



La cuestión de la Sistemática en el Análisis Computacional de la Mente

Vincenzo P. Lo Monaco
Universidad Central de Venezuela
sanchezb@camelot.rect.ucv.ve

ABSTRACT: Fodor, Pylyshyn and McLaughlin reject connectionism and argue that connectionists are unable to explain systematicity without implementing a classical architecture. In this paper, I contend that the traditionalist conclusion only seems to follow if they are able to sketch a neutral account of systematicity. But in absence of such an explanation, connectionists bear no special burden in this matter. In support of this view, I set out three specific weaknesses: circularity, epistemological insufficiency, and atomism/reductionism, which affect the classical argument of systematicity. I conclude that there may be other ways to accommodate systematicity in a scientific framework, for instance, in the manner of Davidson's holistic semantics, and indeed connectionist representation is one of these legitimate ways for doing that.

1. INTRODUCCIÓN

Durante la década pasada se ha logrado un considerable avance en el de la actividad computacional como la función de una red conexionista que reemplaza símbolos y reglas de símbolos con patrones de actividad numérica sobre grupos de unidades y patrones de pesos distribuidos en paralelo. De ahí que el conexionismo haya surgido como rival no desestimable de la aproximación que simula la cognición a través de procesos simbólicos, en la medida en que se ha reconocido que la distribución en paralelo de los modelos conexionistas los hace idóneos para ejecutar determinadas tareas -reconocimiento, categorización y analogía- que resulta difícil realizar en procesos simbólicos clásicos [17].

Sin embargo, a pesar del éxito del conexionismo, existen todavía distintas objeciones a las teorías conexionistas, especialmente en referencia a la dimensión filosófica. Lo que algunos filósofos -como Fodor, Pylyshyn y McLaughlin [10][11][17]- encuentran confuso en el conexionismo es la pretensión de atribuir a la distribución en paralelo la capacidad de proporcionar una explicación de la representación cognitiva. La objeción es, para parafrasearla muy esquemáticamente, la siguiente: según la teoría computacional de la mente, los sistemas tienen estados mentales porque implantan las representaciones en códigos y las disponen según una relación dada. En esta teoría, la cognición como actividad es un proceso de operaciones formales que han lugar en representaciones sintácticamente estructuradas. Los modelos conexionistas, en cambio, no están capacitados para proveer ningún estado mental o cognitivo, en razón de la ausencia de una estructura sintáctica que relacione representaciones; "...una teoría cognitiva adecuada desde el punto de vista

empírico debe reconocer no sólo relaciones causales entre los estados representacionales, sino también relaciones de constitución de tipo sintáctico y semántico; por ende, [...] la mente no puede ser, en su estructura general, una red conexionista"[11].

Sostengo que esta afirmación descansa en un error: la conclusión tradicionalista sólo parece seguirse debido a la confusión que ha caracterizado al debate, dado que lo que los filósofos entienden normalmente por 'representación semántica' ha sido manipulado y convertido en algo trivial. Puesto que la teoría computacional clásica es semánticamente referencialista y atomista en el sentido filosófico ordinario, el atomismo se encuentra en serios problemas a la hora de explicar cómo el concepto en la cabeza llega a ser el tipo de cosa que puede representar al mundo. Además, debido a que el conexionismo propone un modo de representar los contenidos semánticos significativamente diferente del análisis composicional de la cognición [16], los conexionistas están legítimamente facultados para cambiar lo que vale como reconocimiento de patrones de constitución sintáctica y semántica..

2. EL DILEMA DE FODOR

Según Fodor, existe un dilema que engendra una paradoja que afecta al conexionismo. El dilema que presenta Fodor es el siguiente:

"...si el conexionismo no puede dar cuenta de la sistematicidad, falla entonces en proporcionar una base adecuada para una teoría de la cognición; pero si su explicación de la sistematicidad requiere de procesos mentales que son sensitivos a la estructura constituyente de las representaciones mentales, entonces la teoría de la cognición que ofrece será, en el mejor de los casos, la implementación de la arquitectura de un modelo 'clásico' (de lenguaje del pensamiento)..."[11].

