

LOS ORÍGENES DEL PROBLEMA DE LA ARTICULACIÓN DE NIVELES EN EL COGNITIVISMO: NEWELL Y PYLYSHYN

José Antonio de la Rubia Guijarro
I.E.S. "Virgen de la Caridad"
Loja (Granada)

Resumen: Presentamos una exposición del cognitivism en la hipótesis del sistema de símbolos físicos de Allen Newell y el desarrollo posterior de Zenon Pylyshyn. En esos textos pioneros, el cognitivism es expuesto como una síntesis entre monismo y dualismo a partir del establecimiento de tres niveles: nivel físico, nivel sintáctico-funcional y nivel semántico. La tesis consiste en que cada nivel es reductible al inferior pero partiendo de presupuestos explicativos diferentes. Nosotros defenderemos la idea de que dicha tesis es contradictoria porque el nivel semántico es irreductible a una concepción computacional de la mente humana.

Palabras clave: Símbolos físicos, representación, computación, nivel semántico, reduccionismo.

Abstract: We present an exposition of cognitivism in the physical symbols hypothesis by Allen Newell and his later development by Zenon Pylyshyn. In these pioneer texts, cognitivism is exposed as a synthesis between monism and dualism by establishing three levels: physical level, syntactic-functional level and semantic level. The thesis involves that each level is reducible to the lower but starting from different explanatory assumptions. We will defend the idea that such thesis is contradictory because semantic level is irreducible to a computational concept of human mind.

Keywords: Physical symbols, representation, computation, semantic level, reductionism.

El cognitivismo se plantea, en principio, como un paradigma global de investigación cuando no como una auténtica revolución (“la revolución cognitiva” como la denomina Howard Gardner). Este paradigma de investigación se reclama explícitamente como interdisciplinar, integrando filosofía, psicología, inteligencia artificial, lingüística, antropología, neurociencia y, en general, cualquier disciplina que directa o tangencialmente tenga como objeto de estudio el conocimiento, humano o no¹. No vamos a hacer aquí una exposición global del cognitivismo o de su historia, su nacimiento como fruto de diversas teorías o su enfrentamiento con el conductismo, sino a centrarnos en los problemas que intenta solucionar.

Los dos pilares fundamentales del cognitivismo son la *representación* y la *computación*.

“La ciencia cognitiva se basa en la creencia de que es legítimo –más aun, necesario– postular un nivel separado de análisis, al que podría denominarse el ‘nivel de la representación’. El hombre de ciencia que opera en este nivel comercia con entidades representacionales como símbolos, reglas, imágenes –la materia prima de la representación, que encontramos entre lo que afluye y lo que efluye, entre lo que entra en la mente y lo que sale de ella–, y explora la forma en que estas entidades representacionales se amalgaman, transforman o contrastan entre sí. Este nivel es indispensable para explicar toda la variedad de las conductas, acciones y pensamientos humanos.

Al optar por el nivel de la representación, el científico cognitivista considera inadecuadas ciertas maneras tradicionales de dar cuenta del pensar humano. El neurocientífico puede preferir hablar de la célula nerviosa; el historiador o antropólogo, de las influencias culturales; el hombre corriente o el literato, del nivel de la experiencia o fenomenológico. El científico cognitivista no cuestiona la utilidad de estos niveles para diversos propósitos, pero funda su disciplina en la premisa de que, a los fines de la ciencia, la actividad cognitiva humana debe ser descrita en función de símbolos, esquemas, imágenes, ideas y otras formas de representación mental”².

El computacionalismo es fundamental para el cognitivismo, así dice Gardner: “el grado en que una disciplina está próxima a la ciencia cognitiva puede medirse, con seguridad, por el grado en que está ligada a las computadoras. Estas ocupan un lugar central en la esfera de la inteligencia artificial, y sólo unos pocos descontentadizos cuestionan su utilidad como modelo de la cognición humana”³. Representación y computación se integran cuando añadimos conceptos accesorios como “información”, “símbolo” o “contenido”.

¹ Existen diversos diagramas en los que se muestran las relaciones entre las ciencias cognitivas. Quizá la sociología sea la última que se ha incorporado a la corriente. Cf. a este respecto Rafael PARDO AVELLANEDA, “Sociología y ciencia cognitiva”, en *Revista de Occidente* 119 (1991) 151-174. Este número de la revista está dedicado a “la mente y sus representaciones: cognición e inteligencia artificial”.

² Howard GARDNER, *La nueva ciencia de la mente*, Barcelona, Paidós, 1988, pág. 55.

³ *Ibid.*, pp. 57-58.

