces, reemplazar la *metáfora del computador* por la *metáfora del cerebro*, lo que señala una crítica a la arquitectura clásica computacional de la mente.

En consecuencia podríamos decir, que el conexionismo surge como una reacción negativa frente a la falta de plausibilidad biológica del cognitivismo, dado que este último no resulta ser un buen modelo del cerebro. Por el contrario el conexionismo, pretende seguir el modelo del cerebro, aun cuando en definitiva podamos decir que se encuentra muy lejos de alcanzar una real plausibilidad biológica. Rumelhart, señala los siguientes componentes básicos de los modelos conexionista:

1. Un conjunto de unidades de procesamiento. Cualquier sistema conexionista parte con un conjunto de unidades que interactúan entre si de forma parecida a como lo hacen las neuronas, esto es, enviando señales de excitación e inhibición. No existe por consiguiente, un procesador central que realice las tareas de control como en la arquitectura clásica, sino sólo simples unidades que mandan señales de excitación e inhibición, realizando diversos cómputos en forma *paralela* (muchas unidades llevan a cabo sus computaciones al mismo tiempo). Existen tres tipos de unidades dentro del sistema conexionista: *unidades de entrada* (aquellas que reciben información de fuentes externas a la propia red), *unidades de salidas* (aquellas que ofrecen las señales o información al exterior de la red; dan la respuesta del sistema) y *unidades escondidas* (aquellas que no tienen una relación directa ni con la información de entrada ni con la de salida; no son *visibles* al ambiente exterior de la red).

2. Un estado de activación. Al igual que las neuronas, cada una de estas unidades posee un *estado de activación* previo, es decir, un valor que indica cuando está en estado de reposo.

3. Una función de salida (output) para cada una de las unidades. Las unidades interactúan transmitiendo información a sus vecinos. La fuerza de esta interacción depende del estado de activación.

4. Un patrón de conectividad entre las unidades. Las unidades están conectadas unas a otras y el patrón es equivalente a la suma de los pesos de los *inputs*.

5. Una regla de propagación, para propagar la activación a través de la red de conexiones.

6. Una regla de activación , que calcule nuevos niveles de activación para cada unidad a partir de las entradas que recibe de otras unidades y de su estado previo de activación.

7. Una regla de aprendizaje. Es posible modificar los patrones de conexión entre las neuronas como resultado de la experiencia. Esta modificación puede ser de tres tipos; desarrollando nuevas conexiones, por pérdida de conexiones existentes o bien por modificación de la fuerza de las conexiones que ya existen. Estas estructuras pueden aprender por sí solas, gracias a la regla de *retro-propagación*, propia de un sistema conexionista. Estas reglas de aprendizaje le permiten a la red, modificar los pesos de las conexiones para acomodar de modo cada vez más perfecto la información de entrada a la salida requerida.

8. Un ambiente dentro del que debe operar el sistema. Es esencial en el desarrollo de este modelo tener una representación del *ambiente* en el que el modelo existe. De esta manera, las unidades de entrada y salida serán diferentes si se quieren simular diferentes tareas.

Resumiendo, podemos decir, que el comportamiento dinámico de una red conexionista está determinado básicamente por los estados iniciales de activación de sus unidades y por los pesos de sus conexiones. En un sistema conexionista, no hay un ejecutor central, ni tampoco un programa que determine lo que ocurre en el sistema, lo que marca una gran diferencia con la arquitectura clásica, en la cual todo el sistema es controlado por el programa que existe en la unidad de procesamiento central. En los sistemas conexionistas las conexiones son locales y cada nodo sólo codifica la información que obtiene de los otros nodos con los cuales está conectado. De este modo, *nada* en el sistema controla lo que el sistema como totalidad hace. Además, existe cierta independencia entre lo que ocurre en las distintas partes del sistema, de manera tal que cada nodo está determinado sólo por su actual estado y por su *input*.

2. VENTAJAS DEL CONEXIONISMO FRENTE AL MODELO CLÁSICO

Podemos señalar que el conexionismo tiene algunas ventajas, frente al modelo computacional. Una primera ventaja, es que logra modelar la mente de modo neurobiológicamente más real, permitiendo con ello, explicar de mejor manera el tiempo real de procesamiento de la información en la cognición humana. Como ya hemos señalado, las unidades básicas presentes en la arquitectura clásica y en la conexionista, se caracterizan por operar a escalas de tiempo totalmente distintas. Si bien es cierto, que las neuronas son más lentas, el hecho de que ellas *computen* de modo paralelo, permite mostrar cómo unidades tan lentas, pueden llevar a cabo tareas tan complejas, en un corto tiempo. Este aspecto, no queda explicado por el procesamiento secuencial y serial, de la arquitectura clásica.

Por otra parte, los sistemas conexionistas permiten dar cuenta de mejor forma de la degradación grácil (deterioro biológico del sistema). La *degradación grácil*, tiene que ver con la variedad de formas en que puede deteriorarse de manera gradual el funcionamiento de los sistemas naturales. Esto significa, que una lesión o enfermedad cerebral puede afectar sólo el funcionamiento del sistema, pero no causar la pérdida total de recuerdos o funciones específicas. En el caso del conexionismo, vemos que este sistema permite generar soluciones o *outputs* similares aún cuando las condiciones iniciales no sean exactamente iguales, por ejemplo pueden faltar ciertos *inputs* o bien ciertas memorias dentro del sistema y éste sigue funcionando igual de acuerdo a patrones ya obtenidos. Por esta razón se dice que los *outputs* de este sistema son flexibles y no quebradizos, permitiendo dar cuenta, de alguna manera de la degradación grácil. Las señales recibidas son procesadas simultáneamente, de modo paralelo, por

millones de conexiones diferentes y además de modo distribuido, lo que implica que la responsabilidad del almacenamiento de la información no recae sobre una sola unidad, vale decir, sobre una neurona, sino sobre amplios conjuntos de unidades. Además el conocimiento está implícito en el patrón global de la conexión, más que explícitos en las unidades mismas, por lo que si hay pérdida de información en una unidad esto no afecta al conocimiento total del sistema. En el modelo clásico de *reglas y representaciones*, el procesamiento es naturalmente *frágil*, dado que cualquier lesión causa la pérdida de funciones particulares y de todo lo que dependa de ellas.

Otra ventaja del conexionismo, frente a la arquitectura clásica, consiste en que este sistema permite cierto nivel de aprendizaje, gracias a la interacción que realiza con el medio ambiente. Esto significa que es posible enseñar a las redes conexionistas ya sea cambiando el valor de umbral de los nodos ocultos o de salida, o bien cambiando el peso de conexión. El procedimiento anterior se hace a través de algoritmos, básicamente de la siguiente manera: por *algoritmo de patrones* (con los mismos valores de entrada se trata de producir distintos valores de salida) o por *algoritmo de retropropagación* (se producen valores de salida idénticos a partir de valores de entrada diferentes). El aprendizaje se debe al éxito de la *retropropagación* (técnica de entrenamiento), que muestra claramente cómo los sistemas conexionistas son capaces de un autoaprendizaje. La técnica o procedimiento de *retropropagación* (back propagation) consiste en una red multicapa en la que, como su nombre indica, una vez obtenida una salida en la propagación hacia adelante, se comprueba el error cometido, es decir, la diferencia entre la salida real y la deseada. Este error se propaga hacia atrás por toda la red distribuyéndolo entre todos los nodos de cada capa. La distribución consiste en variar los pesos de cada conexión sináptica hasta conseguir después de muchas interacciones, una cuota de error inferior a la estipulada previamente. Así, después de pasar de forma repetida por este procedimiento de *retropropagación*, finalmente el sistema produce el *output* correcto para todo *input*. En consecuencia, el sistema habrá aprendido las respuestas, siendo capaz además de *generalizar* a otros casos que no ha visto.

En comparación con el modelo clásico, el procedimiento de *retropropagación* de los sistemas conexionista constituye una ventaja dado que muestra que los sistemas conexionistas son sistemas altamente adaptativos y flexibles, dado que son capaces de aprender por sí mismo mediante la conexión de los pesos de conexión (siempre y cuando la estructura de la red sea adecuada y se aplique una regla correcta de aprendizaje). En los modelos clásicos, dado su carácter formal, no resulta posible el aprendizaje. Los cambios de estado son inferencias lógicas desde una representación simbólica a otra y son especificados por reglas formales (reglas de transición, programas y algoritmos). Por esto, los sistemas clásicos son caracterizados como sistemas automáticos formales, dado su rigidez.

3. LA REPRESENTACIÓN EN LOS SISTEMAS CONEXIONISTAS

El conexionismo y el computacionalismo clásico consideran que las representaciones mentales son entidades psicológicamente reales en los sujetos y que son objetos que poseen una estructura característica. En este sentido, puede afirmarse, que ambos enfoques son **representacionistas**, dado que postulan estados representacionales *intencionales* o *semánticos* que cuantifican sobre estados del mundo. La diferencia que existe entre ellos, radica únicamente en la naturaleza precisa del sistema de representaciones internas que cada uno postula. El conexionismo cree en un tipo de representación interna que difiere del modelo clásico. Los conexionistas sustituyen las cadenas de símbolos sólidos y manipulables del sistema clásico por vectores numéricos complejos y operaciones básicas de reconocimiento y transformación de patrones, que pueden entenderse como **representaciones**. Nos encontramos en los sistemas conexionistas, con representaciones *esparcidas* que corresponden a la actividad simultánea de un cierto número de unidades. A las representaciones que involucran muchas unidades se las llama representaciones distribuidas. Esto significa que en los sistemas conexionistas, las representaciones, no tienen, típicamente, ninguna estructura sintáctica, como en los sistemas clásicos. El conexionismo plantea una especie de asociacionismo (las redes neuronales no son otra cosa que asociaciones entre unidades), las redes comienzan su aprendizaje dependiendo de la experiencia y del ambiente, siendo éste el que provoca la aparición de los pesos adecuados para el procesamiento y almacenamiento de la información. La **representación** en los sistemas conexionistas, es equivalente al patrón de activación de la red, vale decir, a la actividad simultánea de un cierto número de unidades, que pueden estar excitadas o inhibidas.

Podemos decir en consecuencia, que en ambos enfoques de la CC, se recurre a la **representación mental** para explicar la cognición y que, lo que varía en ellos, corresponde a la presencia o ausencia de una estructura constituyente asignada a ésta. En el conexionismo, nos encontramos con una representación distribuida, que no tiene típicamente ninguna estructura sintáctica (diferencia central con el cognitivismo). Sólo cuando el sistema necesita una estructura, ésta se construye (dentro de éste) con diferentes grupos de nodos que están dedicados a partes diferentes de la estructura que se necesita. Así por ejemplo, si pensamos en un sistema de reconocimiento de palabras, existirá un grupo de nodos para la primera letra, otro para la segunda, etcétera. En los sistemas conexionista sólo se logra una estructura de representación atómica y no de representaciones estructuradas como en el cognitivismo. Estas representaciones pueden ser de codificación local (nodos aislados) o de codificación gruesa (conjunto de nodos).

Cabe señalar además, que en un sistema conexionista, la información se representa como un patrón de activación al estar en uso, y que cuando no se utiliza esta información, dicho patrón no está presente en ningún lugar del sistema. Esto significa que la información no está almacenada como estructura de datos y que, por lo tanto, no existen **símbolos** en un sistema conexionista. Sólo existen representaciones activas que pueden entenderse como símbolos. De este modo, sólo puede decirse que la información está en el sistema, almacenada en los pesos y en las fuerzas de las conexiones entre los nodos, que crean representaciones activas apropiadas cuando se las necesita, en respuesta a los estímulos internos o externos.

4. EL DEBATE EN TORNO A LA ESTRUCTURA CONSTITUTIVA DE LAS REPRESENTACIONES MENTALES

Como ya señalamos anteriormente, una importante diferencia entre conexionismo y cognitivismo, radica en la estructura constitutiva que cada uno asigna a la **representación mental**, pero no en objetar su existencia misma. Una ardua polémica al respecto, fue protagonizada por Fodor y Pylyshyn (1988), quienes criticaron duramente al conexionismo, aduciendo que estos sistemas, no poseen una estructura de constituyentes, siendo incapaces de explicar propiedades significativas de la actividad mental. En opinión de estos autores, los procesos cognitivos dependen de **representaciones estructuradas sintácticamente**, propiedad que no tienen las representaciones del conexionismo.

4.1. LA CRÍTICA DE FODOR Y PYLYSHYN AL CONEXIONISMO

Fodor y Pylyshyn (1988), cuestionaron el enfoque del conexionismo sobre la estructura de la representación mental, considerando dos afirmaciones centrales que definen la perspectiva clásica de la cognición. Estas son:

(1) Los pensamientos tienen una estructura compuesta. Esto significa que el pensamiento *Juan ama a María* no es atómico, sino que corresponde a un estado mental equivalente a una instanciación de una representación mental compleja y estructurada. Las teorías simbólicas postulan un *lenguaje del pensamiento* (LDP), asumiendo con ello representaciones con sintaxis y semántica combinatoria. Además sostienen que es posible distinguir entre representaciones estructuralmente atómicas y estructuralmente moleculares. Estas últimas tienen constituyentes sintácticos que pueden ser ellos mismos estructuralmente moleculares o estructuralmente atómicos. El contenido semántico de una representación (molecular) es una función de los contenidos semánticos de sus partes sintácticas, junto con su estructura constitutiva.

(2) Los procesos mentales son sensibles a esa estructura compuesta. Esto significa que los principios por medio de los cuales los estados mentales se transforman, se definen sobre las propiedades estructuradas (formales) de las representaciones mentales. Así por ejemplo, a partir de cualquier pensamiento de la forma p y q (sin importar lo que p y q sean), es posible deducir p .

Una vez establecidas las afirmaciones (1) y (2) como definitorias de la perspectiva clásica, Fodor y Pylyshyn sostienen que hay argumentos decisivos a su favor. A partir de la idea de que los límites del *lenguaje del pensamiento* son los límites del sistema cognitivo y que en consecuencia la mente real está sometida a las mismas restricciones que tiene este lenguaje; estos autores consideran que es posible emprender una crítica contra la arquitectura conexionista. Considerando esta relación lenguaje- mente, estos

autores afirman que los estados mentales tienen las propiedades de productividad, sistematicidad, composicionalidad y coherencia inferencial. Dentro de estas propiedades, cobra mayor relevancia *la sistematicidad*, dado que mediante ella los computacionalistas clásicos, logran explicar la estructura constituyente de los estados representacionales. El argumento de la sistematicidad, podemos resumirlo brevemente mediante el siguiente ejemplo: Si consideramos los pensamientos *Juan ama a María* y *María ama a Juan*, podemos decir que existe una relación intrínseca entre ambos, dado que cualquier mente capaz de pensar el primero es capaz de pensar el segundo. Esta propiedad puede ser explicada si consideramos que el medio en el que se desarrollan los pensamientos posee una estructura de constituyentes ('Juan', 'ama' y 'María') que corresponden a cada uno de los elementos identificables en los pensamientos anteriores. Estos constituyentes se combinan entre sí, de forma tal que el significado del compuesto es determinado por el significado de los componentes más simples que están presentes en el pensamiento. Lo anterior significa suponer el carácter *componencial* de las representaciones y su composicionalidad en el nivel semántico.

Fodor y Pylyshyn argumentan que el conexionismo no puede dar cuenta de la sistematicidad, dado que las representaciones mentales en los sistemas conexionistas **carecen de estructura sintáctica y de semántica combinatoria**. El hecho de que los sistemas conexionistas vean las representaciones como neurológicamente distribuidas hace pensar que estos puedan tener partes. Sin embargo, no se sigue que lo distribuido deba tener constituyentes, puesto que en opinión de Fodor y Pylyshyn, sólo tenemos estructura constitutiva cuando las partes de las entidades semánticamente evaluables son ellas mismas posibles de evaluar semánticamente. Aclaran en consecuencia, que las relaciones constitutivas existentes en los sistemas clásicos, rigen entre objetos que existen en el nivel representacional y que existen **dentro** de las relaciones de ese nivel. Lo que ocurre en el conexionismo en cambio, en la distributividad neuronal es una relación **entre** los "nodos" y las "unidades" (relación entre niveles), en la cual sólo los nodos cuentan como **representaciones**. Por más compleja que sea una realización espacialmente distribuida de los sistemas conexionistas, ésta no implica necesariamente una estructura constitutiva, según estos autores.

En síntesis, Fodor y Pylyshyn aseveran que en el conexionismo (estándar): (a) Todas las representaciones son atómicas (puesto que los estados mentales no tienen estructura compuesta) y (b) El procesamiento conexionista es asociativo (siendo sensible sólo a la estadística y no a la estructura). De este modo vemos que (a) viola la afirmación 1 (los pensamientos tienen una estructura compuesta) y que (b) viola la afirmación 2 (los procesos mentales son sensibles a esa estructura compuesta), propia de los sistemas clásicos.

Estos autores aseveran además, que el conexionismo estándar es fatalmente defectuoso y propugnan en consecuencia que los conexionistas deberían perseguir un conexionismo no estándar. Esto significa que el conexionismo, en opinión de Fodor y Pylyshyn, debería admitir los principios (1) y (2) aceptando con ello la perspectiva clásica. Según estos autores, el conexionismo debería ser usado como implementación de la arquitectura clásica, dado que no proporciona una nueva arquitectura cognitiva, sino sólo una implementación para la vieja arquitectura cognitiva clásica. El conexionismo no sería

entonces un nuevo paradigma, sino más bien una nueva información sobre *la ciencia de la implementación*.

Según el computacionalismo clásico, (véanse Fodor, 1976, Newell, 1980 y Pylyshyn, 1980 y 1984), **las representaciones mentales** deben tener una estructura simbólica. Esto significa afirmar que las representaciones mentales deben ser símbolos (más o menos complejos) regulados por una sintaxis combinatoria y que deben estar dotadas de una semántica composicional.

4.2. LA RESPUESTA DE PAUL SMOLENSKY A FODOR Y PYLYSHYN

En su artículo *La estructura constitutiva de los estados mentales conexionistas: una respuesta a Fodor y Pylyshyn (1987)*,²⁷ Smolensky intenta responder al punto central de la crítica realizada por estos autores al conexionismo. En su opinión, Fodor y Pylyshyn están equivocados en su afirmación de que los estados mentales conexionistas **carecen de una estructura constitutiva**. Argumenta que el origen del error de esta crítica radica en no apreciar la significación de las **representaciones distribuidas**, propias de los sistemas conexionista.