El dilema surge al explicar la sistematicidad en una teoría de la cognición. Pero la paradoja que afecta al conexionismo tiene implicaciones para explicar los poderes representacionales de una explicación semántica de la cognición. En lo sucesivo deseo examinar estas implicaciones. En particular, quiero explorar los dos cuernos del dilema de manera separada, a fin de mostrar lo que hay de problemático en la aproximación de Fodor y McLaughlin. Ya he mostrado en otra parte [16] que el argumento de la implementación -segundo cuerno del dilema- es incoherente, porque los modelos semánticos conexionistas son holísticos: en razón de la *plasticidad* de su semántica, ellos carecen de una caracterización semántica de la función computada, puesto que la fuerza de la interpretación reside holísticamente en las propiedades globales de la actividad que relaciona patrones y conexiones [16]. En consecuencia podemos afirmar que, dentro de un marco conexionista concebido en términos de la teoría del significado de Davidson, no existe dificultad para explicar intrínsecamente la representación semántica sin implementar una arquitectura clásica. [16]. Pasemos entonces directamente al primer cuerno del dilema de Fodor -i.e., la afirmación de que el conexionismo no puede dar cuenta de la sistematicidad.

3. LA DEFINICION COMPOSICIONAL DE LA SISTEMATICIDAD

A diferencia de los conexionistas, los computacionalistas clásicos hacen un énfasis desmedido en la sistematicidad para explicar la estructura constituyente de los estados representacionales [5]. Sus razones comprenden usualmente el carácter *componencial* de la representación. Ellos arguyen que los estados cognitivos "heredan" sus propiedades semánticas y su intencionalidad *de manos* de las propiedades semánticas de las representaciones mentales, de modo tal que éstas deben tener una estructura interna compuesta de constituyentes simples. Para explicar esta propiedad como la capacidad que poseen los estados mentales de interrelacionar sus contenidos en determinados modos

sistemáticos, ellos proponen la Teoría Representacional de la Mente como una construcción representacional de los estados intencionales en términos de la arquitectura clásica [7][8]. Ésta construye los contenidos semánticos echando mano de un código de un lenguaje simbólico para representar objetos, relaciones, eventos, acciones, etc., y empleando dispositivos "de entrada" para formar representaciones simbólicas, las cuales son entonces relacionadas con una estructura de símbolos en términos de correspondencia biunívoca entre símbolos y contenidos semánticos. Luego, si se toma en cuenta que las manipulaciones simbólicas son regidas por reglas de inferencia válidas, las interacciones de símbolos y contenidos semánticos preservarán los valores veritativos y el marco semántico resultará definido causalmente de un modo tal que el significado de un símbolo consistirá en "...el conjunto de cambios que éste provoca que el sistema efectúe, o como entrada o bien como respuesta a algún estado (interno o externo)". En definitiva, puede decirse que la arquitectura clásica ejecuta sus operaciones sobre representaciones simbólicas a guisa de una sintaxis composicional que sostiene una semántica composicional [2][21].

Ahora bien, ¿satisface la arquitectura clásica el requisito de la sistematicidad? En verdad, una respuesta definitiva a esta cuestión es en la actualidad controversial. Fodor, Pylyshyn y McLaughlin han sugerido al menos tres formas de definir tentativamente la sistematicidad:

(i) la propiedad de tener capacidades cognitivas relacionadas para estados intencionales [11], o

(ii) la propiedad de tener capacidades cognitivas relacionadas para estados intencionales cuya posesión implica la posesión de 'bases constitutivas' en virtud de las cuales quien típicamente posea una de estas capacidades posee también la otra (u otras)[10][17], o

(iii) la posesión de determinadas capacidades sustantivas ("*molar*") cuyas condiciones de satisfacción constituyen la posesión de determinadas capacidades constitutivas y de sus interacciones [17].