En el cognitivismo, la mente humana se concibe como un sistema manipulador de símbolos o como una máquina simbólica. No es necesariamente dualista en el sentido de que postule la existencia de una sustancia mental (el nivel representacional está implementado en un soporte físico de la misma forma en que el *software* lo está en el *hardware*) pero sí es un dualismo metodológico, ya que el plano del conocimiento se estudia independientemente de los diversos soportes físicos (cerebros u ordenadores) en los que está implementado. J. García Albea presenta al cognitivismo como el fruto de una “negociación” entre dualismo y mecanicismo en la que ambos hacen sus respectivas concesiones. Así el dualismo renuncia al sustancialismo y al sobrenaturalismo, mientras que el mecanicismo renuncia al reduccionismo fisicalista y admite la posibilidad de la no descripción física de las operaciones de las máquinas⁴. Fue este dualismo metodológico el que permitió superar las restricciones observacionales del conductismo y recuperar los conceptos habituales de la psicología popular.

El concepto de *representación* también permitía enlazar con otro de los conceptos desterrados por el conductismo, el de *intencionalidad*, la característica fundamental de la mente según F. Brentano⁵. Ahora bien, tanto la intencionalidad como la representación, el significado o la información han de ser entendidos siempre computacionalmente⁶.

La hipótesis del sistema de símbolos físicos, de Newell y Simon, fue considerada durante muchos años como la gran aportación teórica de la inteligencia artificial a la comprensión de la mente humana. Ya desde una primera aproximación a dicha hipótesis se produce una cierta tensión entre lo que podríamos denominar su ultraabarcabilidad y su ultrarrestricción. La ultraabarcabilidad procede del hecho de que por “sistema de símbolos físicos” se entiende cualquier sistema que pueda procesar información y, por tanto, manipular símbolos ya que, como dice Simon, se puede implementar un símbolo con cualquier cosa⁷. La ultrarrestricción estriba en que el concepto de “símbolo” es interno y restringido a los sistemas computacionales⁸.

Los símbolos integran la mente en el mundo debido a su carácter físico, pero a la vez *nunca serán descritos físicamente* porque sus implementaciones físicas pueden ser innumerables y diversas. Pero la tensión no acaba ahí ya que, por un lado, los seres humanos son simplemente un caso más de sistema simbólico pero, por otro lado, se deja fuera de la hipótesis todo el ámbito del simbolismo específicamente humano:

⁴ Cf. José E. GARCÍA ALBEA, “La mente como máquina simbólica”, en *Revista de Occidente* 119 (1991) 52-53.

⁵ Cf. Zenon PLYSHYN, *Computación y conocimiento*, Madrid, Debate, 1988, p. 11.

⁶ Cf. *Ibid.*, p. 12.

⁷ Herbert A. SIMON, “Ciencia cognitiva: la más nueva ciencia de lo artificial”, en Donald NORMAN (ed.), *Perspectivas de la ciencia cognitiva*, Barcelona, Paidós, 1987, p. 26.

⁸ Cf. Allen NEWELL, “Sistemas de símbolos físicos”, en Donald NORMAN, *op. cit.*, p. 52.

“Hasta ahora he tenido el cuidado de referirme siempre a un sistema de símbolos *físicos*, con el fin de poner de relieve dos hechos: primero, un sistema de este tipo es realizable en nuestro universo físico; segundo, su concepto de símbolo es *a priori* distinto del concepto de símbolo que ha surgido al describir directamente las actividades humanas lingüísticas, artísticas y sociales”⁹.

Por tanto, se deja al margen de la hipótesis todo simbolismo humano no caracterizable computacionalmente aunque Newell insiste en que los símbolos físicos son los que “utilizamos cada día en nuestra vida”.

La tensión que hemos mencionado es la consecuencia de esa negociación entre dualismo y mecanicismo de la que hablaba García Albea: aunque nunca sean caracterizados físicamente, los símbolos obedecen a las leyes de la física debido a su realización en dispositivos construidos (*engineered systems*)¹⁰. La ultrarrestricción no procede de ninguna virtud teórica sino que responde más bien a una necesidad impuesta por el modelo epistemológico del cognitivismo. En la investigación empírica en ciencia cognitiva, la construcción de programas de ordenador ejerce a la vez el triple papel de hecho, teoría y demostración, sobre todo en el caso de los investigadores provenientes de la inteligencia artificial, normalmente más preocupados en construir sistemas inteligentes que en definir la inteligencia, así dicen Newell y Simon: “computer science is an empirical discipline. We would have called it an experimental science, but like astronomy, economics, and geology, some of its unique forms of observation and experience do not fit a narrow stereotype of the experimental method. Nonetheless they are experiments. Each new machine that is built is an experiment”¹¹. La hipótesis se reduce a símbolos físicos porque son los únicos que se pueden construir computacionalmente.