Para justificar su perspectiva, Smolensky presenta de modo resumido un contexto meta-teórico, que le permite bosquejar los distintos enfoques que se han adoptado sobre la cognición. Smolensky se refiere en este artículo a la **paradoja de la cognición**, como un problema que surge frente a la disyuntiva entre adoptar el formalismo simbólico o el formalismo conexionista. Esta *paradoja* nos muestra que: (A) Por un lado, podemos decir que la cognición es dura (hard) y que puede ser caracterizada por las reglas de la lógica, como por las reglas del lenguaje. Y (B) Que por otro, podemos decir que la cognición es blanda (soft), lo que significa que frente a la realización de las reglas lógicas, los sistemas formales automáticos (programas de IA) producen sistemas que no tienen un comportamiento suficientemente fluido, ni fuerte para constituir una inteligencia verdadera.

Por esta razón, en opinión de Smolensky, al tratar de caracterizar las leyes de la cognición seremos empujados en dos direcciones diferentes: cuando enfocamos la competencia de alto nivel, somos empujados hacia representaciones y procesos simbólicos estructurados; mientras que, cuando enfocamos el detalle complejo y variante del comportamiento inteligente real, somos empujados hacia descripciones numéricas estadísticas. Smolensky bosqueja cinco posturas diferentes que se han adoptado frente al dilema de esta paradoja:

1) La primera postura consiste en negar lo blando, estableciendo que la esencia de la inteligencia es la lógica (Chomsky, 1965).

2) La segunda postura niega en cambio lo duro, considerando que la esencia de la inteligencia real consiste en no seguir reglas (Dreyfus, 1986).

3) La tercera postura rechaza las dos posiciones anteriores y considera que el cerebro contiene una parte dura y otra blanda (cerebro dividido). Según ésta, el cerebro

²⁷ Smolensky. Artículo en RABOSI (compilador). *Filosofía de la mente y Ciencia cognitiva*.

derecho corresponde al conexionista y el izquierdo al modelo Von Neuman.

4) El cuarto enfoque corresponde al *difuso* (Gupta, Ragade y Yanger, 1979), que básicamente toma una máquina dura y la reviste con blandura. En esta respuesta a la paradoja, la blandura es definida como grados de dureza. Esto significa que, para este punto de vista la máquina cognitiva es en el fondo una máquina dura (puesto que trabaja en base a reglas), pero que a su vez, es una máquina compleja que parece blanda cuando uno la observa desde un nivel superior. En definitiva en este enfoque, *la blandura emerge de la dureza*.

5) El quinto enfoque corresponde al sustentado por Smolensky (1988), llamado *enfoque subsimbólico*²⁸, que en su opinión constituye la base del tratamiento adecuado del conexionismo. Según este enfoque el sistema cognitivo es fundamentalmente una máquina blanda, que es tan compleja que a veces parece dura, cuando es vista desde niveles superiores. En definitiva, es el nivel inferior el blando y el nivel superior el duro, por tanto, *la dureza emerge de la blandura*²⁹, según esta perspectiva.

Smolensky es enfático en señalar que la decisión de adoptar un formalismo computacional o uno conexionista está enraizada en una de estas posturas sobre la *paradoja*, vale decir, sobre la problemática en asumir el *lado blando* o *duro* de la cognición, como prioritario. Smolensky tiene absoluta claridad de las dificultades que conlleva asumir su *enfoque subsimbólico*, vale decir, **adoptar el lado blando** de la cognición como sustrato fundamental sobre el cual construir el *duro*. Citando a Kirsh (1987), señala los costos que tendría asumir un *enfoque blando* de la cognición. La dificultad radica justamente, en tener que construir a partir del *material blando* los símbolos y la computación simbólica, vale decir los *aspectos duros* de la cognición. En este sentido está muy de acuerdo con la crítica de Fodor y Pylyshyn sobre este punto, cuando ellos dicen que:

El precio que uno tiene que pagar por ser conexionista es el fracaso de dar cuenta de ciertas regularidades del lado duro, regularidades que el formalismo simbólico le da a uno prácticamente gratis.³⁰

Sin embargo, no está de acuerdo con la crítica realizada por estos autores al conexionismo, cuando afirman que éste nunca podrá llegar a proveer las capacidades de la computación simbólica. En su opinión, vale definitivamente la pena, el esfuerzo de seguir insistiendo en este enfoque, puesto que gracias a él, sería posible finalmente dar una explicación sobre cómo el *cerebro* construye la computación simbólica. O bien, aún cuando no se logre esto, sería posible al menos explicar cómo a partir de algo que se asemeja remotamente al cerebro, puede ser hecha la computación. Smolensky señala además, que el *lado duro de la cognición*, ha tenido prioridad por varias décadas, con resultados decepcionantes. Por tanto, sería tiempo de dar al lado blando unas pocas décadas, aunque sea para producir resultados decepcionantes por sus propios medios.

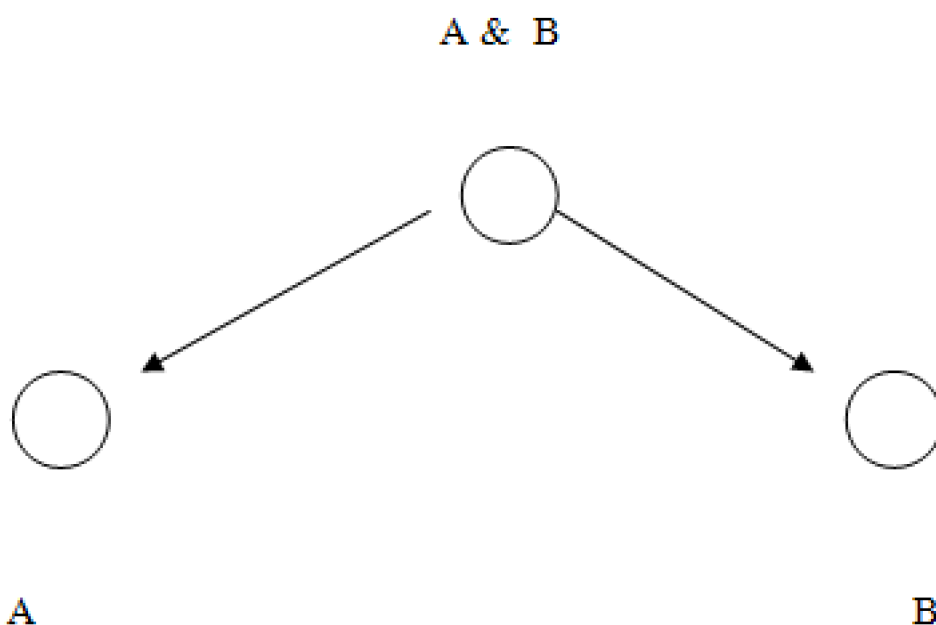
²⁸ Postura que será explicitada con más detalle en la sección siguiente.

²⁹ Smolensky en Rabossi, p. 385

³⁰ *Ibíd.*, p. 186.

Otro aspecto importante a considerar, si el enfoque subsimbólico triunfa, sería el hecho de tener una solución unificada de la paradoja, puesto que no tendríamos posibilidad de obtener una negación de uno de los dos aspectos de la cognición (como en los enfoques 1 y 2), como tampoco una división profunda del cerebro (enfoque 3).

El marco-teórico expuesto, nos permite continuar con la argumentación realizada por Smolensky contra la crítica de Fodor y Pylyshyn. El punto de partida de su defensa está centrado en la siguiente pregunta, *¿Qué es lo que hace que Fodor y Pylyshyn digan que las representaciones conexionistas son atómica?* En opinión de Smolensky ellos se basan en el siguiente esquema (tomado de Ballard y Hayes, 1984).



Esquema 1: La red de Fodor y Pylyshyn (Esta figura sólo encarna una representación hiper-local)

Se supone que en el esquema anterior, la red ilustra el enfoque conexionista estándar de la inferencia 'A y B' a 'A' y a 'B'. En opinión de Smolensky, ello implica un grave error, dado que este tipo de representaciones conexionistas ultra-locales (en las que las proposiciones enteras están representadas por nodos individuales), está lejos de ser un esquema típico de los modelos conexionistas. Por esta razón, este tipo de representaciones no puede ser tomada como la forma definitoria del enfoque conexionista. El contra-argumento central, presentado por Smolensky contra estos autores, sostiene que cualquier crítica que se presente al enfoque conexionista, debe considerar las consecuencias de usar *representaciones distribuidas*. Es importante señalar, que los sistemas conexionistas utilizan tanto *representaciones locales* como *representaciones distribuidas*. La diferencia fundamental entre ambas representaciones radica en que en la *representación local* cada unidad corresponde a una única representación, mientras que en la *representación distribuida*, cada representación está

constituida por la acción de un elevado número de elementos simples. Las representaciones distribuidas tienen importantes propiedades que no poseen las representaciones localizadas y que son propias de los sistemas conexionistas. Por ejemplo, estas representaciones poseen la propiedad de autogeneralización, esto quiere decir que cuando existen representaciones con contenidos semejantes, éstas pueden dar lugar automáticamente a patrones semejantes de activación. Además ellas tienen también la propiedad de *poder completar un patrón*, puesto que la red puede rellenar activaciones no existentes en algunas de las conexiones de los nodos. Finalmente, la red conexionista es capaz de *tolerar pequeños errores*, permitiendo con ello que la actuación del sistema no sea totalmente perfecta. Esto último marca una diferencia importante con los modelos computacionales simbólicos.

Fodor argumenta sin embargo, que la cuestión sobre si la representación es local o distribuida es algo irrelevante, dado que atañe a una cuestión de *nivel entre* los nodos y las unidades, que no implica necesariamente una estructura constitutiva. En cambio, la composicionalidad de los estados mentales es una cuestión de *nivel dentro de*. (Ver sección anterior, pág 31). El argumento de Fodor y Pylyshyn supone que debido a esta diferencia, una cuestión de *nivel entre*, no puede influenciar a otra de *nivel dentro*. Para Smolesky el argumento anterior es una falacia, dado que se asume que en las representaciones distribuidas la relación de *nivel entre* no puede tener ninguna consecuencia en la estructura de *nivel dentro de*, de las relaciones entre las representaciones de A y B y la representación de A. Smolensky considera que esto último es falso, y que por el contrario hay implicaciones de las representaciones distribuidas para la composicionalidad. En consecuencia, el objetivo fundamental que persigue la respuesta de Smolensky a Fodor y Pylyshyn consiste en mostrar que los modelos conexionistas que usan representaciones distribuidas adscriben a los estados mentales la clase de estructura composicional, que la afirmación (1) demanda.³¹

Su argumentación, gira en torno a un análisis realizado por el mismo Zenon Pylyshyn en el *Encuentro de Ciencia cognitiva* (Boulder, 1984) sobre la naturaleza de las representaciones mentales conexionistas. En este encuentro, se produjo un interesante debate, a partir de la siguiente pregunta enunciada por Pylyshyn a Rumelhart: *¿Se puede representar una taza de café en las redes conexionistas?* Frente a esta pregunta, Rumelhart contestó de forma positiva y Pylyshyn entonces agregó esta otra pregunta: *¿se puede representar una taza sin café, en las redes conexionista?* Rumelhart nuevamente contestó de forma positiva. Pylyshyn sin embargo, argumenta que la diferencia entre estas dos representaciones (de taza con café y taza sin café), radica justamente **en la representación de café**. Para justificar su respuesta, propone el siguiente ejemplo hipotético:

1) Si tomamos una representación distribuida de **taza con café** y consideramos que esta tiene las siguientes Unidades Microcaracterísticas más un patrón de actividad que señala las unidades que están activas (viñetas en negro).

³¹ Esto es, que los pensamientos tienen una estructura compuesta.

- recipiente vertical
- líquido caliente
- vaso contactando madera
- superficie curva de porcelana
- olor a quemado
- líquido marrón contactando porcelana
- superficie curva porcelana
- objeto oblongo de plata
- asa del tamaño del dedo
- líquido marrón con lados curvos y fondo

2) Si luego abstraemos de la representación anterior, la representación **de taza sin café**, que contiene las siguientes Unidades Microcaracterísticas, más un patrón de actividad con sus unidades activas (viñetas en negro).

- recipiente vertical
- líquido caliente
- vaso contactando madera
- superficie curva de porcelana
- olor a quemado
- líquido marrón contactando porcelana
- superficie curva porcelana
- objeto oblongo de plata
- asa del tamaño del dedo
- líquido marrón con lados curvos y fondo

3) Obtenemos como resultado (según Pylyshyn) sólo la representación conexionista de **café**, representada por las siguientes Unidades Microcaracterísticas (más su patrón de actividad).

- recipiente vertical
- líquido caliente
- vaso contactando madera
- superficie curva de porcelana
- olor a quemado
- líquido marrón contactando porcelana
- superficie curva porcelana
- objeto oblongo de plata
- asa del tamaño del dedo
- líquido marrón con lados curvos y fondo

La conclusión de Pylyshyn frente a este ejemplo, es que la representación conexionista de **taza con café** es sólo la representación de **taza sin café** combinada con la representación de **café**. Smolensky objeta esta argumentación, considerando que el ensamblaje que propone Pylyshyn es una combinación bastante extraña, que resulta de una situación anterior (puesto que la representación compuesta está construida a partir de *la representación de café* extraída de la situación *de taza con café*, junto con *taza sin café* extraída de la situación *de taza con café*, junto con su interacción). Por esta razón, postula que la estructura composicional sintáctica propuesta por Pylyshyn, está allí sólo en un sentido aproximativo (y en un contexto particular de estar dentro *de una taza de café*) y que no es equivalente a tomar una *representación de café* independiente del contexto y una *representación de taza* independiente del contexto. En opinión de Smolensky, si realmente tomáramos la fórmula propuesta por Pylyshyn para obtener la representación conexionista de *café*, deberíamos en consecuencia de igual manera, poder obtener ésta, al abstraer de la representación conexionista de *taza con café*, la representación conexionista de *taza sin café*; como también, si tomáramos la representación de *árbol con café* y abstrajéramos de ella, la representación de árbol sin café. Sin embargo, si realmente hiciéramos lo anterior, obtendríamos en el primer caso (taza de café) la representación conexionista de café, como *una representación de gránulos marrones semejantes a tierra con aroma a quemado, apilados en una forma cilíndrica, junto con gránulos contactando hojalata*. Y en el segundo caso (árbol de café) como una representación de *granos marrones suspendidas en el aire de forma graciosa*. Por esta razón, Smolensky concluye frente a esta polémica discusión, que la representación conexionista de *taza de café* en este esquema distribuido, corresponde a una *familia de patrones de activación de actividad distribuida*. Esto significa, que lo que une a todas esas representaciones particulares de *café* es un *parecido de familia*. La

representación conexionista de *café* es en consecuencia, una colección de vectores enlazados por una familia de parecidos y por tanto la relación entre la constitutividad conexionista y la constitutividad sintáctica no es de implementación literal.

Los argumentos anteriormente presentados por Smolensky, están encaminados a mostrar que efectivamente las representaciones distribuidas del conexionismo, **poseen relaciones de constitutividad**. Estas relaciones, sin embargo, tienen ciertas características que debemos señalar y que el enfoque simbólico de Fodor y Pylyshyn no consideran. Por ello, frente a la conclusión anterior, Smolensky agrega algunas moralejas que valen la pena señalar:

1. Las representaciones complejas (diferentes de las representaciones distribuidas del caso hiper- local, mostradas en el *esquema 1* (ver pág. 35), **poseen una relación de constitutividad**, que es efectivamente una relación de nivel *dentro de*, tal como Fodor y Pylyshyn demandan. El patrón o vector que representa *taza con café* está compuesto por un vector que puede ser identificado como una representación distribuida de *taza sin café* junto con un vector que puede ser identificado como una representación distribuida particular de *café*. Sin embargo, Smolensky señala que al caracterizar los vectores constituyentes del vector que representa el compuesto (vale decir, de *taza con café*), no necesariamente éste será un vector comprendido por la actividad de unidades de sus microcaracterísticas individuales. Por tanto, la relación de nivel *entre* del vector y sus elementos numéricos individuales, no es una relación de constitutividad y en consecuencia, la crítica realizada por Fodor y Pylyshyn (1988) resulta irrelevante, puesto que apunta a un error que no se ha cometido.

2. La **relación de constitutividad** entre las representaciones distribuidas es importante para el análisis de los modelos conexionistas y para explicar su conducta, pero no es una parte causal dentro del modelo. Esto significa que **para procesar** el vector que representa *taza con café*, la red no tiene que descomponerlo en sus componentes. Para procesar, lo que importa es la relación de nivel *entre* y no la de nivel *dentro de*.

3. La descomposición de estados compuestos en sus componentes no es precisa ni está definida unívocamente. Esto significa que si bien la **noción de constitutividad** es importante, los intentos por formalizarla involucran probablemente sólo una aproximación. Las nociones que provienen de la computación simbólica proporcionan instrumentos importantes para construir enfoques de nivel superior del comportamiento de modelos conexionistas, pero dichas nociones sólo proporcionan enfoques aproximados e imprecisos.

4. Por las razones anteriores, Smolensky afirma que aun cuando las redes conexionistas usen representaciones distribuidas que describen estados con el tipo de constitutividad requerida por el principio (1), ellas no proporcionan una implementación literal del lenguaje del pensamiento. De lo anterior se desprende, que para el enfoque conexionista, hay un lenguaje del pensamiento, pero sólo aproximativamente, dado que este no proporciona una base para la descripción formal exacta de la estructura o de los procesos mentales y en consecuencia no puede proveer un enfoque formal preciso de la arquitectura cognitiva.

La conclusión global obtenida por Smolensky en este artículo, nos señala que la diferencia entre el enfoque clásico y el conexionista no radica en si estos aceptan los principios (1) y (2), sino más bien en cómo los instancian formalmente. En el enfoque clásico, estos principios son formalizados usando estructuras sintácticas para los pensamientos y manipulación simbólica para los procesos mentales. En cambio, en la perspectiva conexionista, estos principios *son formalizados usando representaciones vectoriales distribuidas para los estados mentales y la correspondiente noción de composicionalidad, junto con procesos mentales basados en la asociación que derivan su sensibilidad a la estructura de la sensibilidad a la estructura de las representaciones vectoriales comprometidas en esos procesos.*³² En resumen, en opinión de Smolensky, las representaciones distribuidas si proveen una descripción de los estados mentales con componentes interpretables semánticamente, pero no permiten presentar un enfoque formal preciso de la construcción de compuestos a partir de componentes interpretables semánticamente, independientes del contexto.