Hay una larga tradición en filosofía que intenta articular los lineamientos esenciales de una lógica de la definición. Dado el alcance teórico de la definición clásica de la sistematicidad, me dispongo a comentarla a la luz de aquellos lineamientos, pues sospecho que la definición de marras es o bien trivial o, en el mejor de los casos, contextual. Para explicar por qué el argumento de la sistematicidad levantado contra el conexionismo es demasiado débil, es importante aclarar previamente sendas cuestiones. La primera es ésta: es claro que hay algunas diferencias importantes entre las tres definiciones de la sistematicidad referidas: la definición (i) es simple y básica, mientras que las definiciones (ii) y (iii) añaden los conceptos de 'constitución' y 'satisfacción', progresiva y respectivamente. De ahí que, segunda cuestión, aunque es confundente pensar que las definiciones expresan lo mismo, porque la definición (i) podría ser complementada de muchos otros modos distintos de lo que afirman las definiciones (ii) y (iii), parece razonable asumir que, al menos a los efectos de la crítica al conexionismo, la definición (iii) las incluye a todas y, acorde con los computacionalistas clásicos, podría ser capaz de caracterizar a la sistematicidad de manera cabal y satisfactoria.

4. TRES DEBILIDADES EN LA DEFINICIÓN CLÁSICA DE LA SISTEMATICIDAD

Sin perder de vista estas clarificaciones, puedo ahora regresar a mi hipótesis acerca de la debilidad del argumento clásico de la sistematicidad y ofrecer en su sostén las siguientes consideraciones.

Una primera razón atañe a un requisito básico de toda teoría de la definición, a saber, la no circularidad. Una explicación de la sistematicidad basada en la posesión de determinadas

capacidades substantivas sería circular si la justificación de las bases constitutivas se identificara automáticamente con la posesión de capacidades cognitivas. Considérese el caso de las emisiones lingüísticas, donde la relación de sistematicidad es mediada por los estados cognitivos de hablantes y oyentes. En este ámbito, determinadas capacidades cognitivas podrían ser sistemáticamente relacionadas si, y sólo si, *como cuestión de necesidad nomológica* [17], un parlante de un lenguaje L que comprende una oración L dada, comprende también otra oración L relacionada con aquélla. El problema de cómo es que *necesariamente* un parlante de L puede comprender oraciones L-relacionadas, es suplantado entonces por un problema acerca de las capacidades cognitivas en el "lenguaje del pensamiento". La Teoría Representacional de la Mente proporcionaría un modelo para explicar la sistematicidad por esa vía. He aquí algunos detalles de la paráfrasis: dado un parlante S y un par de L oraciones, P y Q (por ejemplo, "Juan ama a la chica" y "La chica ama a Juan", respectivamente), puesto que *significar (que) P* entraña un estado mental A de S y *significar (que) Q* entraña un estado mental B de S, existe una representación estructurada $RP \leftrightarrow RQ$ tal que S tiene A si, y sólo si, S tiene Q -y a la inversa-, pues se supone que S posee la misma capacidad cognitiva constitutiva que construye al unísono tanto A como B [7][8]. Pero si ésta es la caracterización, entonces la estrategia para explicar la sistematicidad es realmente circular. La sistematicidad y la significatividad de las capacidades cognitivas son explicadas por recurso a la interacción de significados de las oraciones, mientras que la interacción de significados de las oraciones es explicada por recurso a los estados mentales implícitos en la asignación de intencionalidades relacionadas a la misma capacidad cognitiva.