Centremos nuestra atención en la concepción que ofrece la hipótesis del nivel semántico. En el origen de la formulación del concepto de sistema de símbolos físicos se encuentra ya la polémica entre sintaxis y semántica que ha acompañado siempre a la inteligencia artificial y, en general, a todo el cognitivismo. El concepto de función computable es la base de todo el sintacticismo, un sistema de reglas constructivas que generan transformaciones. Desde la máquina de Turing hasta la gramática generativa de Chomsky, las funciones computables son el instrumento lógico más poderoso para ejemplificar una conducta transformacional, una conducta basada en una serie de reglas entrada-salida. Aunque el cognitivismo pretende postular un nivel de análisis de la mente más realista que el corsé estímulo-respuesta del conductismo,

⁹ Cf. *Ibid.*, p. 58.

¹⁰ Cf. Allen NEWELL y Herbert SIMON, “Computer Science as Empirical Enquiry: Symbols and Search”, in Margaret BODEN (ed.), *The Philosophy of Artificial Intelligence*, Oxford, Oxford University Press 1990, p. 109.

¹¹ *Ibid.*, pp. 105-106.

las instrucciones de la máquina son simplemente modos de dar nombre a una conducta¹². La utilización de un concepto semántico como “símbolo” en un sistema computable conlleva la definición de los términos semánticos en un nivel conductual, así *designación* (el concepto que utiliza Newell como representante de una familia de términos semánticos como *referencia*, *denotación*, *representación*, *significación* etc.) se define de esta forma:

“*Designación*: Una entidad X designa una entidad Y relativa a un proceso P, si, cuando P recibe X como entrada, su conducta depende de Y”¹³.

Otro concepto semántico, *interpretación*, se concibe como “el acto de aceptar como entrada una expresión que designa un proceso y luego ejecutar ese proceso”¹⁴. La utilización de un concepto como “símbolo” produce cierta desorientación a quien intenta comprender la naturaleza del nivel semántico. Un símbolo es un *objeto físico*, una entidad, y se sitúa en el mismo plano que su referente que es, a su vez, otro objeto (Newell no considera relevantes las diferencias entre unos conceptos semánticos y otros). El proceso de *designación* es la asignación de relaciones entre objetos, es decir, instrucciones de transformación de símbolos en sus referentes. *Asignar* un símbolo a su objeto designado por él es uno de los operadores de un sistema de símbolos físicos:

“La función del operador *asignar* consiste en crear una relación de acceso, por tanto, de *designación*, entre un símbolo y una entidad. Por supuesto, se limita a las relaciones de acceso sostenidas por la maquinaria subyacente del SS: entre los símbolos del SS y las expresiones, los roles y los operadores del SS.

Asignar implica varias propiedades importantes:

En cualquier momento, un símbolo designa una única entidad.

Muchos símbolos pueden designar la misma entidad.

Un símbolo puede utilizarse para designar cualquier entidad”¹⁵.

A esta variabilidad posible en la asignación de designaciones Newell la considera una capacidad que convierte a los símbolos en abstracciones¹⁶. En su texto, Newell identifica en realidad los “símbolos que utilizan los seres humanos” con lo que cualquier teoría semiótica denominaría un “ícono”, que no son sino un subgrupo de una variedad inmensa de símbolos que no tienen ningún vínculo con lo representado más allá de la intención de relación que le asignan los agentes racionales en su comunidad transcendental (una bandera, por ejemplo, no representa a un país más allá del acuerdo entre los agentes). El que un símbolo, tal y como lo entiende Newell, pueda ser consi-

¹² V. Allen NEWELL, “Sistemas de símbolos físicos”, p. 66.

¹³ *Ibid.*, p. 75.

¹⁴ *Ibid.*, p. 77.

¹⁵ *Ibid.*, pp. 79-80.

¹⁶ Cf. *Ibid.*, p. 80.

derado algo abstracto, teniendo en cuenta que se trata de un objeto, contrasta vivamente con la concreción que Newell atribuye a los símbolos humanos cuya relación de designación sí es claramente abstracta. Newell hace una utilización excesiva de los conceptos semánticos, comprensible si se tiene en cuenta que una de las pretensiones primigenias de la inteligencia artificial era elevar a los ordenadores a un nivel superior al de meras máquinas numéricas, pero incomprensible si de lo que se trata es de averiguar cómo cualquier sistema, computacional o no, puede ser una representación de *algo externo al sistema*. En la hipótesis del sistema de símbolos físicos, tanto objetos simbólicos como objetos designados por los símbolos pertenecen a la estructura interna del sistema en pie de igualdad porque los objetos designados pueden convertirse a su vez en símbolos de otros objetos. A pesar de que Newell considere a los sistemas simbólicos como sistemas adaptativos, en realidad presenta un marco solipsista en el que no se comprende cómo puede establecerse una relación entre el sistema y el entorno externo a él. El problema del conocimiento consiste en dilucidar cómo puede la mente representar algo que no es la mente. Las apariciones de la realidad en el texto de Newell parecen no dar importancia al que es el principal problema de la representación:

“Dada la universalidad de los sistemas de símbolos, el alcance de esta capacidad de designación es sumamente amplio y apenas ha sido explorado todavía. Repitiendo una observación anterior, el poder de una capacidad designatoria depende enteramente de los procesos simbólicos a los que está acoplada. Si estos procesos son restringidos, es poco lo que el sistema total puede llevar a cabo; si son universales, entonces el sistema total puede ser capaz de hacer todo lo requerido en la simbolización humana.