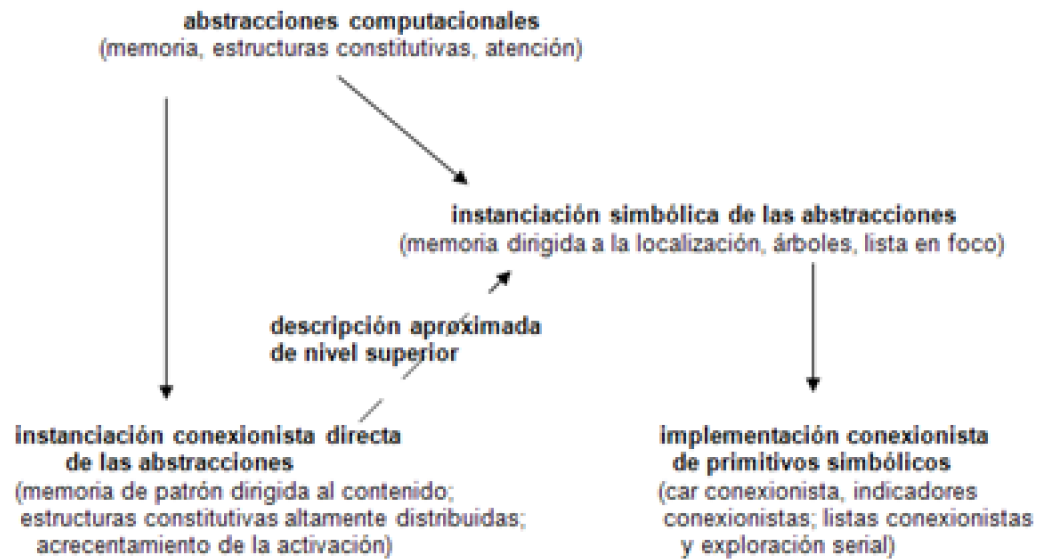
De las consideraciones anteriores se desprende, que en términos de una metodología de investigación, la agenda del conexionismo no debería ser el desarrollo de una implementación conexionista *del lenguaje del pensamiento simbólico* sino más bien *el desarrollo del análisis formal de las representaciones vectoriales de estructuras complejas y de las operaciones en aquellas estructuras que son lo suficientemente sensitivas a la estructura como para realizar el trabajo requerido.*³³ Por ello, en la parte final de este artículo, Smolesky discute el hecho de que si el conexionismo respeta los principios (1) y (2) tiene que ser necesariamente una implementación de una arquitectura clásica.

En su opinión (1988), lo que se necesita es una estrategia distinta a la agenda implementacionista propuesta por Fodor y Pylyshyn, más bien un Tratamiento Adecuado del Conexionismo. Smolensky denomina TAC (Proper Treatment of Connectionism) a su propuesta, y esta básicamente consiste en tener presente *el principio de la correspondencia cognitiva* (cognitive correspondence principle). Según este principio, *cuando los sistemas conexionistas son analizados en niveles superiores, los elementos de la computación simbólica aparecen como propiedades emergentes.*³⁴ Somlensky presenta el siguiente esquema para ilustrar su propuesta:

³² Smolensky, en Rabossi, p. 399

³³ Ibid anterior, p. 400

³⁴ Smolensky, en Rabossi. Sección 3: Conexionismo e implementación, p. 401.



Esquema 2: TAC vs. implementacionismo

El principio de correspondencia cognitiva está indicado en el esquema anterior por la flecha punteada y establece que:

La relación general entre los principios formales conexionistas y los principios formales simbólicos –dado que ambos son instanciaciones de nociones comunes no formales- consiste en que si uno toma un nivel superior de análisis de lo que está sucediendo en los sistemas conexionistas, encuentra que encaja, en algún grado de aproximación, con lo que está sucediendo en el formalismo simbólico.³⁵

En consecuencia, la agenda TAC, difiere de la de Fodor y Pylyshyn en que la metodología de implementación procede desde arriba hacia abajo, pero no directamente *via* la rama izquierda, sino indirectamente, *via* la rama derecha. Según Smolensky los conexionistas deberían tomar las instanciaciones simbólicas de los principios no formales (parte superior del esquema) y deberían encontrar maneras de implementarlas computacionalmente. En el argumento de Fodor y Pylyshyn, los principios no formales relevantes eran sus principios de la composicionalidad (1) y (2).

Sintetizando entonces los argumentos presentados por Smolensky como respuesta a Fodor y Pylyshyn, podemos decir que su propuesta radica, por una parte, en adoptar el lado *blando de la cognición* como sustrato fundamental sobre el cual construir el duro. Es decir, Smolensky aboga por un programa de reconstrucción de las entidades simbólicas por medios de una dinámica subsimbólica. Esto significa, que los modelos conexionistas sirven para refinar los modelos simbólicos, eliminando las distancias que los separan por medios de una *traducción* de los sistemas de símbolos discretos en computaciones continuas. Smolensky asume un *enfoque subsimbólico*, que explicitaremos en la sección siguiente.

³⁵ *Ibid anterior*, pp. 402-403.

5. LA ARQUITECTURA SUBSIMBÓLICA DE SMOLENSKY

Ante la problemática enunciada anteriormente, algunos partidarios del conexionismo como Smolensky (1988), intentaron desarrollar metodologías para dotar a la redes conexionistas de constituyentes y mostrar además que los sistemas obtenidos no son implementaciones de modelos clásicos.

En *Connectionist Modeling: Neural computation / Mental Connections* (1989), Paul Smolensky presenta una propuesta alternativa al paradigma simbólico (propio de la arquitectura clásica), denominada por el mismo, paradigma subsimbólico. Su propuesta, pretende solucionar algunas inadecuaciones que ofrece la arquitectura clásica como comprensión de la cognición. Smolensky, compara mediante dos gráficos la arquitecturasimbólica y subsimbólica, señalando claramente sus diferencias.

1. La arquitectura simbólica. Se parte aquí de la *estructura mental (Psicología popular)* que incluye metas, creencias, conceptos, esquemas, conocimiento, inferencia, percepción y acción. Los conceptos mentales, de este nivel son formalizados en término de un *lenguaje del pensamiento* (como estableció Fodor, 1975), siendo este lenguaje el soporte de la psicología popular. Las reglas para operar en este lenguaje son esencialmente *leyes del pensamiento*. Esta estructura simbólica se sostiene en un *sistema de símbolos físicos* (que computa y manipula símbolos) y que a su vez, descansa en un nivel más bajo de implementación, correspondiente al dispositivo computacional. En este paradigma, se supone además, que si seguimos descendiendo nos encontraríamos en algún momento con el *nivel neural* (vías nerviosas, neuronas, membranas, etc.). Según lo anterior, es posible visualizar dos niveles diferentes en esta arquitectura: el nivel simbólico y el nivel neural, siendo la descripción del nivel simbólico, la determinante en este paradigma, el nivel neural se deja de lado. La conexión entre ambos niveles no queda explicada en este modelo.

2. El paradigma subsimbólico. Se parte aquí de la *estructura neural*, la cual puede ser modelada formalmente y abstractamente por el sistema subsimbólico, pero en forma aproximada. El nivel subsimbólico, representa un tercer nivel intermedio entre la estructura neural y el nivel simbólico y tiene por tarea, formalizar el tipo de procesamiento que tiene lugar a nivel neural. El nivel subsimbólico, posee estructuras menores que los símbolos, mediante las cuales es posible acercarse a éstos (diferencia con el paradigma clásico). En este modelo no hay representaciones mentales propiamente tales. Finalmente se pasa a la *conducta cognitiva real*, que es aquella que sustenta el nivel de *representación mental*. De este modo vemos que lo mental vendría a ser en este modelo un epifenómeno, lo que significa que no tendría un rol causal sobre lo neuronal. La estructura mental aparece como *emergencia* en este modelo, gracias a la estructura neural y al paradigma subsimbólico.

La cognición descrita en el nivel subsimbólico, corresponde a una descripción propia

de un sistema conexionista, la cual se encuentra más cerca del nivel neural que del nivel simbólico. La estructura neural, es la que permite dar una descripción precisa de la inteligencia, suministrando la base para la formalización de este proceso (expresada en el nivel subsimbólico). La estructura mental en cambio, formalizada por los conceptos de la psicología popular (creencias, metas, conocimiento, inferencia, etc.), sólo proporciona una descripción aproximada de la inteligencia.

Además de las diferencias descritas anteriormente, Smolensky considera que el contraste más fundamental entre estos dos paradigmas corresponde a la interpretación semántica que ofrecen ambos modelos. En el paradigma simbólico, los símbolos (átomos) son usados para denotar la interpretación semántica de los conceptos. Así, estos símbolos son manejados por las reglas que definen el sistema. Por tanto, las entidades que son interpretadas semánticamente en el paradigma simbólico, son siempre entidades gobernadas por leyes formales que definen este sistema, las cuales son activadas de manera individual en la red. En el paradigma subsimbólico, en cambio, las entidades interpretadas semánticamente son *patrones de activación*, que actúan sobre un gran número de unidades en el sistema. El paradigma subsimbólico usa *representaciones distribuidas*, propias de los sistemas conexionistas, como opuesto a las representaciones locales.

Los conexionistas separan las unidades de procesamiento de las unidades de interpretación semántica. De este modo, es posible distinguir dos capas o estratos en la estructura de este paradigma en contraste con el paradigma simbólico: un nivel inferior al nivel de la interpretación semántica, propio la descripción formal del sistema (nivel de la manipulación), y el nivel semántico más alto (*the higher semantic level*) propio de la denotación. Ambos niveles son esenciales; el nivel más bajo define lo que el sistema es y el nivel más alto es esencial para entender qué significa el sistema.

Cabe señalar por último, el carácter computacional del lenguaje subsimbólico y el modo en que se concibe la inferencia en este modelo conexionista, que es totalmente distinto a como la conciben los cognitivistas clásicos. La inferencia básica aquí (nivel subsimbólico), es *asociativa estadística*, vale decir, las microunidades permiten realizar inferencias, mediante relaciones de co-presencia (generalizaciones tipo Hume), mostrando con ello que se puede dejar de usar leyes generales, como se hace en el paradigma clásico.

En opinión de Smolensky, el paradigma subsimbólico ofrece varias **ventajas frente al paradigma simbólico**. Estas pueden resumirse en lo siguiente:

1. El paradigma subsimbólico provee una valiosa comprensión sobre la arquitectura real del cerebro, cosa que no hace el paradigma simbólico. Mediante el modelamiento conexionista, es posible mostrar cómo se lleva a cabo el procesamiento cognitivo, mediante un enfoque biológico real, que considera el funcionamiento de las neuronas.

2. Desde la perspectiva del modelamiento de la cognición humana, se observa que el paradigma simbólico es deficiente en la descripción de la estructura más fina de la cognición, logrando sólo una descripción satisfactoria de tareas más gruesas. Según Smolensky, el paradigma subsimbólico permite llevar a cabo esta tarea.

3. El sistema subsimbólico (conexionista) posee mayor flexibilidad en sus

operaciones, puesto que no posee reglas lógicas, aplicadas en la manipulación de los símbolos. Por esta razón, el modelo simbólico en cambio, se transforma en un sistema demasiado rígido.

6. EVALUACIÓN CRÍTICA DEL CONEXIONISMO

Como vimos en este capítulo, el modelo conexionista considera necesario dejar atrás la metáfora del ordenador, propia del paradigma cognitivista, para reemplazarla por la metáfora neuronal. Sostiene que los modelos computacionales están inspirados neuronalmente, es decir, que los ordenadores se conciben bajo la imagen del cerebro y no inversamente como sostenía el paradigma clásico.

Una de las razones fundamentales de este cambio, se debe al inicio de una crítica al concepto mismo de **representación** sustentada por el paradigma clásico. Este modelo discute la idea de que el cerebro está regido por reglas, que posee un procesador central y que existen en él, lugares precisos para almacenar la información. El modelo conexionista se caracteriza por construir un sistema cognitivo, que deja atrás **los símbolos** y las reglas formales propias del paradigma clásico y por utilizar componentes simples que se conectan de forma dinámica entre sí más allá de su actividad local. Por ello, se habla en este enfoque de propiedades emergentes o globales de redes de unidades interconectadas o bien de sistemas complejos. Cabe destacar que en este enfoque, **los símbolos** no juegan ningún papel, lo que significa abandonar el supuesto cognitivista básico, de que tiene que haber un **nivel simbólico** en la explicación de la cognición. La cognición es entendida por el conexionismo como *emergencia de estados globales en una red de componentes simples y esta funciona a través de reglas locales que gobiernan las operaciones individuales y de reglas de cambio que gobiernan la conexión entre los elementos.*³⁶

Los modelos conexionista son por tanto, modelos **no simbólicos** de la mente, en los que las representaciones surgen de la fuerza de las conexiones entre las unidades del sistema y no de los símbolos codificados dentro del sistema, siendo los pesos de estas conexiones los que determinan el funcionamiento de la red.

Considerando lo anterior y teniendo presente el objetivo central de esta investigación, podemos preguntarnos entonces si la **representación mental** tiene o no una relevancia cognitiva en este enfoque de la CC? Según lo expuesto, podemos resumir que para el conexionismo, toda **representación mental** (ya sea local o distribuida), consiste en tener tales o cuales unidades excitadas o inhibidas en cierto grado. **La representación** en los modelos conexionistas es numérica (en lugar de ser simbólica como en el cognitivismo), ya que las unidades de procesamiento de estos modelos se comunican unas con otras mediante señales numéricas en lugar de señales simbólicas. Tanto el conexionismo como el cognitivismo clásico asignan contenido semántico a la representación, la diferencia radica en que el conexionismo lo asigna a los nodos (unidades o agregados de

³⁶ Varela, *Conocer*, pp. 76 - 77.

unidades) y los cognitivistas lo asignan a las expresiones simbólicas.

En un modelo conexionista típico, el input del sistema viene dado por los valores de activación de las unidades inputs de la red. Estos valores numéricos suponen cierta representación del input que se propaga a lo largo de las conexiones hasta que aparece algún conjunto de valores de activación de las unidades output.³⁷

Esto significa, que son los valores de activación los que codifican el “output” que el sistema ha computado a partir del *input*. En consecuencia, **la representación** se distribuye a lo largo de toda la red, así por ejemplo si un modelo conexionista instanciara la proposición “Si A entonces B”, la representación de A o de B no estaría en ninguna parte, sino que la proposición estaría representada de manera holista en toda la red.

Según lo expuesto podemos ver que las representaciones mentales pueden tener una **relevancia cognitiva** en los sistemas conexionistas. Esto significa que hay algunas representaciones que efectivamente **desempeñan un papel causal en la etiología de la conducta**. Esto se explica en función del contenido semántico que tienen las representaciones mentales de los sistemas conexionistas. Dicho contenido semántico resulta ser una función, más o menos compleja, de lo excitados o inhibidos que estén los nodos de la red que estructura el sistema: *una representación mental puede ser un vector en un espacio de activaciones de cierto número de unidades que, a su vez, representan información del entorno.*³⁸

Las representaciones mentales (contrario a lo que opina Fodor) tienen en consecuencia, una estructura semántica no simbólica, que a veces se solapa directamente con su estructura causal y otras veces no. De igual modo como la sintaxis orienta la causalidad en el computacionalismo clásico, en el conexionismo es la propia distribución y cambio de los pesos lo que consigna el efecto modulador. Esto significa que de alguna manera la semántica podría encontrar en el conexionismo (de igual modo como en el cognitivismo) un lugar natural a través de la causalidad.

³⁷ Soledad Ballesteros, *La representación del conocimiento en los sistemas conexionistas*, p. 348.

³⁸ Manuel Liz, *La estructura de las representaciones mentales: una perspectiva integradora*, p. 152.

CAPÍTULO TERCERO: LA REPRESENTACIÓN EN LAS TEORÍAS ALTERNATIVAS DE LA CIENCIA COGNITIVA

Como vimos en el capítulo primero, el cognitivismo clásico sostiene que la cognición es equivalente a la computación (equiparando la mente a un tipo de computador; poseedora de un procesador central, con una base de datos simbólica y algunos módulos periféricos *sensoriales*) y estableciendo que los procesos cognitivos son equivalentes a manipulaciones internas de representaciones simbólicas.

Este paradigma (cognitivismo clásico) ha sido la concepción predominante en la ciencia cognitiva. Sin embargo, este enfoque se ha topado con una variedad de dificultades y limitaciones que han sido cada vez más evidentes y que han permitido desarrollar otros modos para entender la cognición. Como resultado de esto, en la actualidad se han desarrollado muchos programas de investigación (**teorías alternativas en CC**) opuestos al enfoque computacional tradicional, como por ejemplo el enfoque de los sistemas dinámicos, como también la perspectiva de la cognición corporalizada, en sus variados enfoques: la robótica situada de Brooks, el enfoque enactivo de Varela y la postura ecuménica de Andy Clark, entre otros.

Para la mayoría de los partidarios de las **teorías alternativas**, la cognición es un

fenómeno que emerge de relaciones causales más complejas entre elementos del cerebro, el cuerpo y el mundo. De este supuesto, se desprende claramente una visión crítica hacia la doctrina del individualismo metodológico sustentada por el paradigma clásico (ver capítulo 1, p. 15). Las teorías alternativas rescatan la complejidad causal subyacente a los fenómenos cognitivos, cosa que evidentemente una psicología individualista no hace.

En lo que sigue, se presentarán de modo resumido algunas de estas propuestas, con el objeto de evaluar cuál es el rol que se le asigna a la **representación mental** en cada uno de estos enfoques, si es que se le asigna algún rol. En primer lugar se expone la aproximación de Van Gelder a los sistemas dinámicos, dado que a partir de ella podemos graficar de forma sintética las principales diferencias entre computacionalismo, conexionismo y sistemas dinámicos. Posteriormente se presenta la hipótesis de la cognición corporalizada en la Ciencia Cognitiva, desde dos autores principales: el enfoque enactivo de Varela y la postura ecuménica de Andy Clark.

1. LA COGNICIÓN COMO SISTEMAS DINÁMICOS, SEGÚN VAN-GELDER

En *Dynamics and Cognition* (1996), Timothy Van Gelder, intenta aclarar qué es la cognición. Para ello, expone críticamente la visión del paradigma simbólico de la cognición, señalando sus dificultades y limitaciones. A partir de ello considera necesario revisar algunos programas de investigación, que en la actualidad han intentado solucionar las deficiencias del paradigma clásico, como por ejemplo, el enfoque conexionista, la robótica situada, entre otros. La pregunta central, radica entonces en evaluar si ¿hay una alternativa real frente a entender la cognición como computación?, es decir, frente a la concepción computacional. La afirmación central realizada por Van Gelder en este artículo, radica justamente en decir que hay efectivamente una alternativa viable frente al enfoque computacional. En su opinión, es posible concebir a la cognición como un *sistema dinámico*.