Hay otra gran debilidad, sugerida por el análisis del concepto de sistematicidad de Fodor que lleva a cabo R. Matthews [18]. Según éste, la razón por la cual la sistematicidad resulta supuestamente un argumento intranscendente es que "...no existen explicaciones clásicas de la sistematicidad que merezcan ese nombre, en particular en el constructo de la sistematicidad de Fodor y otros como la propiedad de tener capacidades cognitivas sistemáticamente relacionadas para los estados intencionales"[18]. Aunque se ha sostenido que la sistematicidad puede explicarse en términos funcionales por recurso a la Teoría Representacional de la Mente [14], Matthews insiste en que la sistematicidad es epistemológicamente intranscendente porque los computacionalistas clásicos no están actualmente en condiciones de proporcionar una sintaxis composicional y una psicosemántica naturalista [18]. Esta última aseveración encuentra un sostén más preciso en filósofos como Loewer y Rey [19], pero la idea básica es la siguiente: la mayor parte de la investigación que se realiza en psicosemántica en perspectiva computacional, encuentra serios tropiezos a la hora de mostrar que sus modelos proporcionan algo semejante a una explicación exitosa de la cognición con base en una sintaxis composicional. Matthews establece este punto insistiendo en que lo que los clásicos describen como explicaciones de la sistematicidad son en realidad meras estipulaciones, "...algunas ideas muy generales de cómo podría construirse tal explicación"[18]. Mantengo el mismo punto al insistir no sólo en que los modelos computacionales clásicos nada tienen a nivel conceptual que pueda contar como explicación de las capacidades cognitivas e indique sistematicidad, sino también en que hay algo que permanece normalmente oculto en la discusión generada en torno al debate conexionismo/clasicismo -una discusión en la cual el desafío de Fodor de explicar la racionalidad está cayendo velozmente en una mera confusión-, a saber, la naturaleza problemática de las categorías de "significado" y 'referencia', y esta situación nos plantea a su vez un desafío que trasciende el escenario de una teoría empírica de la sistematicidad tal como Fodor la imagina.

Para que se comprenda a plenitud una tercera razón en contra del argumento de la sistematicidad, necesito ahora referirme a la cuestión de cómo un lenguaje del pensamiento explica lo que es representar para un lenguaje natural en términos de lo que sus oraciones realmente representan. Un modo de hacer esto es ilustrado por Fodor y Lepore [9]

argumentando a favor de una explicación atomista de las representaciones mentales como la única forma de establecer las condiciones de identidad de los contenidos mentales sin describir el estado mental total del organismo. Empero, dado que se explica la verdad de una oración en términos de la referencia de sus partes, la explicación composicional de la representación está necesitada de una explicación reductiva de la referencia de los términos que constituyen una oración, De ahí que el análisis composicional clásico sea atomista en un doble sentido. Es atomista porque explica lo que es para un estado mental representar un objeto por recurso a relaciones causales entre el mundo y la mente, donde los *relata* son análogos mentales de las palabras y partes del mundo. Pero es también atomista en tanto asocia los análogos de las palabras mentales con sus extensiones por el procedimiento de correlacionar, por un lado, palabras mentales y objetos, y por el otro predicados mentales y propiedades. En fin, es reductivo, pues emplea los recursos disponibles para explicar la verdad de una oración en términos de la referencia de sus componentes más simples [2].

No obstante, todo esto no parece haber inquietado en demasía a Fodor, pues en escritos recientes [8][9] no ha tenido empacho alguno en conceder que la explicación composicional de la representación que propone es atomista. En defensa del atomismo semántico, entendido como "...la idea de que lo que se significa es enteramente independiente de lo que se cree..."[8], ha afirmado que ése es la única alternativa al holismo [9].

Por mucho que esta afirmación pueda resultar verdadera, tiende no obstante a ocultar el hecho de que el atomismo tiene dificultades por su propia cuenta. Un rastro de tales dificultades puede encontrarse en la descripción que hace Davidson de cómo dar cuenta del significado en un lenguaje natural en términos suficientes para comprender las emisiones de los hablantes de ese lenguaje [6]. Ahí Davidson concibe el atomismo como 'el método del bloque constructivo', aquel que "...empieza con lo simple y construye hacia arriba...", y lo rechaza porque pretende "...dar una caracterización no lingüística de la referencia, pero no parece haber chances de ello"[6]. Dice que "no parece haber chances" porque está asumiendo, siguiendo a Quine, que "...la totalidad de la evidencia disponible para un oyente no determina una forma única de traducir las palabras de un hombre en las de otro..."[6]. Una vez abrazada la tesis de la 'inescrutabilidad de la referencia', Davidson concluye que la referencia misma "...quedará en el camino. Ella no desempeña una función esencial en la explicación de la relación entre lenguaje y realidad"[6]. Pienso que Davidson está en lo cierto cuando afirma que el atomismo es una postura indefendible, pero creo además que el argumento que él ofrece en su contra se queda corto, en especial si se conecta el atomismo con el *representacionalismo*. De hecho, cuando se lo combina con una aproximación semántica representacional, el atomismo es insostenible a causa de dos importantes consideraciones.