Esta capacidad simbólica general que se extiende al mundo externo depende de la capacidad de adquirir en la memoria expresiones que registren rasgos del mundo externo. Esta capacidad depende a su vez de los operadores de *entrar y actuar* (...) pero que limitan en cierta forma el acceso al mundo exterior. Estos límites no afectan la capacidad del sistema de símbolos para designar entidades arbitrarias, aunque podrían limitar el grado de utilización de tales capacidades por parte de un sistema dado”¹⁷.

Bajo la forma “expresiones que registren rasgos del mundo externo” se encuentra agazapado todo el problema de la representación. La concepción computacional de la mente no da respuesta al problema de la “apertura” de la mente al mundo ni al de la normatividad de esa apertura, como dice Newell: “la enunciación de la designación no describe las propiedades de la relación, por ejemplo, entre la palabra ‘gato’ y el animal gato”¹⁸. En el texto de Newell, la relación simbólica no se da entre símbolo y mundo externo pero Newell no considera que eso sea un problema¹⁹. Los sistemas de símbolos

¹⁷ *Ibid.*, p. 76.

¹⁸ *Ibid.*, p. 89.

¹⁹ Cf. *Ibid.*, p. 90.

físicos, pese a todo, son la base para definir el concepto de acción inteligente. La hipótesis general se enuncia así:

“Hipótesis del sistema de símbolos físicos: la condición necesaria y suficiente para que un sistema físico exhiba una actividad inteligente general es que sea un sistema de símbolos físicos.

Necesario significa que cualquier sistema físico que exhiba una inteligencia general será un caso particular de un sistema de símbolos físicos.

Suficiente significa que cualquier sistema de símbolos físicos puede organizarse de manera que exhiba una actividad inteligente general.

Acción inteligente general significa lo mismo que la inteligencia percibida en la actividad humana: en situaciones reales se da una conducta adecuada a los fines del sistema y adaptada a las exigencias del medio, dentro de algunas limitaciones físicas²⁰.

¿Cómo entiende entonces Newell la representación? Para empezar, Newell distingue entre “representación” y “conocimiento”. Aunque dice que los dos conceptos tienen un papel claro dentro del concepto del sistema de símbolos físicos, admite que las teorías formales de ambos conceptos son caóticas. Representación es simplemente otro término para referirnos a una estructura que designa y se define como sigue:

“X representa Y, si X designa aspectos de Y; esto es, si existen procesos simbólicos que pueden tomar a X como entrada y comportarse como si tuviera acceso a algunos aspectos de Y”.

La representación y lo representado no han de ser isomórficos, basta con que compartan algunos aspectos computacionalmente relevantes²¹. Amparado en el requisito de universalidad, no se especifican lo que denomina “estructuras intermedias” que trasladan las funciones constructoras del objeto representado, “distal” en sus palabras, dentro del sistema que lo representa. El problema de cómo algo puede ser una representación y el problema normativo relativo a la validez o invalidez de las representaciones permanece intacto. Newell adopta la posición logicista en el sentido de que considera a la lógica formal como el vehículo adecuado para la representación, no obstante renuncia a una formulación teórica del concepto²².

La idea de “conocimiento” es más restringida que la de “representación”. Tenemos conocimiento cuando accedemos a la conducta producida por la utilización de representaciones. El conocimiento es un “conjunto abstracto de todas las expresiones derivadas posibles”:

“Esta noción, que corresponde al conjunto de todas las implicaciones de un conjunto de proposiciones, tiene su historia en la filosofía, como candidata

²⁰ *Ibid.*, p. 91.

²¹ Cf. *Ibid.*, pp. 98-99.

²² Cf. *Ibid.*, p. 99.

para la definición de conocimiento. Ha parecido insatisfactoria porque difícilmente podría decirse que una persona conoce todas las implicaciones de un conjunto de proposiciones; sin embargo, su posición dentro de una teoría explícita del procesamiento presenta una situación del todo diferente. En este caso, tener conocimiento se distingue de tenerlo disponible para un uso particular cualquiera, especialmente cuando ello depende de los detalles del sistema de procesamiento. Esta formulación, de hecho, se corresponde con el uso concreto de este término en inteligencia artificial, donde es necesario hablar de lo que está disponible en una estructura de datos y que podría extraerse mediante un procesamiento mayor, o diferente²³.