Para introducir la concepción dinámica de la cognición, Van Gelder se remonta a la revolución industrial de 1788 en Inglaterra, considerando el problema de la máquina a vapor. El problema de la máquina a vapor, consistía en que se necesitaba diseñar un dispositivo, que permitiera mantener a la máquina trabajando en forma constante. Lo anterior significaba, implementar un dispositivo de ajuste automático de la válvula de aceleración (del volante de la máquina a vapor), que permitiera mantener constante su velocidad. Sin este mecanismo regulador, la velocidad del volante variaba dependiendo de las fluctuaciones del vapor, de las alteraciones del volumen de trabajo y de otros factores. Por estos motivos, James Watt diseñó un dispositivo, conocido justamente con el nombre *reguladores de Watt* (en honor a su nombre).

La estructura, componentes y funcionamiento de este regulador (equivalente a un **gobernador dinámico**) puede resumirse en lo siguiente:

El regulador se basa en un eje vertical conectado con el volante principal. El eje tiene dos brazos, unidos por goznes, con una esfera metálica en el extremo de cada uno. Cuando el volante gira, los brazos se van separando del eje a medida que aumenta la velocidad de rotación. Los brazos actúan directamente sobre una válvula reguladora que reduce el flujo de vapor cuando los brazos se separan (cuando aumenta la velocidad del volante) y que lo aumenta cuando los brazos se acercan (cuando disminuye la velocidad del volante).³⁹

Lo óptimo de este regulador de Watt, es que permite que efectivamente el volante mantenga una velocidad de rotación constante, lo que resulta muy necesario para muchas aplicaciones industriales. Van Gelder, contrasta este regulador de Watt (similar para él, a un gobernador dinámico) con el funcionamiento de un hipotético gobernador computacional, para intentar evaluar por qué el regulador dinámico ideado por Watt, resultaría ser mucho más eficaz que éste. Efectivamente, el gobernador computacional, nunca se implementó en época de Watt por la falta de tecnología existente. La estrategia de Van Gelder consiste entonces, en presentar de modo hipotético cómo sería el funcionamiento de un gobernador computacional, para el problema de la máquina a vapor. Según él, este gobernador computacional, funcionaría obedeciendo a los siguientes algoritmos:

(1) Fase de inicial, en la cual: a) Habría que medir la velocidad del volante y b) comparar la velocidad real con la velocidad deseada.

(2) Fase de decisión, en la cual: Si no hay discrepancia tendríamos que retornar a la fase (1) y si la hay entonces: a) Habría que medir la presión actual del vapor. B) Calcular la alteración deseada en la presión del vapor. C) Calcular el ajuste necesario para la válvula reguladora. D) Ajustar la válvula reguladora.

(3) Fase de recurrencia, mediante la cual se vuelve a la fase 1.

A partir de esta comparación hipotética, Van Gelder distingue entre dos tipos de gobernadores: el computacional y el dinámico. **El gobernador computacional**, propio de la arquitectura clásica, posee según Van Gelder las siguientes características; su propiedad más distintiva es la **dependencia de la representación**. Esto se evidencia de manera significativa considerando las tres fases descritas anteriormente. La primera cosa que hace este gobernador es medir su medio (velocidad de la máquina) y luego hace una representación interna de la actual velocidad de la máquina. Posteriormente, realiza una serie de representaciones y finalmente emite un output representacional. Aplicado esto a la máquina a vapor, significaría que gracias a esta representación final, sería posible activar el mecanismo de la válvula de aceleración para hacer el cambio correspondiente. Lo anterior implica una **descripción computacional**, en un sentido más restringido, puesto que literalmente lo que la máquina hace es computar el deseo de cambiar la válvula de aceleración manipulando **símbolos**, de acuerdo a un esquema de reglas. La manipulación simbólica es la que permite el éxito del gobernador en su tarea. La manipulación de estos símbolos es mediante pasos discretos, vale decir de forma **secuencial**. Esto significa que para realizar un cambio apropiado en la válvula de aceleración, es necesario chequear primero si existe discrepancia entre la velocidad

³⁹ Clark, pp. 139- 140.

deseada y la real (ver esquema anterior, fase de decisión). Además, siguiendo este mismo esquema, vemos que el dispositivo completo opera de una manera **cíclica**: primero mide o percibe su medio, entonces internamente computa un cambio en la válvula de aceleración y posteriormente lleva a cabo su cambio, actuando en su medio. Así, después que estos cambios se han realizado (y dado el tiempo para afectar la velocidad del motor), el gobernador computacional funciona nuevamente a través de otro ciclo completo. Este gobernador está construido de modo **homuncular**, es decir, posee subdivisiones en partes o componentes, siendo cada una de estas partes del sistema, responsable de una subtarea. Estas partes o departamentos, interactúan comunicándose, traspasándose mensajes significativos. Vemos así, que la naturaleza **representacional y computacional** de este sistema, son esenciales en esta construcción homuncular. Si el sistema (como un todo), no operara mediante representaciones no sería posible para sus componentes, interactuar comunicándose.

Sintetizando, obtenemos las siguientes propiedades del gobernador computacional: representacional, computacional, secuencial, cíclico y homuncular. Estas forman un grupo mutuamente interdependiente, si el mecanismo posee una las posee todas.

El **gobernador dinámico de Watt**, en cambio, no exhibe este grupo de propiedades como un todo ni a ninguna de ellas individualmente. Este gobernador opera con cantidades y no con representaciones (sólo hay magnitudes físicas y no es necesario que éstas se expresen en un nivel simbólico). La relación entre las cantidades se describe en lenguaje matemático. No es computacional, puesto que opera de modo mecánico. Podemos decir entonces, que en este gobernador no habría ningún tipo de manipulación de representaciones ni de computaciones, puesto que sus cantidades establecen entre sí, relaciones de **acoplamiento** (relación causal compleja entre múltiples cantidades que harían ser dinámico al sistema). Además este gobernador no es secuencial, puesto que no hay secuencias discernibles (lo único discernible es la conducta global del sistema), ni tampoco cíclico. Por último, vemos que no es un sistema homuncular, porque sus componentes no se coordinan a partir de la comunicación de la información, como ocurre en el sistema computacional.

Finalmente podemos señalar, una diferencia esencial entre el gobernador computacional y el dinámico: **su temporalidad**. En el gobernador computacional, no se señala ningún tipo de restricción respecto del tiempo en que se realiza cada subtarea (no se especifica cuánto debe tardar cada paso, ni cuánto debe transcurrir entre un paso y otro). La temporalidad de las operaciones internas de este gobernador, resulta arbitraria, considerando el comportamiento total del sistema. En el gobernador dinámico en cambio, la conducta ocurre en el mismo tiempo real en que tienen lugar los cambios de la máquina. Por ello, la temporalidad de este gobernador no puede ser arbitraria, más bien es equivalente a la temporalidad de la máquina.

Las comparaciones establecidas por Van Gelder, entre el gobernador computacional y el dinámico, permiten realizar una evaluación de ambos modelos, como eventuales ejemplificaciones de la arquitectura clásica y de la hipótesis dinámica de la cognición. Podemos decir al respecto, que el gobernador dinámico, mostraría ciertas ventajas con respecto al gobernador computacional, en tanto satisface de mejor manera el **criterio de parsimonia ontológica**. El gobernador dinámico y la **simplicidad** de sus componentes,

muestran que es posible explicar la cognición sin aludir a representaciones, computaciones (no hay manipulaciones de símbolos en este modelo), procesos secuenciales y cíclicos, más departamentos de sub tareas (construcción homuncular). La descripción de la cognición se mantiene en un ambiente puramente físico, como resultado de un proceso mecánico (relaciones de acoplamiento). Sin embargo, este reduccionismo extremo nos lleva a preguntarnos ¿hasta qué punto logra el modelo dinámico de la cognición, proporcionar una explicación adecuada de la cognición? Si consideramos la crítica realizada por Eliasmith⁴⁰ a los sistemas dinámicos, vemos que éstos sólo ofrecen una buena metáfora para describir los procesos cognitivos, pero que no son adecuados para ofrecer un modelo científico de la cognición. Considerando esto, podemos decir que el modelo computacional, aún siendo más complejo en sus elementos, logra ofrecer una explicación más efectiva de los procesos cognitivos, en tanto da cuenta de la realidad representacional y simbólica de la mente humana.

El gobernador dinámico, satisface de mejor manera el **criterio de realidad psicológica**, que el gobernador computacional, con respecto a la temporalidad en que se llevan a cabo las operaciones. Estas ocurren en el tiempo real de la máquina (en los sistemas dinámicos), mientras que en el modelo computacional quedan indeterminadas en este aspecto, siendo arbitrarias con respecto al comportamiento total del sistema.

En el punto cinco de su artículo *Dynamics and Cognition* (1996), Van Gelder realiza una comparación entre los tres tipos de sistemas de la cognición, que hemos estado considerando: computacional (sistema manipulador de símbolos, equivalente a una máquina de Turing o máquina LISP), conexionista (sistema de red de unidades neurales, como Perceptron o la red Hopfield) y dinámico (sistema de sets de magnitudes acopladas, similar a un sistema solar o gobernador de Watt). Para llevar a cabo esta comparación, en sus propiedades más características, Van Gelder considera cuatro aspectos centrales que sirven de contraste entre estos tres sistemas: los tipos de variables comprometidas, el modo en que los estados cambian, las herramientas utilizadas para describir estos cambios y la naturaleza o carácter distintivo de cada sistema. Veremos entonces cómo se presentan, estos cuatro aspectos en cada sistema.

1. Modelos computacionales: Estos modelos siempre utilizan variables digitalizadas (a menudo sintácticas), lo que significa que debe haber un grupo de valores discretos tales que en cualquier momento relevante estas variables no sean ambiguamente tomadas en uno u otro de estos valores. Por consiguiente, cada cambio de estado, se lleva a cabo desde una configuración digital a otra (paso secuencial). Esto significa, que la transición es esencialmente discreta, (no hay tiempo -ni estado- teóricamente relevante, entre el primero y el siguiente). Estas propiedades son reflejadas en la naturaleza de las reglas que describen el comportamiento del computador. Estas reglas (programa), siempre especifican cual es el siguiente estado, mediante una operación discreta que transforma el actual estado en el siguiente estado. Las herramientas utilizadas por el modelo computacional, para describir estos cambios, corresponde a reglas de transición, programas y algoritmos. Por esta razón, los computadores han sido caracterizados, dada la naturaleza de sus variables, sus cambios

⁴⁰ Ver Eliasmith, *The Third Contender: A Critical Examination of the Dynamicist Theory of Cognition*.

de estados y cómo estos (cambios de estados) son especificados, como un sistema automático formal. Estos sistemas automáticos formales, tienen una estructura que soporta una correspondencia (sistemática y razonable) con algún dominio (como por ejemplo la aritmética, el baseball, u otro). Los estados relevantes del sistema son configuraciones estructuradas de *tokens* interpretados como representaciones simbólicas del dominio. Los cambios de estados son inferencias desde una representación simbólica a otra. Esto último evidencia el carácter representacional del sistema computacional.

2. Conexionismo: Las variables del conexionismo (al igual que en los sistemas dinámicos), se expresan en términos cuantitativos, difiriendo así del modelo computacional. Lo que marca la diferencia esencial en este sistema es que estas variables son modeladas de modo genérico por neuronas biológicas y así toda unidad de la red viene definida por un nivel de activación, que es expresado de forma cuantitativa. Los cambios en los estados se deben entonces a la propagación e interacción de neuronas (las unidades interactúan entre si mediante las conexiones que las asocian). Toda conexión viene caracterizada por un valor de fuerza del peso de la conexión, que se expresa también de forma cuantitativa. Las herramientas que se utilizan para describir estos cambios, son especificados por una función simple, correspondiente a ecuaciones de suma ponderadas. Finalmente podemos agregar, que los sistemas conexionistas poseen alta dimensionalidad (gran cantidades de variables) y que son homogéneos (todos los cambios de variables ocurren de la misma forma).

3. Sistemas dinámicos: Las variables en este sistema no son digitales, sino que se expresan en **términos cuantitativos** (son sujetos de medida), dado que se refieren a cantidades que miden los estados y las trayectorias. Aquí no se pueden individuar estados para unidades, ni niveles de activación, sólo se consideran las interacciones. El cambio en sus estados, se lleva a cabo en *un tiempo real*, lo que señala la imposibilidad de poder individuar sus propiedades causales. Las herramientas utilizadas por los sistemas dinámicos, corresponden a ecuaciones diferenciales (empleo de la matemática). El rasgo general distintivo de este sistema, consiste en la noción de *acoplamiento* de las variables, incluyendo también el ambiente.

De acuerdo a lo expuesto, podemos decir que la teoría de la cognición como sistema dinámico busca mostrar que existe un tipo de modelamiento matemático que puede rescatar en forma sistemática la compleja red de interacciones causales que subyacen a la emergencia de la cognición. Desde este enfoque, se discute en forma crítica la concepción clásica que hace equivaler la cognición a computación, y esta a manipulación interna de *representaciones simbólicas*. Los sistemas dinámicos operan con cantidades (magnitudes físicas) y no con representaciones. Según estos, no hay ningún tipo de representación ni de computación, sino sólo relaciones de acoplamiento. Esto significa, que los sistemas dinámicos no otorgan ningún rol a la representación mental, sino que más bien tienden a rechazarla como innecesaria.

2. LA COGNICIÓN CORPORALIZADA:

ANTECEDENTES FILOSÓFICOS

La hipótesis de la cognición como experiencia corporalizada en la CC (sustentada por autores como Brooks, Varela, Clark, entre otros), que expondremos a continuación, tiene algunos antecedentes filosóficos, que vale la pena señalar. Dentro de estos, destaca como punto de partida, la reflexión realizada por Heidegger (1927) en *Ser y tiempo*, donde se señala la estructura unitaria del *Ser-en-el-mundo*. Heidegger escribe sobre la importancia de Dasein (*ser-ahí*), expresión que Clark toma en su expresión inglesa *being there*, que significa tanto como *ser ahí* como *estar ahí*, y que corresponde justamente al título de su libro (*Estar ahí*, 1997). En su planteamiento Heidegger señala un modo de *ser- en -el- mundo*, caracterizado por la existencia de *acoplamientos funcionales* entre los *utensilios* (es decir, las cosas que nos rodean) y las relaciones prácticas que establecemos con ellos. Estos *acoplamientos funcionales* entre sujeto y mundo no suponen la existencia de representaciones separadas (por ejemplo entender a un martillo como un objeto rígido), sino en correlación con su función (utilizamos al martillo para clavar clavos). La existencia humana se presenta, como un *estar en el mundo*, con las cosas del mundo, pero no considerando estas en cuanto cosas, sino en cuanto *situaciones*. El mundo resulta ser así, junto con el ente que existe *una unidad radical* que caracteriza la existencia en tanto existencia.

En opinión de Clark, *el trabajo de Heidegger prefigura el escepticismo hacia lo que se puede denominar tipos de representación interna -neutrales respecto de la acción- y hace eco de nuestro énfasis en el empleo de instrumentos y en los acoplamientos entre organismo y mundo orientados hacia la acción.*⁴¹

Otro antecedente filosófico, lo encontramos en la reflexión realizada por Merleau-Ponty, en sus obras *La estructura del comportamiento* (1942) y *Fenomenología de la percepción* (1945), en las cuales se destaca que la conciencia está verdadera y efectivamente comprometida con el mundo, oponiéndose con ello a toda concepción de la conciencia como *interioridad*, como también a las concepciones que entienden al cuerpo como una *cosa* (cartesianismo). Por el contrario Merleau -Ponty, describe la actividad inteligente cotidiana, mediante la integración entre el organismo, el cuerpo y el mundo, destacando además, que la percepción se dirige hacia el control de la conducta en el mundo real y en tiempo real.

Según Merleau-Ponty la percepción no es un fenómeno causal (síntesis de sensaciones, valoraciones o juicios frente a la realidad), sino un proceso que se enriquece con la experiencia de vida de una persona. La percepción no consiste entonces en una fotografía estable de la realidad (verdad universal), sino que es una construcción que emerge de un cuerpo. El cuerpo, es el que permite activar la comunicación humana, gracias a su interacción con el entorno.

Merleau-Ponty plantea que no hay separación entre cuerpo y espíritu o mente. Para él, la existencia se encarna, *toma cuerpo* en la base biológica del individuo, mediante la *animación* del cuerpo. De este modo, la mente es una autoconstrucción que emerge del cuerpo y que se forja a través de la comunicación existente entre un entorno variable, una

⁴¹ Clark, *Estar ahí*, p. 224

historia, un lugar dado y un momento dado.

De las ideas de estos filósofos, se extrae entonces, la hipótesis que considera al cuerpo como *el lugar donde se realiza la experiencia*, sin distinción previa entre sujeto/objeto, entre yo/mundo. La noción de experiencia, cobra singular importancia y se toma como punto de partida en esta visión; y luego se busca evidencia para tratar de apoyarla.

Bajo esta hipótesis, **la cognición** se concibe como un *fenómeno emergente*, de la compleja relación entre el organismo y su ambiente, no siendo posible en consecuencia, especificar claramente las fronteras entre el sujeto y el mundo. La cognición, por tanto, ya no puede ser entendida como aquella relación externa que se establece entre estos dos elementos (sujeto y mundo), como ocurre en el paradigma clásico. En este sentido, podemos decir que la hipótesis de la cognición corporalizada comparte con los planteamientos de los sistemas dinámicos la afirmación de que la cognición es una propiedad emergente de relaciones causales complejas entre propiedades perteneciente a ámbitos de propiedad diversos.

Si tenemos presente, por ejemplo, los planteamientos realizados por Andy Clark en *Estar ahí* (1997) vemos que él comparte esta idea de indisolubilidad entre la mente y el cuerpo. Clark, nos presenta la idea de una mente esencialmente corporeizada (*embodied*), la cual además se encuentra encajada o embebida (*embedded*) en el mundo en el que actúa. En *Estar ahí*, Clark propone que la interacción cerebro, cuerpo y mundo es constitutiva de la mente. Por su parte, Francisco Varela en sus libros *Conocer* (1990) y *De cuerpo presente* (1997), postula que las verdaderas unidades de conocimiento son de naturaleza eminentemente concreta, incorporadas, encarnadas, vividas; y que la cognición en consecuencia depende de los tipos de experiencia que están determinados por el hecho de tener un cuerpo con determinadas habilidades sensorio-motrices.

Revisaremos en lo que sigue, los planteamientos de estos autores.

3. EL ENFOQUE ENACTIVO Y ANTIRREPRESENTACIONALISTA DE VARELA

*La mayor aptitud de la cognición viviente...consiste en la capacidad de plantear, dentro de amplias restricciones, los problemas relevantes que se deben encarar en cada momento. Estos problemas no son pre-dados, sino que se enactúan a partir de un trasfondo de acción, donde lo que cuenta como relevante está determinado contextualmente por nuestro sentido común.*⁴²

Francisco Varela (1946- 2001), propuso un nuevo punto de vista en la ciencia cognitiva, frente a los planteamientos desarrollados por el cognitivismo clásico y el conexionismo. Su propuesta, **el enfoque enactivo**, postula que la cognición es **enacción**, esto es, *una historia de acoplamiento estructural que alumbró o hace surgir un mundo.*⁴³ La

⁴² Varela, *De cuerpo Presente*, p. 173

designación *enactivo*, corresponde a un neologismo que proviene del verbo inglés *to enact*, que significa *poner en ejecución*, pero también *representar* o *actuar* en el sentido *desempeñar un papel*.