En primer lugar, el atomismo representacional da por sentado que existen los estados mentales (las representaciones), los cuales en algún sentido contienen partes significantes (conceptos) que son establecidas sólo por inferencia lógica a partir de la estructura del lenguaje. Dado que la evidencia necesaria para establecer la existencia de estados mentales simples (conceptos), considerados en el modelo atomista como significados de las palabras, depende de la inferencia a partir del lenguaje, y puesto que todas las inferencias lógicas a partir del lenguaje presuponen únicamente una caracterización completa de los rasgos semánticos de sus partes (las oraciones), aunada a las relaciones formales que rigen las oraciones T de Tarski como vehículos para dar indirectamente un contenido empírico a las relaciones entre nombres -o predicados- y objetos, concluimos que la existencia de representaciones mentales como correlatos psicológicos de oraciones no puede ser establecida con algún grado de certidumbre. La razón de ello es que la evidencia obtenida de las relaciones lingüísticas es sólo y siempre *evidencia lingüística*, en cuyo caso la cuestión de cómo hacer uso del contenido mental que una oración presuntamente entraña

como relación entre una representación mental y un hecho extralingüístico, queda sin respuesta.

Y aquí van nuestra segunda y tercera consideraciones. La segunda es que si se admite que los significados de palabras son 'objetos mentales' o 'conceptos en la cabeza' susceptibles de construcción, de retención y de evocación, en el supuesto de que son inmanentes a los actos intencionales y cognitivos de la mente misma, la postulación de contenidos mentales como "significados correctamente identificables de los palabras" está necesitada de una historia plausible de cómo los valores semánticos de los conceptos simples se combinan unos con otros en la mente para finalmente formar los valores semánticos de las representaciones mentales. La otra consideración es que el ardid de igualar los significados de las palabras (nombres y predicados) y los objetos mentales presupone que las palabras que se usan en el lenguaje están causalmente conectadas con sus análogos mentales por medio de cierta técnica. De hecho, se necesita esta conexión causal para poder mantener que los estados cognitivos son componentes referenciales de las representaciones semánticas introducidas en el marco sintáctico. El primer caso constituye esencialmente un recurso al lenguaje del pensamiento para explicar y validar la diferencia entre tener creencias y tener meramente una lista de conceptos en la mente. El segundo es un recurso a una caracterización reductiva de la referencia de los términos que constituyen una oración en conexión con una esquematización funcional del significado de los conceptos que constituyen una representación. En ambos casos se trata de cuestiones que el atomista representacional tiene dificultad para explicar. La primera se autoaniquila, pues la Teoría Representacional de la Mente está aún lejos de proporcionar un modo de explicar cómo se combinan los valores semánticos de las representaciones mentales sin sencillamente presuponer la combinación misma. La segunda es realmente una cuestión problemática, en vista de que el atomista representacional no dispone de una explicación causal que permita construir una función vicaria que cubra la trayectoria que va de los objetos/valores de verdad a los conceptos/representaciones mentales, y a la inversa. En consecuencia, si en una teoría del lenguaje que proporciona las condiciones bajo las cuales todas sus oraciones representan fielmente el mundo no hay nada que indique la necesidad de recurrir a estados mentales representacionales para fijar la relación lenguaje-mundo, entonces una explicación atomista de la representación puede tornarse del todo irrelevante.