Newell es consciente de todas las objeciones posibles a la hipótesis del sistema de símbolos físicos, aunque centra esas objeciones en los problemas relativos al diseño de los sistemas, problemas como el de la limitación de memoria, la necesidad de operar en tiempo real o la explosión combinatoria si se usan funciones no computables, etc. La hipótesis es simplemente una base de estudio desde la que se podría ascender hasta poder caracterizar computacionalmente, incluso con otra tecnología, el resto de restricciones de la mente, empezando por la racionalidad e incluso el variado panorama de las experiencias fenoménicas²⁴, basta con continuar el estudio de los “niveles de sistema” desde el nivel macromolecular del cerebro hasta el nivel lógico de la “arquitectura”. No obstante, la utilización que Newell y Simon hicieron de los conceptos semánticos en la formulación de la hipótesis tuvo que ser ampliada estableciendo un nuevo nivel semántico, el “nivel del conocimiento”, al que Newell dedicó un artículo posterior²⁵.

La difícil integración de sintaxis y semántica se muestra más claramente en la obra de un computacionalista ortodoxo como Zenon Pylyshyn. La articulación de los distintos niveles en el cognitivismo (el nivel físico, el nivel funcional y el nivel semántico) ha pretendido ofrecer una explicación psicológica, “real y empírica”, del nivel intencional a partir del nivel funcional; ése es el caso de Pylyshyn. El puente hacia la intencionalidad se construye desde una teoría representacional, computacional, de la mente. El punto de partida de Pylyshyn es, en principio, similar al de Daniel Dennett en su clásico trabajo “Sistemas intencionales”²⁶. Los niveles físico, funcional e intencional de Dennett podrían equipararse en Pylyshyn con un nivel de arquitectura funcional relativo a la naturaleza del mecanismo, un nivel de estructuras de símbolos relativo a la naturaleza de los códigos y, por último, un nivel relativo al contenido semántico. La equivalencia entre los niveles de cada autor no es exacta. El punto de partida de Pylyshyn es el de la explicación del comportamiento. Cada nivel explicativo utiliza un vocabulario específico: cuando dis-

²³ *Ibid.*, p. 100.

²⁴ Cf. *Ibid.*, p. 102.

²⁵ Cf. Allen NEWELL, “The Knowledge Level”, en *Artificial Intelligence*, 18 (1982) 87-127.

²⁶ Cf. Daniel DENNETT, “Sistemas intencionales”, *Cuadernos de Crítica*, U.N.A.M., México 1985.

tintas ciencias explican los mismos fenómenos crean marcos semánticos diferentes que parten de sus propios presupuestos. Pues bien, aunque sea materialista, la psicología cognitiva parte de la base de que es necesario un vocabulario cognitivo:

“El principal mensaje (...) es que, a pesar de que muchos de nosotros adoptamos una postura materialista, defendiendo la causalidad física y otros razonables principios del mismo estilo (como aquel según el cual el cerebro tiene una importancia crucial en la producción de la conducta), el problema de la explicación en psicología nos obliga a adoptar un tipo determinado de taxonomía de la conducta. Más concretamente, puesto que las explicaciones intentan captar generalizaciones, y puesto que la utilización de vocabularios distintos pone de manifiesto generalizaciones distintas, a la hora de revelar algunas pautas fundamentales de la conducta inteligente, en gran medida racional, de determinadas criaturas y, podríamos añadir, de determinados artefactos, nos vemos obligados a recurrir a lo que yo llamo un *vocabulario cognitivo*. También descubriremos, y quizás esto no sea sorprendente, que el vocabulario cognitivo es muy parecido al utilizado por lo que, sin duda, constituye el esquema predictivo de mayor éxito de entre los disponibles en el ámbito de la conducta humana, la psicología popular”²⁷.

Tanto cognitivismo como psicología popular crean un vocabulario cognitivo para contestar a preguntas como *por qué, qué y cómo* un determinado comportamiento se produce²⁸. Pylyshyn analiza el ejemplo de un hombre que pide ayuda por teléfono ante la visión de un accidente. Ninguna posible explicación que se situara en el marco de las ciencias físicas o biológicas respondería a esas preguntas²⁹. El vocabulario cognitivo, empero, nos permite captar generalizaciones entre situaciones que, siendo físicamente distintas, se puede considerar que responde a una misma estructura de reglas, objetivos y sistema de creencias. Aunque la situación física sea idéntica tenemos que cambiar la explicación si el peatón *cre*e que el accidente es la escena del rodaje de una película, por ejemplo. Ahora bien, la psicología popular se empieza a separar del cognitivismo cuando Pylyshyn considera que las regularidades de un sistema físico que consideramos “sometido a reglas” sólo son posibles partiendo de la base de una hipótesis computacionalista³⁰, debido al hecho de que no queremos crear teorías informales sobre el comportamiento de los sujetos sino que “deseamos poder relacionar dichas generalizaciones con mecanismos posibles, físicamente realizables, y es en este punto en el que la computación puede desempeñar un importante papel de enlace”³¹.