En su libro *Conocer*, Varela señala por una parte, que la orientación de este *enfoque enactivo*, surge de una profunda insatisfacción frente a los planteamientos del cognitivismo clásico y del conexionismo. Esta insatisfacción radica en el olvido que hacen estos enfoques del *sentido común*, intentando eliminar progresivamente el *know-how* ('saber como') contextual, mediante el descubrimiento de reglas elaboradas. En opinión de Varela, el *sentido común* es un conocimiento imposible de empaquetar en un conocimiento explícito y proposicional, como plantea el modelo clásico.

Por otra parte, en opinión de Varela, tanto el conexionismo como el cognitivismo utilizan un criterio de cognición que se traduce en una **representación de un mundo externo que está dado de antemano**. Para Varela en cambio, el mundo en que vivimos no está predefinido, sino que va surgiendo o es modelado, por la actividad diaria del ser humano. Por esta razón, Varela cuestiona la noción de representación utilizada por el cognitivismo, considerando que ésta no debería desempeñar un papel protagónico en la ciencia cognitiva: *se trata de una crítica a la noción de representación como núcleo de las ciencias cognitivas, ya que sólo se puede representar un mundo que está pre-definido. En cambio, si el mundo en que vivimos va surgiendo o es modelado en vez de ser predefinido, la noción de representación ya no puede desempeñar un papel protagónico.*

44

Varela, citando a Rorty⁴⁵ intenta presentar una visión alternativa frente a la tradición occidental. Para ello, toma como referencia a dos de los autores presentados en la sección precedente: Heidegger y Merleau Ponty, junto con la tradición fenomenológica. Su idea es defender un antirrepresentacionalismo. De la fenomenología destaca que el conocimiento se relaciona con el hecho de estar en un mundo, resultando inseparable de nuestro cuerpo, de nuestro lenguaje y de nuestra historia social. Basándose en estos planteamientos, Varela postula que la cognición no se puede entender adecuadamente sin considerar nuestra historia corporal y social (*sentido común*) y nos conduce además a afirmar que *conocedor* y *conocido*, *sujeto* y *objeto*, se determinan el uno al otro y que surgen simultáneamente.

Lo que Varela discute es un supuesto que se encuentra muy arraigado en nuestra tradición científica, este es: *que el mundo tal como lo experimentamos es independiente de quien lo conoce*. En su libro *De cuerpo presente* Varela desarrolla una crítica, a lo que él considera como el supuesto tácito de las diversas formas de *realismo cognitivo*⁴⁶

⁴³ Varela, *Conocer*, p. 109

⁴⁴ Varela, *Conocer*, p. 88.

⁴⁵ Rorty (1979), *La filosofía y el espejo de la naturaleza*.

⁴⁶ Varela entiende por *realismo cognitivo*, aquella postura que entiende la cognición como recuperación de un mundo pre-dado, pre-existente antes de toda actividad cognitiva.

(cognitivismo, emergencia, entre otros) que suponen que el mundo se puede dividir en regiones de elementos y tareas discretas; postulando que la cognición consiste en *resolución de problemas* siendo ésta posible en tanto se pueda distinguir entre los elementos, propiedades y relaciones de regiones pre-dadas, en un *espacio del mundo*.

Para ejemplificar su crítica, Varela toma como ejemplo un robot móvil que debe conducir un automóvil dentro de la ciudad y lo contrapone con lo que ocurre en el juego de ajedrez. En este último, el enfoque de la cognición como *resolución de problemas* resulta efectivo, dado que es posible definir los componentes del *espacio de ajedrez*, como por ejemplo la posición en el tablero, las reglas de movimiento, los turnos para jugar y otros. En su diferencia, en el caso del robot móvil, Varela señala que este enfoque resulta poco productivo, dado que aún cuando podamos distinguir ítems discretos en este *espacio de conducir*, tales como ruedas, ventanillas, otros coches o luces rojas; no resulta posible delimitar el movimiento entre estos objetos. Por ejemplo, frente a la pregunta, ¿debe el robot prestar atención a los peatones o no? O bien ¿debe tomar en cuenta las condiciones meteorológicas o el país donde está situada la ciudad? Varela considera que esta lista de preguntas podría resultar eterna, lo que nos indicaría que el *mundo de la conducción* no terminaría concretamente en ninguna parte, sino por que consta de *una estructura de incesantes niveles de detalle que no se funden con un trasfondo específico*.⁴⁷ Lo que realmente ocurre, es que el éxito de un movimiento como conducir un automóvil depende de habilidades motrices ya adquiridas y de un uso continuo del *sentido común*, o de un *know-how* ('saber cómo') acerca del trasfondo.

En opinión de Varela, el conocimiento por *sentido común*, resulta casi imposible de empaquetar en un conocimiento explícito y proposicional como postula el cognitivismo, dado que se trata en gran medida de una "*disposición*" o más bien de un *conocimiento práctico*, basado en la acumulación de experiencia a partir de un gran número de casos. Varela discute la idea de querer capturar el sentido común en una representación, puesto que para él, el mundo vivido no tiene límites pre-definidos. Por el contrario su propuesta señala la importancia de recuperar el *sentido común*, lo que implicaría en su opinión, invertir la actitud representacionista.

Si deseamos recuperar el sentido común, debemos invertir la actitud representacionista: en vez de tratar el know how contextual como un desecho que se debe eliminar progresivamente mediante el descubrimiento de reglas más elaboradas, considerarlo la esencia misma de la cognición creativa.⁴⁸

En consecuencia, el gran desafío que enfrentaría la ciencia cognitiva consistiría en opinión de Varela, en dejar atrás este supuesto tácito, presente en todas las formas de *realismo cognitivo*, que afirma: ***que el mundo es independiente del conocedor***. La siguiente cita resulta esclarecedora respecto a este punto:

Si estamos obligados a admitir que la cognición no se puede comprender sin sentido común, y que el sentido común no es otra cosa que nuestra historia corporal y social, la conclusión inevitable es que el conocedor y lo conocido, la mente y el mundo, se relacionan mediante una especificación mutua o un

⁴⁷ Varela, *De Cuerpo presente*, p. 174.

⁴⁸ *Ibíd.*, p. 176

coorigen dependiente. Si esta crítica es válida, no habrá progreso científico en la comprensión de la cognición a menos que partamos de una base que no sea la idea de un mundo pre-dado que existe 'ahí afuera' y se recobra internamente mediante la representación.⁴⁹

Los planteamientos de Varela, se enfrentan con gran parte de la tradición filosófica presente en el desarrollo de la ciencia cognitiva y fundamentalmente con aquellas perspectivas reacias a aceptar esta orientación no objetivista del conocimiento y de la cognición. Su antirrepresentacionalismo, discute en esencia la noción de representación postulada por el cognitivismo clásico y por el enfoque conexionista.

En su análisis de la representación, Varela considera necesario distinguir entre un sentido débil y un sentido fuerte de representación.⁵⁰ Según el sentido débil, la representación se entiende como *interpretación* y la cognición consiste entonces en interpretar o representar el mundo *como si fuera de cierta manera*. Este sentido débil, señala el ámbito puramente semántico de la representación, dado que se refiere a todo lo que se pueda interpretar como *siendo acerca de algo*. Por ejemplo, un mapa es acerca de una zona geográfica, representa ciertos rasgos del terreno y gracias a él es posible interpretar que el terreno es de una cierta forma. Varela enfatiza que este concepto de representación es débil, puesto que no conlleva necesariamente ninguna implicación epistemológica u ontológica fuerte. Agrega además, que este concepto débil de representación es *pragmático*, dado que lo usamos constantemente sin preocuparnos. Por ejemplo, resulta totalmente aceptable decir que un mapa representa el terreno sin tener que preguntarnos cómo adquieren los mapas su significado.

El sentido fuerte de la representación surge en cambio, cuando se generaliza sobre la base del concepto débil de la representación, para elaborar una teoría cabal acerca del funcionamiento de la percepción, del lenguaje o de la cognición. Según este sentido fuerte, la cognición se debe explicar mediante la hipótesis de que un sistema actúa sobre la base de representaciones internas, lo que conlleva fuertes implicaciones epistemológicas y ontológicas. Estas implicaciones son básicamente dos:

1. Da por sentado que el mundo está predefinido (es decir, que los rasgos del mundo están predefinidos antes de toda actividad cognitiva).

2. Plantea además la existencia de representaciones mentales dentro del sistema cognitivo, tales como imágenes, símbolos o patrones subsimbólicos de actividad distribuidos por una red, que permiten explicar la relación entre la actividad cognitiva y el mundo predefinido. En opinión de Varela, de estas implicaciones se deduce una teoría cabal que establece que: a) el mundo está predefinido; b) nuestra cognición aprehende este mundo, aunque sea en forma parcial; y c) la forma en que conocemos este mundo predefinido consiste en representar sus rasgos y luego actuar sobre la base de estas representaciones.

Su crítica, está encaminada justamente en demoler este segundo sentido fuerte de la representación. Su antirrepresentacionalismo discute la idea de un mundo pre-existente.

⁴⁹ *Ibid.*, p. 178.

⁵⁰ *Ibid.*, pp. 96 - 98

Contrario a esta visión, Varela postula que las aptitudes cognitivas están entrelazadas con una historia vivida *tal como si fuera una senda que no existe pero que se hace al andar* (recordando a Machado). De este modo, la cognición no puede ser ya entendida como un dispositivo que resuelve problemas mediante el uso de representaciones, sino que consiste en una *acción efectiva: una historia del acoplamiento estructural que hace emerger un mundo*.

La propuesta de Varela intenta buscar una vía media entre dos modos de entender la cognición: el realismo cognitivo (que entiende la cognición como recuperación de un mundo externo pre-dado) y el idealismo (que entiende la cognición como proyección de un mundo interno pre-dado). Su idea, consiste en estudiar la cognición *ni como recuperación, ni como proyección, sino como acción corporizada*.⁵¹ Con los términos *acción corporizada*, Varela pretende enfatizar lo siguiente:

1. Al hablar de corporizada desea subrayar dos elementos: primero, que la cognición depende de las experiencias originadas en la posesión de un cuerpo con diversas actitudes sensorio-motrices y segundo, que estas aptitudes sensorio-motrices están encastradas en un contexto biológico, psicológico y cultural más amplio.⁵²

2. Al usar el término acción desea enfatizar que los procesos sensorio-motrices, tales como: la percepción y la acción, son fundamentalmente inseparables de la cognición vivida.

Teniendo presente el *enfoque enactivo* de Varela recién expuesto, podemos ver que en este enfoque *la representación* no juega ningún rol en la explicación de la conducta inteligente. Como ya hemos visto Varela pretende estudiar la cognición como *acción corporizada*, considerando que el cuerpo y los elementos sensorio-motrices son inseparables de ésta. Su crítica al sentido fuerte de la representación, vale decir, a la afirmación de que un sistema cognitivo actúa sobre la base de representaciones internas, nos muestra claramente su postura *antirrepresentacionista*. Varela discute la idea de un mundo pre-existente o pre-dado ante de toda actividad cognitiva. En su opinión, la cognición consiste en una *historia del acoplamiento estructural que hace emerger el mundo*. En consecuencia, *la representación mental* (al igual que en los sistemas dinámicos) resulta una noción innecesaria para entender la cognición y por tanto no juega ningún rol en este enfoque.

4. LA POSTURA ECUMÉNICA DE ANDY CLARK

¿No sería más fructífero pensar en los cerebros como controladores de la actividad corpórea? Este pequeño cambio de perspectiva tiene grandes consecuencias para nuestra manera de construir una ciencia de la mente. De hecho, exige una reforma

⁵¹ Varela, *De cuerpo Presente*, p. 202

⁵² *Ibíd.*, p. 203

radical de nuestra manera de concebir la conducta inteligente. Nos exige abandonar la idea (común desde Descartes) de lo mental como una esfera distinta del ámbito del cuerpo, abandonar la idea de unas limpias líneas divisorias entre percepción, cognición y acción; abandonar la idea de un centro ejecutivo donde el cerebro lleva a cabo el razonamiento de alto nivel; y, sobre todo, abandonar los métodos de investigación que separan artificialmente el pensamiento de la actuación corpórea.⁵³

Como el mismo Clark enuncia en su libro *Estar ahí*⁵⁴, su propuesta (frente a tres décadas de ciencia cognitiva clásica), intenta rescatar una imagen de la cognición que haga justicia tanto a los postulados del conexionismo, como a los del cognitivismo clásico y que a su vez implique una renovación de algunas de sus ideas centrales. La tesis central postulada por Clark en este libro, señala que la interacción *cerebro, cuerpo y mundo* es constitutiva de la mente, adheriéndose en este sentido a la hipótesis de la cognición corporalizada, pero guardando grandes diferencias con la postura antirepresentacionista de Varela.

Estar ahí, se divide en tres importantes secciones: *Excursión por la mente, Explicando la mente extendida y Más allá*, proporcionando un interesante material de análisis, que da cuenta de las investigaciones más recientes en ciencia cognitiva. En la primera parte de este libro *Excursión por la Mente*, Clark sintetiza varios campos de investigación que se presentan como aparentemente dispares, pero que finalmente se relacionan entregando al lector un cuadro coherente de los estudios actuales sobre la mente y la cognición. Esta sección comienza presentando la investigación de los agentes autónomos (Brooks, 1991), posteriormente considera los estudios del desarrollo cognitivo que aplican la teoría de los sistemas dinámicos (Thelen y Smith, 1994), luego la perspectiva de las redes conexionistas (discutidas en Clark, 1989, 1993), teniendo presente además las ideas de la *acción epistémica* (Kirsh y Maglio, 1994), la psicología de Vygotsky (1986), los estudios sobre la percepción de Gibson (de 1979) e incluso algunos de los planteamientos de la fenomenología (Heidegger, 1927; Merleau-Ponty, 1945).

Cada uno de estos enfoques ha intentado ser presentado por sus autores como *un cambio de paradigma* en su área de competencia; sin embargo Clark pretende demostrar que para rivalizar con el paradigma clásico de las CC, se necesita de la integración y complementación de todas estas perspectivas para dar cuenta del funcionamiento adecuado de la mente y para lograr integrar el cuerpo, el cerebro y el ambiente en la descripción de ésta. La síntesis propuesta por Clark, consiste justamente en la necesidad de poner el foco en la *encarnación*, es decir, en la inclusión del cuerpo en la cognición.

Una de las investigaciones que inspira y apoya a su vez esta tesis, corresponde como ya hemos señalado, a los trabajos realizados por Rodney Brooks (del MIT Mobile Laboratory) sobre los agentes autónomos. Estos robots móviles creados por Brooks (*mobile robots* o *mobots*) son criaturas simples capaces de funcionar en entornos reales,

⁵³ Clark, *Estar ahí*, p. 33

⁵⁴ Para el desarrollo de esta sección, se consideran los planteamientos de Clark en su libro "*Estar ahí*".

desordenados e imprevisibles, lo que significa que son capaces de realizar acciones robustas en un ambiente dinámico. Por ejemplo el robot *Attila*, que pesa un poco más de kilo y medio se caracteriza por usar *minicerebros* especializados para controlar un conjunto de conductas locales, que al estar unidas permiten que el robots se desplace con gran habilidad moviendo cada pata por separado, logrando detectar las diferencias de pendiente del terreno como también los objetos cercanos (gracias al empleo de sensores infrarrojos). De este modo, *Attila* es capaz de desplazarse por terrenos abruptos y hasta es capaz de ponerse de pie cuando se cae de espaldas. En opinión de Brooks, *Attila* encarna algo parecido a la inteligencia de un insecto.

Lo interesante de los Robots como *Attila* es que poseen dispositivos que son cuasi independientes y que cada uno tiene un acoplamiento autónomo. Estos dispositivos se excitan e inhiben uno a otro, sin pasar ningún mensaje complejo, permitiendo así una inteligencia simple que emerge sin necesidad de un ejecutivo central que controle la acción. En opinión de Clark, gran parte del comportamiento de los seres humanos y de otros animales, es como el de los "*mobots*" a este respecto.

Para ilustrar mejor estas ideas, Clark cita los estudios realizados por Thelen y Smith en su libro *A Dynamic Systems Approach to the Development of Cognition and Action* (1994). Estos autores critican la concepción del desarrollo como un progreso lineal (concepción predominante en los años treinta y cuarenta) dado que consideran que ésta suele incurrir en un error, al tomar un fenómeno complejo (como por ejemplo caminar) y buscar sólo un factor determinante para explicarlo. Clark postula en cambio, que las capacidades de actuar hábilmente (como por ejemplo caminar) resultan posible gracias a interacciones entre varios factores tales como: el crecimiento de la persona, los factores ambientales, el aprendizaje o bien la maduración del cerebro.

En opinión de Clark, los fenómenos complejos presentan una gran autoorganización entre varios factores que interactúan, sin la existencia de un sólo factor controlador. Así pues, al igual que el comportamiento de los *mobots*, el desarrollo humano ocurriría sin el control de un ejecutivo central que decide qué hacer y cuando; lo que nos En ambos casos, el comportamiento de la persona o de los *mobots* se organiza, *emergiendo* (en un cierto plazo) de la interacción de varios componentes.

Teniendo presente ambas investigaciones (más los estudios de la percepción de Gibson, que no detallaré en este trabajo), Clark concluye que para explicar la cognición se debe explotar de varias maneras la acción real, reduciendo con ello la carga computacional que el enfoque del cognitivismo clásico, le ha asignado. Enfatiza además, que esta nueva perspectiva conduce a un profundo cambio en nuestra comprensión de la mente y de la cognición, cambio caracterizado *como la transición desde los modelos de la representación como reflejo o codificación, hasta los modelos de la representación como control.*⁵⁵ Agrega además que, *La idea es que el cerebro no se debería contemplar como la sede de unas descripciones internas de situaciones externas; más bien se debería considerar la sede de estructuras internas que operan en el mundo mediante su papel en la determinación de acciones.*⁵⁶

A pesar de estar de acuerdo con este cambio de mirada, Clark no se apresura en

⁵⁵ Clark, p. 88

acoger del todo las diferencias que se establecen desde la revolución de la *nueva robótica* hacia la imagen clásica de la mente. Estas diferencias radican, en el rechazo presentado por las investigaciones en robótica (ver Brooks, 1991) hacia la imagen de un planificador central, como poseedor de toda la información disponible en cualquier parte del sistema (cuya función consiste en descubrir las posibles secuencias conductuales para lograr satisfacer metas concretas). Como bien se mostró con *Attila*, la acción de este *mobot* se destaca por la falta de un planificador central. La arquitectura por subsunción desarrollada por Rodney Brooks sitúa múltiples dispositivos cuasi independientes, de los cuales cada uno de ellos constituye una vía autosuficiente que enlaza la entrada sensorial con la acción. Las conductas de estos sistemas, en consecuencia, no están mediatizadas por ninguna base integrada de conocimientos (modelo detallado del mundo) que represente el estado actual del entorno global. En consecuencia, esta arquitectura muestra que *estos seres* (como *Attila*), no dependen de un procesador central (planificador de datos); sino que sólo poseen una colección de conductas que compiten entre sí y que están mediadas por las entradas del entorno. Estas criaturas no poseen *representaciones*, ni *mapas internos* y logran alcanzar un éxito adaptativo sin emplear *modelos detallados del mundo*.