5. CONCLUSIÓN

Si de dar se tratara un veredicto sobre el dilema de Fodor, ésa no sería a estas alturas una tarea demasiado difícil. Irónicamente, en la perspectiva desde la cual hemos enfocado el asunto, el desafío de explicar la sistematicidad que los tradicionalistas proponen a los conexionistas no puede ser acometido por aquéllos, pues los computacionalistas clásicos adolecen de problemas más graves que los que ellos mismos critican en otras formas de explicar la cognición. ¿Qué ocurre entonces con la sistematicidad? Como ya hemos indicado, cabe imaginar otras formas de dar cuenta de la sistematicidad y aún rechazar, no obstante, el dilema de Fodor, pero no ha sido mi propósito explorar esas posibilidades aquí. En efecto, en este trabajo mi interés se ha centrado exclusivamente en un punto específico, a saber: el desafío de Fodor de explicar la sistematicidad no constituye un argumento legítimo contra el conexionismo.

Referencias

[1] W. Bechtel, "Current connectionism", *Minds and Machines*, 3 (1993).

[2] M. Boden (ed.), *The Philosophy of Artificial Intelligence*, Oxford: Oxford University Press, 1990

- [3] M. Boden, "The impact on Philosophy", en *The Simulation of Human Intelligence*, ed. D. Broadbent, Oxford: Blackwell, 1992, pp. 178-197.
- [4] A. Clark, "Special issue: Philosophical issues in connectionist modeling", *Connection Science*, No 4, 1992, pp. 171-173.
- [5] D. E. Bradshaw, "Connectionism and the Specter of Representationalism", en [12].
- [6] D. Davidson, "Reality without reference", en *Reference, Truth and Reality*, ed. M. Platts, 1980, pp. 131-140.
- [7] J. A. Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, MIT Press, 1987.
- [8] J. A. Fodor, *A Theory of Content and Other Essays*, Cambridge, MIT Press, 1990.
- [9] J. A. Fodor y E. Lepore, *Holism: A shopper's guide*, Cambridge, Blackwell, 1992.
- [10] J. A. Fodor y B. P. McLaughlin, "Connectionist and the problem of systematicity: Why Smolensky's solution doesn't work?", *Cognition*, No 35, 1990, pp. 183-204.
- [11] J. A. Fodor y Z. W. Pylyshyn, "Connectionist and the cognitive architecture: A critical analysis", *Cognition*, No 28, 1988, pp. 3-71.
- [12] M. Hogan, "What is Wrong with an Atomistic account of Mental Representation", *Synthese* 100 (1994), pp. 307-327.
- [13] T. Horgan y J. Tienson, "Settling into New Paradigm", *Southern Journal of Philosophy*, 26 (1987), pp. 97-113.
- [14] T. Horgan y J. Tienson (eds.), *Connectionism and the Philosophy of Mind*, Dordrecht, Kluwer, 1991.
- [15] V. P. Lo Monaco, *Necesidad, designación, mundos posibles: La metafísica de la lógica modal*, Trabajo de Ascenso a Profesor Asociado, Caracas: Universidad Central de Venezuela, 1995.
- [16] V. P. Lo Monaco, "The Computational Theory of Mind and Searle's Problem", en N. Callaos (ed.), *Proceedings of ISAS '96*, Orlando, IIS, 1996.
- [17] B.P.MacLaughlin, "The Connectionism/Classicism Battle to Win Souls", *Philosophical Studies*, 71 (1993), pp. 163-190.
- [18] R. J. Matthews, "Three-Concept Monte: Explanation, Implementation and Systematicity", *Synthese*, 101 (1994), pp. 347-63.
- [19] B. Loewer and G. Rey (eds.), *Meaning in Mind: Fodor and Its Critics*, Cambridge, Blackwell, 1991.
- [20] P. Smolensky, G. Legendre y Y. Miyata., *Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition*, Cambridge, M. A.: MIT Press, 1994.
- [21] T. van Gelder, "Compositionality: A Connectionist variation on a classical theme", *Cognitive Sciences*, No 14 (1990), pp. 355-84.