²⁷ Zenon PLYSHYN, *op. cit.*, p. 25.

²⁸ Cf. *Ibid.*, p. 27.

²⁹ Cf. *Ibid.*, pp. 27-28.

³⁰ Cf. *Ibid.*, p. 30.

³¹ *Ibid.*, pp. 35-36.

Los términos cognitivos son intencionales y su utilización parte del supuesto de la racionalidad de la conducta y del sistema de deseos-creencias orientado a objetivos, un sistema que es independiente de los estímulos. No podemos hacer una descripción de un sistema racional en términos físicos o biológicos³². Ahora bien, mientras que en la situación intencional de Dennett la intencionalidad es un atributo que surge de una *actitud* epistemológica del observador de la conducta, en Pylyshyn la intencionalidad, en el nivel del conocimiento o semántico, pretende construirse computacionalmente a través del concepto de “representación” porque la computación opera sobre representaciones³³. El nivel semántico o representacional es un “refinamiento” del nivel simbólico, el cual a su vez es un refinamiento del nivel físico o neurológico³⁴. Hay diferencias de matiz entre Newell y Pylyshyn, no obstante, ya que en Newell la representación era una mera relación de designación entre símbolos y significado: la computación opera sobre símbolos físicos. Pylyshyn, en cambio, distingue el nivel representacional del símbolo y lo equipara al “nivel del conocimiento” de Newell. El nivel representacional es el nivel semántico aunque las representaciones sean físicas debido a que son la base de la computación

¿Qué entiende Pylyshyn por *representación* y cómo se construye el nivel semántico-intencional a partir de las representaciones? Pylyshyn comienza advirtiéndonos de que él no quiere resolver “el rompecabezas del significado” sino que su pretensión sólo es “describir cómo se considera implícitamente la idea del contenido semántico de las representaciones dentro del campo de la ciencia cognitiva, y estudiar por qué esta consideración resulta justificada”³⁵. En primer lugar, la apelación a las representaciones es necesaria porque, mientras que las explicaciones físicas o biológicas operan sobre dominios cerrados, las explicaciones cognitivas hacen referencia a entidades o propiedades que no forman parte de su descripción de estado. En el caso de un ordenador, su estado puede describirse como una serie de entidades y propiedades que interaccionan en un medio físico o electromagnético. Si un ordenador está ejecutando un programa que aplica la función “sucesor” y escribe la secuencia 1, 2, 3, 4, etc. está haciendo referencia a los “números”, que son entidades que no forman parte de la descripción del estado del ordenador. Esto también sucede en el caso de una persona que tiene creencias, como la persona que llama por teléfono para pedir ayuda; ni números ni creencias son entradas del sistema, no están conectadas causalmente con el sistema puesto que los sistemas físicos o biológicos son cerrados y ni un número abstracto ni la expectativa de ayuda (que es un estado de cosas no real sino posible) pertenecen al dominio del sistema. Desde una posición

³² Cf. *Ibid.*, p. 45.

³³ Cf. *Ibid.*, p. 46.

³⁴ Cf. *Ibid.*, p. 48.

³⁵ *Ibid.*, p. 47.

materialista, que sostiene que en el mundo sólo hay entidades físicas que se rigen por leyes físicas, no se puede admitir que operen en el mundo entidades extrínsecas como números o creencias, por eso hay que postular la existencia de “representaciones internas”, códigos simbólicos físicamente instanciados que sí actúan causalmente³⁶.