Considerando estas radicales diferencias entre los planteamientos de la nueva robótica y el paradigma clásico, Clark señala que:

Uno de los mayores peligros que acechan a toda propuesta revolucionaria en la ciencia es que se puede descartar una parte demasiado grande de esta 'visión antigua': es decir, se puede tirar mucha fruta fresca junto con la pocha. Este mismo peligro existe, creo yo, en el rechazo de la nueva robótica frente a las representaciones, los modelos y los mapas internos. Si sólo se interpreta como un aviso para prestar atención a los costes de los modelos centralizados, integrados y simbólicos, la crítica es oportuna e importante. Pero si se interpreta como un rechazo sistemático a las economías internas cuyas complejidades incluyen múltiples representaciones centradas en la acción y múltiples modelos parciales del mundo, eso sería un error al menos por dos razones.⁵⁷

Las razones que da Clark son las siguientes. La primera se debe a que resulta indudable que, en ocasiones, el cerebro humano integra múltiples fuentes de información (como por ejemplo combinar modalidades utilizando el tacto, la vista y el sonido), siendo en estos casos similar en su función a un planificador central. La segunda enfatiza que no siempre constituye un embudo costoso de tiempo (o un *embotellamiento representacional* como decía Brooks) la presencia de modelos internos que intervienen en la entrada y la salida de la información. Por ejemplo dice Clark, si consideramos la tarea de alcanzar una taza, podemos utilizar la *emulación motriz* (muy empleada en los sistemas industriales de control) que consiste en lograr predecir como debe ser la retroalimentación *propioceptiva* (es decir, el sentido interno que nos dice cómo se sitúa nuestro brazo).

Esta *emulación motriz*, es una especie de modelo interno, que no se traduce en un embudo que obstaculice el éxito en tiempo real, sino por el contrario lo facilitan al ofrecer

⁵⁶ Ibíd anterior.

⁵⁷ Clark, p. 61

un tipo de *retroalimentación virtual* que supera la retroalimentación procedente de las periféricas sensoriales reales. Clark señala que esto corresponde a un modelo parcial dedicado a una clase de tarea específica y que, en consecuencia, es compatible con el escepticismo de la nueva robótica acerca de los modelos del mundo detallados y descentralizados y con su insistencia en el éxito conductual en tiempo real.

Teniendo presente lo expuesto, podemos ver que la postura de Clark puede caracterizarse como una **postura ecuménica**, dado que pretende lograr una cierta unidad, frente a los diversos enfoques desarrollados en las ciencias cognitivas, sin desechar por completo ninguno de ellos. Como vimos, aún cuando Clark tiene presente los trabajos presentados tanto en el ámbito de la robótica, como de los sistemas dinámicos del desarrollo cognitivo (críticos del enfoque cognitivista clásico); su propuesta se presenta como mediadora justa, desarrollando una crítica equilibrada y pertinente. La siguiente cita de Clark es enfática en este sentido:

En concreto, no deberíamos tener demasiada prisa en rechazar los aparatos explicativos más tradicionales de la computación y la representación. Las mentes pueden estar esencialmente corporalizadas y embebidas y, aun así, depender fundamentalmente de unos cerebros que computan y representan. Sin embargo para hacer que está postura ecuménica cuaje, necesitamos enfrentarnos específicamente a algunos retos directos relacionados con las nociones mismas de computación y representación (¿tienen estas nociones unas definiciones no triviales con el nuevo marco?)⁵⁸

Esta *postura ecuménica* de Clark, enfatiza dos dimensiones fundamentales presentes en el **paradigma simbólico**: la representacional y la computacional; y señala que estas nociones requieren adecuarse a las nuevas visiones de la cognición. Su perspectiva intenta así, conciliar la dimensión representacional y computacional del paradigma clásico, con su propia visión de la cognición corporalizada; lo que marca una diferencia sustancial con la perspectiva desarrollada por Varela. Clark reformula la noción de computación y de representación propia del paradigma clásico, evitando así, desarrollar una postura de ruptura frente a este paradigma.

4.1. REFINANDO LA NOCIÓN DE REPRESENTACIÓN

En *Estar Ahí*, Clark es enfático al señalar que la noción de representación interna sigue desempeñando un papel importante en la actualidad, pero hace ver que dicha noción está experimentando alteraciones fundamentales en la neurociencia contemporánea y que en consecuencia resulta necesario *refinar esta noción de representación*. Su propuesta está encaminada, en consecuencia, a modificar la noción de representación (y de computación), para que de este modo, puedan estas nociones adecuarse dentro de los nuevos márgenes de la CC.

Clark considera, que los cambios acaecidos en la noción de *representación interna*, constituyen a su vez, un cambio en la imagen del *cerebro representador*, que lo aleja de la vieja idea que lo concebía como un único código simbólico interno, como por ejemplo

⁵⁸ Clark, *Estar Ahí*, p. 193

en la imagen *representacional y computacional de la mente* desarrollada por Fodor⁵⁹ (ver capítulo 1, p. 9). La nueva imagen del *cerebro representador*, lo presenta como descentralizado y considera a la representación de modo multidimensional y distribuido.

En la segunda parte de *Estar ahí: Explicar la mente extendida*, específicamente en el apartado 8.2. *Qué es eso que llamamos representación*, Clark hace ver que comúnmente los científicos cognitivos suelen hablar de los cerebros como si albergaran *representaciones internas*; siendo esta una idea básica y un terreno común, para campos tan opuestos como el del conexionismo y el de la inteligencia artificial clásica. Por esta razón, Clark realiza un análisis comparativo de la noción de *representación interna*, presente en tres perspectivas disímiles: en la clasicista (cognitivism clásico), en la conexionista y en la tesis de la *cognición corpórea radical*.

La perspectiva clasicista, postula una representación interna simbólicamente sólida, en la cual existen contenidos mentales que son almacenados como cadenas de símbolos que pueden ser leídas, copiadas y trasladadas por algún tipo de unidad central de procesamiento interno.⁶⁰ Los conexionistas en cambio, creen en un estilo más implícito de representación interna, que reemplaza las cadenas de símbolos sólidas y manipulables del enfoque clasicista por vectores numéricos complejos y operaciones básicas de reconocimiento y transformaciones de patrones. Clark es enfático en señalar que tanto las representaciones explícitas y sólidas del enfoque clasicista como las representaciones conexionistas distribuidas y vectoriales se consideran propiamente representaciones internas.

La tesis de la *cognición corpórea radical* (Maturana y Varela, 1987; Varela y otros, 1991), apoyada en la idea de sistemas dinámicos postula en cambio, que:

Toda perspectiva estructurada, simbólica, representacional y computacional de la cognición es errónea. Que la mejor manera de estudiar la cognición corpórea es mediante ideas y esquemas explicativos no computacionales y no representacionales.⁶¹

Esta tesis, de la *cognición corpórea radical*, se resiste a incluir el supuesto computacional y representacional de las CC. Clark, rechaza el enfoque de esta última propuesta, señalando que tanto radicalismo es injustificado y además contraproducente, dado que tiende a fomentar la competición cuando lo que se necesita para lograr progreso real en las ciencias cognitivas es cooperación. Su propuesta entonces, pretende *refinar la noción de representación* sin anularla del todo y reconociendo la importancia del mundo (ambiente) en la cognición humana. Clark considera que si hay representación, pero que estas son locales y orientadas hacia la acción:

Una cosa está cada vez más clara. En la medida en que el cerebro biológicamente opere con algo que se pueda describir adecuadamente como "representación interna", una gran parte de esas representaciones serán locales y estarán

⁵⁹ véase Fodor, 1975 y Fodor, 1986.

⁶⁰ Clark, p. 194

⁶¹ Clark pp. 198-199

orientadas hacia la acción en vez de ser objetivas e independientes de la acción.

62

Para ejemplificar esta propuesta de las *representaciones orientadas hacia la acción*, Clark presenta el trabajo de Maja Mataric (del MIT Artificial Intelligence Laboratory). Mataric, desarrolla un modelo inspirado en la neurología sobre cómo se orientan las ratas en su entorno y lo implementa en un robot móvil (una rata móvil). Esta *rata robot* dispone de sensores de sonido y de una brújula, gracias a los cuales consigue el éxito en tiempo real; empleando una *arquitectura por subsunción*⁶³, es decir utilizando un conjunto de *capas* cuasi independientes. Cada una de estas *capas* constituye una ruta de procesamiento completa desde la entrada hasta la salida y sólo se comunican entre sí, pasándose señales muy simples.

A partir de este ejemplo de la rata robot, Clark enfatiza nuevamente las diferencias de su nueva propuesta con la del enfoque clásico de la mente. Señala que un enfoque más clásico postularía la presencia simultánea de un mapa almacenado y de un módulo de control central que accede a ese mapa y lo emplea para planificar los movimientos. En cambio, el robot de Mataric, no emplea ningún dispositivo de razonamiento fuera del mapa mismo (el mapa corresponde a su propio usuario). Para Clark, *este robot es un ejemplo perfecto de la noción de representaciones orientadas hacia la acción: representaciones que describen aspectos del mundo y al mismo tiempo prescriben acciones posibles, y que se encuentran entre las puras estructuras de control y las representaciones pasivas de la realidad externa.*⁶⁴

En consecuencia, las representaciones internas que emplea la mente para orientar acciones se entienden de mejor modo, como *estructuras de control específicas de la acción y del contexto* y no como *recapitulaciones pasivas de la realidad externa* (como postulaba el enfoque cognitivista clásico). En opinión de Clark, los modelos internos detallados y neutrales respecto a la acción que proporcionaban el ámbito para un pensamiento incorpóreo y centralizado (postulado por el enfoque clásico), aparecen en esta propuesta como un lujo lento, costoso y difícil de mantener.

Resumiendo podemos decir, que Clark rechaza la tesis de la *cognición corpórea radical* y que en consecuencia no estaría de acuerdo con los planteamientos de Varela, dado que éste se resiste a incluir el supuesto representacional de las CC, en su explicación de la cognición. La propuesta de Clark en cambio, intenta rescatar el supuesto representacional, presente tanto en el enfoque del cognitivismo como en el del conexionismo, pero enfatizando la necesidad de refinar la noción de representación. Clark cuestiona el carácter neutral que le asignan ambos enfoques a la representación mental. La hipótesis que Clark defiende es, que en términos de cognición, hay representaciones internas pero que éstas no se agotan en explicaciones de estados internos, sino que involucran el entorno mismo y las relaciones dinámicas entre éste y el

⁶² Clark, p. 200

⁶³ Arquitectura descrita por Brooks (1991). Ver Clark, *Estar Ahí*, p. 88, para una descripción completa de este ejemplo.

⁶⁴ Clark, *Estar ah*, p. 90

agente.

4.2. MENTE Y MUNDO: LA FRONTERA ESCURRIDIZA

Entonces, ¿dónde se encuentra la mente? ¿Se encuentra realmente 'dentro de la cabeza' o quizás ahora se ha extendido, de una manera un tanto libertina, por el mundo?⁶⁵

En la primera parte de *Estar ahí*, específicamente en el apartado tres de esta sección, Clark se propone mostrar una imagen integrada de la mente y el mundo. En su opinión, la tarea fundamental que tiene por delante una ciencia del pensamiento corpóreo, consiste en poder comprender la compleja interacción existente, entre nuestros recursos neuronales por un lado y los soportes externos del mundo por otro (tales como, artefactos culturales y lingüísticos que estructuran la vida moderna, incluyendo mapas, textos y planes escritos).

Para mostrar de qué manera se puede lograr esta integración, Clark expone la irrupción y desarrollo de los modelos conexionistas (de redes neurales) con el objeto de evaluar sus ventajas e inconvenientes. En primer lugar señala, que con el surgimiento de estos modelos en los años 80, el campo de investigación de la inteligencia artificial (IA) se fue transformando. Desde una imagen general de la inteligencia humana, postulada por la IA en sus comienzos, como manipulación de símbolos siguiendo reglas (como puede verse en CYC⁶⁶), se fue avanzando hacia un modelo explicativo que fuera biológicamente más atractivo (modelo de redes neurales).

Los modelos de redes neurales (como su nombre lo indica), están inspirado en una reflexión sobre la estructura del cerebro. Como sabemos, el cerebro se compone de muchas unidades simples de procesamiento (las neuronas) que se encuentran enlazadas en paralelo por una gran masa de cables y empalmes (axones y sinapsis). Estas unidades individuales (las neuronas) son sensibles sólo a informaciones locales (esto significa que, cada neurona *escucha*, lo que sus vecinas le cuentan. Lo interesante es que a partir de esta masa de conexiones en paralelo, de procesadores simples e interacciones locales, surge una asombrosa habilidad computacional y de resolución de problemas.

En opinión de Clark, uno de los aportes significativos de las investigaciones de redes neurales, consistió en expandir nuestra visión sobre las maneras en que un sistema físico como el cerebro puede codificar y explotar información y conocimiento. Clark señala además, que los modelos de redes neurales presentan una interesante combinación de ventajas e inconvenientes: dentro de las ventajas se destaca su rapidez (debida al funcionamiento en paralelo), la capacidad de tolerar datos imperfectos e incompletos (capacidad de recrear patrones completos a partir de señales parciales) y su resistencia a daños locales (debida al empleo de múltiples recursos, a nivel de unidad, para codificar cada pauta). Además, son modelos que sobresalen en tareas que suponen la integración

⁶⁵ *Ibid.*, p. 110

⁶⁶ Enciclopedia electrónica, que intenta hacer explícito todo lo que sabemos de nuestro mundo. (Ver Clark, p. 41)

simultánea de muchas pequeñas señales o elementos de información, capacidad que es esencial para el reconocimiento perceptivo y el control motor en tiempo real.

Dentro de las desventajas Clark señala que estos modelos no son intrínsecamente adecuados para la resolución de problemas en forma secuencial y gradual requerida por el pensamiento lógico y la planificación. Por tanto, considera que estos sistemas se pueden caracterizar por *ser buenos al fútbol, pero malos en lógica, por ser rápidos, pero limitados*.

Teniendo presente estas ventajas y desventajas, Clark considera que, *si nuestro objetivo es modelar la cognición humana, se deben favorecer aquellos fundamentos computacionales que produzcan una pauta de puntos fuertes y débiles parecida a la nuestra. Y en general somos mejores al fútbol que en lógica*. Pero de todos modos, también somos capaces, por lo menos en ocasiones, de planificar a largo plazo y de llevar a cabo razonamientos secuenciales.

Clark considera que los trabajos realizados sobre las redes neuronales artificiales, proporcionaron un valioso antídoto contra los que se ha denominado *imagen de archivador* de la mente, imagen que describe a la mente como un almacén de símbolos pasivos, que esperan ser recuperados y manipulados por una especie de unidad central de procesamiento neuronal; y que condujo además a una visión de la mente que la considera como el órgano de una razón intemporal e incorpórea.

Para rebatir esta *imagen de archivador* de la mente, Clark introduce la idea de *andamiaje externo* y resume un ejemplo sencillo dado por Rumelhart en *Parallel Distributed Processing* (la biblia del conexionismo), para explicar esta idea. El ejemplo toma el caso de las multiplicaciones largas, y señala que la mayoría de nosotros podemos aprender a reconocer de un solo vistazo la respuesta a multiplicaciones simples como $7 * 7 = 49$, gracias a un dispositivo básico incorporado para el reconocimiento de un patrón simple. No ocurre lo mismo sin embargo, frente a multiplicaciones largas como por ejemplo frente a $7.222 * 9.422$, que plantean un problema diferente. Para llevar a cabo esta multiplicación, generalmente se echa mano a un lápiz y papel o bien a una calculadora, para conseguir reducir un problema complejo a una sucesión de problemas más simples. En consecuencia, en opinión de Clark, se utiliza el medio, externo (un papel u otro) para almacenar los resultados de estos problemas simples, para así, finalmente llegar a una solución mediante *una serie interrelacionada de compleciones de patrones simples y almacenamientos externos*.⁶⁷

Según Clark, lo que ocurre en estos casos consiste en aprender a manipular un modelo mental de la misma manera en que se manipula el mundo real. Señala además, que este tipo de manipulación interna de símbolos es fundamentalmente distinta de la concepción clásica de los símbolos internos, dado que no dice nada **del sustrato computacional** de estas imágenes. En su opinión, podemos simular mentalmente el entorno exterior, lo que implica que a veces interiorizamos competencias cognitivas. Sin embargo esto no significa que por ello, dejen de estar enraizadas en manipulaciones del mundo exterior.

⁶⁷ Clark, p. 102

Sintetizando, vemos que Clark cuestiona la clásica separación (limpia y tajante) entre datos y procesos, entre estructuras de símbolos y procesadores centrales postulada por la IA en sus comienzos. Para él, la visión de la *mente archivador* no ha reflejado nada más que la separación entre el agente y un andamiaje externo de ideas que persisten sobre el papel (en archivadores o en soportes electrónicos). En palabras de Clark, tenemos que,

El mensaje es claro: controlamos nuestro entorno físico y espacial de maneras que alteran fundamentalmente las tareas de procesamiento de información a las que se enfrenta nuestro cerebro.... Si este enfoque es valioso se debe a la diferencia entre la naturaleza de los cómputos que se pueden realizar parasitando recursos del entorno. Pero, como veremos, este parasitismo pone en duda los mismos límites tradicionales que existen entre mente y mundo.⁶⁸

Clark, aplica las consideraciones anteriores al ámbito de la planificación y resolución de problemas. Cuestiona la visión incorpórea clásica de la planificación (Agre y Chapman, 1990) que sostiene que las secuencias complejas de acciones están determinadas por una versión interiorizada de un conjunto de instrucciones. Un plan, según esta visión, especifica una sucesión completa de acciones que bastaría realizar para lograr una meta (por ejemplo dar una lista completa de instrucciones para cocinar un huevo).