El paso de una situación de diseño a lo que Pylyshyn denomina un “genuino caso epistémico” estriba en que el segundo necesita un *contenido*, no basta con ofrecer una mera descripción funcional: si no se apela a contenidos no se pueden expresar las generalizaciones adecuadas para explicar conductas³⁷. El problema, por tanto, estriba en saber cuál es la relación entre un estado del nivel representacional, semántico, y un estado funcional. Un funcionalista ortodoxo identificaría estados funcionales y estados representacionales y, por tanto, puede construir una teoría sintáctica de la mente. Pylyshyn admite que diferencias en el estado funcional pueden provocar diferencias en el contenido del estado representacional, pero no a la inversa. Aunque Pylyshyn considera al contenido representacional como un nivel de descripción, sosteniendo que hemos de postular estados representacionales distintos de los estados funcionales de la misma manera que postulamos estados funcionales como distintos de los estados físicos³⁸, nunca desvincula a estados representacionales de estados funcionales porque están intermediados por el nivel simbólico: son las expresiones simbólicas las que *codifican* los contenidos³⁹. El estado representacional está conectado con el funcional a través del nivel simbólico, que consiste en computaciones de símbolos físicos con papel causal. Llegados a este punto, nos encontramos ante el mismo problema que veíamos en Newell: ¿cómo surge el contenido semántico en el sistema de símbolos físicos? La respuesta de Pylyshyn es que “estar en un cierto estado representacional consiste en tener en alguna parte de la memoria una expresión simbólica determinada”. Pylyshyn atribuye el origen de esta idea a Jerry Fodor⁴⁰ aunque parece contradecirse con su idea de que el contenido semántico responde a generalizaciones necesarias (“generalizaciones del nivel semántico”) como distintas de las generalizaciones sintáctico-funcionales (“generalizaciones del nivel simbólico”). Él mismo admite implícitamente que quizá establecemos el contenido semántico porque no tenemos nada mejor, esto es, que su establecimiento responde a una estrategia heurística forzada por el inconveniente de no haber sido capaces de establecer las adecuadas generalizaciones funcionales⁴¹.

³⁶ Cf. *Ibid.*, p. 50.

³⁷ Cf. *Ibid.*, pp. 52-53.

³⁸ Cf. *Ibid.*, p. 58.

³⁹ Cf. *Ibid.*, pp. 54-56.

⁴⁰ Cf. *Ibid.*, p. 54, nota 4.

⁴¹ Cf. *Ibid.*, p. 55, nota 5.

Según Pylyshyn, hay dos razones para negarnos a eliminar la adscripción de contenido:

“a) Puede haber expresiones sinónimas, esto es, conjuntos de códigos con el mismo contenido semántico. Podríamos distinguir estos códigos funcionalmente, pero no semánticamente. Por ejemplo, uno de ellos podría tener una estructura más compleja que otro y, por tanto, podría tardar más en procesarse, o en producir determinados errores sistemáticos, que podríamos predecir de acuerdo con ciertos principios funcionales pertenecientes a lo que más abajo denominamos “nivel simbólico”. (Sigue sin resolverse el problema de distinguir las diferencias de contenido de las diferencias estructurales. Según Putnam (...) puede tratarse de uno de esos rompecabezas que es mejor dejar que se resuelvan por sí solos).

b) La mera posesión de una cierta expresión simbólica que codifique un contenido semántico es insuficiente, por sí sola, para producir una conducta. Falta por contar una parte del asunto. Se trata de aquella por la que se han interesado mucho más los informáticos que los filósofos o los psicólogos (...). Lo que nos hace falta es contar con un conjunto de mecanismos que hagan que el sistema ejecute (*run*) o, como a veces se dice, “interprete” los símbolos. Son estos mecanismos los que hacen que las expresiones simbólicas no agoten eso que informalmente hemos estado llamando el “estado funcional” del sistema”⁴².

Desafortunadamente para el cognitivismo, no sólo no se ha resuelto el rompecabezas de distinguir diferencias de contenido de diferencias estructurales, sino que también constituye un duro rompecabezas saber *cómo pueden adscribirse los mismos contenidos a estados funcionales distintos y seguir sosteniendo a la vez que el nivel semántico es producido por las interpretaciones (que no son sino transformaciones) de los símbolos físicos*. En el Hilary Putnam de *Representación y realidad*⁴³ (un texto posterior a *Computación y conocimiento*) ésto supuso una ruptura entre el nivel semántico y el funcional (y una renuncia al funcionalismo) precisamente porque se partía de una filosofía antirreduccionista (lo que también se suponía que era en un principio el funcionalismo). Al no querer renunciar al funcionalismo (que es el soporte filosófico del cognitivismo), Pylyshyn se tiene que enfrentar al problema de la articulación de niveles y a la contradicción básica del cognitivismo: necesita un nivel semántico como nivel de descripción pero utiliza una metodología puramente sintáctica. Quien lea detenidamente el escrito de Pylyshyn comprenderá por qué el problema del contenido semántico es una dura trampa de la que no puede escapar (quizá porque se ha negado desde el principio a una indagación filosófica sobre el significado). Dice que “el que en un nivel haya generalizaciones válidas que no podemos expresar en un nivel inferior constituye el principal argumento a favor de la existencia de niveles”⁴⁴. Sin embargo, al

⁴² *Ibid.*, p. pp. 55-56.

⁴³ Cf. Hilary PUTNAM, *Representación y realidad*, Barcelona, Gedisa, 1990.