En oposición a la visión incorpórea de la planificación, Clark señala que cuando observamos detenidamente las conductas de planificación de agentes reales, vemos que existe una interacción bastante más compleja entre el plan y el entorno de apoyo, que no es considerada en esta visión. Esta interacción, conduce en algunos casos a replantear sobre la marcha la consecución de algunos subobjetivos concretos, lo que señalaría que el plan resulta ser mucho más parcial y dependiente de las propiedades del entorno local.

En apoyo de las ideas anteriores, Clark presenta algunos ejemplos proporcionados por David Kirsh (1995), sobre el empleo inteligente del espacio físico. Kirsh observa que los estudios típicos de IA tratan la planificación como un fenómeno muy incorpóreo, dado que desatienden nuestra manera de utilizar las propiedades espaciales reales del ámbito de trabajo, permitiendo con ello simplificar el cómputo interno. Por el contrario, en la vida diaria abundan ejemplos, en los cuales se ve claramente nuestra manera inteligente de utilizar el espacio físico. Consideremos por ejemplo, la resolución de un puzzle: para ello es necesario agrupar las piezas similares, para así poder compararlas visualmente y de manera detallada (agrupar por ejemplo todas las que sean de un color determinado y posean borde). O bien, cuando reparamos un alternador y al desmontarlo colocamos las piezas en línea o formando grupos para que así la tarea de montarlo nuevamente sea más fácil.

En opinión de Clark, los ejemplos proporcionados por Kirsh hablan a favor de un modelo de recursos internos distintos al clásico, de capacidades operativas que en la vida real sólo emergen a partir de interacciones entre el cerebro y el mundo. Señalan en definitiva, que la resolución sistemática de problemas de los cerebros biológicos no sigue realmente el modelo del plan como programa. Por el contrario, la inteligencia humana consiste en la capacidad de estructurar nuestro entorno y de operar activamente sobre él,

⁶⁸ Clark, p. 104

simplificando así nuestras tareas de resolución de problemas. El flujo de pensamientos y el éxito adaptativo de la razón dependen de las interacciones reiteradas y cruciales con recursos externos.

5. EL ROL DE LA REPRESENTACIÓN EN LAS TEORÍAS ALTERNATIVAS

Existen como hemos visto diferentes enfoques en las teorías alternativas. Cada uno de ellos asigna un rol distintos a la representación mental. En algunos casos, como en los sistemas dinámicos y en el antirepresentacionalismo de Varela, es posible detectar un escepticismo respecto de la relevancia cognitiva de las propiedades mentales. En este sentido, se observa que en algunos de estos enfoques se da pie a un reduccionismo extremo, donde las nociones de representación y computación (como elementos fundamentales para explicar la cognición) son eliminadas.

En consecuencia, resulta necesario precisar en cada uno de estos enfoques cuál es el rol que se le asigna a la representación mental, si es que se le asigna alguno.

5.1. REPRESENTACIÓN Y SISTEMAS DINÁMICOS

Teniendo presente, las características centrales, descritas por Van Gelder de los sistemas dinámicos y en particular el rasgo distintivo de estos sistemas, dado por su noción de **acoplamiento** (causalidad recíproca), podemos señalar que este enfoque conlleva algunas dificultades para intentar modelar un sistema cognitivo, como el humano. Si tenemos presente por ejemplo, los difusos límites que se establecen entre el sistema cognitivo y el entorno, como también la complejidad del sistema mismo (las millones de conexiones neuronales) y los parámetros que influyen en este sistema (todos los factores ambientales y todo el orbe que rodea al sistema), vemos que no resulta claro saber qué condiciones iniciales se deben tener presente a la hora de querer entender la cognición. ¿Qué tipo de respuesta podemos esperar por parte de un sistema con estas características?, ¿sería posible, mediante un cálculo (a través de sus ecuaciones diferenciales), pronosticar un determinado comportamiento del sistema? La amplitud de este sistema, hace que su trayectoria no sea interpretable en términos específicamente cognitivos, lo que señala que no sería un modelo adecuado para dar cuenta del sistema cognitivo humano. Resulta difícil en consecuencia, poder entender qué función cumple la representación en este enfoque, puesto que al existir un acoplamiento entre sistema nervioso, cuerpo y mundo, se hace extremadamente difícil identificar ni siquiera el rol que le cabría a las representaciones.

Si tenemos presente las críticas realizadas por Eliasmith⁶⁹ a los sistemas dinámicos, podemos además señalar una objeción metodológica a este enfoque. Los dinamicistas

⁶⁹ Eliasmith. *The Third Contender: A Critical Examination of the Dynamicist Theory of Cognition*.

tienen dificultades para explicar el comportamiento cognitivo complejo, (alta dimensionalidad), desde un enfoque de baja dimensionalidad. Según los dinamicistas, la explicación de toda conducta multidimensional (compleja) es entendida como una interacción de baja dimensionalidad de los agentes. La dimensionalidad en la teoría dinámica, corresponde a la cantidad de parámetros, vale decir, de variables, contemplados en las ecuaciones diferenciales, que representan la trayectoria de los agentes cognitivos a través de los diferentes estados.

El problema que ve Eliasmith en la teoría de sistemas dinámicos, radica justamente en poder determinar, cuántos parámetros deben ser considerados en estas ecuaciones para lograr un modelamiento adecuado de la cognición compleja. Lo anterior resulta aún más complejo si consideramos que los sistemas dinámicos son sistemas naturales complejos, cuyas variables se encuentran íntimamente relacionadas (acopladas). ¿Cómo aislar, entonces, una variable del sistema?, ¿cómo escoger aquellas variables que sean relevantes de aquellas que no lo son? Frente a este problema, la teoría dinamicista postula la idea de considerar *parámetros colectivos*, vale decir, agrupar varios comportamientos en un solo parámetro. Mediante estos parámetros colectivos, se lograría modelar en baja dimensionalidad el comportamiento cognitivo complejo, según los dinamicistas.

Sin embargo, la idea de *parámetros colectivos*, presenta nuevamente otras dificultades: No permite identificar los mecanismos existentes que determinan el comportamiento del sistema cognitivo, dado que no existen unidades simples que permitan describirlos. Además, siendo esto así, tampoco resulta posible asignar un significado unívoco a estos parámetros ni hacer una interpretación adecuada de ellos, dado que sólo son una colección de varios comportamientos. ¿Cómo corroborar entonces el modelo, si no es posible verificar ningún tipo de hipótesis a partir de la observación del sistema? En definitiva ¿cómo sería posible realizar una explicación de lo simple a lo complejo, si los parámetros colectivos se evidencian como algo complejo e inexplicado? Considerando estos problemas, resulta prácticamente imposible realizar una explicación desde lo simple a lo complejo, como postula el sistema dinámico.

Teniendo presente las críticas y cuestionamientos de Eliasmith a los sistemas dinámicos, podemos decir que efectivamente esta teoría está imposibilitada para individuar sus elementos o componentes, dado que la naturaleza misma de este sistema, se define en función de su carácter global. En consecuencia, como ya señalamos anteriormente, se hace extremadamente difícil poder identificar en los sistemas dinámicos qué rol le cabría a las representaciones mentales. Más aún, si tenemos presente la comparación que hace Van Gelder entre los gobernadores *dinámico* y *computacional*, podemos ver que el gobernador dinámico opera con cantidades y no con *representaciones* como el computacional. En el gobernador dinámico no habría manipulación de representaciones ni de computaciones, sino sólo relaciones de acoplamiento. La representación mental no juega ningún rol causal o etiológico en los sistemas dinámicos.

5.2. COMPARACIÓN CRÍTICA DE LOS ENFOQUES DE VARELA Y

CLARK

Si tenemos presente, las dos perspectivas expuestas en este capítulo (Varela y Clark) enmarcadas dentro del enfoque de la cognición corporalizada, podemos señalar en primer lugar que en ellas existen importantes diferencias. Como hemos visto, la propuesta de Varela, se enmarca dentro de la *tesis radical* de esta hipótesis, mientras que la perspectiva de Clark cuestiona el carácter antirepresentacionista postulado por Varela. Clark considera que el énfasis e interés de la propuesta de Varela es muy distinto al suyo, puesto que las reflexiones llevadas a cabo por él, son empleadas como pruebas para rebatir las perspectivas realistas y objetivistas del mundo. Clark discrepa de la conclusión escéptica, que se desprende de la perspectiva de Varela respecto de los términos de *representación* y *computación* y dice enfáticamente:

Yo evito deliberadamente esta postura porque conlleva el riesgo de oscurecer el valor científico de un enfoque corpóreo y embebido, vinculándolo a la idea problemática de que los objetos no son independientes de la mente. Lo que yo afirmo es, simplemente, que los aspectos de la estructura de la realidad que representan los cerebros biológicos se suelen orientar estrictamente hacia unas necesidades y capacidades sensorio-motrices específicas. Por tanto, el objetivo de gran parte de la crítica actual no es la idea de que los cerebro representen aspectos de un mundo independientemente real, sino más bien la idea de que estas representaciones sean neutrales respecto a la acción y que, en consecuencia, requieran un importante esfuerzo computacional adicional para producir respuestas inteligentes.⁷⁰

El enfoque de Varela se opone a la idea de que la cognición es fundamentalmente representación y, en consecuencia, la noción de representación no tiene una relevancia explicativa para la vida de un agente cognitivo. La propuesta de Clark en cambio, es mucho más cordial con los análisis representacionales, basados en el procesamiento de información, dado que pretendereconceptualizar parcialmente las ideas sobre los contenidos y los formatos de los procesos internos, sin que ello implique rechazar las ideas mismas de procesamiento de información y de representaciones internas.

Vemos así que en los planteamientos de Varela, se destaca la necesidad de incorporar la noción de *sentido común* para entender la cognición, cuestionando con ello las visiones objetivistas en CC. Clark en cambio, postula un mínimo de representacionismo (e incluso de computacionalismo), implicando con ello una modificación del modo en que han sido entendidas estas nociones en los enfoques anteriores. Además, en el proyecto de Clark se destaca una búsqueda de integración de diferentes investigaciones en ciencia cognitiva, tales como las investigaciones sobre robótica en mundo real (teoría de los agentes autónomos), tratando de mostrar cómo encajan las ideas y los análisis que emergen de estas investigaciones, con algunas de las antiguas nociones de representación y computación (presentes en el paradigma clásico), junto con otras propuestas psicológicas, psicofísicas y evolutivas. En consecuencia, el enfoque corporalizado de Clark, no implica un abandono total de la noción representacional y computacional, sino un ajuste de estas nociones dentro de este nuevo

⁷⁰ Clark, p. 227

marco, que incluye el cuerpo y el ambiente en la comprensión de la cognición. En palabras de Clark,

El énfasis incipiente en los papeles del cuerpo y del mundo se pueden considerar complementarios de la búsqueda de interpretaciones computacionales y representacionales.⁷¹

Como hemos visto, la propuesta de Clark señala con insistencia que sería una insensatez rechazar por completo los enfoques anteriores desarrollados en la CC, específicamente la noción de representación interna y de computación, presente en ellos. Al respecto dice Clark,

En la búsqueda de este entrelazamiento, sería un insensatez echar por la borda el núcleo de la comprensión científica cognitiva -que tanto esfuerzo ha costado construir- basada en las nociones de representación interna y de computación. La verdadera lección de nuestras investigaciones sobre la cognición corpórea y embebida no es que tengamos éxito sin representación (o, peor aún, sin computación), sino que los tipos de representación interna y de computación que empleamos han sido seleccionados para complementar los complejos entornos sociales y ecológicos en los que debemos actuar. Así, pues será conveniente que no ignoremos ni minimicemos la importancia de estos entornos más amplios.⁷²

Lo que Clark cuestiona de los enfoques anteriores es el carácter neutral que otorgan a la representación (como vimos, para él la representación se caracteriza por ser local y orientada hacia la acción). En su propuesta Clark mantiene los principios más generales de la concepción conexionista, pero los potencia con las siguientes imágenes: una imagen del entorno como recurso activo cuya dinámica intrínseca pueda desempeñar importantes funciones en la resolución de problemas y una imagen del cuerpo como parte del bucle computacional.

Como evaluación final de ambos enfoques, podemos señalar lo siguiente: La propuesta de Varela, tiene implicaciones negativas para el desarrollo de una ciencia de la mente. El carácter *antirrepresentacionista* y *enactivo* de su enfoque, no permite un estudio científico de la cognición, puesto que no permite proporcionar una explicación causal de ésta. Si efectivamente, como señala Varela, el mundo no está predefinido y va surgiendo conjuntamente con la actividad diaria del ser humano, ¿cómo podemos entonces realizar un estudio científico de la cognición? La propuesta de Varela, se presenta finalmente como una crítica que da cuenta de las deficiencias del enfoque cognitivo clásico, pero que no se pronuncia finalmente sobre ¿cómo podríamos en definitiva estudiar la cognición corpórea?

La postura *ecuménica* de Clark en cambio, en su intento de solucionar las deficiencias de los enfoques anteriores (incluyendo el de Varela), tiene la pretensión de reconciliar paradigmas opuestos. Clark, quiere hacer justicia a varios de los enfoques anteriores, sin desarrollar una crítica radical a éstos. El aspecto relevante de su propuesta, radica en que efectivamente logra dar cuenta de modo muy efectivo, de lo que se necesita para un estudio completo de la cognición, reconociendo así los grandes

⁷¹ Clark, p. 200

⁷² *Ibid.*, p. 280

desafíos que se presentan para el futuro de la CC. Al respecto, Clark nos señala enfáticamente que:

La aspiración de la ciencia cognitiva -dilucidar la cognición biológica real- no puede ser compatible con una estrategia basada en la abstracción continua de los anclajes de la percepción y de la acción en el mundo real.⁷³

En consecuencia, la teoría de Clark postula que la actividad de un organismo en su entorno es parte integrante de sus estados mentales. Esto significa que la *interacción cuerpo-mundo* constituye la mente, por tanto, la mente no se encuentra desconectada del cuerpo y sus correspondientes actividades, sino que *la mente se encuentra extendida por el entorno*. Por tal razón, la propuesta individualista del paradigma clásico debe ser reformulada, incorporando la interacción del agente con el mundo real.

Sumado a lo anterior, Clark agrega, que si aceptamos la muerte de la reconstrucción racional (del enfoque cognitivo), obtenemos con ello una especie de vacío conceptual y metodológico, que es menester llenarlo aún cuando existan varias inquietudes y problemas aún no resueltos.

5.3. EVALUACIÓN DEL ENFOQUE CORPORALIZADO DE LA COGNICIÓN

En la medida que la ciencia corpórea y embebida esté bien encaminada, la ciencia cognitiva ya no se podrá permitir los sesgos individualistas y aislacionistas que la caracterizaron en los primeros decenios de su andadura.⁷⁴

La hipótesis corporalizada sostiene que la cognición es un fenómeno complejo que emerge de relaciones causales complejas entre el cerebro, el cuerpo y el ambiente. Clark (1997) nos señala que en la explicación de la cognición, el cerebro, el cuerpo y el ambiente son socios causales con idénticos derechos. La cognición se entiende entonces, como la operación de un agente encarnado y ambientalmente situado, proporcionando una visión particular, que se aleja de la imagen tradicional de la mente (similar a la de un procesador de representaciones).

El mérito fundamental del enfoque corporalizado, radica en lograr colocar en el lugar que se merece la relación existente entre el organismo y su ambiente, señalando de qué manera, nuestra organización cognitiva refleja nuestra participación física en el mundo. La crítica realizada por este enfoque al paradigma clásico (cognitivismo), nos muestra que no es posible explicar la cognición, a partir de un distanciamiento de las bases biológicas y ambientales de esta misma. El modelo cognitivista hace una abstracción de las realizaciones biológicas de las estructuras cognitivas, mostrando con ello un profundo *prejuicio antifibiológico*. En su intento de dar una explicación científica de la conducta inteligente y de la cognición, los cognitivistas introducen prescripciones metodológicas (tales como el individualismo metodológico y solipsismo metodológico), que conllevan dejar fuera los aspectos sociales y ambientales de la cognición.

⁷³ Clark, p. 100

⁷⁴ Clark, p. 280.

La perspectiva corporalizada de la cognición, nos presenta en cambio (como ya expusimos) una mente esencialmente corporeizada (*embodied*) que se encuentra además encajada o encrustada (*embedded*) en el mundo en el que actúa. A pesar de las diferencias que sustentan los enfoques de Clark y Varela, es evidente que ambos incluyen el cuerpo y el ambiente como un recurso activo en la cognición. El variado número de ejemplos presentados tanto por Clark como por Varela como apoyo empírico para ésta hipótesis (como los agentes autónomos de Brooks, entre otros), proporcionan evidencia bastante significativa como testimonios de conductas emergentes en sistemas simples. En consecuencia, podemos señalar que la estrategia de estudiar modelos simples que exhiban características básicas de agentes cognitivos situados, resulta ser una poderosa estrategia para explorar las implicaciones de una perspectiva encarnada.

Cabe señalar, además, que esta hipótesis pone en primer plano la noción de *causación recíproca* (denominada así por Clark), que ocurre entre *agente y mundo*, a través de sus interacciones sensorio-motrices; cuestionando con ello, la visión en términos de input y output, propia de una *supuesta máquina cognitiva aislada*. La idea de "*causación recíproca*" nos indica que lo importante a tener en cuenta no es el agente cognitivo por sí mismo, sino las interacciones que se dan entre éste y el ambiente.

Sin embargo, aún cuando existan considerables aspectos positivos en la hipótesis corporalizada, resulta necesario señalar algunas inquietudes metodológicas y conceptuales que surgen de los planteamientos fundamentales de este enfoque. Un primer problema que presenta esta hipótesis radica en poder determinar ¿cómo estudiar la mente corpórea y embebida? Si efectivamente, como señala Clark, los procesos cognitivos y computacionales entrecruzan los límites de la piel y el cráneo, podemos preguntarnos, si esto implica que los cerebros y los organismos individuales no sean objetos apropiados para el estudio científico.

Frente a estos problemas Clark señala que una metodología adecuada para estudiar la cognición activa y corpórea debe poseer las siguientes características fundamentales:

1. Debe centrarse en el mundo real, teniendo presente el tiempo real. Esto significa que las tareas deben identificarse en función del mundo real (las entradas son cantidades físicas y las salidas son acciones). La conducta debe limitarse a marcos temporales que sean biológicamente realistas.

2. Debe reconocer las soluciones descentralizadas. Esto significa que no se parte del simple supuesto de que una acción coordinada inteligente requiere de una planificación central detallada, sino que con frecuencia pueden surgir acciones globales inteligentes a partir de múltiples interacciones más elementales en las cuales intervienen individuos, componentes y/o el entorno.

3. Debe considerar una visión extendida de la cognición y la computación. Esto significa que los procesos computacionales se pueden extender más allá de la cabeza de un individuo e incluir transformaciones realizadas mediante el empleo de instrumentos externos y que también pueden incorporar las cabezas y los cuerpos de múltiples individuos en situaciones de resolución colectiva de problemas.

Podemos señalar en consecuencia, que existen importantes retos conceptuales y metodológicos para el estudio de la cognición activa y corpórea. Entre ellos, podemos

señalar: *el problema de la tratabilidad*⁷⁵, concerniente a cómo podemos aislar fenómenos que resulten tratables para poder estudiarlos, dado el carácter escurridizo de la cognición hacia su entorno local; y *el problema de la cognición avanzada*, que plantea la incertidumbre sobre si existe o no algún lugar para la planificación central en la cognición avanzada.

Resumiendo los puntos anteriores, podemos expresar la necesidad de introducir nuevas herramientas analíticas y metodológicas en la CC, que dejen atrás los sesgos individualistas y aislacionistas que la caracterizaron en sus primeros años. Esto significa que la CC deberá asumir los nuevos retos metodológicos que han sido señalados por las teorías alternativas y aceptar en definitiva, que las propiedades mentales (representaciones) no son las únicas propiedades causalmente relevantes para una explicación de la cognición. Esto significa además, que la CC deberá aceptar que las nociones tradicionales de representación interna y de computación son inadecuadas para dar cuenta de modo efectivo de la cognición. En consecuencia la noción de representación debe ser rearticulada, asumiendo que ella está doblemente situada: en el cuerpo y en el ambiente. La noción de representación no puede en definitiva concebirse sólo como un mero símbolo en la cabeza, siendo "neutral" respecto de la acción.

⁷⁵ Ver Clark, p. 125

CONCLUSIONES FINALES

Teniendo presente el objetivo fundamental propuesto para el desarrollo de esta investigación, esto es, lograr esclarecer **cuál es el rol explicativo que desempeña el supuesto representacional** en los diversos enfoques existentes en las CC y considerando las objeciones críticas presentadas al final de cada capítulo, podemos establecer las siguientes conclusiones finales.

Los partidarios de la TRCM suscriben un individualismo metodológico. Según éste, aquello que es causal y explicativamente relevante para la psicología son aquellos estados y procesos mentales que supervienen sobre la estructura del cerebro. El cognitivismo clásico postula que los *símbolos físicos* desempeñan un rol causal en la conducta. Como vimos, la teoría de Fodor parte afirmando que efectivamente los estados mentales, tales como *creencias* o *deseos*, de un individuo si juegan un rol causal en su conducta. Fodor postula que existen estados mentales que corresponden a actitudes proposicionales, que poseen propiedades intencionales genuinas que se implican causalmente en la producción de la conducta. Para mostrar de qué modo ocurre esto, su teoría transita desde una PSC (psicología de sentido común) a una psicología computacional. En opinión de Fodor la CC solo debe preocuparse de explicitar los dispositivos formales que implementan formalmente las leyes intencionales de una psicología mentalista. De este modo, el comportamiento inteligente es explicado en la medida que se logren especificar los dispositivos formales que subyacen a los procesos mentales que producen la conducta inteligente. Por consiguiente el contenido de los estados mentales no desempeña ningún rol explicativo o cognitivo, en la explicación de la conducta humana, en este enfoque de la CC. Fodor (1994) es enfático en señalar que la

semántica no es objeto de preocupación de una ciencia cognitiva, pero que debe ser abordada por la filosofía.

El cognitivismo clásico le asigna sólo un rol formal al supuesto representacional, dado que *es la sintaxis lo que orienta la causalidad de la conducta intencional y no la semántica*. Esto significa que aún cuando el supuesto representacional sea un supuesto sustantivo del enfoque clásico, éste queda constreñido por la estrategia del solipsismo metodológico (enfoque sintáctico de la cognición). El computacionalismo ortodoxo de Fodor, le quita legitimidad a la asignación de un rol causal a la semántica de las representaciones mentales, por parte de la investigación en ciencia cognitiva.

Resulta bastante evidente sin embargo, que el talón de Aquiles del paradigma clásico radica en no dar claras luces acerca de la causalidad eficiente del contenido intencional en la explicación del comportamiento. Aún cuando Fodor sea explícito en formular que el contenido semántico de las representaciones mentales no sea un tema que le concierna a la CC, es evidente que dicha propuesta resulta coja sin proporcionar una explicación sobre qué rol cognitivo le cabe al supuesto representacional en ella. No parece muy probable, explicar la cognición humana desde las actitudes proposicionales (creencias, deseos, etc.) sin hacer referencia a los objetos intencionales que ellas representan. De igual modo, no tiene mucho sentido hablar de computación sin considerar las relaciones existentes entre las expresiones simbólicas. La computación es representacional y, que por consiguiente, debiera tener un valor semántico. *No hay computación sin representación.*⁷⁶

Si consideramos el enfoque del conexionismo, vemos que éste asigna un rol distinto a la representación mental, dado que considera que todas las creencias y los deseos se almacenan en forma superpuesta a la red, por lo que la explicación de la conducta se encuentra en la organización de toda la red cognitiva. Toda la red cognitiva es causalmente responsable de una conducta y no un estado físico específico. La representación mental en este sentido se distribuye a lo largo de toda la red, siendo una especie de función (o vector) que señala lo excitadas o inhibidas que están las unidades o nodos de la red que estructura el sistema y que a su vez representa información del entorno. Las unidades de procesamiento se comunican unas con otras, mediante señales numéricas en lugar de señales simbólicas, por lo que la representación mental conexionista es numérica en lugar de ser simbólica como en el cognitivismo. En términos generales podemos decir, que el *patrón de activación* de la red (representación emergente) representa una solución al problema. En consecuencia, es posible decir, que las *representaciones mentales* tienen una estructura semántica (no simbólica) que juega un rol causal y explicativo respecto de la producción de la conducta inteligente, en este enfoque de la CC.

Las teorías alternativas por su parte, nos obligan a abandonar el individualismo metodológico propio del enfoque clásico. Los partidarios de los enfoques alternativos consideran una simplificación excesiva afirmar que las propiedades mentales sean las únicas que desempeñan un rol causal en la explicación de la conducta, dada la complejidad de los factores que intervienen y subyacen a la cognición. Por ello,

⁷⁶ Martínez Velasco, *La ciencia cognitiva. Una Investigación Interdisciplinar*, p.205.

comparativamente con el paradigma clásico, las teorías alternativas requieren de una psicología heterónoma y no individualista. La cognición, según este enfoque está determinada por propiedades que emergen de una causalidad compleja entre factores no sólo mentales.

Es importante señalar, que si consideramos la crítica realizada por Andy Clark al paradigma clásico, vemos que de ésta no se sigue necesariamente la anulación o aniquilación del rol efectivo de la noción de la representación de las CC. Como vimos en la tercera parte de esta investigación (Sección 5.1.), en opinión de Clark, la noción de representación interna sigue desempeñando un papel importante en las teorías actuales, pero resulta necesario refinar dicha noción. Clark considera que las creencias y deseos no causan literalmente nuestro comportamiento, pero que si pueden explicar algunos tipos de comportamiento.

La visión general de Clark sobre la CC, es que esta no debe sólo referirse a las propiedades abstractas de la cognición, sino que debe considerar la cognición del mundo real y en tiempo real. Según Clark la cognición no puede concebirse como un mero símbolo en la cabeza, *neutral* respecto de la acción, sino más bien sobre una base corporalizada. Si efectivamente el cerebro biológico opera con algo que puede describirse como una *representación interna*, gran cantidad de estas serán locales y orientadas hacia la acción, en opinión de Clark (1997). Cabe señalar que dicho enfoque proporciona una base significativa sobre la cual debe constituirse una nueva noción de representación. La representación en consecuencia, debe concebirse como un producto de nuestro actuar (biológico), vinculado con la interacción entre el cuerpo y el mundo.

Evaluando los tres marcos de investigación en su conjunto, podemos señalar las siguientes conclusiones generales sobre las CC y sobre el rol de la representación mental en ella:

1. Si tenemos en cuenta las críticas presentadas por las teorías alternativas (sistemas dinámicos, la robótica situada o la cognición distribuida), vemos que éstas muestran que el marco de investigación propuesto por la arquitectura computacional / representacional no es capaz de integrar los variados hallazgos empíricos proveniente de estas teorías. De igual manera, es posible, además, afirmar que aún cuando la investigación conexionista resulte ser una buena herramienta de modelación y simulación que es compatible con estas teorías, es muy similar en las orientaciones metodológicas generales de la arquitectura representacional/ computacional clásica. Por consiguiente, la crítica desde las teorías alternativas nos señala que es necesario avanzar hacia un nuevo marco de investigación en la CC, con orientaciones metodológicas nuevas.

2. Frente al diagnóstico anterior, cabe señalar que la propuesta de Clark (1997, 2001) ofrece un interesante camino con nuevas propuestas metodológicas que vale la pena considerar. Clark nos invita a avanzar hacia un nuevo marco de investigación que logre subsumir, dando coherencia y unidad, la diversidad de aspectos que ponen en jaque el marco de investigación del paradigma clásico. Para ello Clark nos presenta *siete desiderata*⁷⁷ que una ciencia de la mente debería satisfacer:

⁷⁷ Estos desiderata han sido extraídos de Vallejos, 2008, *Conceptos y Ciencia Cognitiva*, Apéndice II, *Sobre 'Mindware' de Andy Clark*, pp. 142-143.

I. La cognición humana tiene lugar en organismos biológicos sometidos a la evolución; II. La evolución ha operado en dichos organismos al modo del hojalatero y no del ingeniero de manera que las adaptaciones que han sido seleccionadas, han surgido en virtud de los dispositivos biológicos que históricamente esa especie de organismos ha adquirido, explotando las características de su entorno; III. La cognición es corporalizada, es decir, los dispositivos neurales no son los únicos que determinan el surgimiento de la conducta inteligente, sino también dispositivos sensoriales y motores que conectan directamente al organismo con el ambiente; IV. La cognición corporalizada no solamente es dependiente de los dispositivos fisiológicos internos del organismo, sino que se constituye en virtud de las características ambientales, formando así una suerte de nicho ecológico-cognitivo del cual el organismo es parte; V. Para los organismos humanos la cognición es situada, es decir, los factores culturales ejercen una determinación importante sobre ésta; VI. En tanto situada, la cognición también es distribuida, es decir, la coordinación entre los agentes es necesaria para una acción eficaz que conduzca a solucionar los problemas de supervivencia; VII. La coordinación entre los agentes requiere de instrumentos y tecnologías que extiendan las capacidades cognitivas puramente biológicas de éstos, a los que Clark designa con el término *wideware*. El lenguaje sería el *wideware* por excelencia, y el que haría posible las actividades y productos cognitivos superiores tales como la ciencia y el arte.

3. A partir de la propuesta de Clark, podemos decir que en efecto la noción de representación no debe ser eliminada de la CC, pero si reformulada. La representación mental (y la cognición) no puede ser concebida solamente desde una psicología individualista como plantea Fodor, sino desde una psicología heterónoma que de cuenta del ambiente real en el que operan los agentes cognitivos. Esto trae como consecuencia, un cuestionamiento de los argumentos metodológicos presentados por Fodor (individualismo y solipsismo metodológico), como marco adecuado para una ciencia de la mente. Las consideraciones anteriores nos hacen cuestionar la unidad de la CC. Como hemos visto, los tres marcos de investigación presentados, nos muestran que no hay una clara unidad con respecto a cuáles serían los supuestos compartidos por los variados enfoques que existen al interior de la CC.

Finalmente, frente a estas conclusiones generales podemos agregar que aún cuando la propuesta de Clark, pueda parecer sólo un esquema teórico provisional y no una teoría propiamente tal, como postula Guido Vallejos⁷⁸, precaviéndonos de comparar apresuradamente su propuesta, ésta resulta ser bastante lúcida en cuanto a las deficiencias del enfoque clásico. Sin duda es tarea abierta para la CC seguir avanzando en su estudio de la cognición humana, sin desconocer las críticas que se han realizado desde las teorías alternativas.

⁷⁸ Ver Vallejos, 2008, *Conceptos y Ciencia Cognitiva*, Apéndice II, Sobre 'Mindware' de Andy Clark, p. 137.

REFERENCIAS BIBLIOGRÁFICAS

- BECHTEL, W. 1991. *Filosofía de la mente. Una panorámica de la ciencia cognitiva*. Versión castellana de LUIS ML. VALDÉS VILLANUEVA. Madrid, Editorial Tecnos, S. A.
- BECHTEL, W. and G. Graham (eds). 1999. *A Companion to Cognitive Science*. Oxford: Blackwell Publishers.
- BROOKS, *Intelligence without Representation*. En Haugeland (1997), *Mind Design II*, Cap. 15. A Bradford Book. The Mit Press, Cambridge, Massachusetts. London, England.
- CLARK, A. 1997. *Being There: Putting Brain, Body, and World Together Again*. The Mit Press, Cambridge, Massachusetts. Traducción al español: *Estar ahí: cerebro, cuerpo y mundo en la nueva ciencia cognitiva*. Editorial Paidós.
- CLARK, A. *Embodied, situated, and distributed cognition*. In: W. Bechtel and G. Graham, Editors. *A Companion to Cognitive Science*. Oxford: Blackwell Publishers (1999), pp. 506 -517.
- CHURCHLAND, P.M. 1992. *Materia y conciencia: Introducción contemporánea a la filosofía de la mente*. Barcelona, Traducción Margarita N. Mizraji. Editorial Gedisa.
- FODOR, J. A. 1975. *The language of thought*. New York: Thomas Y. Crowell. Traducción al español: *El lenguaje del pensamiento*. Alianza Editorial.
- FODOR, J. A. 1981. *Representations*. Cambridge, Mass.: MIT PRESS.

- FODOR, J. A. 1983 *The modularity of mind*. Cambridge, Mass. : MIT Press. Traducción al español: *La modularidad de la mente*. Ediciones Morata, Madrid.
- FODOR, J.A. 1985. *Fodor's guide to mental representation*. *Mind*, 94: 77-100.
- FODOR, J. A. 1987 *Psychosemantics. The problem of meaning in the philosophy of mind*. Cambridge, Mass. MIT Press. Traducción al español: *Psicosemántica. El problema del significado en la filosofía de la mente*. Editorial Tecnos.
- FODOR, J.A. 1990. *A theory of content and other essays*. Cambridge, Mass.: MIT Press.
- FODOR; J. A.: 1991. *A argument for narrow content*. *The Journal of Philosophy*. 88, 5-26.
- FODOR, J.A. y E. LEPORE. 1992. *Holism: A shopper's guide*. Oxford: Basil Blackwell.
- FODOR, J. A. 1994. *The elm and the expert. Mentalese and its semantics*. Cambridge, MA: MIT Press. Traducción al español: *El olmo y el experto. El reino de la mente y su semántica*. Ediciones Paidós.
- GARDNER, H. 1987. *La nueva ciencia de la mente: historia de la revolución cognitiva*. Barcelona, Ediciones Paidós.
- HAUGELAN, J. 1997. *Mind Design II*. A Bradford Book. The Mit Press, Cambridge, Massachusetts. London, England.
- LOWER Y REY. 1991. *Meaning in Mind: Fodor and his Critics*, Editors Introduction, p. XV. Blackwell. Cambridge, Massachusetts.
- H. NEWELL, A.; SIMON, H. 1976. *La ciencia de la Computación como investigación empírica: Símbolos y Búsqueda*. En: *La filosofía de la inteligencia artificial*. Boden, M. (Comp.) México, FCE, 1994, p. 127.
- RABOSI, E. (compilador). 1995. *Filosofía de la mente y Ciencia cognitiva*. España, Ediciones Paidós Ibérica, S. A.
- SMOLENSKY. *Connectionist Modeling: Naural Computation/Mental Connections*. En Haugeland (1997), *Mind Design II*, cap 9. A Bradford Book. The Mit Press, Cambridge, Massachusetts. London, England. Traducción al español en Rabossi
- VALLEJOS, G. 1990. *Realismo intencional y naturalización del contenido*. *Lenguas Modernas*, 17: 5-34. Universidad de Chile.
- VALLEJOS, G. 1991. *La teoría representacional de la mente y el rol del contenido intencional*. *Lenguas Modernas*, 18: 5-22. Universidad de Chile.
- VALLEJOS, G. 2008 *Conceptos y Ciencia Cognitiva*. Bravo y Allende Editores.
- VAN GELDER, *Dynamics and Gognition*. En Haugeland (1997), *Mind Design II*, Cap. 16. A Bradford Book. The Mit Press, Cambridge, Massachusetts. London, England.
- VARELA, F. 1997. *De cuerpo presente. Las ciencias cognitivas y la experiencia humana*. Barcelona. Gedisa. Segunda edición.
- VARELA, F. 1990. *Conocer. Las ciencias cognitivas: tendencias y perspectivas: Cartografía de las ideas actuales*. Gedisa. Primera edición.
- VON ECKART, BARBARA. 1993. *What is cognitive science?* Cambridge, MA: MIT Press.
- WILSON, R 1995: *Cartesian Psychology and Pshysical Minds: individualism and the*

Sciences of de mind. Cambridge University Press.

ARTÍCULOS ONLINE

BALESTERO, SOLEDAD. *La representación del conocimiento en los sistemas conexionistas*. Psicothema, 1992, vol 4, N° 2 (1992), pp. 343-354.

En: <http://www.psicothema.com/pdf/836.pdf> Accedido el 10 de abril de 2008.

LIZ MANUEL, 1995. *La estructura de las representaciones mentales: Una perspectiva integradora*. Análisis filosófico XV (1995) N° 1/2 En: <http://webpages.ull.es/users/manuliz/articulos/representaciones.pdf> Accedido el 10 de abril de 2008.

MARTÍNEZ VELASCO, JESUS. *La ciencia cognitiva, una investigación interdisciplinar. Etapas de su construcción (I y II)*. Departamento de Filosofía, Universidad de Valladolid.

En: <http://www3.unileon.es/dp/dfh/ctx/1997/Martinez1.doc> Accedido el 25 de marzo de 2008.

ORTUZAR, MARÍA JOSÉ, 2005. *La Representación Mental en las Teorías "Encarnadas" Contemporáneas de la Cognición*. En Cybertesis Universidad de Chile: http://www.cybertesis.cl/sdx/uchile/notice.xsp?id=uchile.2005.ortuzar_m-principal&qid=pcd-

VINCENZO P. LO MONACO. *La cuestión de la Sistemática en el Análisis Computacional de la Mente*. Universidad Central de Venezuela. En: <http://www.bu.edu/wcp/Papers/Mind/MindLomo.htm> Accedido el 25 de abril de 2008.