⁴⁴ Zenon PLYSHYN, *op. cit.*, p. 61.

especificar el nivel simbólico como un codificador de contenidos se nos está diciendo que el nivel semántico *sí* es reducible ¿por qué seguir entonces hablando de un nivel semántico o un nivel de conocimiento? Es evidente que Pylyshyn sigue hablando del nivel semántico porque no tiene más remedio (no puede hacer generalizaciones válidas sobre la conducta sin apelar al contenido) pero no porque lo considere una postura asumible por principio: si el sistema de símbolos físicos pudiera *realmente* codificar el contenido no añadiría un nuevo nivel al físico-biológico y al funcional. A veces, de hecho, Pylyshyn parece sostener que no cree que todos los procesos sean computacionales:

“Aunque no tenemos una teoría detallada de la arquitectura funcional humana, ya hemos visto que los principios arquitectónicos sólo pueden restringir, de un modo determinado, las transiciones entre estados representacionales, y específicamente las determinadas formas estructurales que son *independientes del contenido* de estos estados representacionales. Por ejemplo, una limitación del tamaño de la memoria de trabajo puede restringir el número de elementos que podemos interrelacionar simultáneamente, o bien es posible que estos elementos requieran ciertas recodificaciones que conduzcan a determinados compromisos entre tiempo y espacio de memoria (...). Es evidente que, por muy minuciosas que sean las restricciones, la arquitectura funcional permite la aplicación de cualquier regla que regule las transacciones entre estados funcionales, siempre que estas reglas se ajusten a determinados principios estructurales o funcionales de carácter muy general. No obstante, cabe pensar que haya otro tipo de restricciones sometidas a principios que no pueden ser especificados en términos de propiedades de mecanismos de procesamiento de la información, que en algunos casos restringen las reglas, incluyendo algunos de los casos más evidentes de lo que podríamos llamar ‘procesos cognitivos de alto nivel’”⁴⁵.

La crítica, sin embargo, podría seguir en sentido descendente: si el nivel funcional realmente codificara el nivel semántico no haríamos generalizaciones semánticas y prescindiríamos del engorroso problema del contenido, pero si el nivel físico-biológico nos permitiera hacer generalizaciones válidas sobre la conducta también prescindiríamos del nivel funcional. Que el sistema de símbolos físicos codifique el contenido es un *desideratum* del cognitivismo que no se cumplirá mientras los cognitivistas tengan que seguir postulando un nivel semántico para hacer generalizaciones válidas sobre la conducta. Pero ¿por qué no decir simplemente que el nivel semántico aparece cuando presuponemos la racionalidad y hacer en consecuencia un análisis filosófico de esa presuposición (adoptar, en definitiva, una posición trascendental)? La vía filosófica está cerrada para un cognitivista por lo siguiente: si partimos del supuesto de racionalidad entendemos que el sistema está sometido a reglas y sólo podemos caracterizar científicamente un sistema de reglas

⁴⁵ Cf. *Ibid.*, p. 63.

si aplicamos el tratamiento sintáctico-funcional del sistema de símbolos físicos. Como decía Pylyshyn, el cognitivista no quiere aplicar un tratamiento “informal” a la racionalidad, sino que tiene que demostrar que ese sistema de reglas es físicamente realizable y no se puede realizar de otra forma más que computacionalmente, es decir, aplicando a la racionalidad el tratamiento lógico-sintáctico del sistema de símbolos físicos. Sin embargo, el problema aparece una vez más desde el instante en que Pylyshyn reconoce que la racionalidad opera sobre el contenido de un estado, no sobre su función, y que la aparición de este contenido exige un tipo de nivel de explicación distinto del funcional⁴⁶. Hay una continua oscilación arriba-abajo: cuando suponemos la racionalidad aparece un nuevo nivel de descripción, surge el contenido, necesitamos un análisis que supere al nivel funcional y ascendemos. Por el contrario, cuando queremos codificar el contenido de las representaciones y caracterizar el sistema de reglas que atribuimos a la racionalidad aplicamos el tratamiento computacional del sistema de símbolos físicos y descendemos un nivel. Por último, todos los niveles están conectados⁴⁷. Al considerar el conocimiento y la conducta humanas como una jerarquía de niveles, cada uno con su propio vocabulario y sus propios presupuestos, la transición entre niveles es lógicamente imposible. Si el nivel funcional codificara el contenido no hablaríamos de un nivel semántico, y si el nivel físico-biológico caracterizara la transición entre estados no hablaríamos de un nivel funcional. El hecho de que cada nivel tenga sus propios presupuestos y vocabularios demuestra que el verdadero problema del análisis de la mente no es simplemente un problema de articulación de niveles sino que afecta a todo el proyecto de naturalización, en el cognitismo y en otras tendencias filosóficas y científicas.

⁴⁶ Cf. *Ibid.*, pp. 59-60.

⁴⁷ Cf. *Ibid.*, p. 65.